Functional and structural analysis of cell-free produced transporters and G-protein coupled receptors: Development of new techniques for the fast and efficient production of integral membrane proteins

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich
Biochemie, Chemie und Pharmazie (FB 14)
der Johann Wolfgang Goethe – Universität
in Frankfurt am Main

von
Christian Klammt
aus Frankfurt am Main

Frankfurt 2006
(D 30)

vom Fachbereich Biochemie, Chemie und Pharmazie (FB 14)
der Johann Wolfgang Goethe – Universität als Dissertation angenommen

Dekan:	Prof. Dr. Harald Schwalbe
Gutachter:	Prof. Dr. Volker Dötsch
	Prof. Dr. Clemens Glaubitz

Datum der Disputation:

Danksagung

Ein herzliches Dankeschön gilt Dr. Frank Bernhard für die intensive und einzigartige Betreuung und Förderung meiner Arbeit. Durch seine ständige Hilfsbereitschaft und tatkräftige Unterstützung ist er sowohl wissenschaftlich als auch menschlich ein Vorbild. Ich danke Ihm sehr für die Einblicke in seine wissenschaftliche Denkweise und für seine Ideenvielfalt.

Meinem Doktor-Vater Prof. Dr. Volker Dötsch möchte ich herzlich für seine Förderung und Unterstützung danken. Sein Engagement und seine Menschlichkeit haben mich schwer beeindruckt. Volker wir schaffen die Kanzelwand unter 4 Minuten!

Herrn Prof. Dr. Clemens Glaubitz danke ich herzlich für seine Unterstützung und sein offenes Ohr für jegliche Problemstellungen. Vielen Dank für die Übernahmen des Zweitgutachtens!

Ich danke Prof. Dr. Heinz Rüterjans für seine stetige Unterstützung im Studium und während meiner Diplom- und Doktorarbeit, für seine Förderung und für die interessanten Gespräche. Ihm verdanken wir maßgeblich die Existenz und die Ausstattung unseres NMR Zentrums.

I like to thank Dr. Vladimir A. Shirokov from the Institute of Protein Research RAS in Pushchino for his help in establishing the individual cell-free expression system. I benefited from his great knowledge about cell-free expression systems.

Ein herzliches Dankeschön an alle Mitglieder der Arbeitsgruppen Dötsch und Glaubitz für die schöne gemeinsame Zeit, für die produktive Zusammenarbeit, für eine freundschaftliche Atmosphäre und für all die gemeinsam verbrachten Stunden. Herrn Dr. Frank Löhr (Murph) danke ich für die Aufnahme der NMR Experimente und gute Zusammenarbeit, Sigrid Oguzer-Fachinger für ihre wunderbare Organisation und Birgit Schäfer für ihre Hilfe im Labor, bei DNA Präparationen und für ihr stets offenes Ohr. Besonders danke ich auch Daniel Schwarz, Birgit Schneider, Sina Reckel, Florian Durst, Dr. Wesley McGinn-Straub, Tobias Weber, Mehdi Talebzadeh, Ines Lehner und Daniel Basting für ihre Freundschaft, für interessante Gespräche und für ihre Unterstützung bei der Arbeit und im privaten Bereich. Ich freue mich auf Euren Besuch in Kalifornien!

Bei meinen Eltern und meiner ganzen Familie bedanke ich mich für ihre immerwährende Unterstützung während meiner gesamten Ausbildung und darüber hinaus. Außerdem möchte ich meinen Freunden und meiner Freundin Miriam für ihre Zuneigung, ihr Verständnis, ihr Vertrauen und ihre Liebe danken.

Contents

1. Summary

Integral membrane proteins (IMPs) account for 20-40% of all open reading frames in fully sequenced genomes and they are target of approximately 60% of all modern drugs. So far, cellular expression systems are often very insufficient for the high-level production of IMPs. Toxic effects, instability or formation of inclusion bodies are frequently observed effects that prevent the synthesis of sufficient amounts of functional protein. I have successfully established an individual cell-free (CF) expression system to overcome these IMP synthesis difficulties.

The CF system was established in two different expression modes. If no hydrophobic compartment is provided, the IMPs precipitate in the reaction mixture. Interestingly, these insoluble proteins are found to differ from inclusion bodies as they readily solubilize in mild detergents and the bacterial small multi drug transporter EmrE, expressed in the insoluble mode was shown to reconstitute into liposomes in an active form. Alternatively, IMPs can be synthesized in a soluble way by supplementing the CF system with detergents. A comprehensive overview of 24 commonly used detergents was provided by analyzing their impact on the CF system as well as their ability to keep three structurally very different proteins in solution. The class of long chain polyoxyethylene-alkyl-ethers turned out to be most suitable for soluble expression of α-helical EmrE, the bacterial β-barrel type nucleoside transporter Tsx and the porcine vasopressin receptor type 2, resulting in several mg of protein per mL of reaction mixture.

So far IMPs have almost completely been excluded from solution nuclear magnetic resonance (NMR) analyses. I could demonstrate that CF expression enables efficient isotopic labeling of IMPs for NMR analysis and further facilitates selective labeling strategies with combinations of ^{13}C and ^{15}N enriched amino acids that have not been feasible before.

Four different G-protein coupled receptors (GPCRs) were successfully CF expressed in preparative scale and for the human endothelin B receptor (ETB), ligand binding ability was observed. A series of truncated ETB derivatives containing nested terminal deletions have been CF produced and functionally characterized. The core area essential for Endothelin-1 binding as well as a central region responsible for ETB oligomer formation was confined to a 39 amino acid fragment including the proposed transmembrane segment 1. The binding constant (K_D) of ETB was determined to 6 nM for circular ET-1 by SPR and 29 nM for linear ET-1 by TIRFS. This data indicate a large potential of the established individual CF expression system for functional IMP synthesis.

2. Zusammenfassung

Funktionelle und strukturelle Analyse von zellfrei produzierten Transportern und G-Protein gekoppelten Rezeptoren: Entwicklung neuer Techniken zur schnellen und effizienten Produktion von integralen Membranproteinen

Etwa 20-40% aller Gene in Organismen codieren für integrale Membranproteine (IMPs) und es wird angenommen, dass über 60% aller Arzneimittel direkt gegen sie gerichtet sind. Diese enorme Vielfalt und Wichtigkeit steht in einem großen Widerspruch zu dem derzeitigen Wissensstand über diese Proteinklasse. Über Erfolg und Misserfolg bei funktionellen und vor allem strukturellen Untersuchungen entscheidet vorerst die Herstellung des Zielproteins. Meist aber ist die Präparation ausreichender Proteinmengen äußerst problematisch. Die Überproduktion von IMPs in zellulären Expressionssystemen wird hierbei durch die Einlagerung in die Zellmembran und den dadurch bedingten toxischen Einfluss stark behindert. So ist es nicht verwunderlich, dass bis heute nur weniger als 100 Strukturen von Membranproteinen bestimmt werden konnten, wobei der Erfolg für diese wenigen Beispiele hauptsächlich durch ihre natürliche Häufigkeit begründet ist [1-3].

Ziel dieser Arbeit war ein individuelles zellfreies Expressionsverfahren zu etablieren, mit Hilfe dessen die Herstellung und somit funktionelle und strukturelle Untersuchung von bisher problematischen IMPs ermöglicht wird. Wie in dieser Arbeit gezeigt wird, konnten auf diesem zellunabhängigen Wege verschiedenste IMPs in einem präparativen Maßstab und in einem sehr kurzen Zeitrum von 12 bis 24 Stunden hergestellt werden.

Die beschriebene zellfreie Proteinsynthese basiert auf einem optimierten Zellextrakt aus *E. coli*, dem alle nötigen Bestandteile, die eine gekoppelte Transkription/Translation erlauben, zugegeben werden, wie z.B. DNA, T7-RNA Polymerase, tRNA, Aminosäuren und Puffer. Ein hocheffizientes Energiesystem, basierend auf Phosphoenolpyruvat und Acetylphosphat in Kombination mit den zugehörigen Kinasen gewährleistet die benötigte Regeneration von Nukleotidtriphosphaten. Die Reaktion selbst wird in einem so genannten „continuous exchange cell-free“ CECF System [4, 5] angesetzt, in dem eine Dialysemembran die Reaktionslösung von einer Versorgungslösung trennt, die einen ständigen Zustrom von Substraten ermöglicht. Als Reaktionsmenge wurden hierbei im analytischen Maßstab 70 µl sowie im präparativen Ansatz 1000 µl verwendet. Im Gegensatz zu konventionellen zellulären Expressionssystemen, bei denen die Proteinsynthese innerhalb der Zelle abgeschottet durch die Zellmembran stattfindet, gewährleisten zellfreie Systeme eine vollständig offene Zugänglichkeit. Dies erlaubt die direkte Kontrolle der Reaktionsbedingungen zu jeder Zeit und ermöglicht erstmals eine genaue Anpassung der Reaktionsbedingungen auf das Zielprotein. So

können Komponenten, wie z.B. Liganden, Detergenzien oder Protease Inhibitoren, die eine Proteinsynthese fördern und stabilisieren direkt zugegeben werden. Damit stellt die zellfreie Proteinsynthese eine viel versprechende Alternative für sonst zellulär nur unzureichend exprimierte Proteine dar, wie z.B. modifizierte Proteine, zytotoxische Proteine, instabile Proteine oder Disulfid-Brücken besitzende Proteine. So ist es z. B. möglich, funktionelle Antikörper herzustellen [6] und bis zu neun Disulfid-Brücken enthaltende Proteine zellfrei zu synthetisieren [7].

Für die zellfreie Expression wurden zwei Verfahren zur Herstellung von IMPs entwickelt. Zunächst konnte gezeigt werden, dass IMPs aufgrund fehlender hydrophober Bereiche im zellfreien System unlöslich exprimiert werden. Interessanter Weise verhalten sich diese Präzipitate völlig anders als Inclusion Bodies, wie man sie aus herkömmlichen *E. coli* Expressionen kennt, da sie sich bereits in milden Detergenzien lösen lassen und somit keine umfangreichen Rückfaltungsprotokolle benötigt werden. Das Detergenz 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LMPG) wurde als generelles Mittel zum Resolubilisieren von IMPs identifiziert. Dass auf diesem neuen Wege nicht nur große Mengen sondern auch aktive IMPs hergestellt werden zeigt ein spezifischer Transport-Test des bakteriellen α-helikalen Small Multi Drug Resistenz (SMR) Transporters EmrE. EmrE, zellfrei hergestellt und in *E. coli* Lipide rekonstituiert, zeigte eindeutig spezifischen Substrattransport.

Eine weitere Möglichkeit zur zellfreien Herstellung von IMPs ist die Expression in Anwesenheit von Detergenzien und somit in löslicher Form. Bei dieser zweiten Methode war jedoch anfangs zu erwarten, dass einige Detergenzien die zellfreien Expressionsbedingungen stören oder aber das IMPs nicht ausreichend solubilisieren können. Um eine umfangreiche Analyse zu gewährleisten, wurden 24 der am häufigsten für IMPs verwendeten Detergenzien sowohl auf ihre Wechselwirkung mit den Reaktionskomponenten hin untersucht, als auch auf ihre Fähigkeit, IMPs löslich zu exprimieren. Zudem wurden drei strukturell sehr unterschiedliche Proteine, EmrE, das bakterielle aus β-Faltblättern bestehende Tsx, als auch der G-Protein gekoppelte Rezeptor (GPCR) Vasopressin 2 (V2R) auf zellfrei lösliche Expression in Gegenwart dieser Detergenzien hin untersucht. Etwa ein Drittel der untersuchten Detergenzien hatten innerhalb der verwendeten Konzentrationen so gut wie keinen Effekt auf die Löslichkeit der drei IMPs. Mindestens ein Detergenz von jeder analysierten Gruppe und insgesamt zehn der Detergenzien ermöglichten jedoch für mindestens eines der Proteine die lösliche Herstellung von präparativen Mengen, d.h. von mehr als 0,5 mg pro ml Reaktionslösung. Außergewöhnlich gute Ergebnisse erzielte jedoch die Gruppe von langkettigen Polyoxyethylen-Alkyl-Ethern, Brij35, Brij58, Brij78 und Brij98, die für

alle drei Proteine die lösliche Expression von mehr als 1 mg pro ml Reaktionslösung ermöglichte.

Eine Vielzahl von verschiedenen IMPs konnte bisher zellfrei hergestellt werden. Neben Proteinen, die in dieser Arbeit beschriebenen werden, konnten auch andere Arbeitsgruppen die zellfreie Synthese von IMPs zeigen. Verschiedene SMR Transporter, wie EmrE [8-11], Hsmr, YdgF, Yfbw [12], Psmr [10], und TBsmr [10, 12, 13] wurden in analytischen bis präparativen Mengen erhalten. Zudem konnte die funktionelle Synthese verschiedener G-Protein gekoppelter Rezeptoren in Fusion mit Thioredoxin [14] als auch die funktionelle Synthese des mechanosensitiven Ionenkanals MscL [15] gezeigt werden. Diese Arbeit demonstriert neben der präparativen zellfreien Synthese der SMR Transporter EmrE, SugE und TBsmr die Synthese von wesentlich größeren IMPs. Der 36 kDa und 10 Transmembran Segmente (TMS) enthaltende Metall-Ionen Transporter TehA und der 6 TMS umfassende Aminosäure-Exporter Yfik, welche in *E. coli* Zellen keinerlei Expression zeigten, konnten zellfrei sogar in Mengen von 3 mg pro ml Reaktionslösung hergestellt werden.

Die absolute Kontrolle über die zellfreien Reaktionszusammensetzungen erlauben einzigartige Möglichkeiten für das gezielte Markieren von überproduzierten Proteinen wie sie speziell für die Kernmagnetische Resonanz (NMR) Spektroskopie von Nöten sind. Die zellfreie Proteinsynthese ermöglicht das einfache Ersetzen von unmarkierten Aminosäuren mit markierten Derivaten, welche ohne jegliches Aminosäure-Scrambling und ohne Expressionseinbußen zellfrei eingebaut werden [16-22]. Die Markierung von IMPs mit ^{2}H, ^{13}C und ^{15}N Isotopen war jedoch bisher, aufgrund extremer Probleme in der Proteinherstellung, fast undenkbar. Ein Problem, welches sich speziell für die Lösungs-NMR von IMPs ergibt ist die Erfordernis von Detergenz Lösungen. IMPs eingebettet in Mizellen, ergeben eine enorme Gesamtgröße, welche sich in deutlichen Linienverbreiterungen und somit Auflösungseinbußen im NMR Spektrum widerspiegeln. Weiter sind viele IMPs hauptsächlich aus α-helikalen Sekundärstruktur-Elementen aufgebaut, die nur eine geringe chemische Verschiebungsverteilung aufweisen.

Diese Arbeit zeigt, dass größere IMPs erstmals auch mit Lösungs-NMR untersucht werden können. Dies wird durch die Synergie von außergewöhnlichen zellfreien Expressionsraten für IMPs in Verbindung mit einzigartigen Markierungsmöglichkeiten und der Verwendung ultra hoher NMR Feldstärken (800 und 900 MHz), welche höhere Sensitivität und Auflösung gewährleisten, ermöglicht. Die Verwendung von LMPG, das als effizientes Detergenz zum Lösen von zellfreien IMPs Präzipitaten gefunden wurde

und dem sehr gute Eigenschaften für die Lösungs-NMR zugesprochen werden [23], ermöglicht sogar eine NMR-Probenzubereitung innerhalb 24 Stunden.

Am Anfang jeder NMR Analyse steht die Zuordnung der einzelnen Signale eines Spektrums. Als Modellbeispiel für eine strukturelle Untersuchung mit Hilfe von Lösungs-NMR kombiniert mit zellfreier IMPs Synthese wurde das 24 kDa und 7 TMS enthaltende c-terminal verkürzte TehA Protein (ΔTehA) ausgewählt. ΔTehA werden gleiche Aktivitätseigenschaften wie dem Gesamtprotein zugesprochen [24]. Obwohl das [^{15}N,^{1}H]-TROSY-HSQC Spektrum von ΔTehA für den Beginn einer Zuordnung der Aminosäure-Hauptkette Proton-Resonanzen eine ausreichende Dispersion besitzt, existiert eine starke Überlagerung von Signalen in bestimmten Bereichen des Spektrums. Für die Zuordnung von IMPs bedurfte es jedoch der Entwicklung neuer Strategien, da die Hürde der Signalüberlappung selbst durch Aminosäure spezifische Markierungen nicht genommen werden konnte.

Eine Lösung brachte die gleichzeitige selektive Markierung mit einigen ^{15}N und anderen ^{13}C angereicherten Aminosäuren. Die Verwendung von [^{15}N,^{1}H]-TROSY-HSQC und [^{15}N,^{1}H]-TROSY-HN(CO) Spektren erlaubte nun die Identifizierung von ^{15}N markierten Aminosäuren, denen eine ^{13}C angereicherte Aminosäure vorausging [25]. Die Verwendung dieses so genannten kombinatorischen Markierens ermöglicht somit die eindeutige Zuordnung von einzelnen Aminosäuren, die als Anhaltspunkte für weitere sequenzielle Zuordnungen verwendet werden konnten. Das kombinatorische zellfreie Markieren von IMPs bildet somit einen Kernpunkt für die Zuordnung dieser Proteine, wenngleich dieses Markieren auch nahezu exklusiv durch zellfreie Synthese möglich ist. In zellulären Systemen wäre dieses Markierungsschema fast undenkbar, da hier für jede erdenklichen Kombinationen doppelt auxotrophe Mutanten benötigt würden.

Aufgrund einer limitierten Verteilung der chemischen Verschiebung in α-helikalen IMPs können Aminosäure-Seitenketten-Protonen nur erschwert zugeordnet werden. Somit ist eine schlechte Basis für klassische ^{1}H-^{1}H-NOE basierte Abstandsermittlungen gegeben, die für eine dreidimensionale Strukturbestimmung unabkömmlich sind. Eine weitere Methode zur Abstandbestimmung ist die Verwendung von paramagnetischen Markierungen, welche in paramagnetisch relaxationsverstärkungs (PRE) Experimenten eingesetzt werden [26, 27]. Mit Hilfe dieser Technik, angewendet auf das verkürzte TehA Protein, war es letztlich doch möglich, genügend Abstandsinformationen zu erhalten, um eine vorläufige dreidimensionale Struktur des 7 TMS enthaltenden IMPs aus 325 Abstandsbedingungen zu errechnen.

Innerhalb der Familie von IMPs bilden die GPCR die größte Gruppe von Zell-Oberflächen-Rezeptoren und über 1000 Gene des menschlichen Genoms kodieren für GPCRs [28]. Fast alle Signalmechanismen im menschlichen Körper werden durch GPCRs kontrolliert und so ist es nicht verwunderlich, dass über 60% aller modernen Medikamente mutmaßlich gegen GPCRs gerichtet sind [29]. Wie allgemein für IMPs gültig, existiert nur ein beschränktes Wissen über strukturelle und funktionelle Eigenschaften von GPCRs, hauptsächlich aufgrund der sehr schwierigen Proteinpräparation in zellulären Systemen. Verschiedene GPCRs, V2R aus Mensch und Schwein, der Corticotropin Releasing Factor Rezeptor (CRF) aus Ratte und humaner Endothelin B Rezeptor (ETB) konnten erfolgreich in präparativen Mengen mit Hilfe des individuellen zellfreien Systems ohne die Verwendung von großen Fusionsproteinen exprimiert werden. Der 12 Aminosäuren kleine c-terminale T7-Tag zeigte sich hierbei ausreichend für die zellfreie Expression von 3-6 mg per ml Reaktionslösung.

Der humane ETB wurde in dieser Arbeit genauer untersucht. Unter Verwendung des fluoreszenz markierten Liganden Endothelin-1 (ET-1) konnte in Koelutions-Analysen gezeigt werden, dass nur für löslich in Anwesenheit von Detergenz synthetisiertes ETB eine Ligandenbindung erreicht wird und diese sogar in Abhängigkeit zu der verwendeten Detergenz steht. Unlöslich zellfrei hergestellter und in LMPG resulubilisierter Rezeptor zeigte keinerlei Ligandenbindung, wobei über 50% des in Brij78 hergestellten Rezeptors den Liganden ET-1 binden. Surface Plasmon Resonanz (SPR) und totale interne Reflektions-Fluoreszenz-Spektroskopie (TIRFS) Untersuchungen ergaben Bindungskonstanten (K_D) von 6 nM für zirkuläres ET-1 und 29 nM für lineares ET-1. Diese Ergebnisse stehen im Einklang mit bereits publizierten Werten, wenn auch diese Interaktionen erstmals in Mizellen gemessen wurden.

Bislang ist nahezu ungeklärt, wie und wo die ET-1 Bindung an ETB stattfindet. Zur Lokalisierung der Ligandenbindung wurden verschiedenste verkürzte ETB Fragmente kloniert und zellfrei exprimiert. Die Untersuchung mit fluoreszenz markiertem ET-1 in Koelutions-Analysen, als auch Pull-Down Experimente mit biotinyliertem ET-1 auf Streptavidin Säulen wurde herangezogen. Nur Fragmente, die das erste TMS enthalten, zeigten Ligandenbindung. Die Region der Ligandenbindung konnte so auf 39 Aminosäuren eingeschränkt werden. SPR Messungen ergaben zudem K_D Werte für diese Konstrukte, die nur drei- bis fünffach schwächer als für den gesamten Rezeptor sind.

Seit kurzem ist bekannt, dass ETB Homodimere ausbildet [30]. Es ist jedoch ungeklärt, welche Bereiche des Proteins an dieser Dimerisierung beteiligt sind. In dieser Arbeit beschriebene Pull-down Experimente von Streptavidin markiertem gesamt ETB und

Histidin markierten ETB Fragmenten identifizieren eindeutig das erste TMS als maßgeblich beteiligt an der Dimerisierung. Für diese Untersuchungen wurden die verschiedenen ETB-Derivate sowohl einzeln zellfrei synthetisiert, aufgereinigt und äquimolar zusammengegeben als auch ko-exprimiert und auf Interaktionen hin untersucht. Die erhaltenen Daten werden durch Single Partikel Analysen, die deutlich die Ausbildung von Dimeren in Brij78 Mizellen zeigen, unterstützt. Beide ETB Monomere berühren sich in verschiedenen gemittelten Klassen eindeutig an einer Seite des Proteins, mutmaßlich der ersten TMS. Diese enge Kolokalisation mit der Liganden-Bindungs-Domäne könnte durch eine spekulative dimermodulierte Ligandenbindung in ETB begründet sein.

Letztlich wurden alle C-terminal verkürzten ETB Fragmente für Lösungs-NMR Untersuchungen zellfrei vollständig ^{15}N markiert. Während das gesamte Protein und größere Fragmente erwartungsgemäß mindere Spektrenqualität zeigten, sticht das Fragment ETB_{93} (TMS 1-3) mit einem relativ gut aufgelösten [^{15}N,^{1}H]-TROSY-HSQC Spektrum hervor. Dieses Ergebnis verdeutlicht erstmals, dass selbst relativ große GPCRs, bzw. deren verkürzte aktive Derivate in Zukunft mit Lösungs-NMR untersucht werden könnten.

IMPs und vor allem GPCRs waren bislang von vielen biochemischen und strukturbiologischen Methoden aufgrund extremer Probleme in der Proteinpräparation nahezu ausgeschlossen. Diese Arbeit zeigt die erfolgreiche Etablierung eines individuellen zellfreien Systems zur präparativen und funktionellen IMPs Herstellung. Durch zellfreie Expression erreichbare IMPs-Mengen dieser begehrten Proteine öffnen nun neue Tore für funktionelle und strukturelle Untersuchungen.

3. Introduction

3.1. Overproduction of integral membrane proteins (IMPs)

Integral membrane proteins (IMPs) are embedded into cellular membranes by multiple hydrophobic transmembrane segments (TMSs). They control numerous essential functions like transport activities, energy regeneration, signal perception and communication of the cell with its environment (Fig. 1).

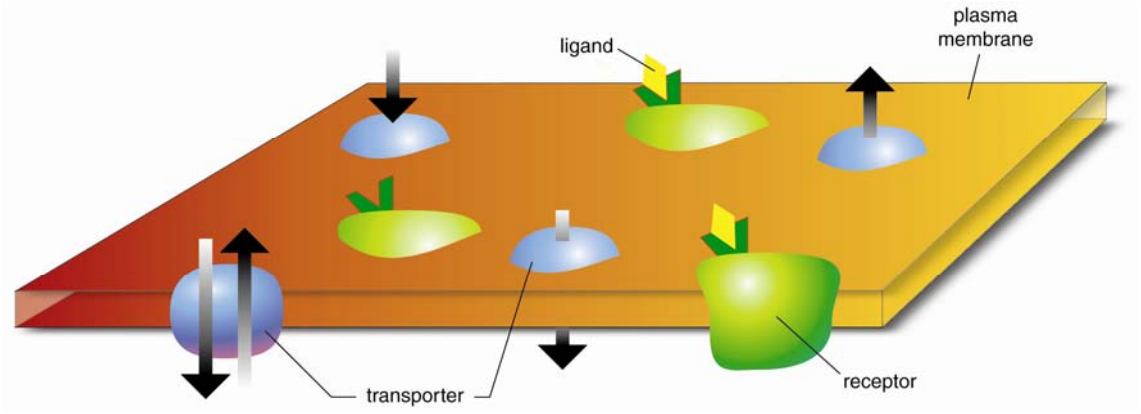


Fig. 1 Integral membrane proteins inside the plasma membrane.
Integral membrane proteins (IMPs) are crossing the plasma membrane allowing the transport of solutes and information. These IMPs can either be transporters, transporting in different directions or receptors allowing signal transduction from the outside to the inside of a cell. (Modified after Forschung Frankfurt 3/2005: Mit neuer Methode lassen sich Membranproteins "knacken". Erstmals können ausreichende Mengen begehrter Proteine produziert werden – Auszeichnung für Christian Klammt (P 12)).

IMPs account for 20-40% of all open reading frames in fully sequenced genomes, and in bacteria half of the IMPs are estimated to function as transporters. The topology of IMPs is generally dominated by α-helical structures while typical β-barrel arrangements are prevalent in IMPs inserted into the outer membrane of Gram-negative bacteria. As IMPs play key roles in a variety of global human diseases, many pharmaceutical studies focus currently on IMPs and they provide thus an estimate of approx. 60% of all modern drug targets.

A basic prerequisite for directed drug design as well as for the understanding of biological functions is the knowledge of the three-dimensional structure of a protein. Relatively high amounts of protein in the range of several 100 mg are often needed for structural approaches by either X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. With few exceptions where natural abundances of a protein offers the opportunity of high yield preparations, like in case of rhodopsin [2], aquaporin [3]

or a Ca^{2+}-ATPase [1], heterologous cellular expression systems of prokaryotic or eukaryotic origin have to be used in order to produce sufficient amounts of protein. Unfortunately, the hydrophobic nature of IMPs and their distinct localization in cellular membranes does often cause tremendous problems upon their synthesis in conventional expression systems. Blocking of cellular protein targeting systems, complex formation with cellular membrane proteins or disintegration of membranes are frequent effects that result in toxicity to living cells leading to low expression rates. These tremendous difficulties obtained in IMP expression might explain why less than 100 different IMP structures are known compared to 38882 overall structures currently posted in the protein data bank.

For successful overproduction the response of host cells to IMP expression and the mechanism of membrane proteins insertion are crucial and have to be understood [31]. Mammalian IMPs frequently cause large difficulties in bacterial systems and require an insect or mammalian cell host for activity or high level expression [32]. Using semliki forest virus vectors for IMP expression more than 50 G-protein coupled receptors (GPCRs) and several ion channels have been expressed at high levels [33]. E. g. the human endothelin B receptor (ETB) was expressed with 20 pmol per mg of membrane protein in baby hamster kidney (BHK) and simian fibroblasts (CV-1 cells) transformed by SV40 (COS) cells [34] or the β2 adrenergic receptor that has been expressed in BHK and COS cells with 5 mg per liter of cultured cells [35]. The baculovirus-insect cell system further constitutes a versatile tool for the maximal production of GPCRs [36], thus ETB could be produced in 100 pmol/ mg of total IMPs [37]. The comparison of seven different heterologous protein expression systems for the production of the serotonin transporter revealed functional expression in the baculovirus system and in three different mammalian cell lines, but not in *E. coli* or *Pichia pastoris* with nonfunctional expression levels of 2-3 mg per liter of cell culture [38]. A detailed review of heterologous GPCR expression is provided by Sarramegna and co-workers [39] whereas Lundstrom compares the GPCR productivity in different expression systems [40]. A remarkable expression result was obtained for the neurotensin receptor. A 200-liter scale *E. coli* expression provided 90 mg of functional receptor fusion protein, corresponding to approx. 0.5 mg per liter *E. coli* cell culture in cells [41] that further will have to be processed by purification and solubilization protocols.

3.2. Cell-free expression techniques

Cell-free (CF) protein synthesis is an attractive and promising alternative to the conventional technologies for protein production using bacterial or eukaryotic cell cultures. In contrast to *in vivo* gene expression methods where protein synthesis is carried out in the cellular context surrounded by cell walls and membranes, CF protein synthesis provides a completely open system. This allows direct control and access to the reaction at any time. Compounds that improve protein production or stabilize recombinant proteins, e. g. ligands, inhibitors, detergents or protease inhibitors, can easily be added without considering side-effects on the cellular metabolism or transport problems through the cell membrane. Most cellular functions with the exception of transcription and translation need not to be maintained during CF protein expression. In principal, applications can therefore be extended to proteins or conditions that would not be tolerated by living organisms. While *in vivo* protein production is often limited by the formation of inclusion bodies [42, 43] or protein instability caused by intracellular proteases [44, 45], the CF system offers new possibilities for the synthesis of complex proteins (Table 1).

Table 1 General pros and cons of cell-free protein expression.

Advantages	Disadvantages
Completely open system and easy access → customized reaction conditions for each individual target protein	Intensive optimization of reaction conditions is often required
Efficient incorporation of labeled, modified or unnatural amino acids	High quality control of reaction components essential
Direct translation of PCR products → high-throughput screening	Relatively high costs
Production of toxic proteins	Preparation of CF extracts necessary
Expression of proteins requiring cofactors	Complexity of system requires experience
Easy addition of chaperones or PDIs	
Miniaturization (e. g. 50 µl reactions) → high throughput	
Working without living organisms → no growth restrictions, no template instability	
Fast isolation of products	

Protein folding and stability can be promoted by direct addition of chaperones or specific cofactors. In contrast, overexpression of chaperones *in vivo* can result in cell filamentation, detrimental for the viability of *E. coli* cells and protein expression [46]. To assist the folding of proteins, chaperones like the GroEL/ ES system [47] or DnaK and DnaJ [6] had been added to the reaction mixture. Usually, lower temperatures will lead to higher yields of recombinant protein in the absence of chaperones, but not necessarily in their presence [48]. Reaction parameters such as pH, redox potential and ionic strength can be modified without harmful effects on the growth and viability of cells. This new opportunity of *in vitro* gene expression allows therefore full control and high flexibility of conditions. It offers new potentials to solve problems associated with cytotoxicity, proteolytic degradation, improper folding or aggregation of synthesized proteins.

3.2.1. Extract sources and preparation

CF expression systems are classified according to the origin of their extract. In principle, functional *in vitro* systems might be prepared from any cell type, but many factors contribute to the protein production efficiency. The most common CF reactions are based on extracts made from *E. coli*, wheat germ or rabbit reticulocyte lysates (Table 2). *E. coli* extracts consist of the so-called S30 supernatant fraction, named after the soluble fraction when centrifuged at 30000 g, containing endogenous ribosomes, enzymes like acetate kinase and factors necessary for transcription and translation, ARSases and tRNA. Endogenous mRNA is removed from the ribosomes during preincubation of the crude cell extract in a “run-off” step and destroyed by endogenous ribonucleases [49]. Another way to deploy bacterial extracts is described in the Gold-Schweiger system [50, 51]. Ribosomes are added to the supernatant fraction of a S100 extract especially purified from endogenous amino acids and nucleic acids by ion-exchange chromatography. This system provides a very low background due to endogenous synthesis and better-controlled conditions at the expense of more complicated preparations.

The *E. coli* system functions well in a temperature range of 24-38°C with an optimum at 30°C [4, 49, 52]. Wheat germ extract possesses a low level of endogenously expressed messengers and therefore can be directly used for expression of exogenous templates. In the wheat germ system, the optimum temperature is in the range 20-27°C [53, 54], but can be increased to up to 32°C for higher expression of some templates [55]. Reticulocyte extract is prepared by directly lysing blood cells of anemic rabbits; this increases the number of proerythrocytes of reticulocytes that are subsequently treated

with micrococcal Ca^{2+}-dependent RNase to remove endogenous mRNA [56]. With regard to the reaction temperature, it should be noted that apart from any effect on the enzymatic process of transcription/ translation and mRNA degradation, the temperature affects the folding of the synthesized protein.

Table 2 Overview of bacterial and eukaryotic cell-free expression systems.

Bacterial systems		**Eukaryotic systems**	
Escherichia coli		Wheat germ	Rabbit reticulocyte (Rr)
S30 extract preparation	**S100 extract preparation**	**Wheat germ extract**	**Rr lysates**
Extract preparation: Supernatant fraction at 30,000 g centrifugation of *E. coli* extract pre-incubated to detract endogenous DNA and mRNA	Extract preparation: Supernatant fraction at 100,000 g centrifugation deprived of all nucleic acids, e.g. by DEAE cellulose treatment	Extract preparation: Directly used for expression of endogenous or exogenous templates	Extract preparation: Treated with micrococcal Ca2+ dependent RNase
Advantages: High efficiency, high translation rate, easy extract preparation, optimized protocols for various applications available, preparative system	Advantages: Deprived of all nucleic acids, ribosome preparation	Advantages: Low nuclease activities, long life-times up to several days, mRNA as template possible, preparative eukaryotic system	Advantages: eukaryotic system
Disadvantages: High variations, shorter life-times, high rate of degradation of genetic messages, proteins, energy	Disadvantages: No preparative system	Disadvantages: Difficult, low efficient extract preparation, complex system, lower translation rate.	Disadvantages: Very difficult extract preparation, very complex non preparative system
References: [4, 49, 52, 57, 58]	References: [50, 51]	References: [53-55]	References: [52, 56, 59, 60]

With regard to the preparative CF expression of proteins, only *E. coli* and wheat germ extracts have been used. Shimizu et al. developed a CF translation system reconstructed from purified poly(His)-tagged translation factors [61]. Their system, termed the "protein synthesis using recombinant elements" (PURE) system, contains 32 individually purified components with high specific activity, allowing efficient protein production. An advantage of the PURE system, apart from the absence of inhibitory substances such as nucleases, proteases and enzymes that hydrolyze nucleoside triphosphates, is a new possibility of synthesized protein purification by removing tagged protein factors with affinity chromatography.

One limitation of the CF system is the degradation of exogenous added mRNA. Various RNase activities present in cell extracts usually restrict the lifetime of mRNA, and subsequently the efficiency of protein synthesis is inhibited. This problem could be solved by the periodical reintroduction of messengers into the reaction mixture in a simple translation system [62]. Coupled transcription-translation systems, where mRNA

is continuously synthesized from DNA templates added to the reaction mixture, can be advantageous to translation systems containing presynthesized messengers. Direct transcription in the reaction mixture is achieved by exogenous phage T7-RNA polymerase in bacterial systems [63] or in eukaryotic systems [52, 59]. To avoid rapid messenger degradation, especially in E. coli cell-free systems, partially ribonuclease-depleted extracts or RNase inhibitors are used. A good choice of template DNA in the prokaryotic system is circular plasmid DNA. In the wheat germ system with lower nuclease activities, both plasmid DNAs [64] and linear PCR fragments function well [65-67], but more recently even bacterial systems have been optimized for efficient expression of PCR products [68, 69].

3.2.2. Components and design of cell-free expression

Components

In CF systems all components involved in gene expression and protein synthesis are combined in a reaction mixture (Fig.2). A significant problem in CF protein synthesis is the high consumption of biochemical energy provided by ATP and GTP. Creatine phosphate concomitant with creatine kinase is usually used for ATP and GTP regeneration in eukaryotic CF systems, whereas the combination of phosphoenol pyruvate (PEP) and pyruvate kinase, acetyl phosphate and acetate kinase or a combination of both has been applied for bacterial *in vitro* protein synthesis. The acetyl phosphate energy system may have the advantage that the ATP level is maintained twice as long as in the presence of PEP. Since acetate kinase is present at sufficient levels in bacterial extracts, it does not need to be added exogenously in the *E. coli* system [70]. Studies of the biochemical energy levels in different CF systems observed a high rate of triphosphate hydrolysis to mono- and diphosphates during protein synthesis in wheat germ extracts [71] as well as in *E. coli* S30 extracts [72]. It is reported that more than 80% of ATP and GTP hydrolysis in the wheat germ system initially occurs independently of protein synthesis and it was suggested that acid phosphatases are responsible for the nonspecific hydrolysis of the nucleotide triphosphates [53, 71].

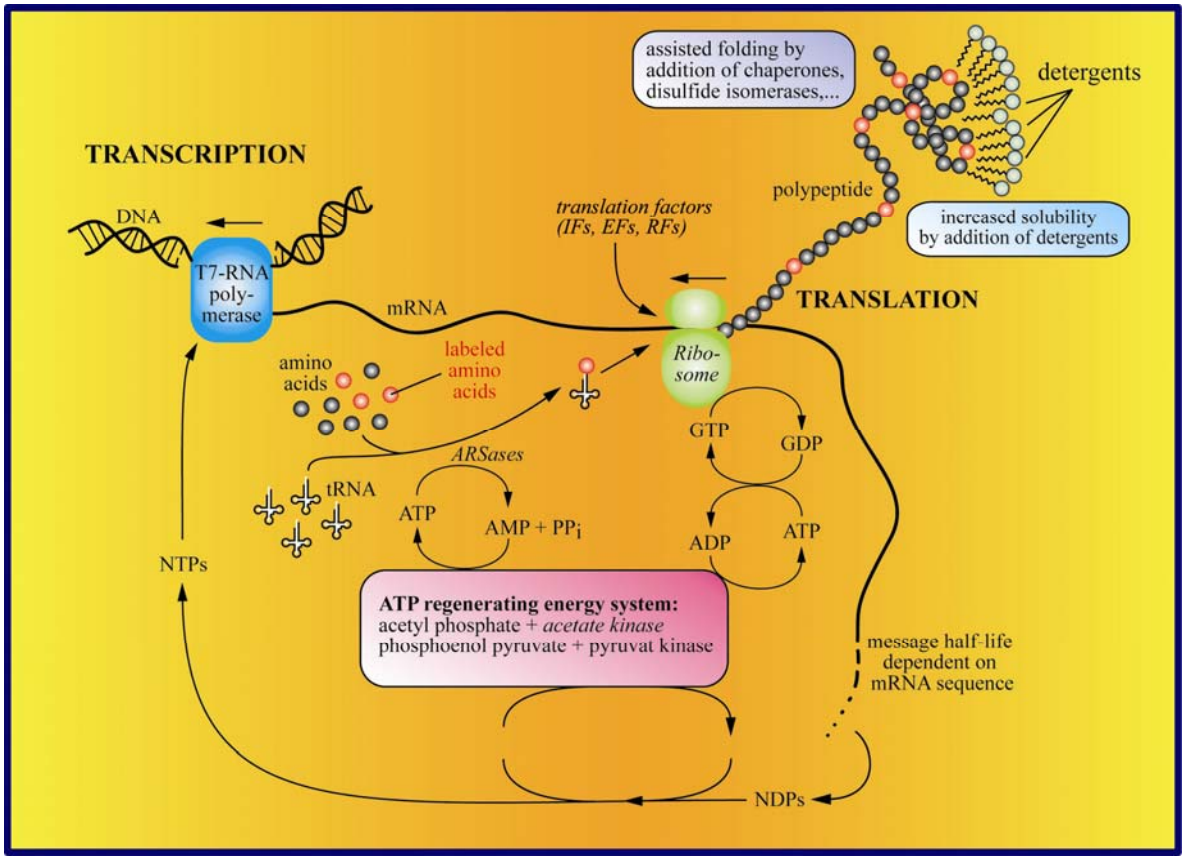


Fig. 2 Components of the cell-free system.
All high molecular-weight compounds necessary for transcription and translation are present in the reaction compartment. The CF reaction is based on a crude cell-extract containing ribosomes (green), translation factors like initiation factors (IFs), elongation factors (EFs) and release factors (RFs), acetate kinase and aminoacyl-tRNA synthetases (ARSases). Substrates like amino acids, the energy-regenerating system components or nucleoside triphosphates (NTPs) and salts are added to the extract, and protein synthesis is initiated by adding the template DNA. The DNA is transcribed by an added RNA polymerase (blue). Added tRNA is loaded with amino acids by ARSases and they are used in the translation of mRNA. The incorporation of some stable isotope-labeled amino acids (red) can easily be done in the CF system, leading to a selective isotope-labeled protein. Regeneration of ATP and GTP and even the NTPs in the cell-free system is achieved by an ATP-regenerating energy system based on the hydrolysis of high-energy substrates in the presence of their kinases. Chaperones can easily be added to the reaction mixture to assist the folding of the target protein and in order to increase the solubility of membrane proteins, detergents can be supplemented directly to the reaction- and feeding mixture. (Modified after (P 08)).

Design

In CF reactions carried out in a “batch” mode, the reaction conditions change as a result of substrate consumption and the accumulation of products. Translation stops as soon as any essential substrate is exhausted, or any product or by-product reaches an inhibiting concentration. Actually, the bacterial CF systems are active only for 10-30 min at 37°C. Systems based on rabbit reticulocyte lysates or wheat germ extract are typically capable

of working for up to 1 h. However, *in vitro* protein-synthesizing systems in batch mode work well for most analytical purposes, but short lifetimes and low productivities limit their application for the synthesis of preparative amounts of protein. The reasons for the low yield are degradation of mRNA, depletion of nucleotide triphosphates and accumulation of their hydrolysates. Prolonged reaction times in CF expression were first achieved by Spirin and co-workers [60, 73, 74] by using a continuous-flow CF (CFCF) translation device (Fig. 3 A).

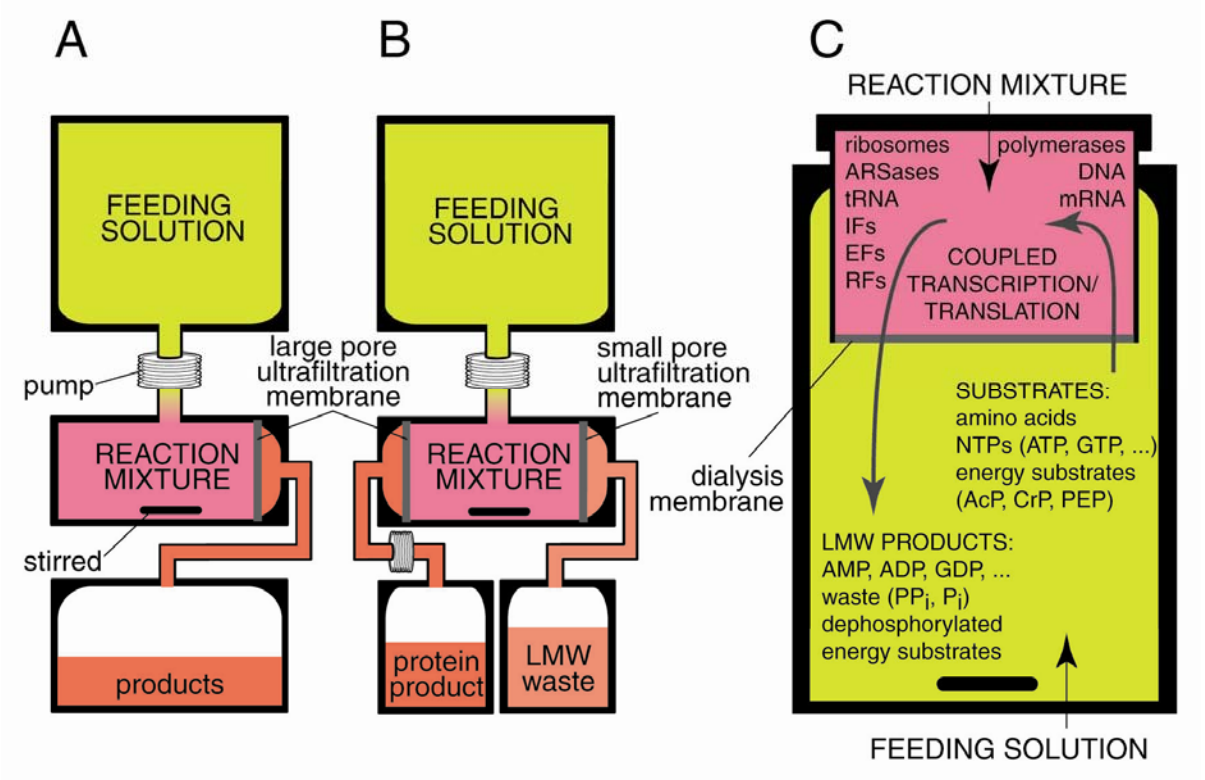


Fig. 3 Cell-free reactors.
Illustration of cell-free reactors for continuous flow (A and B) and continuous exchange (C) cell-free expression systems. **A**, Schematic drawing of a direct flow CFCF reactor where substrates are supplied and products are removed by the flow of a feeding solution, forced by a pump. **B**, Y-flow CFCF reactor, containing two ultrafiltration membranes of different pore size, separating protein product and low molecular weight (LMW) waste outflows [75]. **C**, CECF reactor design with explanation of feeding solution (light green) and reaction mixture components (purple). Compartments are separated by a semi permeable dialysis membrane, providing the diffusion exchange of substrates and LMW products, and the retention of reaction mixture components. Taken from (P 08), modified after Shirokov *et al.* [76].

The basic idea of CFCF systems is to continuously supply amino acids, energy-regenerating components (AcP, CrP or PEP) and NTPs from a feeding solution, and to remove small byproducts of triphosphate hydrolysis like inorganic phosphate and nucleotide monophosphates from the reaction mixture (RM), by active exchange across

an ultrafiltration membrane. The molecular weight cut-off of these membranes typically ranges 10-300 kDa (Fig. 3 A). Permanent stirring of the RM and in some set-ups upright flow of the feeding solution are applied to minimize membrane clogging [77]. The CFCF system can function for more than 20 hours and results in preparative protein expression of about 0.1-1 mg protein per mL of RM. The advantage of the CFCF system is the continuous removal of synthesized proteins from the RM, which can result in 80-85% purity of the protein product. To reduce the dilution effect, the so-called Y-flow reactor with a split outflow has been proposed [75]. The Y-flow reactor (Fig. 3 B) has two membranes with different pore sizes. Initially the low molecular weight (LMW) products are removed through a small-pore membrane at a high rate and the synthesized protein is subsequently collected through a large-pore membrane at a low rate. Here, the flow of the protein product is controlled by a separate pump.

Continuous-exchange CF system (CECF) [5] or semi-continuous-flow CF (SFCF) system [4] are alternatives to obtain prolonged protein synthesis, by separating RM from a feeding mixture (FM) through a dialysis membrane (Fig. 3 C). The simplest CECF device is a dialysis bag [5]. The pore size of these dialysis membranes is usually in the range of 10-50 kDa and better performance of larger pore membrane has been reported [20]. Stirring of either the FM or both, the FM and RM, is necessary for efficient exchange of compounds.

3.2.3. Applications

Due to its independence of a living organism CF expression systems in principle can be individually optimized to allow the production of all those proteins causing problems in conventional expression systems (Fig. 4). Thus, the production of cytotoxic proteins [78] or functional antibodies using PDI and chaperones [6] became possible. Disulfide bridged proteins are hardly expressed in cellular systems due to the reducing environment inside the cell. CF systems have been optimized for the production of those proteins by using modified *E. coli* extracts, by supplementation of DsbC disulfide isomerase and by using an optimized redox potential [79]. Thus the nine disulfide bridges of recombinant plasminogen activator protein could successfully be formed in the bacterial system [7]. Optimized eukaryotic wheat germ systems, supplemented with protein disulfide isomerase further allowed the production of a single-chain antibody variable fragment [80].

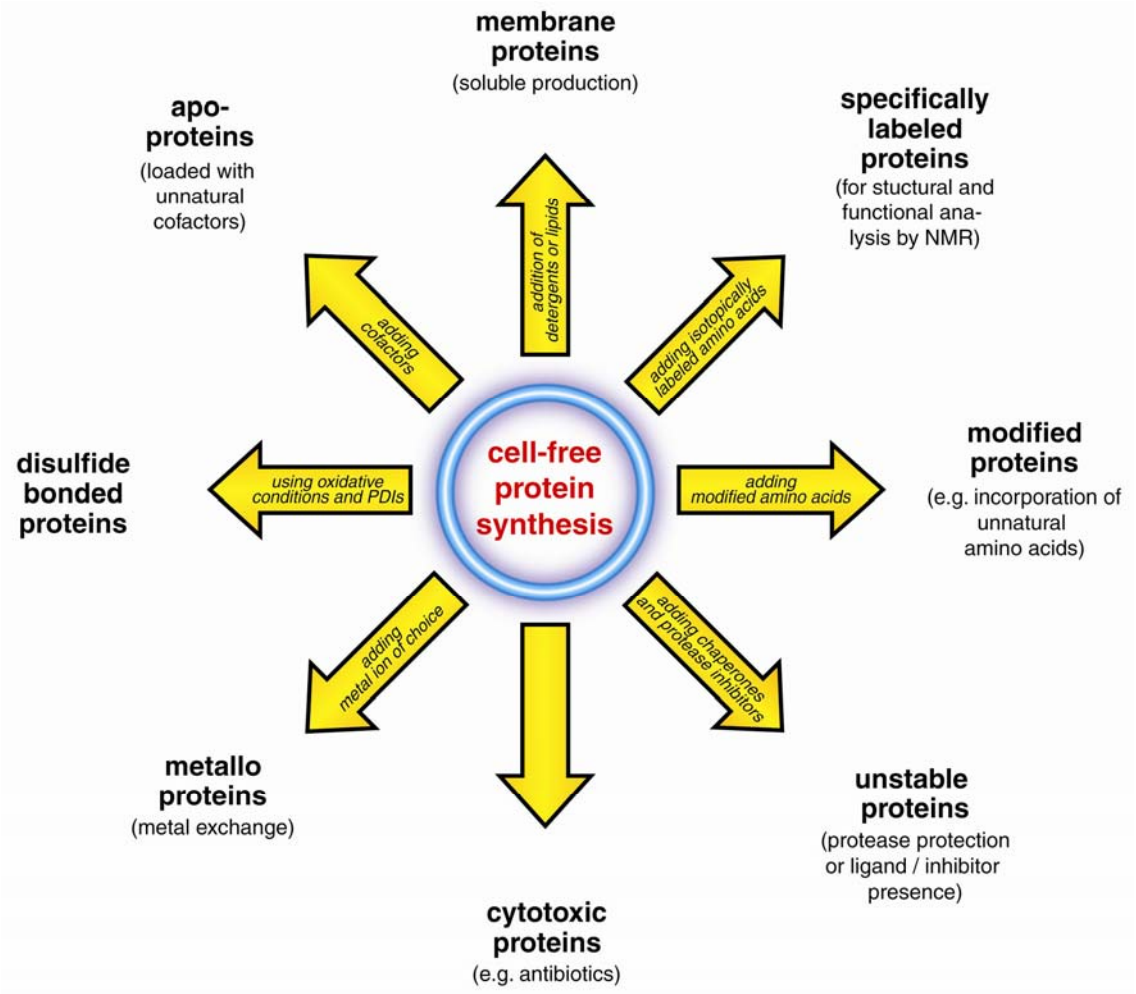


Fig. 4 Expression of difficult proteins by using an individual cell-free system.

Another promising potential for CF synthesis can be found in its suitability for high-throughput expression of proteins by direct translation of linear polymerase chain reaction (PCR) products [68, 69, 81]. Bacterial as well as eukaryotic system have thus been optimized for high-throughput (HT) applications of soluble proteins [82, 83]. Expression screening in 96 well formats [84] as well as in even smaller nanowell chip formats is possible [85]. Problems, related to instability of linear DNA templates can be addressed by preparing CF extracts from nuclease deficient strains [69, 86]. Alternatively, modifications of the mRNA, like the addition of poly(G) tails or mini-hairpin sequences, or the positioning of a stem-loop structure at the 3' end can substantially help to increase the half-life of the transcripts [68, 86, 87]. Further a platform that utilizes wheat germ CF technology to produce protein samples for NMR structure determination in structural genomics has recently been described by Vinarov and co-workers [88-90], showing the potential of CF HT approaches.

The open nature of CF systems provides unique labeling possibilities. Nonlabeled amino acids in the CF mixtures are simply exchanged against the labeled derivatives. No other changes of conditions are required and labeled proteins are synthesized in same amounts like unlabeled ones. Moreover, CF production provides significantly reduced scrambling or background labeling problems [16-22]. Besides the incorporation of isotopically labeled amino acids for nuclear magnetic resonance (NMR) analysis, unnatural and chemically modified amino acids can efficiently become incorporated into CF synthesized proteins [91-93]. This allows new possibilities in functional studies, protein engineering and pharmaceutical research.

4. Results and Achievements

– High level cell-free expression of integral membrane proteins –

4.1. Set-up of an individual cell-free system

4.1.1. Cell-free protocol

The individual CF system is based on a crude cell-extract, prepared form *E. coli* strain A19 (*E. coli* Genetic Stock Center CGSC) by a procedure after Zubay [49] with several modifications. In the modified protocol, the cells were grown in terrific broth medium up to an optical density (OD_{600}) of 3 and washed in washing buffer (10 mM Tris-acetate, pH 8.2, 14 mM $Mg(oAc)_2$), with 6 mM 2-mercaptoethanol and 0.6 mM KCl. Rapid cooling prior harvesting the cells turned out to be crucial. The lysis buffer was the washing buffer supplemented with 1 mM dithiotreitol (DTT) and 0.1 mM phenylmethanesulfonyl fluoride. The extract was dialyzed in washing buffer supplemented with 0.5 mM DTT and 0.6 mM KoAc. The run-off procedure was optimized by removing endogenous mRNA from the ribosomes by incubating the extract with 400 mM NaCl at 42°C for 45 min. Aliquots of the S30 extract were frozen in liquid nitrogen and stored at -80°C. The extract preparation was carried out in one day instead of freezing the harvested cells before continuing with cell-disruption by French press. Approximately 60 ml of extract were obtained from a 10 liter fermentor.

The CF reaction conditions have been optimized for high level protein expression by titration of each component and by using the expression of green fluorescent protein (GFP) as a monitor, yielding more than 4 mg soluble folded GFP per mL of RM (Fig. 5). The most critical parameters appeared to be the concentrations of Mg^{2+}-, K^{+}-Ions and the amino acids, and the quality of the prepared S30 extract. The energy regenerating system was most efficient if a combination of phosphoenol pyruvate, acetyl phosphate and pyruvate kinase was used. The final protocol containing an optimized stock-solution concentration as well as the optimal final concentration of reaction components (Table 3) is capable to synthesize protein up to 4 mg per mL of reaction mixture (RM) and almost 80% of the protein is synthesized during the first 7 h of incubation (Fig. 5).

Table 3 Components of the individual CF system for soluble MP expression.

Component	stock concentration	final concentration
in RM		
S30-extract / S30-buffer	100%	35 %
plasmid DNA	0.3 mg/ml	≥ 15 µg/ml
RNAguard	39.8 U/µl	0.3 U/µl
T7-RNA polymerase	40 U/µl	≥ 3 U/µl
E. coli tRNA	40 mg/ml	500 µg/ml
pyruvate kinase	10 mg/ml	40 µg/ml
in RM + FM		
amino acids (aa)	4 mM	1 mM
aa RCWMDE	16.7 mM	2 mM each
acetyl phosphate	1 M	20 mM
phosphoenol pyruvate	1 M	20 mM
ATP	360 mM	1.2 mM
CTP, GTP, UTP	240 mM	0.8 mM each
1.4-dithiothreitol (DTT)	500 mM	2 mM
folinic acid	10 mg/ml	0.2 mM
complete protease inhibitor	50 x (1 tablet /1 ml)	1 tablet /50 ml
HEPES-KOH pH 8.0	2.5 M	100 mM
magnesium acetate	2 M	15 mM
potassium acetate	4 M	290 mM
polyethylenglycol 8000	40 %	2 %
sodium azide	10 %	0.05 %
detergent (Brij78)	10%	1%

Note: Concentrations of Mg^{2+} and K^{+} are highly critical and should be subject of optimization. 9.1 mM magnesium acetate and 150.8 mM potassium acetate are added, 4.9 mM Mg^{2+} results from the S30-extract and 139.2 mM K^{+} results from other reaction components. Amino acids and T7-RNA-polymerase are limiting compounds and only minimal concentrations are given.

For CECF a RM and FM are needed. In order to avoid pipetting errors and save time, the reaction components are combined to a master mixture, subsequently divided into FM and RM (Table 4). Whereas the RM is kept on ice, the FM is preincubated at 30°C in order to immediately start the reaction after bringing both together.

Table 4 Pipetting protocol for a preparative 1 ml cell-free reaction.

Stock solution	master-mix FRM
10% NaN_3	93 µl
40% PEG8000	926 µl
4 M KoAc	698 µl
2 M $Mg(oAc)_2$	94 µl
25x Buffer	741 µl
50x Complete + EDTA	370 µl
10 mg/ml folinic acid	185 µl
0.5 M DTT	74 µl
75x NTP-mix	247 µl
1 M PEP	370 µl
1 M AcP	370 µl
4 mM aa-mix	2315 µl
16.7 mM aa RCWMDE-mix	1111 µl
10% Brij-78	1852 µl
	9447 µl

	FM	RM
Master-mix FRM	8927 µl	520 µl
4 mM aa-Mix	2188 µl	-
S30-buffer	6125 µl	-
10 mg/ml pyruvate kinase	-	4 µl
40 mg/ml *E. coli* tRNA	-	13 µl
420 U/µl T7-RNA-Polym.	-	15 µl
21.5 U/µl RNAsin	-	14 µl
S30-Extract	-	357 µl
0.3 mg/ml Plasmid-DNA	-	51 µl
H_2O	261 µl	46 µl
	17500 µl	1020 µl

Note: All volumes less then one µl have been rounded up.

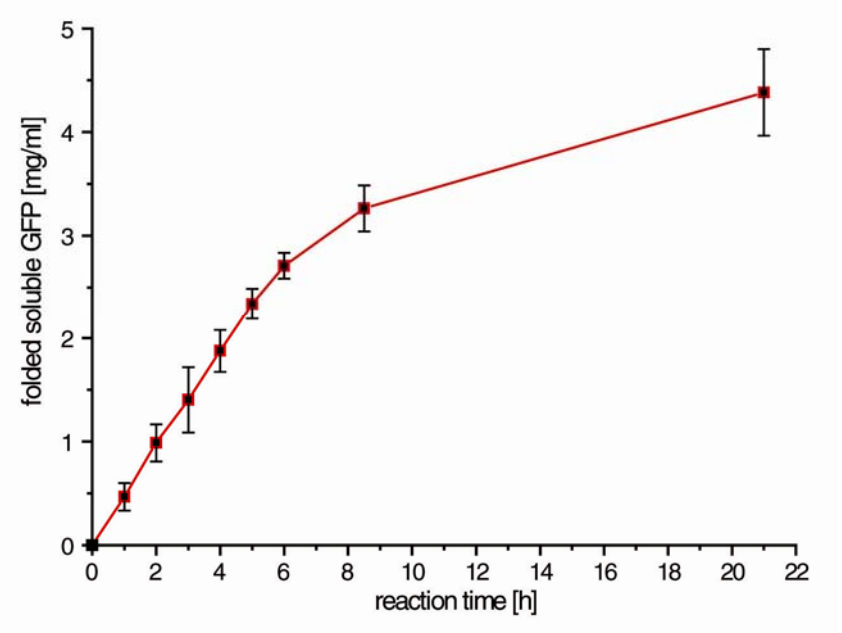


Fig. 5 Protein production kinetics in the cell-free system. Soluble GFP production in a standard cell-free reaction with a membrane cut-off of 25 kDa and an RM/FM ratio of 1 : 17 was monitored by fluorescence at an emission at 509 nm and after excitation at 395 nm. Data are averages of at least three determinations. (Modified after (P 01))

4.1.2. Analytical and preparative scale set-up

Generally, the reaction can be set up in analytical design for optimization and screening reactions as well as in preparative configuration for the production of mg amounts of recombinant protein. As a device for analytical scale reactions with a RM volume of 70 µl, Microdialysers® (MDs) (Spectrum, Rancho Dominquez, USA) that are commercially available with different molecular weight cut-offs (MWCOs) are used with 15 kDa or 25 kDa MWCO (Fig. 6 A). Those MDs holding the RM are placed in a suitable tank holding a 14 fold excess volume of FM. The MD reactions are incubated over night in a lab-shaker at 30°C.

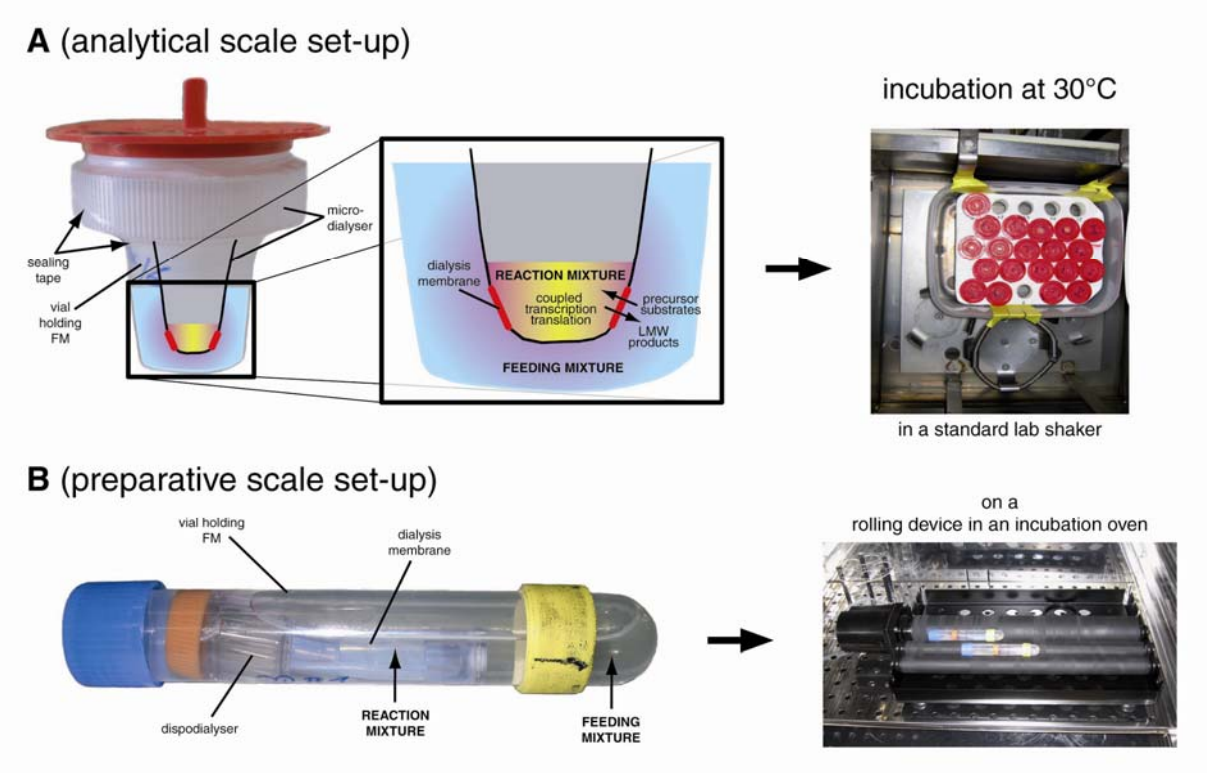


Fig. 6 Design of cell-free reaction containers for analytical and preparative scale set-up. **A,** for optimization and screening reactions, commercially available Microdialyser® units holding up to 100 µl RM where placed in a suitable vial containing a 14 times volume excess of FM, fixed with a sealing tape and incubated in a standard lab-shaker at 30°C over night. **B**, for preparative scale cell-free protein expression, commercially available Dispodialysers® holding 500 to 1000 µl RM where placed into a plastic vial containing a 17 times volume excess of FM. The set-up is incubated on a roller at 30°C. (Modified after (P 03)).

For preparative reactions with RM volumes of 500-1000 µl, larger Dispodialysers® (DDs) (Spectrum, Rancho Dominquez, USA) are used with MWCOs of 15 kDa and 25 kDa. Those closable dialysis tubes are placed in a suitable vial containing a 17 fold excess of FM (Fig. 6 B), which was proven to give the highest efficiency. In order to allow sufficient diffusion between the RM and FM in DDs, the preparative scale set-up is rolled at 40 RPM on a roller device at 30°C in an incubation oven over night. In DDs, the use of regenerated cellulose membranes is preferred to that of ester-cellulose membranes.

4.1.3. Cell-free modes of IMP expression

Most cellular membranes have been removed during extract preparation and only spurious amounts of lipids might remain. Standard CF reactions therefore result consequently in the production of IMPs as precipitates as no hydrophobic compartments are present in the RM. The supplementation of CF reactions with detergents or lipids generates preformed micelles or liposomes in the RM and could enable the direct synthesis of IMPs into micelles composed of the desired detergents or into liposomes of defined compositions. In principle one can think of three different CF expression modes of IMPs: (A) the production of IMP precipitates without any hydrophobic additives [10, 11], (B) the synthesis of IMPs into micelles in presence of detergents [9, 10, 14, 15] and (C) the synthesis of IMPs into liposomes in presence of lipids which has not been proved yet (Fig. 7).

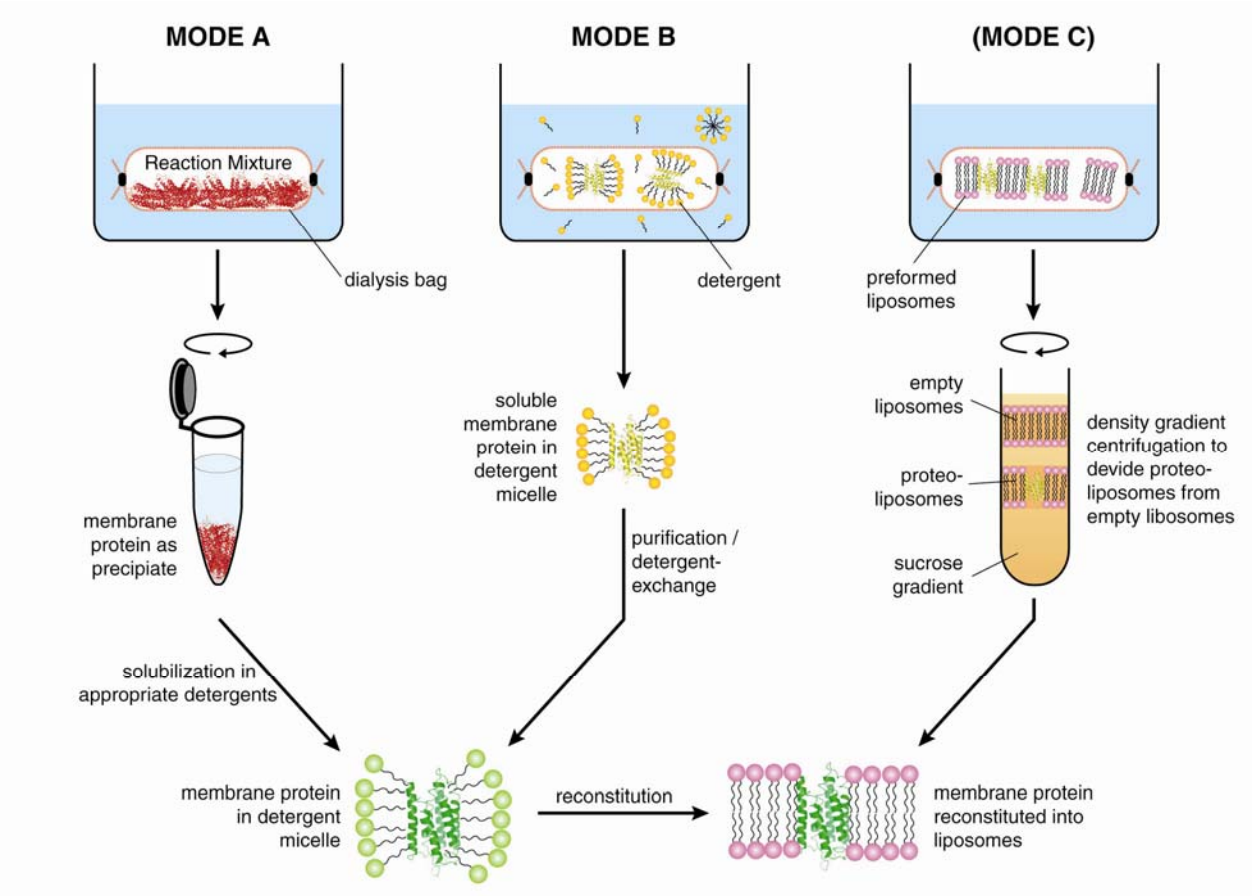


Fig. 7 Cell-free expression modes.
The CF expression of IMPs can be run in three different modes: IMPs can be expressed as precipitate (**Mode A**), solubilized in an appropriate detergent and further reconstituted into proteoliposmes, alternatively the IMPs can be soluble expressed in presence of detergents (**Mode B**), purified and reconstituted into proteoliposomes, or expressed in presence of preformed liposomes (**Mode C**) and isolated from the RM by density gradient centrifugation. (Modified after (P 05)).

This option is an attractive feature specific to CF expression and it is not possible with any other expression system. Critical steps in common IMP preparation protocols like the disintegration of cellular membranes or repeated transfers of IMPs into micelles of different compositions can thus be reduced or even completely avoided. In the best case, the IMPs can become inserted in the desired environment right after translation and the resulting proteomicelles or proteoliposomes could be directly used for further analysis.

However, diverse IMPs have to be expected to behave different in the CF expression modes. To give a comprehensive overview about the variety of IMP behavior, we analyzed three structurally very different IMPs (Fig. 8). The synthesis of the bacterial four transmembrane segments (TMS) containing α-helical small multi drug transporter EmrE [8], the bacterial β-barrel type nucleoside transporter Tsx [94] and the eukaryotic α-helical and seven TMS containing GPCR Vasopressin 2 Receptor (V2R) has been analyzed by CF expression in presence of detergents.

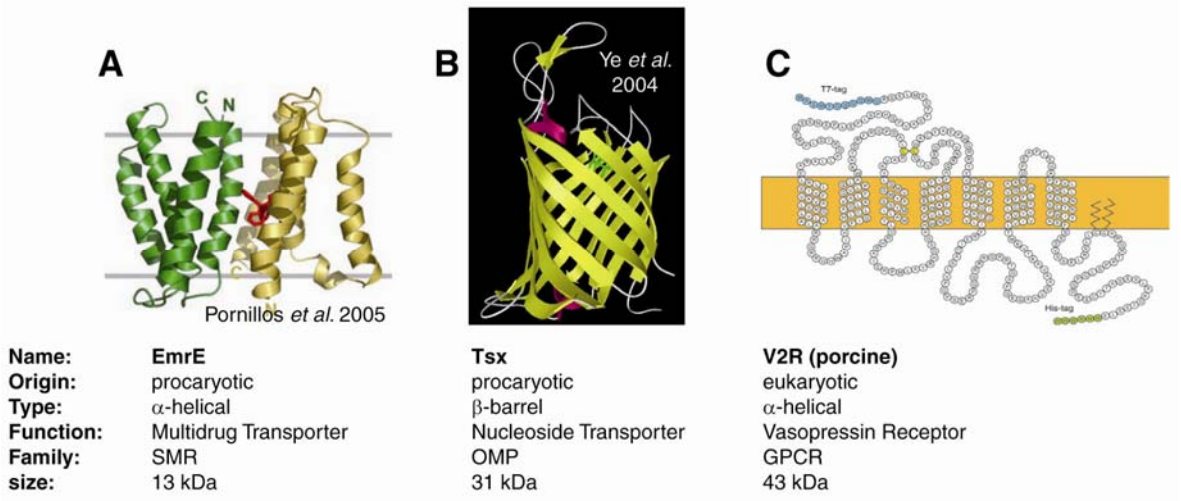
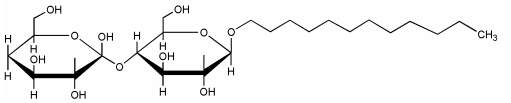
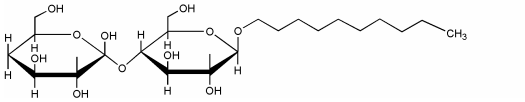
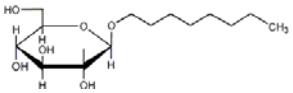
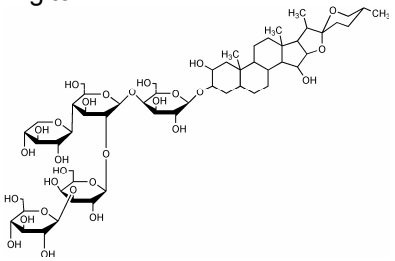
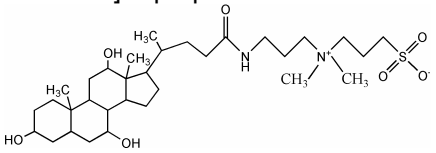
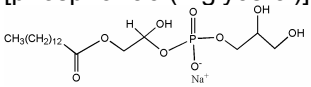
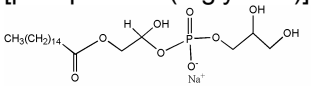


Fig. 8 Overview of analyzed IMPs for soluble cell-free expression in presence of detergents. The structurally very diverse IMPs have been analyzed for their cell-free expression in presence of detergents. Information of these proteins is given below the figures.

It has to be expected that some detergents are detrimental to the CF system whereas others might not be tolerated enough to keep the different IMPs in solution. In order to give a comprehensive overview, 24 different commonly used detergents of six different classes, have been evaluated for their CF system properties. Some physicochemical properties of these detergents are summarized in Table 5.

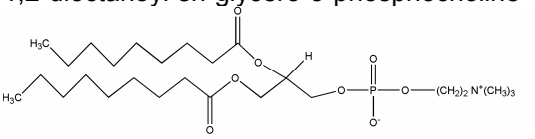
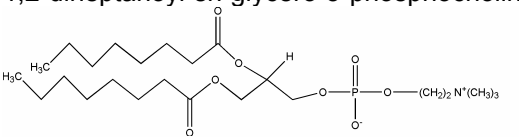
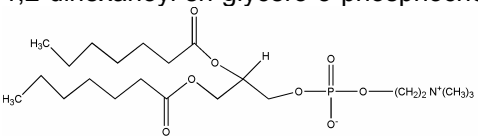
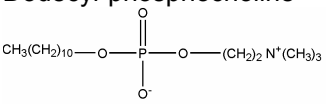
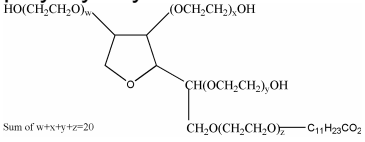
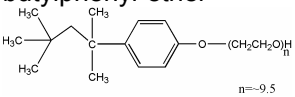
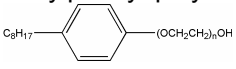
Table 5 Physicochemical properties of analyzed detergents.

Detergent	short name	charge[a]	MW [Da]	CMC[b] [mM]
Alkyl-glucosides:				
n-dodecyl-β-D-maltoside 	DDM	N	511	0.12
n-decyl-β-D-maltoside 	DM	N	483	1.8
n-octyl-β-D-glucopyranoside 	β-OG	N	292	19
Steroid-derivatives:				
Digitonin 	Digitonin	N	1229	0.73
3-[(3-cholamidopropyl)dimethly-ammonio]-1-propansulfonat 	CHAPS	Z	615	2.4 – 8.6
Long chain-phosphoglycerols:				
1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)] 	LMPG	A	479	0.05
1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)] 	LPPG	A	507	n.a.

[a] A: anionic; N: nonionic; Z: zwitterionic.

[b] estimated CMCs, averaged values are given if different CMCs for a detergent have been published; n.a., not available

Table 5 (continued)

Detergent	short name	charge[a]	MW [mM]	CMC[b]
Mono-/Bi-chain-phosphocholines:				
1,2-dioctanoyl-sn-glycero-3-phosphocholine 	diC$_8$PC	Z	510	0.22
1,2-diheptanoyl-sn-glycero-3-phosphocholine 	DHPC	Z	482	1.4
1,2-dihexanoyl-sn-glycero-3-phosphocholine 	diC$_6$PC	Z	454	14
Dodecyl-phosphocholine 	DPC	Z	352	0.9 - 1.5
Polyoxyethylene-alkyl-ether:				
polyoxyethylene-(23)-lauryl-ether, ($C_{12/23}$) $C_{12}H_{20}(O\text{-}CH_2CH_2)_nOH$ n= ~23	Brij35	N	1200	0.08
polyoxyethylene-(10)-dodecyl-ether, ($C_{12/10}$)	Genapol C-100	N	627	0.1
polyoxyethylene-(10)-cetyl-ether, ($C_{16/10}$)	Brij56	N	682	0.035
polyoxyethylene-(20)-cetyl-ether, ($C_{16/20}$)	Brij58	N	1123	0.075
polyoxyethylene-(2)-stearyl-ether, ($C_{18/2}$)	Brij72	N	359	n.a.
polyoxyethylene-(20)-stearyl-ether, ($C_{18/20}$)	Brij78	N	1152	0.046
polyoxyethylene-(10)-oleyl-ether, ($C_{18\text{-}1/10}$)	Brij97	N	709	0.217
polyoxyethylene-(20)-oleyl-ether, ($C_{18\text{-}1/20}$)	Brij98	N	1150	0.025
polyoxyethylene-sorbitan-monolaurate 20 	Tween 20	N	1228	0.059
Polyethylene-glycol derivatives:				
polyethylene-glycol P-1,1,3,3-tetramethyl-butylphenyl-ether 	Triton X-100	N	647	0.23
polyethylene-glycol 400 dedecyl-ether $C_{12}H_{25}(O\text{-}CH_2CH_2)_nOH$ n=~9.5	Thesit	N	583	0.1
nonylphenyl-polyethylene-glycol 	NP40	N	603	0.05 - 0.3

Mode A, insoluble expression

For the CF production of IMP precipitates, the standard protocol can instantly be used and no time-consuming evaluation of detergents is needed. High yields of synthesized IMPs are ensured as no inhibitory effects of supplied detergents are present. In addition, the IMPs can be obtained in highly purified form by only few washing steps in detergents inefficiently solubilizing the IMPs [see publication (P 01)]. This expression mode mostly resembles conventional *in-vivo* approaches of the production of IMPs in form of inclusion bodies. However, structural differences might exist between cellular inclusion bodies and CF produced precipitates. IMP precipitates obtained by CF reactions usually solubilize rapidly upon addition of relatively mild detergents and they do not require intensive denaturation and refolding steps as it is known from the solubilization of inclusion bodies. Denaturating agents like Guanidinium HCl or excessive amounts of urea could thus be avoided and already gentle mixing of the IMP precipitate with a suitable detergent at a final concentration of 2 % at 30°C is often sufficient for a quantitative solubilization as analyzed for the three structural very different proteins Tsx, V2R and EmrE. Whereas the small multi drug transporter EmrE was readily resolubilized in high quantities in most of the detergents analyzed, the quantitative resolubilization of the GPCR protein V2R was only possible with the long-chain phosphocholine, 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LMPG). Other detergents had no or only minor resolubilization effects. In addition, only LMPG and dodecyl-phosphocholine (DPC) were highly effective for the resolubilization of the β-barrel type transporter Tsx [see publication (P 02)]. LMPG and probably similar long chain-phosphoglycerols appear therefore to be the primary choice for a general and efficient resolubilization of even structurally diverse types of IMPs in mild detergents.

Mode B, soluble expression in presence of detergents

For direct soluble CF expression into detergent micelles, the detergents have to be added above their specific critical micellar concentration (CMC) in order to become effective for the solubilization of IMPs. The quantitative analysis in analytical scale CF reactions of the proteins in the soluble and insoluble fractions turned out that only few detergents become inhibitory to the CF system at low concentrations and resulted in an at least considerably reduced production of all three IMPs. In particular, some of the phosphocholine derivatives seemed to be problematic as additives in CF systems, while the mono-chain phosphocholines were clearly more critical than the bi-chain phosphocholines. DPC almost completely inhibited general protein synthesis. However,

the bi-chain phosphocholine 1,2-dioctanoyl-*sn*-glycero-3-phosphocholine (diC_8PC), still resulted in the production of up to 0.6 mg of soluble EmrE per mL of RM. Negative effects on the CF system were further observed with the nonionic alkyl-glucoside n-octy-β-glucopyransoide (β-OG), the steroid Chaps and the long-chain phosphoglycerol LMPG. Chaps and β-OG are generally considered to be relatively mild detergents, but their high CMC, resulting in relatively high final detergent concentrations in the RM, might be a reason for the observed inhibitory effect on protein expression. LMPG, which was found to be highly suitable for the resolubilization of CF produced precipitates, resulted only with Tsx in the production of some 100 µg of soluble protein, while the total yield of protein was considerably affected for all three IMPs.

Approximately one-third of the tested detergents did not have any, or only a marginal effect on the solubility of any of the expressed IMPs. But they did not cause marked decreases of the total protein production in the range of the supplied concentrations. This group includes polyethylene glycol 400 dedecyl-ether (Thesit), nonylphenyl-polyethylene-glycol (NP40), 1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)] (LPPG) and the polyoxyethylene derivatives polyoxyethylene-(10)-cetyl-ether (Brij56), polyoxyethylene-(2)-stearyl-ether (Brij72), polyoxyethylene-(10)-oleyl-ether (Brij97), polyoxyethylene-(10)-dodecyl-ether (Genapol C-100), polyoxyethylene-sorbitan-monolaurate 20 (Tween 20). These detergents are obviously tolerated by the CF system, even at higher CMC concentrations, and they might still be beneficial for the soluble expression of other IMPs. However, they are clearly not suitable for general usage.

Several detergents from almost all analyzed groups were found to be suitable for the soluble CF expression of IMPs, yielding protein rates of more than 0.5 mg of soluble IMP per mL of RM. With a total of 10 detergents one obtains preparative scale soluble expression yields for at least one of the analyzed MPs, while the individual proteins behave quite differently. The most outstanding group of effective detergents covered the long chain polyoxyethylene-alkyl-ethers, polyoxyethylene-(23)-lauryl-ether (Brij35), polyoxyethylene-(20)-cetyl-ether (Brij58), polyoxyethylene-(20)-stearyl-ether (Brij78) and polyoxyethylene-(20)-oleyl-ether (Brij98), and they resulted in the high-level soluble expression of all three IMPs analyzed. The steroid derivative Digitonin was identified as a second relatively general solubilizing detergent. However, the yields obtained in the case of EmrE and V2R remained mostly below 0.5 mg per mL in the RM and were thus lower by comparison with the above mentioned Brij derivatives. The alkyl glucosides *n*-dodecyl-β-D-maltoside (DDM) and *n*-decyl-β-maltoside (DM) resulted in the soluble expression of only EmrE and Tsx in the range between 0.5 mg and 1.5 mg per mL of RM. The beneficial effect of polyethylene glycol P-1,1,3,3-

tetramethyl-butylphenyl-ether (Triton X-100) was specific for the Tsx protein and yielded soluble protein at a concentration of nearly 1 mg per mL in the RM.

4.2. Cell-free expressed IMPs

Various IMPs have been successfully synthesized in preparative amounts using CF expression techniques and reviewed in the FEBS Mini Review [13]. Besides my contribution to this field, different small multi drug resistance transporters, EmrE [8-11], Hsmr, YdgF, Yfbw [12], Psmr [10] and TBsmr [10, 12] have been expressed in analytical to preparative amounts. In addition the mechanosensitive ion channel MscL [15] and a bacterial light-harvesting membrane protein [95] have been CF expressed in preparative amounts using the Roche RTS™ system. Besides these bacterial IMPs, the rat neurotensin receptor, the muscarinic acetylcholine M2 receptor and the β2-adrenergic receptor, all members of the GPCR family, have been CF synthesized in fusion to thioredoxin [14].

A variety of different bacterial transporters were successfully synthesized in the individual CF system using the insoluble expression mode (Fig. 9). SugE, EmrE and TBsmr have been synthesized in amounts of 2, 3 and 1 mg per mL of RM, respectively. In addition, the six TMS comprising amino acid transporter YfiK and the putative tellurite transporter TehA, containing 10 TMS were CF expressed in amounts of 3 mg per mL of RM each. Beside this α-helical IMPs of the inner bacterial membrane, the β-barrel type nucleoside outer membrane transporter Tsx was synthesized in amounts up to 4 mg per mL of RM. For preparative CF expression the amino acid composition in RM and FM was optimized for each protein regarding their amino acid distribution (Table 6). The least abundant amino acids (present at ≤ 3% in the protein) were added at 1.25 mM, medium abundant (between 3 and ≤ 8%) at 1.8 mM and highly abundant (more than 8%) at 2.5 mM final concentration. Besides these bacterial transporters four different GPCRs were CF expressed at preparative amounts as will be discussed later in section 6 (unpublished results). All of the CF produced IMP precipitates could be solubilized in mild detergents like LMPG.

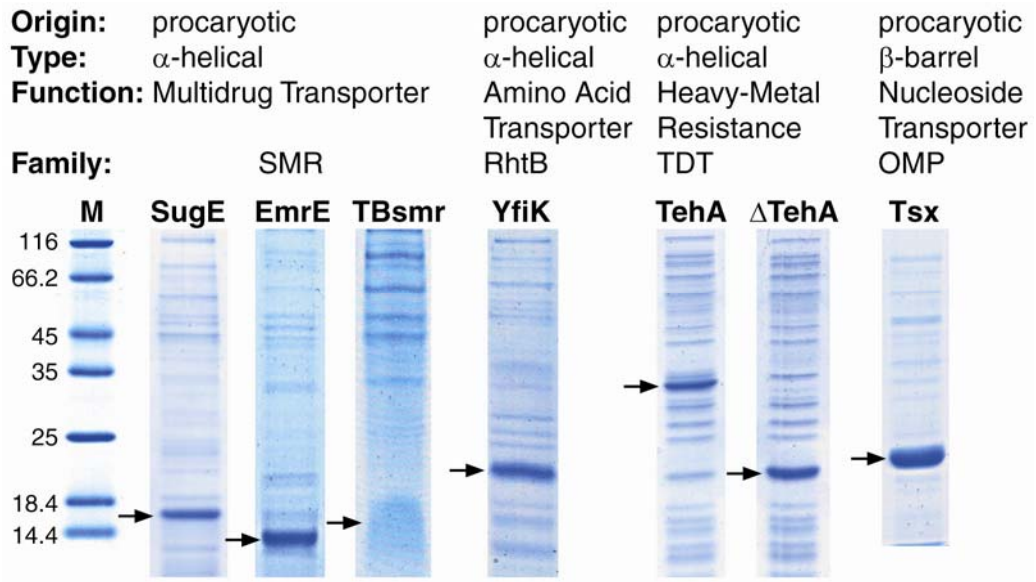


Fig. 9 IMPs from different families that have been produced as a precipitate by CF expression. The precipitate of an analytical reaction was suspended in the corresponding volume of buffer and the proteins in 1 µl of each suspension were separated by electrophoresis in a 12 % SDS-polyacrylamide-gel. The arrows indicate the overproduced recombinant IMPs. OMP: outer membrane proteins, SMR: small multidrug transporter. (Modified after (P03)).

Table 6 Information about preparative CF expressed IMPs.

name	EmrE	SugE	TBsmr	YfiK	TehA	ΔTehA	Tsx	V2R (h)	V2R (p)	CRF (rat)	ETB (h)
origin	procaryotic	procaryotic	procaryotic	procaryotic	procaryotic	procaryotic	procaryotic	eukaryotic	eukaryotic	eukaryotic	eukaryotic
type	α-helical	α-helical	α-helical	α-helical	α-helical	α-helical	β-barrel	α-helical	α-helical	α-helical	α-helical
function	multidrug transport	multidrug transport	multidrug transporter	amino acid transporter	heavy-metal transporter	heavy-metal transporter	nucleoside transporter	vasopressin receptor	vasopressin receptor	corticotropin releasing facor receptor	endothelin-1 receptor
familiy	SMR	SMR	SMR	RhtB	TDT	TDT	OMP	GPCR	GPCR	GPCR	GPCR
predicted TMS	4	4	4	6	10	7		7	7	7	7
size	13.02 kDa	10.9 kDa	11.08 kDa	21.25 kDa	37.0 kDa	24.66 kDa	32.48 kDa	43.59 kDa	42.85 kDa	47.76 kDa	52.51 kDa
aa number	118	105	107	195	338	227	280	392	391	414	467
theoretical pI	7.04	9.82	6.69	9.72	8.49	7.87	5.32	9.49	10.00	9.58	9.15
tags	c-H6	no tag	no tag	no tag	c-H6	c-H6	c-H6	n-T7, c-H6	n-T7, c-H6	n-T7, c-H6	n-T7, c-H10
aa content											
Ala (A)	9 (7.6%)	14 (13.3%)	14 (13.1%)	24 (12.3%)	35 (10.4%)	26 (11.5%)	15 (5.4%)	48 (12.2%)	51 (13.0%)	26 (6.3%)	32 (6.9%)
Arg (R)	3 (2.5%)	2 (1.9%)	1 (0.9%)	6 (3.1%)	13 (3.8%)	9 (4.0%)	9 (3.2%)	30 (7.7%)	37 (9.5%)	29 (7.0%)	21 (4.5%)
Asn (N)	2 (1.7%)	1 (1.0%)	1 (0.9%)	5 (2.6%)	7 (2.1%)	4 (1.8%)	25 (8.9%)	6 (1.5%)	5 (1.3%)	22 (5.3%)	18 (3.9%)
Asp (D)	1 (0.8%)	0	1 (0.9%)	2 (1.0%)	6 (1.8%)	3 (1.3%)	23 (8.2%)	10 (2.6%)	12 (3.1%)	7 (1.7%)	14 (3.0%)
Cys (C)	3 (2.5%)	0	2 (1.9%)	3 (1.5%)	3 (0.9%)	3 (1.3%)	0	11 (2.8%)	10 (2.6%)	13 (3.1%)	20 (4.3%)
Gln (Q)	2 (1.7%)	0	1 (0.9%)	8 (4.1%)	10 (3.0%)	7 (3.1%)	8 (2.9%)	11 (2.8%)	10 (2.6%)	14 (3.4%)	14 (3.0%)
Glu (E)	3 (2.5%)	2 (1.9%)	2 (1.9%)	1 (0.5%)	10 (3.0%)	6 (2.6%)	10 (3.6%)	12 (3.1%)	12 (3.1%)	15 (3.6%)	18 (3.9%)
Gly (G)	12 (10.2%)	11 (10.5%)	11 (10.3%)	13 (6.7%)	31 (9.2%)	20 (8.8%)	30 (10.7%)	28 (7.1%)	29 (7.4%)	22 (5.3%)	25 (5.4%)
His (H)	7 (5.9%)	2 (1.9%)	2 (1.9%)	3 (1.5%)	12 (3.6%)	10 (4.4%)	13 (4.6%)	15 (3.8%)	12 (3.1%)	16 (3.9%)	15 (3.2%)
Ile (I)	15 (12.7%)	11 (10.5%)	8 (7.5%)	13 (6.7%)	16 (4.7%)	11 (4.8%)	13 (4.6%)	12 (3.1%)	11 (2.8%)	28 (6.8%)	34 (7.3%)
Leu (L)	17 (14.4%)	16 (15.2%)	17 (15.9%)	30 (15.4%)	54 (16.0%)	33 (14.5%)	16 (5.7%)	50 (12.8%)	49 (12.5%)	39 (9.4%)	58 (12.4%)
Lys (K)	1 (0.8%)	3 (2.9%)	2 (1.9%)	4 (2.1%)	5 (1.5%)	1 (0.4%)	11 (3.9%)	4 (1.0%)	3 (0.8%)	13 (3.1%)	29 (6.2%)
Met (M)	4 (3.4%)	5 (4.8%)	3 (2.8%)	5 (2.6%)	10 (3.0%)	5 (2.2%)	5 (1.8%)	12 (3.1%)	13 (3.3%)	9 (2.2%)	12 (2.6%)
Phe (F)	5 (4.2%)	1 (1.0%)	4 (3.7%)	12 (6.2%)	25 (7.4%)	13 (5.7%)	18 (6.4%)	14 (3.6%)	15 (3.8%)	27 (6.5%)	21 (4.5%)
Pro (P)	5 (4.2%)	3 (2.9%)	2 (1.9%)	7 (3.6%)	15 (4.4%)	12 (5.3%)	8 (2.9%)	26 (6.6%)	26 (6.6%)	13 (3.1%)	28 (6.0%)
Ser (S)	8 (6.8%)	8 (7.6%)	7 (6.5%)	15 (7.7%)	25 (7.4%)	18 (7.9%)	18 (6.4%)	37 (9.4%)	33 (8.4%)	31 (7.5%)	35 (7.5%)
Thr (T)	7 (5.9%)	8 (7.6%)	7 (6.5%)	14 (7.2%)	18 (5.3%)	11 (4.8%)	12 (4.3%)	18 (4.6%)	14 (3.6%)	22 (5.3%)	21 (4.5%)
Trp (W)	4 (3.4%)	4 (3.8%)	2 (1.9%)	7 (3.6%)	11 (3.3%)	9 (4.0%)	14 (5.0%)	11 (2.8%)	12 (3.1%)	12 (2.9%)	10 (2.1%)
Tyr (Y)	5 (4.2%)	2 (1.9%)	4 (3.7%)	6 (3.1%)	7 (2.1%)	5 (2.2%)	21 (7.5%)	7 (1.8%)	8 (2.0%)	18 (4.3%)	14 (3.0%)
Val (V)	5 (4.2%)	12 (11.4%)	16 (15.0%)	17 (8.7%)	25 (7.4%)	21 (9.3%)	11 (3.9%)	30 (7.7%)	29 (7.4%)	38 (9.2%)	28 (6.0%)

4.2.1. Reconstitution of cell-free expressed IMPs into proteoliposomes

Two different modes for CF expression of IMPs have been established so far. The soluble expression of IMPs in presence of detergents as well as the insoluble expression in combination with easy detergent solubilization results in proteomicelles. For analyzing IMPs in their natural environment of a lipid bilayer, they have to be further reconstituted into proteoliposomes (Fig. 10).

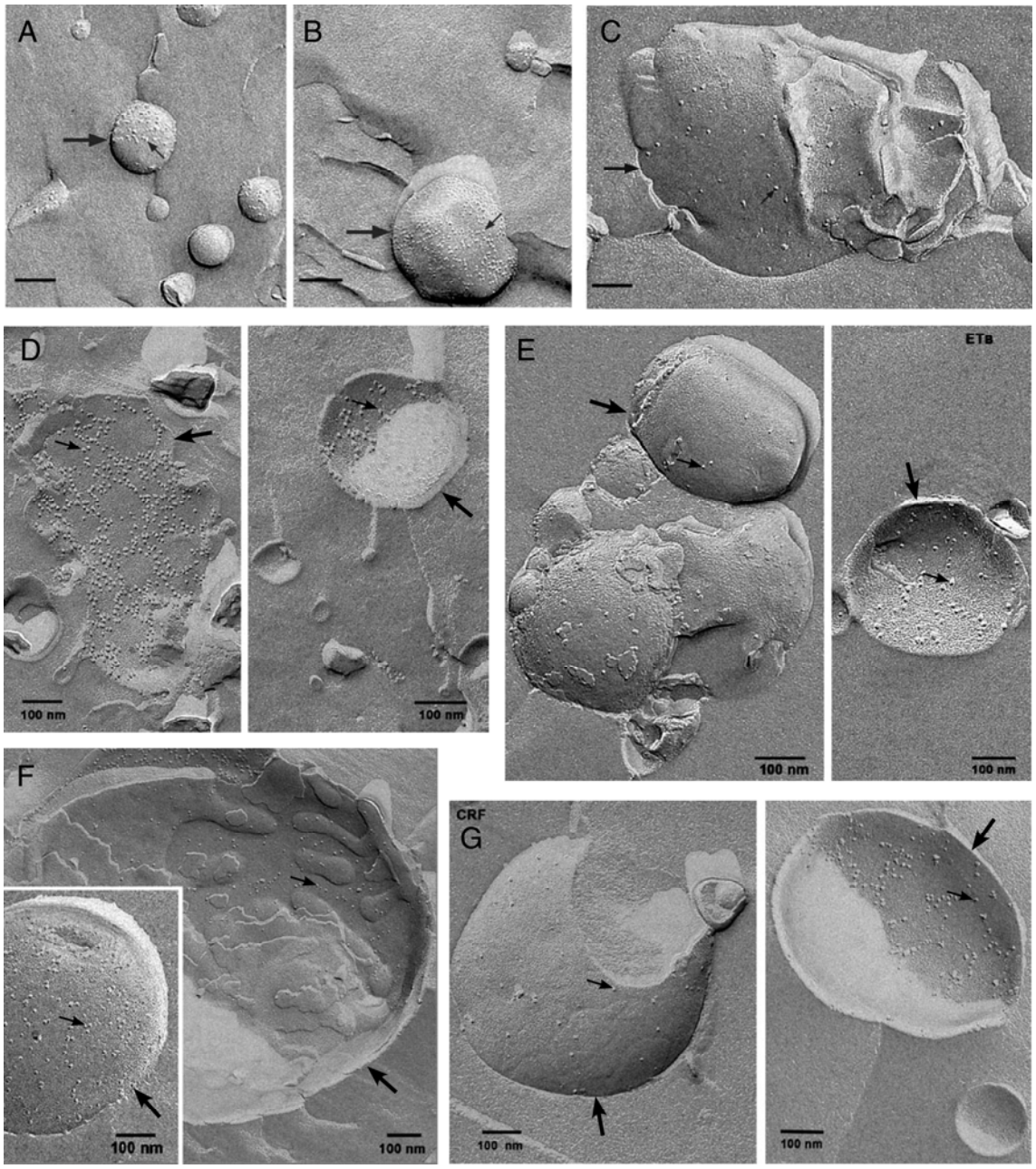


Fig. 10 Freeze-fracture analysis of proteoliposomes with reconstituted cell-free produced IMPs. Detergents were removed with biobeads and the proteins were reconstituted into *E. coli* lipid mixture. Big arrows indicate the liposome; small arrows show the reconstituted IMPs. **A**, EmrE; **B**, SugE; **C**, TehA, **D**, V2Rp; **E**, ETB; **F**, V2Rh; **G**, CRF. The pictures where provided by W. Haase.

In principle, the reconstitution procedure includes the addition of lipids and subsequently the removal of detergents. In these studies *E. coli* lipids where added at a molar ratio of protein : lipid between 1 : 500 and 1 : 2500 to the proteomicelles. After 1 h incubation at 30°C the detergents were removed by the incubation with biobeads SM-2 (Bio-Rad, Hercules, USA), hydrophilic absorbing compounds, that have been added in 10-fold excess to the detergent and have been exchanged within 36 hours for 2-3 times. The successful reconstitution of the CF insoluble expressed and detergent resolubilized IMPs, EmrE, SugE, TehA, porcine vasopressin receptor type 2 (V2R), human endothelin B receptor (ETB), human V2R and rat corticotropin releasing factor receptor (CRF) were analyzed by freeze fracture electron microscopy (Fig. 10). The evenly and homogeneously distributed IMP particles in the membranes gave evidence for a structural folding of those proteins.

4.2.2. Activity assay of the α-helical transporters EmrE

The small multi drug transporter EmrE was cell-free expressed as precipitate, solubilized in 1% DDM and reconstituted into *E. coli* lipids. One substrate of the proton antiporter EmrE is ethidium bromide and the uptake of this dye into proteoliposomes can be measured by a specific transport assay [96]. In this assay a proton gradient between the DNA containing liposome and the solute is generated. Intercalation of ethidium into DNA causes an effect on the quantum yield of its fluorescence. Active EmrE protein should therefore generate a significant increase in the fluorescence intensity, by pumping ethidium against the pH gradient into the proteoliposomes where it is accumulated in the DNA molecules. Approximately 140 nM EmrE embedded in *E. coli* lipids were assayed in a total volume of 1 mL. Once the baseline was established, proteoliposomes where added, followed by ethidium bromide after 10 s to a final concentration of 2.5 µM and an immediate large biphasic increase in the fluorescence was monitored (Fig. 11). The first phase of the increase can be attributed to the binding of ethidium to the residual DNA in the extraliposomal space [96], while the second phase represents the accumulation of ethidium inside the liposomes due to the transport activity of EmrE. This ethidium uptake could be inhibited by preincubation of the proteoliposomes with an excess of 50 µM of the high affinity substrate tetraphenylphosphonium (TPP^+), probably through competition with the ethidium binding site at EmrE. In addition, the collapse of the pH gradient upon supplementation of nigericine also prevented the accumulation of ethidium in the proteoliposomes, resulting only in the single phase increase of fluorescence after addition of ethidium bromide. The results clearly demonstrate that the ethidium/ H^+ antiport was responsible

for the observed increase in fluorescence, indicating the functional reconstitution of cell-free expressed EmrE in *E. coli* lipids.

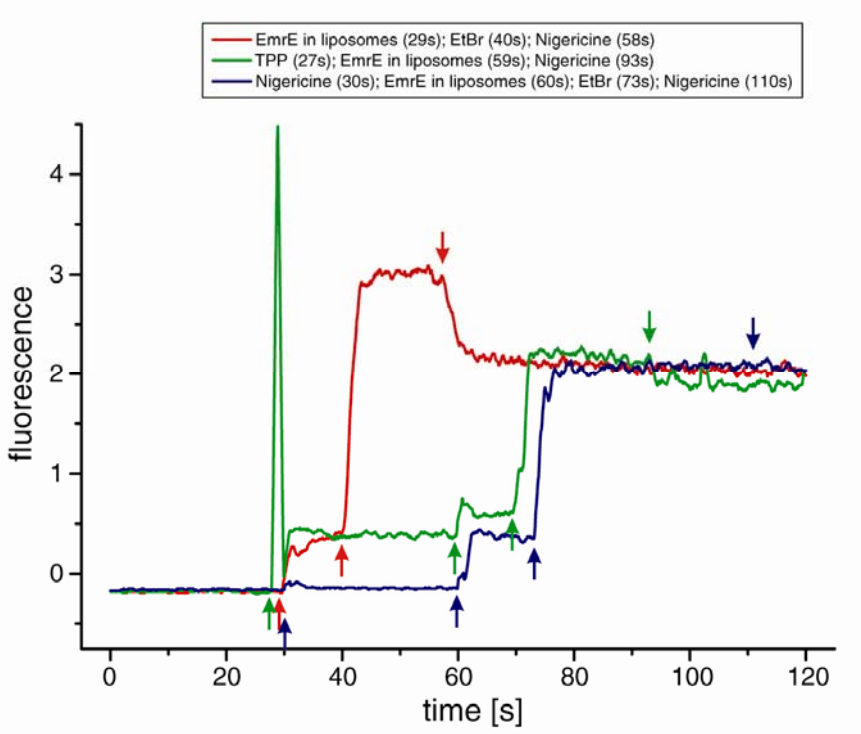


Fig. 11 Ethidium transport assay of EmrE proteo-liposomes. Transport of ethidium (EtBr) into reconstituted EmrE proteoliposomes in 15 mM Tris/Cl, pH 9.5, 2 mM dithiothreitol, 150 mM KCl was measured by an increase in fluorescence at excitation and emission wavelengths of 545 and 610 nm, respectively. Ten micro liters of proteoliposomes, approximately 140 nM EmrE, were added after 30 or 60 s. If appropriate, substances were added at the following final concentrations: TPP (50 µM), ethidium bromide (2.5 µM) and nigericine (5 µg/mL). Arrows indicate the time points of addition. (Modified after (P01)).

4.3. Synergies of cell-free expression and high field NMR spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool to investigate the structure and function of proteins. An indispensable prerequisite for NMR spectroscopy is the large scale preparation of samples that have been labeled with the stable isotopes ^{2}H, ^{13}C and ^{15}N. Depending on the complexity of the spectra and the strategy chosen for their assignment, NMR samples may be labeled uniformly or selectively with ^{13}C and

^{15}N and, in particular for larger proteins, with ^{2}H. The uniformly labeling of a protein requires the growth of the expression strain in defined medium supplemented with precursors that are labeled with the desired isotope. For amino acid specific labeling approaches, usually *E. coli* strains containing auxotrophic mutations corresponding to the provided labeled amino acid have to be employed. Thus, conventional cellular labeling techniques are often accompanied by low protein yields due to retarded growth in a minimal medium. CF protein expression overcomes those drastic expression limitations and especially in the case of IMP synthesis it allows the fast and efficient NMR sample preparation. Complete label incorporation is ensured as no unlabelled amino acids are present in the reaction. Moreover, no auxotrophic mutations are needed as any amino acid can be simply replaced in the reaction mixtures by a labeled derivative.

For NMR spectroscopy the often large size of IMPs resulting in slow rotational tumbling and broad line width constitutes a significant problem. This problem is aggravated by the fact that membrane proteins have to be solubilized in micelles, which contribute considerably to the overall molecular weight of the protein/ micelle complex. Furthermore, α-helical proteins tend to display narrower chemical shift dispersions, resulting in peak overlap, and the transmembrane sections of these proteins consist predominantly of hydrophobic amino acids often leading to clustering of identical amino acids with very similar chemical shifts. Solution NMR using ultra high-field spectrometers (800 and 900 MHz ^{1}H frequencies) reduces those difficulties, by increased sensitivity and higher resolution.

The size limitation of detergent solubilized proteins was recently addressed by using specific detergents. A systematic screen for liquid-state NMR compatible detergents with regard to maximal signal detection, high spectral resolution as well as long sample lifetimes, distinguished the class of lyso-phosphoglycerols including LMPG and LPPG, providing outstanding benefits [23, 26]. LMPG micelles obviously do not restrict the tumbling rates of the inserted solubilized IMP and therefore do not result in excessive line-broadening. Interestingly we could show that LMPG and its derivatives are in addition highly suitable for the solubilization of CF produced IMP precipitates. Obviously several limitations in IMP analysis by solution NMR can now be addressed by virtue of CF expression, giving the structural analysis of IMP by high field NMR analysis new hope.

4.3.1. [^{15}N,^{1}H]-TROSY-HSQC NMR measurements of IMP samples

The CF expression technique for IMPs allows the fast production of several mg of isotopically labeled samples. Here a typical NMR sample can be produced within 24 h since no complicated cell disruption and purification schemes are involved. In fact, since the synthesized target protein is the only labeled macromolecule in the RM, NMR spectra can, in principle, be measured without any chromatographic purification steps directly in the RM [19]. However, the insoluble CF expression mode of IMPs results in almost pure precipitates. Those precipitates can be solubilized in organic solvents or by using mild detergents like LMPG. The small multi drug transporter EmrE was selectively ^{15}N-alanine and ^{15}N-glycine labeled, insoluble CF expressed and dissolved in a mixture of deuterated chloroform ($CDCl_3$), deuterated methanol (CD_3OH) and water in a ratio of 6 : 6 : 1 and analyzed by NMR (Fig. 12 A).

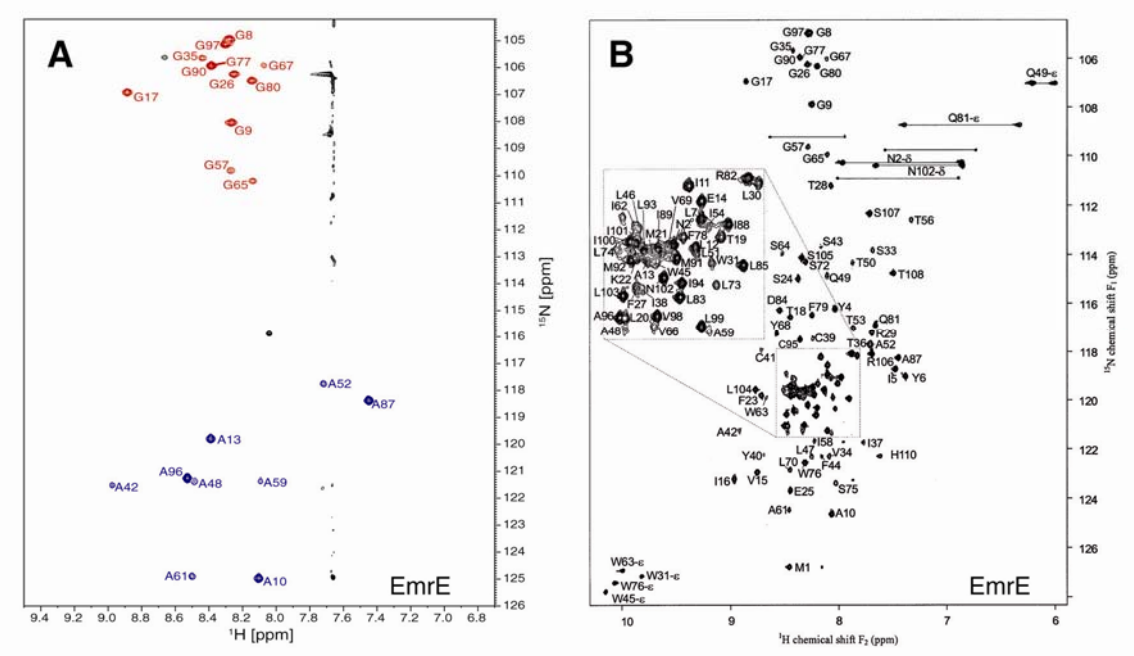


Fig. 12 Liquid-state NMR of EmrE.
A, [^{15}N,^{1}H]-TROSY-HSQC spectrum of selectively ^{15}N-alanine (blue) and ^{15}N-glycine (red) labeled EmrE, CF expressed as precipitates and solubilized in $CDCl_3/CD_3OH/H_2O$ (6:6:1) plus 200 mM ammonium acetate (pH 6.2) 10 mM dithiothreitol. The spectra of 0.1 mM EmrE was measured at 15°C with a 600 MHz spectrometer equipped with cryogenic ^{1}H [^{13}C/^{15}N] triple-resonance probe. **B**, *in vivo* expressed EmrE (taken from [97] recorded under the same conditions), for comparison.

The chemical shift dispersion of the corresponding glycine and alanine residues in a [^{15}N,^{1}H]-TROSY-HSQC spectrum could be assigned as they are completely identical to

the published spectrum of EmrE that has been uniformly labeled by *in vivo* expression and measured under the same conditions [97] (Fig. 12 B). This demonstrates that CF expressed and *in vivo* expressed EmrE are completely identical.

The combination of CF IMP expression levels with the use of high field NMR spectrometers for the first time allowed to record spectra of proteins that have not even been expressed in cellular systems yet. The 26 kDa amino acid transporter YfiK comprising 6 TMS was CF uniformly labeled with ^{15}N-amino acids and analyzed in LMPG micelles (Fig. 13 A). The bacterial protein TehA is a 36 kDa membrane protein that shows limited homology to the SMR family [24]. *In vivo* it confers resistance to tellurite compounds as well as to lipophilic cation dyes. *In vivo* experiments have demonstrated that a 24 kDa fragment of TehA (ΔTehA), containing seven out of 10 predicted TMS, shows the same biological effects as the full length protein [24]. The NMR sample preparation for ΔTehA was systematically optimized. Whereas the LMPG solubilized CF ΔTehA precipitate in standard phosphate buffer showed only a reasonable good spectral quality, the soluble CF expression of ΔTehA in presence of Brij78 followed by buffer exchange to LMPG during Ni-chelate purification resulted in enhanced spectral quality.

However, insoluble expressed protein that was additionally purified by Ni-chelate chromatography and subsequently buffered in MES-TRIS buffer pH 6, optimized for NMR measurements [98] resulted in so far best NMR spectral quality (Fig. 13 B). This sample behavior of ΔTehA enabled collection of sufficient data to start the assignment for a structural analysis. However, there arise extreme difficulties as α-helical IMPs show a relatively narrow chemical shift dispersion and peak overlap in the centre of the spectrum. This overlap might be overcome by using selectively amino acid labeled samples, indicating all one type amino acids in this region (Fig. 13 C-H). In addition the unique possibility of the combinatorial labeling possibility by CF expression, explained previously, can be used to identify distinct amino acids in the protein backbone.

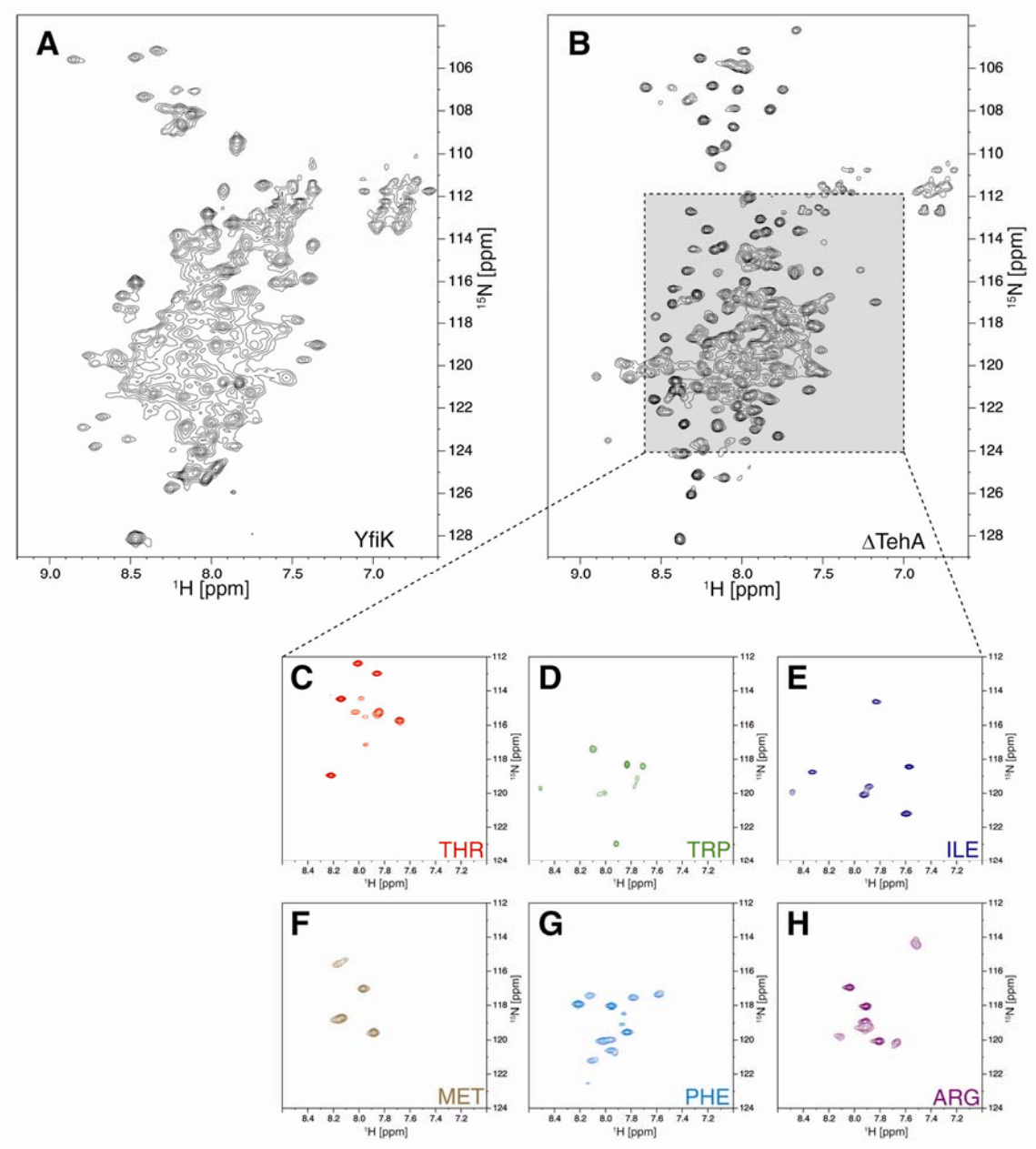


Fig. 13 Liquid-state NMR of the CF produced transporters YfiK and ΔTehA.
[^{15}N-^{1}H]-TROSY-HSQC spectra of cell-free expressed 15N labeled transporters, expressed as precipitates and dissolved in 25 mM potassium phosphate buffer (pH. 6.0) containing 5% LMPG (A, C-H) or solubilized in 1% LMPG, purified by Ni-NTA chromatography and equilibrated in 20 mM MES/Bis-Tris (pH 6.0) with 2% LMPG (B). **A**, U-^{15}N labeled YfiK; **B**, U-^{15}N labeled ΔTehA; **C**, ^{15}N-threonine labeled ΔTehA; **D**, ^{15}N-tryptophane labeled ΔTehA; **E**, ^{15}N-isoleucine labeled ΔTehA; **F**, ^{15}N-methionine labeled ΔTehA; **G**, ^{15}N-phenylalanine labeled ΔTehA; **H**, ^{15}N-arginine labeled ΔTehA. The spectra were taken with protein concentrations of 1 mM for YfiK (A), 0.1 mM for ΔTehA (B) and 0.9 mM for selectively ΔTehA samples (C-H) at 40°C on a 700 MHz (A), 900 MHz (B) or on a 800 MHz NMR spectrometer (C-H), all equipped with cryogenic ^{1}H [^{13}C/^{15}N] triple-resonance probes.

4.3.2. Selectively cell-free isotopically labeling of IMPs and combinatorial labeling

The first step in structural analysis of a protein by solution NMR is the assignment of the protein backbone resonances. Unfortunately α-helical proteins tend to display narrower chemical shift dispersion [26, 99, 100] as compared to that of proteins containing β-sheets. Since the majority of IMPs consist exclusively of α-helices, their NMR spectra tend to show a significant degree of peak overlap. Along with line broadening due to protein/ detergent complexes sizes, these disadvantages of IMPs pose a considerable challenge for the chemical shift assignment, suggesting that new strategies might be necessary in order to make backbone assignment of IMPs as routine.

A simple way to overcome the overlap problem is selective isotopically amino acid labeling of the target protein identifying all signals of one amino acid type. However, this does not usually result in sequence specific assignments. This difficulty can now be approached by a mixture of amino acid specific and combinatorial labeling strategies [101]. Here, the IMP is simultaneously labeled with selected ^{15}N-enriched amino acids in combination with distinct ^{13}C-labeled amino acids (Fig. 14). The application of a transverse relaxation optimized spectroscopy (TROSY) version of a [^{15}N,^{1}H] heteronuclear single quantum correlation (HSQC) experiment in combination with two-dimensional [^{15}N,^{1}H]-TROSY-HN(CO) (HNCO) spectra helps to identify only those ^{15}N-labeled amino acids that were N-terminally preceded by a ^{13}C-labeled amino acid type. This strategy enables the unambiguous identification of consecutive amino acids pairs that can be subsequently used as anchor points for further backbone resonance assignments (Fig. 14). This combinatorial labeling strategy of IMPs is just possible due to the unique properties of CF expression. Whereas amino acid type selective labeling is also possible in auxotrophic bacteria, for combinatorial labeling multiple auxotrophic strains in almost any combination would be needed. The combination of enormous expression levels of IMPs, unique labeling possibilities coming along with no cross labeling problems [19], and the fact that just small amounts of isotopically labeled amino acids are needed makes CF expression promising for the structural analysis of IMPs by high field NMR.

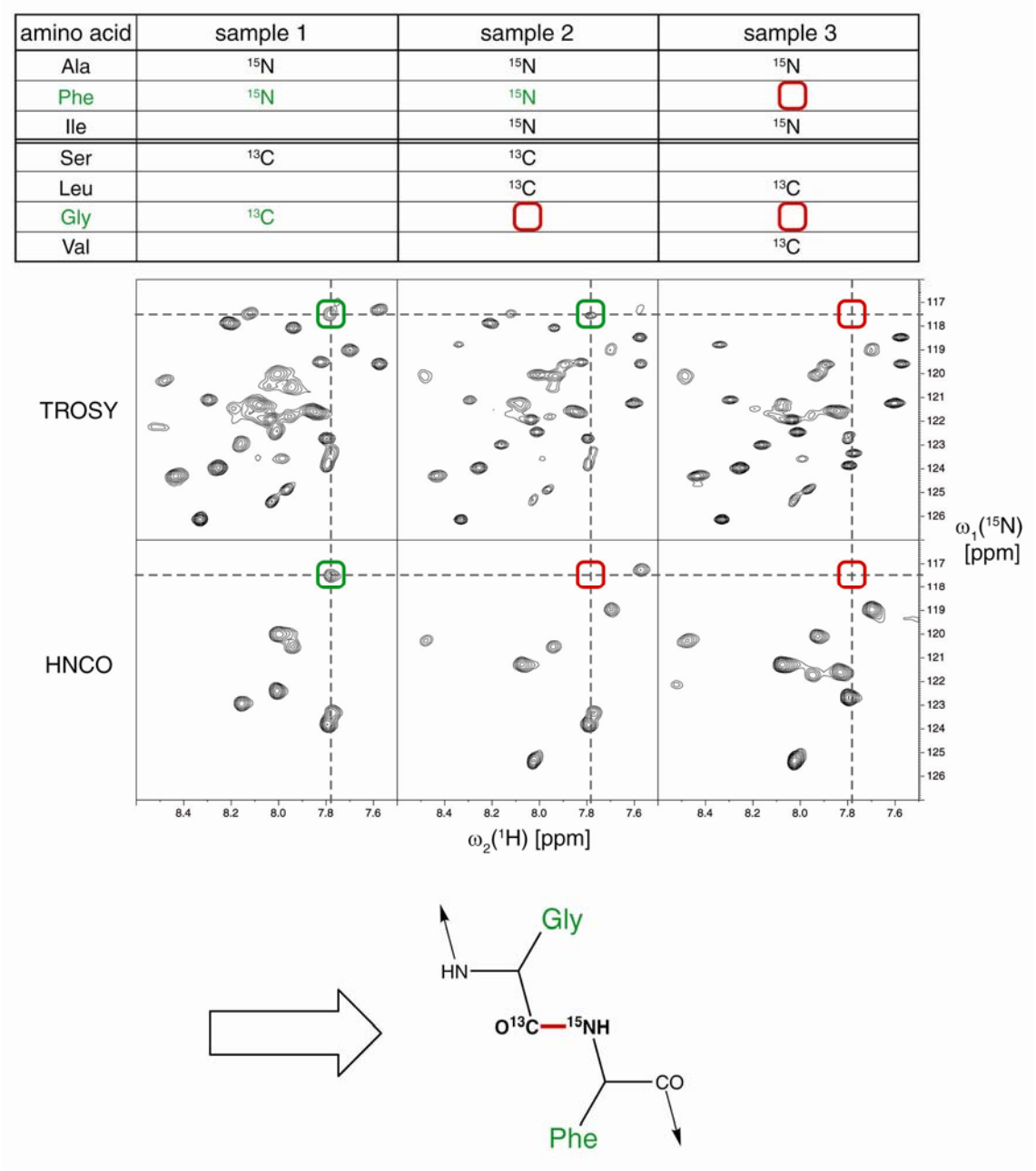


Fig. 14 Combinatorial labeling approach.
Example of a combinatorial labeling scheme with three differentially ^{15}N and ^{13}C labeled samples of the tellurite transporter ΔTehA. The labeling scheme is shown in the table on top. In the [^{15}N,^{1}H]-TROSY spectra, correlations of the ^{15}N-labeled amino acids are visible. The crossed lines indicate the peak to become identified. The peak at position 117.5/7.78 must be due to a phenylalanine as it is present in samples 1 and 2 (green circles) but not in sample 3 (red round edge square). The corresponding HNCO spectra show the amide cross peaks after carbonyl transfer (big arrow) and indicate that the preceding residue of this phenylalanine must be a glycine, as no cross peaks are visible in sample 2 and 3 (red round edge squares) without ^{13}C-glycine. The analyzed phenylalanine residue can now be localized in a Gly-Phe pair in the primary sequence which in that example was identified as Phe97 of ΔTehA. All spectra were recorded at a Bruker Avance 600 MHz spectrometer. (Modified after (P 05)).

4.3.3. Protein-detergent interactions analyzed by NMR

Once the amino acids of the protein backbone resonances are assigned, NMR spectroscopy offers versatile possibilities for further analyzes like the determination of protein-detergent interactions. Nuclear Overhauser enhanced and exchange spectroscopy (NOESY) allows the determination of spatial proximity of protons. Thus, amino acids of ΔTehA that are close to atoms of the detergent LMPG could be identified (Fig. 15).

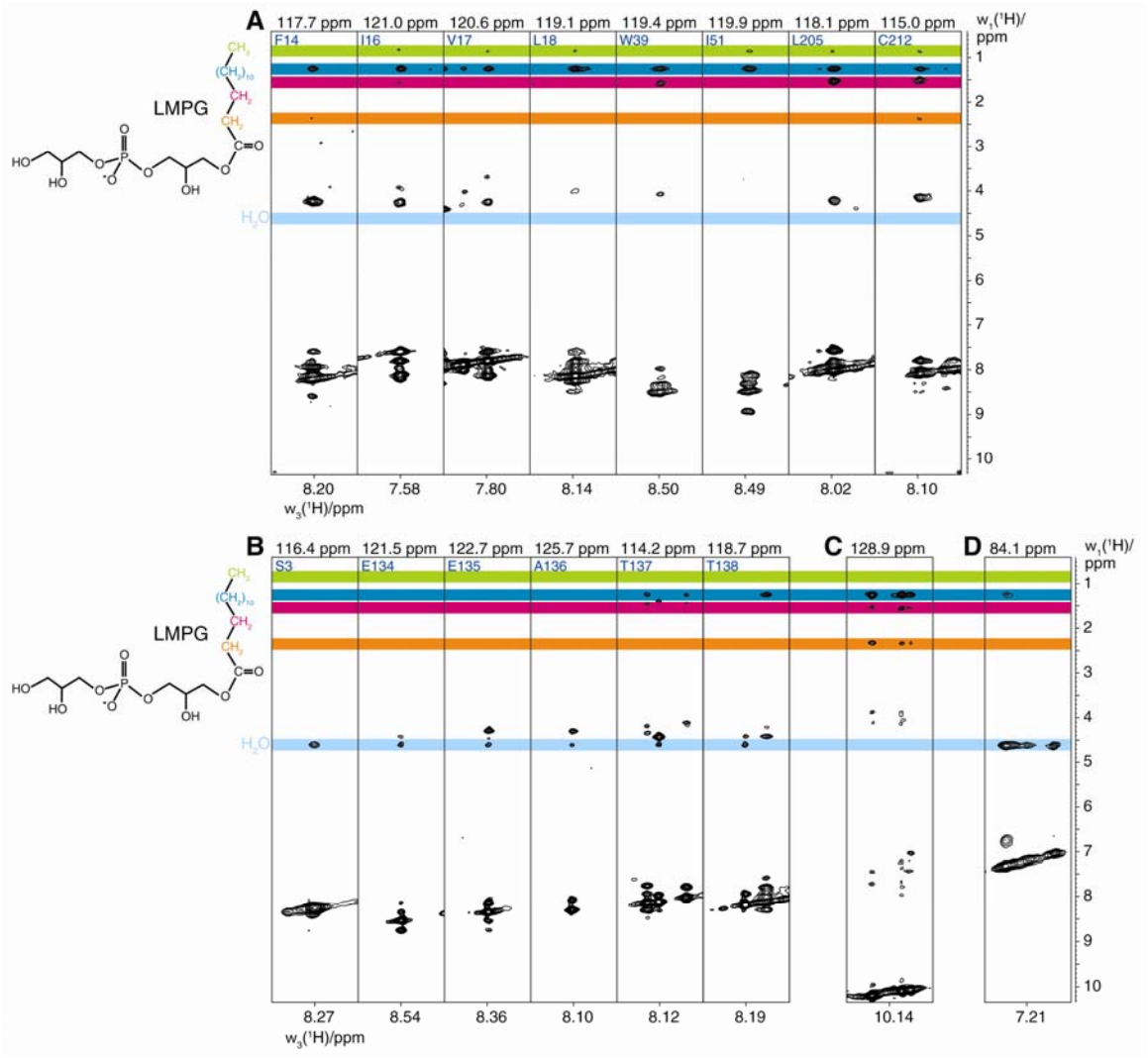


Fig. 15 Protein-LMPG interactions.
F1-F3 Strips are taken from a 3D NOESY-[^{15}N,^{1}H]-TROSY spectrum (250 ms mixing time, 900 MHz) of U-^{2}H^{15}N-labeled ΔTehA in 5% LMPG at ^{15}N positions indicated at the top of each slice. Strips are grouped according to: **A**, backbone amides of residues in helical regions; **B**, backbone regions of residues in loop regions; **C**, tryptophan side-chain N^{e1} protons; **D**, arginine side-chain N^{e} protons. Cross beaks outside the marked positions are intraresidual NOEs either involving exchangeable protons or incompletely deuterated α-carbons.
(The figure was kindly provided by Frank Löhr).

It can further be demonstrated that backbone amides of residues in helical regions are in direct proximity to the hydrophobic part of LMPG (Fig. 15 A), whereas backbone residues in loop regions show NOEs to water (Fig. 15 B). This information supports the determination of α-helical and loop regions of the protein.

4.3.4. Paramagnetic Resonance Enhancement (PRE) measurements

Whereas the secondary structure is mainly based on the $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts as well as on the pattern of sequential and medium-range NOEs in the ^{15}N-edited [^{1}H,^{1}H]-NOESY spectrum, investigations of the three-dimensional (3D) structure relies on different strategies. A limited chemical shift dispersion observed in the spectra of α-helical IMPs limits the assignment of side chain resonances to very few residues only. In addition, the high molecular weight of the protein/micelle complexes requires deuteration, which eliminates most of the protons that can be used in classical ^{1}H-^{1}H NOE-based structure determination procedures. Another possibility getting distance constrains for IMPs is the use of paramagnetic relaxation enhancement (PRE) experiments [8]. This type of distance constraints has been developed with soluble proteins that are too big for classical structure determination procedures [27]. In these experiments the line broadening effect of a paramagnetic tag, like 1-oxy(-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSL)) that is attached to a protein through a specific cysteine side chain on the surrounding amide proton is analyzed. Since the line broadening is proportional to $1/r^6$, a quantitative measurement of the line broadening can be used for obtaining distance constraints. The structure of the membrane-associated protein Mistic [26] as well as the refinement of the outer membrane β-barrel protein OmpA [102] already took advantage of PRE derived distance information.

In order to use PRE to obtain distance constrains of ΔTehA, all three naturally occurring cysteins where mutated to alanins and at specific locations, mostly at the beginning or the end of helices, cysteine residues have been reintroduced by quick change PCR. As uniformly ^{15}N amino acid labeled ΔTehA show significant peak overlap in the [^{15}N,^{1}H]-TROSY-HSQC spectrum, several selectively amino acid labeled samples where prepared by CF expression. Figure 16 shows as a comparison of the [^{15}N,^{1}H]-TROSY-HSQC spectra of ΔTehA, labeled with a paramagnetic tag at amino acid position 11, 98 and 193 and the same sample after reducing the spin label to a diamagnetic compound with ascorbic acid. Some signals, in case of paramagnetic labeled protein, are completely absent in the paramagnetic species whereas others show significant line broadening.

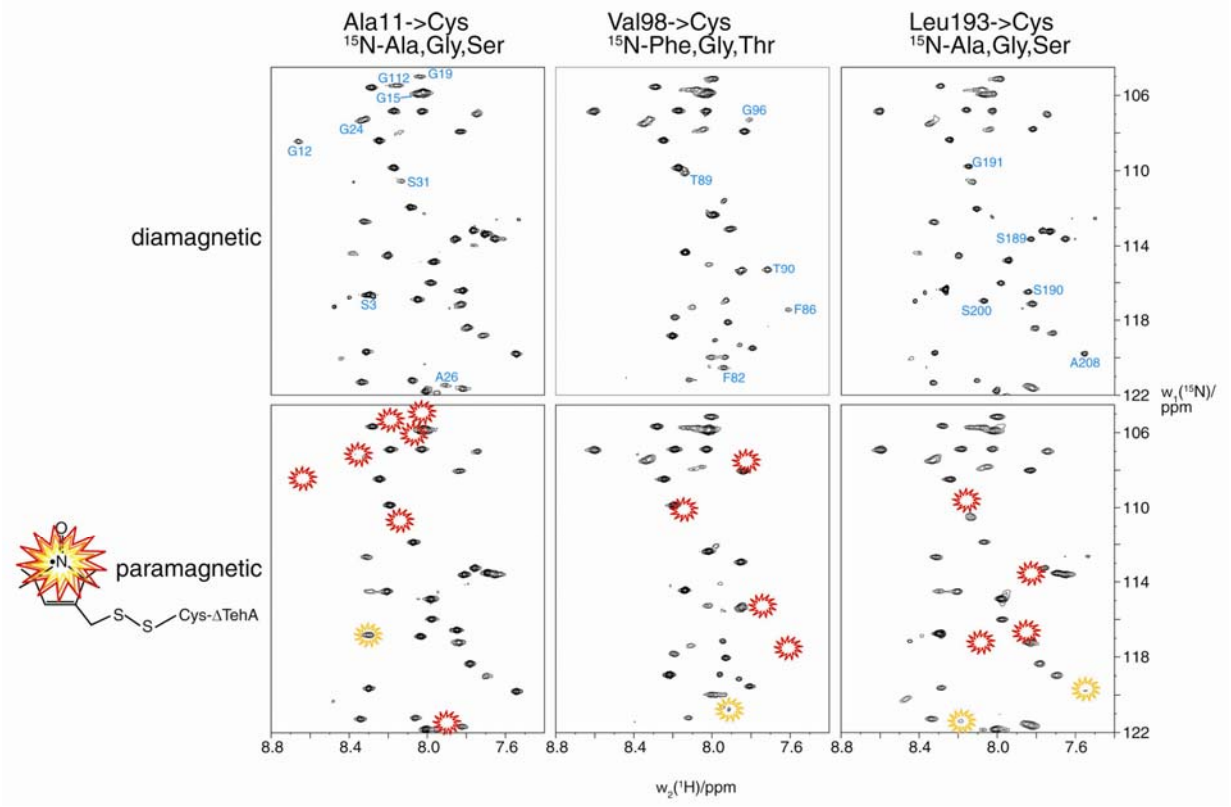


Fig. 16 Paramagnetic labeling of ΔTehA.
Comparison of [^{15}N,^{1}H]-TROSY-HSQC spectra of ΔTehA in its reduced, diamagnetic form (upper spectra) and in its paramagnetic form with a spin label (MTSL) attached (lower spectra). The significant overlap of the fully labeled spectrum makes amino acid-type selective labeling necessary. The site of the spin label attached as well as the type of labeling pattern is indicated on the top. Peaks that are completely absent in the paramagnetic spectrum are marked with a red symbol and those that show significant line broadening with a yellow symbol. The spectra were measured on an 800 MHz instrument. The assignment of these affected beaks is shown in the diamagnetic spectrum. (The figure was kindly provided by Frank Löhr and has been modified after (P 10)).

4.3.5. Preliminary structural data of ΔTehA

The unique advantages of CF IMP expression and accompanied labeling possibilities in combination with high field NMR spectroscopy allowed the assignment of 85% of the amide backbone resonances of ΔTehA [Trbovic, Diplomarbeit 2005]. Using the PRE technique allowed the determination of approx. 120 meaningful distance constrains as a pair of upper and lower limit for an individual defined amide-amide distance. This information combined with 205 distance constraints based on unambiguous NOEs from 4D-[^{15}N,^{1}H]-HMQC-NOESY-[^{15}N,^{1}H]-TROSY and 3D-[^{15}N,^{1}H,^{15}N]-NOESY-[^{15}N,^{1}H]-TROSY spectra, defining inter-helical distances, allowed the calculation of a preliminary backbone model of ΔTehA [Fig. 17 B]. The model comprises all seven

predicted α-helices [Fig. 17 A], of which helix no. 4 (green) seems to be flexible and not transmembrane oriented.

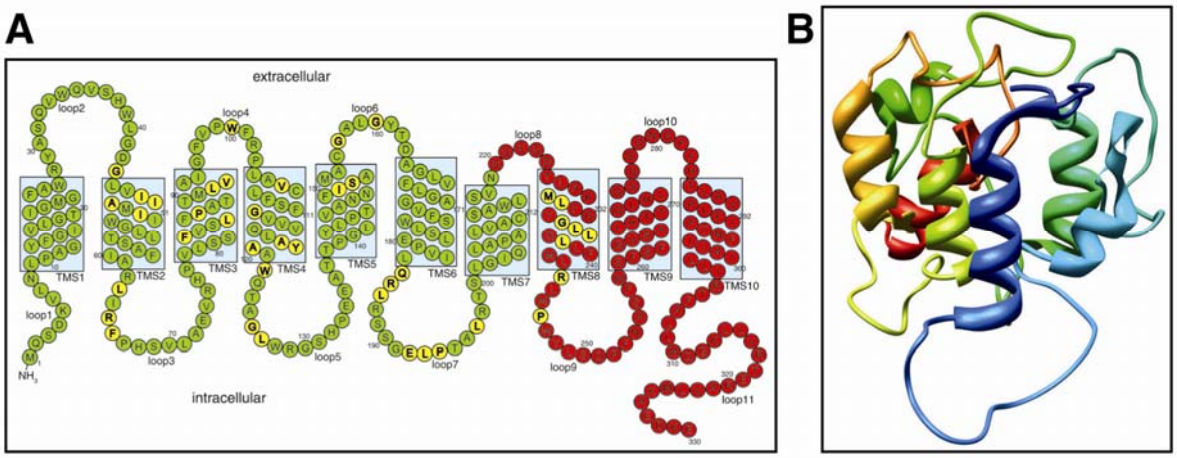


Fig. 17 Secondary and preliminary structure of the putative tellurite transporter ΔTehA.
A, The proposed secondary structure of TehA is shown. The 10 transmembrane segments of TehA (green, red) are shown. The C-terminal truncated version ΔTehA is indicated in green. A homology alignment to the related multi-drug-transporter EmrE regarding conservative replacements is indicated in yellow. **B**, Preliminary model of ΔTehA as derived from dyana 1.05 simulated annealing protocol. A first model of a protein with seven transmembrane helices could derive from 325 distance constraints using dyana 1.05 simulated annealing in vacuum. This calculated model symbolizes the proposed averaged topology of ΔTehA in LMPG micelles. N-terminus (blue) and C-terminus (red) are oriented on opposite sites of a assumed conformational space of a membrane. (This model was calculated by Alexander Koglin).

This shows the first model of an α-helical IMP of this size, containing seven TMS and analyzed by solution NMR in detergent micelles.

5. Summary of publications

5.1. Establishment of high level cell-free expression of IMPs

In the following section, the publications that contribute to this work will be reviewed shortly. In addition I will highlight my contribution to these publications. The aim of this work was to establish a CF system for preparative IMP expression for functional and structural analysis. The individual CF expression method used was further optimized from the system that has been established in Frankfurt during my Diploma Thesis in the year 2002 (Development and set-up of a cell-free expression system for preparative protein synthesis).

5.1.1. Method development

5.1.1.1. High level cell-free expression and specific labeling of IMPs (P 01)

C. Klammt, F. Löhr, B. Schäfer, W. Haase, V. Dötsch, H. Rüterjans, C. Glaubitz and F. Bernhard: **High level cell-free expression and specific labeling of integral membrane proteins.** *Eur. J. Biochem.*, **271**, 568-580. (2004)

This article was judged to be the best one published in the FEBS Journal during the year 2004 and honored with the FEBS Journal Prize for Young Scientists award 2004. In this article we demonstrate for the first time the high level expression of IMPs in a CF system. Using an optimized protocol the *E. coli* small multi drug transporters EmrE and SugE and even the heavy metal transporter TehA and the cysteine transporter YfiK, both not overexpressed in *E. coli*, were produced at high levels of up to 2.7 mg per single mL of RM. As no hydrophobic compartment was supplemented to the CF system all IMPs were insoluble synthesized in the RM. Those precipitates are found to solubilize in already mild detergents, and modifications to the solubilization procedure yielded in almost pure protein. Solubilized IMPs were successfully reconstituted in proteoliposomes, demonstrated by freeze fracture electron microscopy (EM). The function of CF expressed, solubilized and reconstituted EmrE was analyzed by the specific binding of ethidium. In addition the CF expression technique allowed the efficient amino acid specific labeling of the synthesized proteins and NMR analysis indicated a correctly folded conformation of the proteins.

Major results:

- Protocol for an optimized individual CF expression system.
- The preparative scale expression of diverse IMPs (EmrE, SugE, TehA and YfiK) using an individual CF system is demonstrated for the first time.
- It is observed that CF IMP synthesis produce insoluble proteins that can easily be resolubilized in already mild detergents, and purified by modifications of the solubilization procedure. This makes these precipitates completely different to inclusion bodies, known from cellular expression systems.
- IMP precipitates solubilized in mild detergents can functionally be reconstituted into proteoliposomes, analyzed by freeze fracture EM and a specific ethidium transport assay for EmrE.
- The secondary structure of IMPs is analyzed by circular dichroism (CD) spectroscopy, solubilized in various detergents.
- CF ^{15}N-isotopic labeling allows the solution NMR analysis of IMPs in detergent micelles.

My contribution to this publication:

I have established the CF system in Frankfurt and optimized for the high level expression of IMPs. I have labeled the analyzed IMPs with ^{15}N-isotopes for NMR analysis, performed by Frank Löhr. I reconstituted the CF expressed and detergent solubilized proteins, I carried out the EmrE ethidium transport assay and have done the CD analyzes. The freeze fracture EM analysis was done by Winfried Haase.

5.1.1.2. Soluble cell-free expression of IMPs in presence of detergents (P 02)

C. Klammt, D. Schwarz, K. Fendler, W. Haase, V. Dötsch and F. Bernhard: **Evaluation of detergents for the soluble expression of α-helical and β-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system.** *FEBS Journal,* **272**, 6024-6038. (2005)

This paper evaluates 24 of the most commonly used detergents for their suitability as supplement to CF systems for the direct soluble expression of IMPs. In order to give a comprehensive overview, the soluble CF synthesis of three structurally very diverse IMPs, the α-helical small multi drug transporter EmrE, the β-barrel type nucleoside transporter Tsx and the eukaryotic vasopressin 2 receptor (V2R), a member of the G-protein coupled receptor (GPCR) family, were analyzed regarding to their soluble expression. Some members of the long chain poly-oxyethylene-alkyl-ether family were identified to be most suitable for the soluble expression of all three types of membrane proteins. Furthermore, the yield of soluble expressed membrane protein was found to remain relatively stable above a certain detergent threshold concentration. In addition, for insoluble CF IMP synthesis, several detergents were analyzed for their ability in solubilizing precipitates of these structurally divers proteins. Here the detergent 1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)] (LMPG) was found to be most effective for the resolubilization. Finally, the high-level CF expression of a β-barrel type membrane protein in a functional form was reported for the first time and a structural and functional variation of the analyzed membrane protein in dependence of the CF expression mode and the supplied detergent was demonstrated.

Major results:

- The evaluation of 24 different commonly used detergents for their suitability for soluble CF expression of structurally very different IMPs revealed that some members of the Brij-family (Brij35, Brij58, Brij78 and Brij98) give outstanding results.
- Above a certain detergent threshold concentration the yield of soluble expressed protein remains relatively stable.
- The detergent LMPG is the most effective for the resolubilization of diverse IMP precipitates, obtained form insoluble CF IMP expression.
- The functional and preparative expression of a β-barrel type protein is demonstrated for the first time and the functionality can be related to the mode of CF expression and the type of detergent used.
- The GPCR V2R is soluble CF expressed in amounts up to 6 mg of one mL of the RM in the presence of Brij58 and Brij78. This is the highest expression rate of a GPCR protein ever reported.
- CF insoluble produced V2R show homogenous insertion in reconstituted proteoliposomes, analyzed by freeze fracture EM.

- The secondary (2D) structure of the insoluble and soluble expressed proteins Tsx and V2R, analyzed by CD-spectroscopy, differ from each other.

My contribution to this publication:

I have cloned and performed all analyzes regarding the *E. coli* small multi drug transporter EmrE and the GPCR V2R, whereas my coauthor Daniel Schwarz took care of all Tsx experiments, including the activity assay. The freeze fracture EM was performed by Winfried Haase.

5.1.2. Reviews

Several review articles about CF expression of IMPs have been prepared highlighting different aspects of CF IMP synthesis.

5.1.2.1. Cell-free production of IMPs on a preparative scale (P 03)

C. Klammt, D. Schwarz, V. Dötsch and F. Bernhard: **Cell-free production of integral membrane proteins on a preparative scale.** *Meth. Mol. Biol.*, accepted. (2006)

This article focuses on the high level CF expression of IMPs by using an individual coupled transcription/ translation system. It describes in detail the set-up and optimization of the CF expression technique in order to obtain the maximum yield of recombinant proteins. The protocol can be used for the expression of soluble IMPs as well as for their production as precipitates. In addition, it provides detailed protocols for the solubilization of CF expressed detergent precipitates and for further reconstitution.

Major results:

- Detailed description for the establishment of an individual CF expression system for high level expression of IMPs, including stock concentration tables and a pipetting protocol.
- Protocols for solubilization of CF produced IMP precipitates and further reconstitution.
- Guide for the soluble CF expression of IMPs and the resolubilization of the three proteins EmrE, V2R and Tsx, including 11 different detergents.

- Successful reconstitution of CF insoluble expressed YfiK solubilized in detergent, analyzed by freeze fracture EM.

My contribution to this work:

I have supplied the detailed protocols for the high level expression of IMPs, the solubilization of precipitates and the reconstitution. I provided all of the figures and reconstituted YfiK for freeze fracture analysis. Regarding the soluble CF expression of IMPs, I have analyzed EmrE and V2R, whereas Daniel Schwarz investigated the β-barrel type nucleoside transporter Tsx.

5.1.2.2. Cell-free expression as an emerging technique for the large scale production of IMP (P 04)

C. Klammt, D. Schwarz, F. Löhr, B. Schneider, V. Dötsch and F. Bernhard**: Cell-free expression as an emerging technique for the large scale production of integral membrane protein.** *FEBS Journal*, **273**, 4141-4153. (2006)

This article gives an overview about the current state of the art of CF IMP expression with special focus on technical aspects as well as on the functional and structural characterization of CF produced IMPs. Besides the structural analysis of CF produced IMPs, applications like oligomerization studies of CF expressed IMPs and high throughput expression screening of IMPs are discussed.

Major results:

- Comparison of the two expression modes of insoluble and direct soluble synthesis of IMPs in CF systems.
- Detailed compilation of IMPs synthesized in preparative scale CF expression systems.
- Detailed overview comparing the activities of CF synthesized and conventionally *in vivo* expressed IMPs.
- High resolution [^{15}N,^{1}H]-TROSY-HSQC NMR spectra of the CF uniformly ^{2}H^{15}N-labeled cysteine transporter YfiK in LMPG micelles measured at a 900 MHz spectrometer.

My contribution to this work:

I provided all the figures and tables for the manuscript and have uniformly $^{2}H^{15}N$-labeled YfiK for the high field NMR analysis in LMPG micelles, performed by Frank Löhr.

5.1.2.3. Cell-free synthesis of IMPs for structural studies (P 05)

C. Klammt, D. Schwarz, B. Schneider, F. Löhr, I. Lehner, C. Glaubitz, V. Dötsch and F. Bernhard: **Cell-free expression of integral membrane proteins for structural studies.** In: *Cell-free expression techniques,* A. Spirin (ed.), Wiley-VCH, Weinheim, submitted. (2006)

This manuscript compares the two already established CF IMP expression modes of insoluble and direct soluble IMP expression with a third possible but not yet sufficient analyzed mode of direct supplementation of preformed liposomes. In addition it demonstrates the variety of different successfully high level CF expressed IMPs. In addition several high resolution solution NMR spectra, including the proteins Yfik, SecE and ΔTehA of uniformly and single amino acid labeled IMPs and first solid state NMR experiments of selectively labeled CF produced EmrE are shown.

Major results:

- Comparison of the two already established CF IMP expression modes and a potential third method using supplemented preformed liposomes for direct reconstitution of IMPs.
- Detailed description of detergents used for the solubilization of IMPs and the soluble expression in presence of detergents.
- Detailed presentation showing the variety of different high level expressed IMPs including four different GPCRs.
- Discussion of state of the art achievements of combined approaches CF expression and NMR. Several high resolution [^{15}N,^{1}H]-TROSY-HSQC spectra of uniformly and selective amino acid labeled SecE, YfiK and ΔTehA are presented.
- Description of combinatorial labeling schemes using CF expression of IMPs.

My contribution to this work:

I supplied all figures except of the solid state NMR information (6 B), that was supplied by my coauthor Ines Lehner. The detailed detergent table was prepared by Daniel Schwarz. The proteins SecE and YfiK were expressed by Solmaz Sobhanifar and the NMR analysis was done by Frank Löhr.

5.1.2.4. Preparative scale cell-free expression systems for functional and structural studies (P 06)

D. Schwarz, C. Klammt, A. Koglin, F. Löhr, B. Schneider, V. Dötsch and F. Bernhard: **Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies.** *Methods*, in press. (2006)

This manuscript presents detailed protocols for the CF production of IMPs in different modes and summarizes the current knowledge of this technique. Here, the production of soluble and functionally folded IMPs in presence of detergents is emphasized. In addition, the advantages of CF expression for the structural analysis of IMPs especially by liquid state NMR is highlighted and new strategies for structural approaches are discussed.

Major results:

- Detailed protocols for CF expression including CF extract preparation, design of DNA templates and reaction conditions.
- Detailed protocols for CF IMP expression in both modes as insoluble protein or direct soluble synthesized in detergent micelles.
- Comprehensive summary of all published results with CF expressed IMPs in the presence of detergents.
- CF expression kinetics of the β-barrel nucleoside transporter Tsx.
- Comparison of ΔTehA expressed in both modes and buffered in LMPG by [^{15}N,^{1}H]-TROSY-HSQC NMR spectra.
- Optimization of soluble CF Tsx expression regarding the detergent concentration.

- Discussion of combinatorial labeling strategies for the rapid assignment of CF produced IMPs.

My contribution to this work:

I supplied the detailed table of CF soluble expressed IMPs reported in the literature, prepared the ΔTehA samples for NMR analysis and drew the figure explaining the combinatorial labeling scheme. The rest of the manuscript was done by the first author Daniel Schwarz.

5.1.2.5. DNA recombination and protein expression systems (P 07)

F. Bernhard, C. Klammt, and H. Rüterjans: **DNA recombination and protein expression.** In: *Comprehensive Medicinal Chemistry II*, in press. (2006)

This book chapter reviews the currently most promising expression systems in bacterial and eukaryotic cells as well as in CF environments. It intends to serve as a guideline for making decisions in order to receive the highest yields of recombinant proteins. Critical details and features that have to be considered upon the *in-vitro* recombination of expression vectors and during the process of protein production are addressed. In addition, key elements of expression systems are discussed individually and modified techniques that have been approved for the production of distinct and difficult groups of proteins; e. g. membrane proteins or disulfide bonded proteins are especially highlighted. A special focus is also the preparative scale CF production of recombinant proteins as this relatively new technique provides several unique characteristics with a high potential for the future.

Major results:

- Reviews the variety of promoters and control mechanisms of gene expression in detail in combination with codon usage and regulatory DNA sequences important for protein expression.
- Comprehensive overview of translation fusion constructs for optimized protein expression.
- Discussion and comparison of the currently most promising expression systems including the *E. coli*, yeast, mammalian, viral (baculovirus and semliki forest virus expression vector systems) and CF expression systems in detail.

- Description of applications in *E. coli* expression, including inclusion body formation, co-expression of chaperones, solubility tags, expression of IMPs and disulfide bonded proteins.
- Comparison of all preparative scale protein expression techniques for the optimized expression of recombinant proteins.
- Detailed presentation of CF expression systems, including the preparation of CF extracts and the advantages in CF IMP synthesis and the expression of disulfide bridged proteins.

My contribution to this work:

I supplied all figures and reviewed the mammalian expression system, (like baby hamster kidney (BHK), Chinese hamster ovary (CHO), human embryonic kidney (HEK) and simian fibroblasts (CV-1 cells) transformed by SV40 (COS) cell lines), the viral expression systems, including the baculovirus and semliki forest virus expression systems, and the CF expression system paragraphs.

5.2. Unique labeling possibilities for IMPs by cell-free expression

Nuclear magnetic resonance (NMR) spectroscopy has experienced a tremendous growth over the past decade, and has been developed as a widely used technique with enormous potential for the study of the structure and dynamics of biological macromolecules. In particular; structural analysis of proteins has greatly benefited from recent advancements of high field solution NMR. Improvements in the hardware and the development of NMR spectrometers with field strength up to 900 MHz considerably increased the sensitivity and reduced the requirements for high amounts of samples. However, IMPs have almost been completely excluded form this studies as they are rarely expressed in cellular expression systems and even more difficult to label with stable isotopes, an indispensable prerequisite for multi-dimensional NMR analysis.

CF expression methods provide unique labeling possibilities. These advantages in combination with high level IMP expression and ultra-high field NMR spectrometers open a new avenue for the functional and structural analysis of IMPs.

5.2.1. ^{13}C- and ^{15}N-isotopic labeling of proteins (P 08)

C. Klammt, F. Bernhard and H. Rüterjans: **^{13}C- and ^{15}N-Isotopic Labeling of Proteins.** In: *Molecular Biology in Medical Chemistry*, T. Dingermann, G. Folkers, H. Steinhilber (eds.), pp. 269-292, Wiley-VCH, Weinheim. (2004)

This book chapter reviews recently developed strategies for the incorporation of ^{15}N and ^{13}C labels into proteins. Conventional labeling strategies require the overproduction of proteins in bacterial or eukaryotic expression systems and we give an overview of the most advanced systems as they are indispensable prerequisites for the efficient incorporation of isotopes into proteins. Furthermore, we focus on newly developed *in vitro* techniques, utilizing CF extracts for the generation of protein samples. While CF production of proteins in an analytical scale has been possible for several years, preparative protein synthesis using extracts from *E. coli* or other sources has become one of the most powerful tools for the production of labeled samples suitable for NMR analysis. We therefore especially address the recent advantages in CF expression and labeling strategies.

Major results:

- Review of all commonly used expression systems for the production of ^{13}C- and ^{15}N-labeled protein samples, including *E. coli* expression systems, the *P. pastoris* system and the Chinese Hamster Ovary (CHO) cell system.
- Description of the main strategies for the production of selectively ^{13}C- and ^{15}N-labeled proteins. Here selective labeling of amino acids, specific isotope labeling with ^{13}C isotopes and segmental isotope labeling are included.
- Discussion of advantages and disadvantages of CF expression systems. Comprehensive review of this method regarding bacterial and eukaryotic systems, reaction set-ups, extract preparation and preparative CF protein synthesis.
- Demonstration of unique labeling possibilities that arise from CF protein expression and comparison of *in vivo* uniformly labeled protein with CF selectively labeled one.
- Protocols for the expression of up to 4 mg soluble CF expressed GFP per mL of RM.

My contribution to this work:

I have prepared all of the figures and reviewed the CF paragraph. In addition I have selectively labeled the soluble protein cRcsB and prepared the NMR sample.

5.2.2. Efficient Strategy for the Rapid Backbone Assignment of IMPs (P 09)

N. Trbovic, C. Klammt, A. Koglin, F. Löhr, F. Bernhard, and V. Dötsch: **Efficient Strategy for the Rapid Backbone Assignment of Membrane Proteins.** *J. Am. Chem. Soc.*, **127**, 13504-13505. (2005)

This JACS communication describes the efforts in assigning the NMR backbone resonances of the completely α-helical IMP TehA using unique advantages of CF expression by a combinatorial labeling strategy. The 24 kDa fragment of the putative tellurite transporter TehA was CF isotopically labeled and solubilized in LMPG micelles for solution NMR analysis. Although the obtained NMR spectra behaves extremely well for an IMP of this size incorporated in detergent micelles, severe overlap in some regions of the spectra allowed the assignment only of 55% of the protein's backbone resonances. Selectively labeling, usually performed to solve this problem, using 10 different amino acid types obtained only further 10% of unambiguous assignments. Finally a specific labeling procedure, unique for the CF system, based on the simultaneous labeling of certain amino acid types with ^{15}N and other amino acid types with ^{13}C on the backbone carbonyls helped to overcome this problem. The measurement of a two-dimensional [^{15}N,^{1}H]-TROSY-HN(CO) (HNCO) experiment now allowed selecting only those ^{15}N-labeled amino acids that are N-terminally preceded by a ^{13}C-labeled amino acid type. In order to optimize this procedure and reduce the sample preparation, a combinatorial labeling scheme has been developed. The preparation of 3 combinatorial labeled samples measured with [^{15}N,^{1}H]-TROSY-HSQC spectra as well as two-dimensional HNCO spectra obtained new distinct amino acid assignments that served as specific starting points fore more residues, bringing the total backbone assignment to 85%.

Major results:

- CF expression of IMPs LMPG solubilized in combination with high field NMR spectrometry allows the generation of NMR samples sufficient to start the assignment and structural investigations on even large IMPs of 24 kDa and comprising 7 TMS.

- Program for combinatorial labeling provided on the internet (http://www.biophyschem.uni-frankfurt.de/AK_Doetsch/projects/download/combilabel.m).
- First high resolution [^{15}N,^{1}H]-TROSY-HSQC spectra of a 24 kDa α-helical IMP in detergent micelles.
- CF expression provides unique labeling possibilities like the use of a combinatorial labeling strategy that has not been able in this timescale in cellular systems.
- CF expression allows the fast generation of isotopically labeled IMP NMR samples within 24 hours.

My contribution to this work:

I have cloned the c-terminal truncated TehA construct, optimized the CF protein expression of this protein and prepared all NMR samples. This includes all uniformly and selectively and combinatorial labeled ΔTehA samples. The NMR analysis was done by Frank Löhr whereas Nicola Trbovic assigned the ΔTehA backbone resonances during his Diploma work.

5.2.3. Combination of CF expression and NMR for structural assignment of IMPs (P 10)

A. Koglin, C. Klammt, N. Trbovic, D. Schwarz, B. Schneider, B. Schäfer, F. Löhr, F. Bernhard and V. Dötsch: **Combination of cell-free expression and NMR spectroscopy as a new approach for structural investigation of membrane proteins.** *Magn. Res. Chem.*, **44**, 17-23. (2006)

This Mini-review article discusses the combination of CF membrane protein expression and liquid state NMR spectroscopy. Recently, the application of CF expression systems to membrane proteins has demonstrated that this technique can be used to produce quantities sufficient for structural investigations form many different membrane proteins. In particular for NMR spectroscopy, CF expression provides major advantages since it allows for amino acid type selective and even amino acid position specific labeling.

Major results:

- Reviews CF expression systems showing the successful expression of diverse IMPs like the cysteine transporter YfiK, TehA, EmrE and SugE as precipitates that can easily be resolubilized in mild detergents and purified by modified solubilization strategies.
- Reviews NMR approached for the investigations of α-helical membrane proteins, in particular the assignment and structural analysis of TehA.
- Description of the paramagnetic resonance enhanced (PRE) technology, using paramagnetic labels in order to receive additional distance constrains that can be used along NOE information for the structural analysis of IMPs in detergent micelles.
- Description of the [^{15}N,^{1}H]-TROSY-HSQC spectrum of the soluble CF expressed ^{2}H-^{15}N-labeled 24 kDa fragment of TehA.

My contribution to this work:

I have provided the SDS gel figures of CF expressed and detergent purified IMPs. I overexpressed solubilized and purified all of the mentioned IMPs. I optimized the expression of ΔTehA for successful NMR sample preparation. I have further prepared all NMR samples including the mutant samples used for PRE measurements. The NMR measurements were done by Frank Löhr, whereas Nicola Trbovic assignment the ΔTehA backbone resonances and performed site specific mutagenesis to provide PRE samples during his diploma work.

5.2.4. Incorporation of Fluorescence Labels into cell-free produced proteins (P 11)

K. Sengupta, C. Klammt, F. Bernhard and H. Rüterjans: **Incorporation of fluorescence labels into cell-free produced proteins.** In: *Cell-Free Protein Expression*, J.R. Swartz (ed.) pp. 81-88, Springer Verlag, Berlin, Heidelberg, New York. (2003)

This book chapter describes the incorporation of unnatural amino acids into proteins by using a CF expression system. Here spectrally enhanced tryptophan residues were incorporated in the transcription regulator protein RcsB. This incorporation did not affect the protein activity but changed the protein tryptophan spectral properties. Steady state fluorescence spectroscopy analysis of differently tryptophan labeled proteins allowed studying the interaction between two RcsB proteins titrating non-labeled and

labeled RcsB. In addition, the use of N- and C-terminal RcsB fragments gave evidence for an involvement of the C-terminal effector domain in the oligomerization of RcsB.

Major results:

- Unnatural amino acids like tryptophan analogues can easily be incorporated into proteins by using CF expression.
- The transcription regulator protein RcsB was found to act as a multimer, analyzed by steady state fluorescence spectroscopy of differently tryptophan labeled proteins. Further the C-terminus of the protein gave evidence for an involvement in the oligomerization process.

My contribution to this work:

I have performed the steady state fluorescence titration measurements of differently tryptophan labeled RcsB proteins and constructs and provided the fluorescence spectra figure. The proteins were CF expressed in the commercially available RTS™ system (Roche, Penzberg, Germany) by Frank Bernhard. Kaushik Sengupta further purified those soluble proteins out of the RM and performed the activity assay for CF synthesized RcsB.

5.3. FEBS Journal prize award of 2004 for Christian Klammt (P 12)

Mit neuer Methode lassen sich Membranproteine "knacken". Erstmals können ausreichende Mengen begehrter Proteine produziert werden – Auszeichnung für Christian Klammt**.** *Forschung Frankfurt*, **3**, 9-10. (2005)

This article describes the FEBS Journal Prize of 2004 award given to me for being the first author of the paper judged to be the best one published in the FEBS Journal during the year 2004 for the paper:

C. Klammt, F. Löhr, B. Schäfer, W. Haase, V. Dötsch, H. Rüterjans, C. Glaubitz and F. Bernhard: ***High level cell-free expression and specific labeling of integral membrane proteins.*** *Eur. J. Biochem., 271, 568-580. (2004)*

6. Unpublished Results – Structural and functional analysis of the human endothelin B receptor (ETB)

6.1. Introduction

6.1.1. G-protein coupled receptors (GPCRs)

The G-protein coupled receptors (GPCRs) are the largest class of cell-surface receptors and are encoded by more than 1000 genes in the human genome [28]. GPCRs are involved in the recognition and transduction of messages as diverse as odorants, small molecules including amino-acid residues, nucleotides and peptides, proteins, Ca^{2+} as well as light. They control the activity of ion channels, enzymes and transport of vesicles via the catalysis of the GDP-GTP exchange on heterotrimeric G proteins (Gα-βγ). Almost all physiological regulatory mechanisms in the human body are controlled by GPCRs and thus more than 60% of modern drugs are supposed to be directed against GPCRs [29]. Their structural characterization has been inhibited due to extreme difficulties in the preparation of sufficient amounts of functionally folded protein. While mammalian expression systems can be optimized in order to produce highly functional GPCRs, the final yields are still low and do not warrant structural approaches in most cases [103]. CF expression might overcome this problem as it was successful in synthesizing large amounts of bacterial transporters that do not overexpress in *E. coli*.

GPCRs are divided into six Groups according to the GPCR database online (www.GPCR.org), whereas they comprise three major families (Fig. 18). Family A, also referred to as the rhodopsin-like family is by far the largest subgroup and contains receptors of odorants, small molecules, peptides and glycoprotein hormones. This family is further classified into 3 sub classes with regard to their different ligand binding site locations (Fig. 18). All Family A class GPCRs are activated by small ligands. Members of the first sub class bind ligands in a cavity formed by the third and fourth TMS. In the case of the light-activated receptor rhodopsin, the target of photons, retinal, is covalently linked in this cavity and its change in conformation induced by light activates the receptor. The second sub class is activated by short peptides which interact with the extracellular loops and the N-terminal domain. However, the C-terminal end of these peptides has been proposed to interact within a cavity similar to that of the first sub family [104]. Family B is characterized by a relatively long amino terminus that contains several cysteins, which presumably form a network of disulfide bridges and plays a role in the binding of the ligand [105]. Family C receptors are characterized by a long amino terminus, which is often described as being like a “Venus fly trap”. A unique characteristic of the family C receptors is that the third intracellular loop is short and highly conserved.

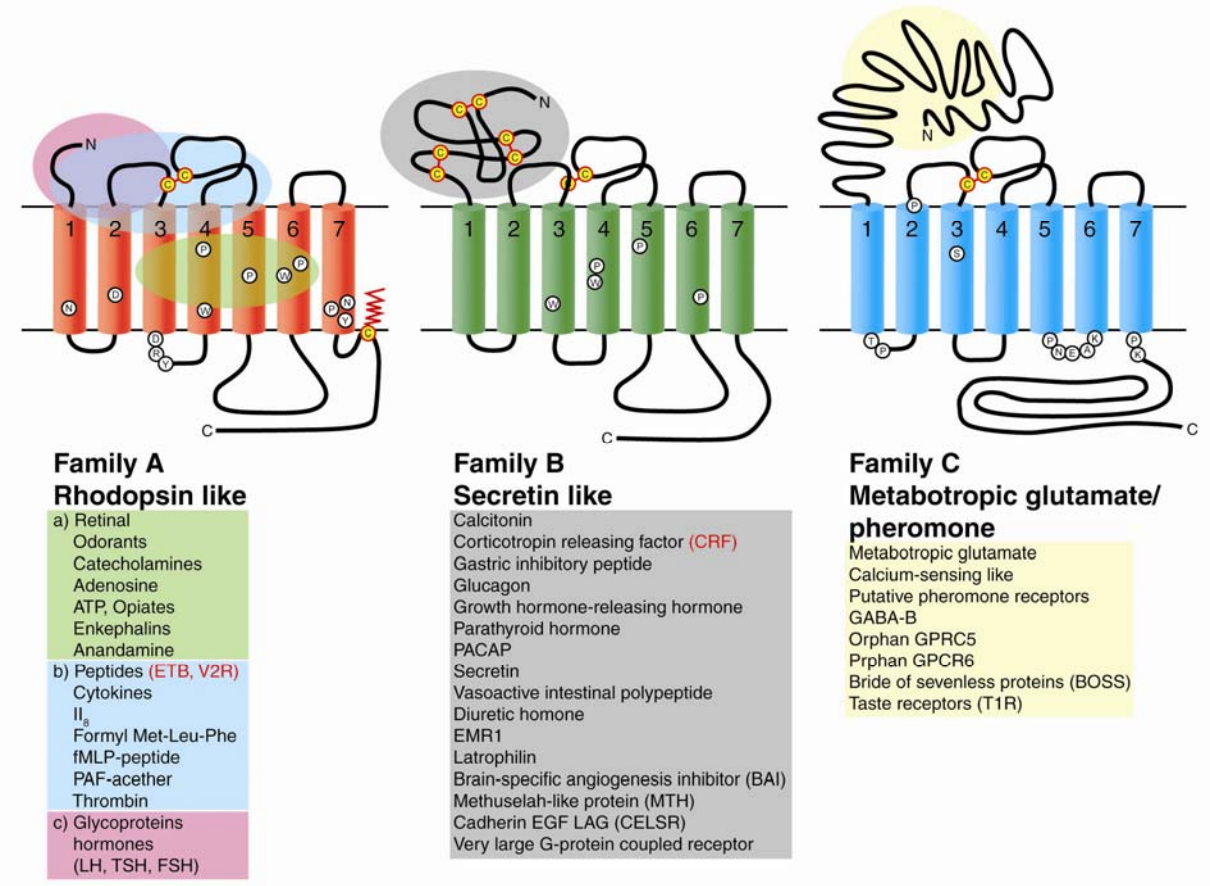


Fig. 18 Overview of the G-protein coupled receptor family
G-protein-coupled receptors (GPCRs) can be divided phylogenetically into six families. A schematic representation of the three main families is shown, illustrating some mayor differences between these families besides highly conserved amino acid residues. In addition the main receptor targets of the corresponding families are listed. Receptors of **family A** are characterized by several highly conserved amino acids, indicated in the diagram, and a disulfide bridge that connects the first and second extra cellular loops and most of these receptors also have a palmitoylated cysteine in the carboxy-terminal tail. In addition the **family A** is sub classified into **a**, **b** and **c**. **Family Aa** contains GPCRs for small ligands including rhodopsin and ß-adrenergic receptors. The binding site is localized within the seven TMS. **Family Ab** contains receptors for peptides whose binding site includes the N-terminal, the extracellular loops and the superior parts of TMS. **Family Ac** contains GPCRs for glycoprotein hormones. (Modified after [106, 107]).

All GPCRs share a similar architecture of seven membrane-spanning α-helices connected by three intracellular and three extracellular loops [108]. GPCR families show no sequence homology to one another, indicating that they might be unrelated phylogenetically and that the similarity of their TMS structure might be the result of common functional requirements. Two cysteine residues, one in the first and one in the second extracellular loop, which are conserved in most GPCRs, form a disulfide link which is probably important for the packing and for the stabilization of a restricted

number of conformations of these seven TMS [107]. Besides sequence variations, GPCR families differ in the length and function of their N-terminal extracellular domain, their C-terminal intracellular domain and their intracellular loops (Fig. 18).

Virtually all therapeutics that are directed towards GPCRs have been designed using assays that presume that these receptors are monomeric [106]. However, over the past years it has widely been accepted that GPCRs form oligomers. For several receptors a homooligomerization have been shown, whereas others are able to even oligomerize as heterooligomers (reviewed by [106]). It is further assumed that heteromeric receptor complexes have functional characteristics that differ form homogeneous populations of their individual constituents. Regarding the number of oligomers, many GPCR oligomerization studies do not make a clear distinction between dimers and larger receptor complexes.

6.1.2. Endothelin B receptor (ETB)

Endothelin-1 (ET-1) is identified as a potent vasoconstrictor peptide secreted from vascular endothelial cells [109]. The entire endothelin family comprises endogenous isoforms of 21-aminoacid peptides named ET-1, ET-2 and ET-3, each one coded by a different gene [110]. All isopeptides are synthesized from biological inactive precursors by the endothelin-converting enzymes that belong to a subgroup of membrane-bound zinc metalloproteases [109]. The active isopeptides are distributed in many tissues and cell types, and play important roles in the modulation of hemodynamics, cardiac, pulmonary, renal, and neural function, and in mitogenesis and neural development [111-114]. Most abundantly expressed along those three endothelins is ET-1, the only isoform constitutively released by vascular endothelium [115].

The action of ET-1 in humans occurs via two endothelin receptor subtypes: The endothelin A receptor (ETA) and the endothelin B receptor (ETB) [116, 117]. Both receptors, belonging to the large family of GPCRs, are responsible for the whole range of physiological effects mediated by ET-1. Whereas the ETA is predominantly expressed in smooth muscle cells, and its stimulation causes a long-lasting vasoconstriction [118], the mainly in endothelial cells expressed ETB's activation results in a transient vasodilatation [119], including NO release upon stimulation [120]. Another difference between both receptor subtypes that comprises 63% homology in primary structure is their affinity to ET isopeptides. ETA possesses higher affinity for ET-1 than for ET-2 and ET-3, whereas ETB binds the three ET isoforms with similar affinity [121]. Besides controlling the vascular tone, the endothelin system is involved

in many other physiological processes such as neurotransmission, embryonic development, renal function, and regulation of cell proliferation. It thus plays an important role in physiopathological disorders like congestive heart failure, diabetes, arteriosclerosis or primary pulmonary hypertension [122-124].

In contrast to a wealth of information form *in vivo* approaches, the detailed *in vitro* characterization of functional and structural properties of ETB has not been feasible so far due to the restricted availability of protein samples. CF expression overcomes these limitations and further allows the fast and efficient preparation of ETB protein samples. Here the human ETB can be synthesized in a soluble form in micelles of selected detergents and besides small terminal peptide tags that facilitate detection and purification no large fusion proteins are needed for the stabilization of the receptor.

Ligand binding studies and pull down assays demonstrate that the CF produced ETB protein is functionally folded. Electron microscopy and single particle analysis of negatively stained samples indicate a dimeric state. Construction of terminal deletions in combination with functional analysis revealed a relatively small region including TMS1 as absolutely essential for ligand binding as well as for oligomerization. Further, recorded solution NMR spectra of a truncated ETB construct containing TMS1-3 show reasonable spectral quality compared to the full-length receptor. The presented results demonstrate a new and highly efficient approach for the rapid production and functional characterization of GPCRs like ETB.

6.2. Results

6.2.1. Preparative cell-free expression of various GPCRs

Four different GPCRs have been analyzed for the high level CF expression. The human and porcine V2R and the human ETB, belonging to the first class of Rhodopsin like GPCRs (family Ab) and the rat corticotropin releasing factor receptor (CRF) that belongs to the second GPCR class, have been tested in the CF system. Using standard unmodified expression, no expression has been found at al. However, using the eleven amino acids containing T7-tag at the N-terminus, all four GPCRs could successfully be CF expressed as precipitates in amounts of 3 mg per mL of RM (Fig. 19). It could be shown that all GPCR precipitates dissolve in 1% LMPG and can be purified by Ni-chelat chromatography (Fig. 19).

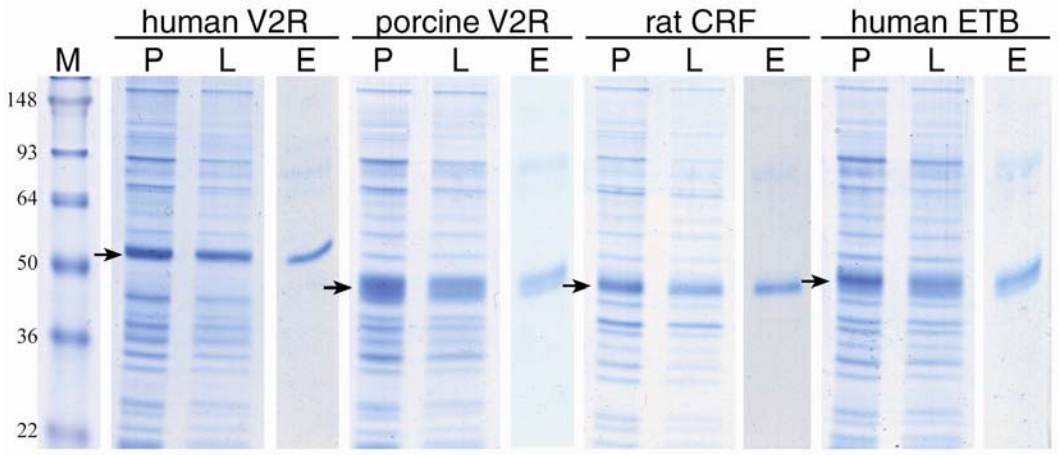


Fig. 19 SDS gel of cell-free expressed GPCRs.
The human vasopressin receptor type 2 (V2R), the porcine V2R, the rat corticotropin releasing factor receptor (CRF) and the human endothelin B receptor (ETB) have been expressed in the cell-free system in the insoluble mode, resolubilized in 1% LMPG and purified by His-tag purification. 1 µl of the CF expressed precipitate (P), the resolubilized protein in LMPG (L) and the elution (E) from the Nickel-column was analyzed on a 12% SDS-gel.

The CF expression results obtained for those GPCRs again reflect the great potential of this alternative method. It further demonstrates that besides bacterial transporters, complex eukaryotic IMPs can be synthesized in large amounts. Here the addition of the T7-tag at the N-terminus was proven to be highly beneficial for expression of the GPCRs. A nonoptimal initiation of translation in the *E. coli* system, caused by the eukaryotic codon sequences of the GPCR genes, might account for that observation.

6.2.2. Expression and purification of ETB and various ETB truncations

The CF reaction protocol was optimized in order to obtain the preparative scale expression of His tagged ETB (ETB_{cHx}). Protease inhibitors were supplied in order to prevent ETB_{cHx} degradation and the concentration of Mg^{2+} and K^{+} ion concentration was optimized in the range between 12-16 mM and 250-340 mM, respectively.

Besides the full-length receptor a series of truncated ETB derivatives containing nested deletions have been designed (Table 7). For successful expression the construction of a translational fusion with the small 12 amino acid containing T7-tag at the N-terminus of ETB (Fig. 20) was crucial as essentially no expression of the non-fused full-length $ETB_{\Delta T7}$ was detectable.

Table 7 Structural characteristics of CF produced ETB derivatives.

			Included domains[2]							C-terminus								
Fragment[1]	Region	[kDa]	ND	T1	C1	T2	E1	T3	C2	T4	E2	T5	C3	T6	E3	T7	CDO	Tag
$ETB_{\Delta T7}$	M1-S443	50.7	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	cH6
ETB_{cHx}	Q_2-S_{443}	52.5	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	cHx
ETB_{strep}	E_{27}-S_{443}	49.5	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	strep
ETB_{131}	E_{27}-C_{131}	14.4	x	x	/	-	-	-	-	-	-	-	-	-	-	-	-	cHx
ETB_{168}	E_{27}-P_{168}	18.5	x	x	x	x	/	-	-	-	-	-	-	-	-	-	-	cHx
ETB_{203}	E_{27}-V_{203}	22.2	x	x	x	x	x	x	/	-	-	-	-	-	-	-	-	cHx
ETB_{306}	E_{27}-G_{306}	34.1	x	x	x	x	x	x	x	x	x	x	/	-	-	-	-	cHx
ETB_{132}	M_{132}-S_{443}	38.6	-	-	/	x	x	x	x	x	x	x	x	x	x	x	x	cHx
ETB_{204}	A_{204}-S_{443}	30.8	-	-	-	-	-	-	/	x	x	x	x	x	x	x	x	cHx
ETB_{307}	M_{307}-S_{443}	18.8	-	-	-	-	-	-	-	-	-	-	/	x	x	x	x	cHx
ETB_{93}	P_{93}-V_{203}	15.3	/	x	x	x	x	x	/	-	-	-	-	-	-	-	-	cHx

[1], with exception of $ETB_{\Delta T7}$, all proteins contain additionally a N-terminal T7-tag.
[2], ND, N-terminal domain; T1-7, transmembrane segment 1-7; C1-3, cytoplasmatic loop 1-3; E1-3, extracellular loop 1-3; x, included; -, deleted; /, partially truncated.

At optimized conditions, ETB_{cHx} could be synthesized at yields of up to 3 mg per ml reaction mixture (RM). The calculated molecular weight (MW) of ETB_{cHx} is 52 kDa but a prominent band of approx. 48 kDa was visible after separation of the RM by SDS-PAGE (Fig. 21 A).

ETB_{cHx} was CF produced either as soluble protein in presence of detergents or as precipitate in the absence of detergents. In a previous screen comprising a variety of 24 detergents most commonly used in membrane protein analysis, only digitonin and long chain Brij derivatives have been effective in the soluble expression of the porcine V2R [9]. I focused therefore primarily on these two detergent classes for the soluble expression of ETB_{cHx}. The presence of 1 % Brij78 and 1.5 % Brij58 resulted in the complete solubilization of ETB_{cHx} with final yields of approx. 3 mg soluble protein per ml RM. With the detergents Brij35 (0.1 %) and digitonin (0.4 %) only 500 µg and 100 µg of soluble ETB_{cHx} per ml RM was obtained, respectively, in addition to ETB_{cHx} precipitate. The majority of the synthesized protein thus still precipitated after translation in presence of the two latter detergents. Soluble produced ETB_{cHx} was purified in one step by Ni-chelate chromatography as described in the methods section. At the average, approx. 60 % of the synthesized ETB_{cHx} in the RM could be recovered

after purification and buffer exchange. The purified ETB_{cHx} could be homogenously reconstituted into proteoliposmes with *E. coli* lipid mixtures as evaluated by freeze fracture electron microscopy (Fig. 10 E).

Besides the direct soluble CF expression mode, ETB_{cHx} precipitate which has been CF produced in the absence of detergents could be completely solubilized in 1 % LMPG. The detergent DPC at a final concentration of 2 % was slightly less effective with solubilization of approx. 80 % of the precipitated ETB_{cHx}. With the detergents Brij35 (1 %), DDM (2 %), DM (2 %), Chaps (5 %) and DHPC (2 %) only 5-20 % of the ETB_{cHx} precipitate could be solubilized. The detergents Thesit (1 %), Tween20 (1 %), Triton X-100 (2 %) and digitonin (2 %) were completely ineffective for solubilization of the CF produced ETB_{cHx} precipitate.

In order to confine the ETB region essential for binding of the ligand ET-1 and for oligomerization, a series of eight plasmids encoding for truncated ETB fragments starting from either end and containing different secondary structural elements were constructed (Table 7, Fig. 20). All fragments could be overexpressed in amounts of at least 1 mg per ml RM in the individual CF system as soluble proteins in presence of 1% Brij78 (Fig. 21 B). The fragments were purified subsequently after expression in one step by Ni-chelate chromatography and the purity was evaluated by SDS-PAGE analysis (Fig. 21 B). In addition all ETB-fragments were successfully reconstituted into *E. coli* lipids (Fig. 21 C).

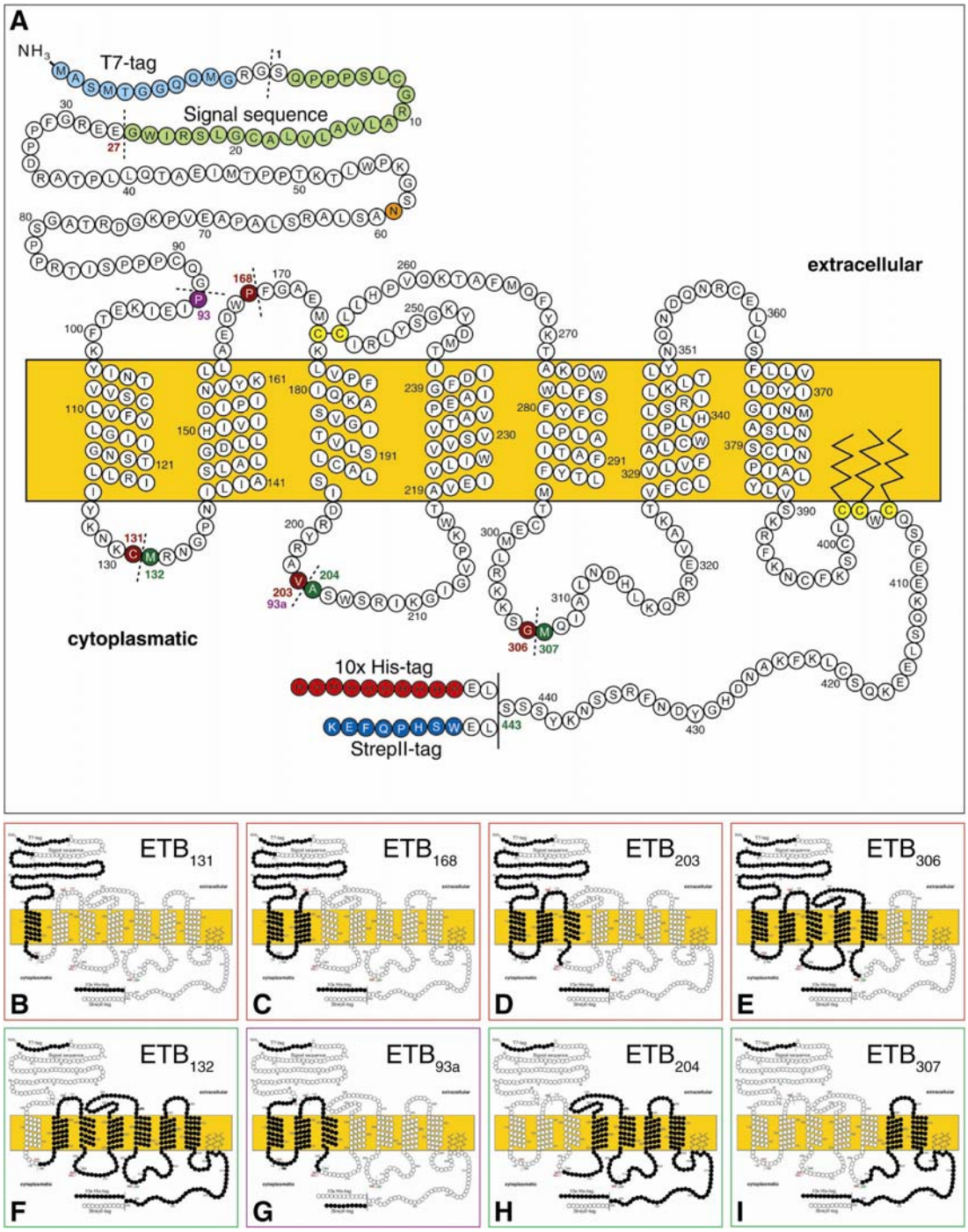


Fig. 20 Proposed secondary structure of ETB and overview of ETB truncations.
Besides the GPCR-typical seven transmembrane topology, ETB is predicted to have a glycosylation site at asparagine 59 (orange), a disulfide bridge connecting the extracellular loops two and three and several palmitoylated cysteine residues in the C-terminal loop (yellow). For cell-free expression ETB was N-terminally supplemented with the 12 amino acid containing T7-tag (light blue) and c-terminally tagged by either a deca histidine tag (red) or a strepII-tag (blue). The last amino acid of C-terminal truncations (dark red) and start of N-terminal truncations (green) and the start of the ETB_{93a} truncation (purple) is indicated. ETB truncated fragments all containing a N-terminal T7-tag and a C-terminal deca histidine tag are illustrated in black (B-I). **A**, ETB; **B**, ETB_{131}; **C**, ETB_{168}; **D**, ETB_{203}; **E**, ETB_{306}; **F**, ETB_{132}; **G**, ETB_{93}; **H**, ETB_{204}; **I**, ETB_{307}.

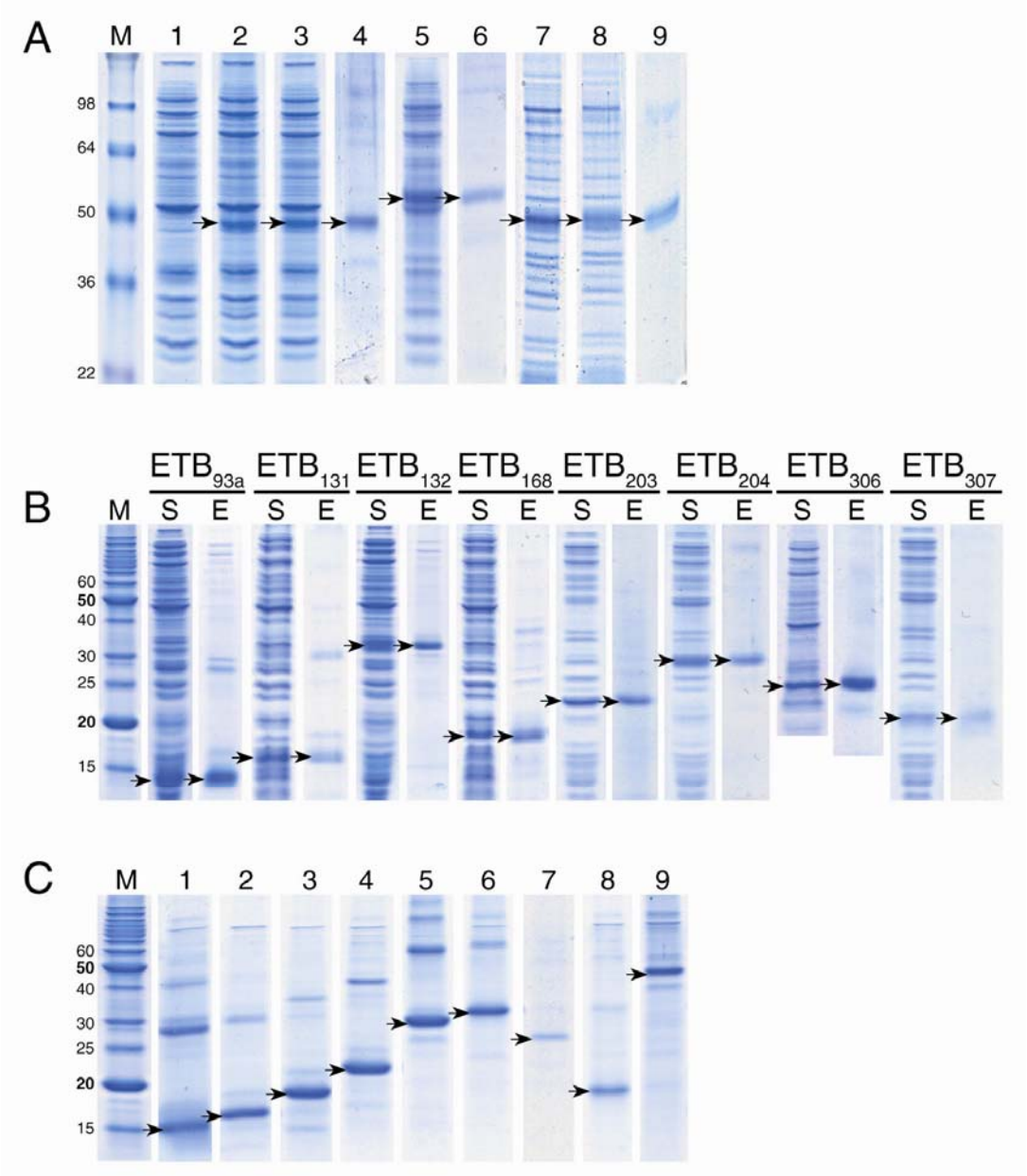


Fig. 21 SDS gels of cell-free expressed ETB and ETB truncations on standard 16.5% SDS gels. **A**, Cell-free expression of ETB in presence of Brij58 (2), Brij78 (3-6) and as insoluble precipitate (7). M, marker; 1, reaction mixture before incubation; 2, soluble fraction of RM after ETB_{cHx} expression in presence of Brij58; 3, soluble fraction of RM after ETB_{cHx} expression in presence of Brij78; 4, Elution fraction of ETB_{cHx} after His-tag purification; 5, soluble fraction of ETB_{strep} expression in presence of Brij78; 6, elution fraction of Strep-Tactin™ purification; 7, insoluble fraction of the RM in the insoluble mode; 8, resolubilized precipitate in 1% LMPG; 9, Elution fraction of resolubilized ETB in LMPG after His-purification. The overproduced and purified receptors are marked with arrows. **B**, Analysis of the cell-free expression of ETB fragments in presence of Brij78. 0.7 µl of the supernatant after 20 hrs reaction (S) and 2 µl of the elution fraction (E) after His-purification have been analyzed. The overproduced and purified ETB truncations are indicated by arrows. **C**, soluble expressed ETB fragments where reconstituted into proteoliposomes and 9 ml where analyzed on an SDS-gel. 1, ETB_{93a}; 2, ETB_{131}; 3, ETB_{168}; 4, ETB_{203}, 5, ETB_{306}; 6, ETB_{132}; 7, ETB_{204}; 8, ETB_{307}; 9, ETB_{cHx}.

6.2.3. Functional analysis of full-length protein and ETB-truncations

The affinity to its natural peptide ligand Endothelin-1 (ET-1) as one of the two main functions of a GPCR was analyzed by using different ET-1 ligands. Besides the non-tagged circular wild type ET-1, circular ET-1 labeled with one Cy3-dye at lysine 9 (cET-1), double biotinylated ET-1 (bET-1) at position cysteine 1 and lysine 9 and linear 1,3,11,15 alanine ET-1, N-terminally labeled with fluorescein (fET-$1_{A(1,3,11,15)}$), have been used (Fig. 22).

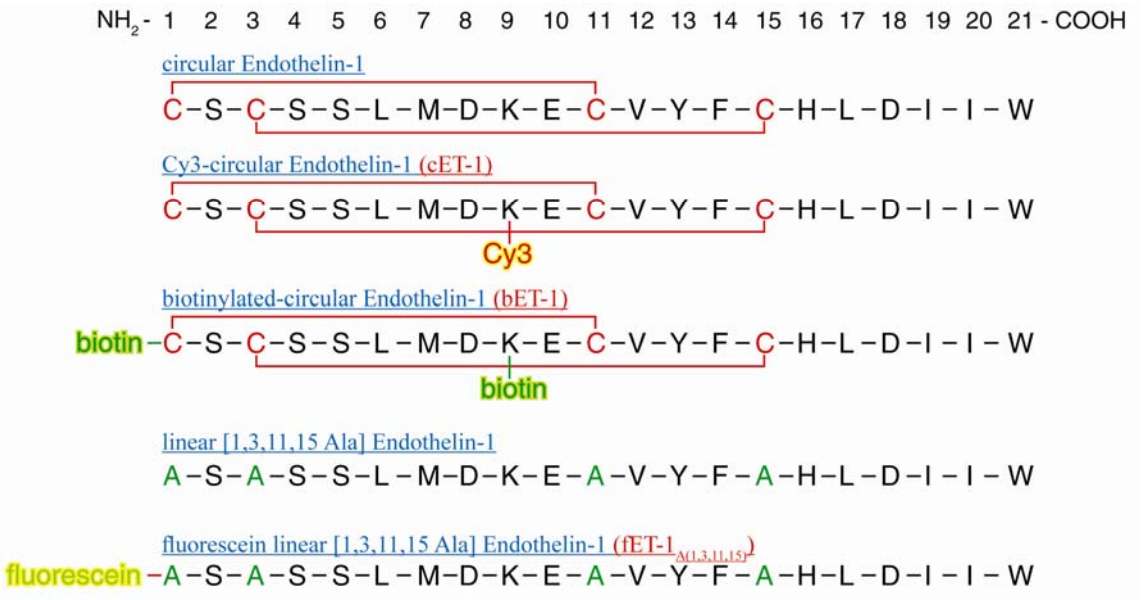


Fig. 22 ETB-ligands overview.
Illustration of used ETB ligands Endothelin-1 (ET-1), explaining the sites where labels are situated. The circular ET-1 was used Cy3 labeled and biotinylated, whereas the linear ET-1 was used in the fluorescein labeled form.

6.2.3.1. Fluorescent ligand coelution

The ETB_{cHx} protein can be CF produced at high levels in a variety of different conditions. However, the mode of CF expression, i.e. the expression as precipitate or as soluble protein, as well as the type of supplied detergent could have a significant impact on the folding of ETB_{cHx} into a functional conformation. The affinity to its natural peptide ligand ET-1 was therefore analyzed with purified ETB_{cHx} samples that have been produced under different conditions. Mixtures of the cET-1 with purified ETB_{cHx} either CF produced as precipitate and solubilized in 1 % LMPG or directly produced as soluble protein in presence of 0.1% Triton X-100, 0.2% Brij35, 1.5 % Brij58, 1 % Brij78 or 0.4 % digitonin, respectively, were separated by gelfiltration and the elution

fractions were analyzed by taking advantage of the different absorbencies of the two compounds (Fig. 23). The 52 kDa ETB_{cHx} elutes at a retention volume of 1.6 ml whereas the 21-mer cET-1 will start to elute at a volume of 2.1 ml. Co-elution of cET-1 with ETB_{cHx} therefore indicates complex formation of the receptor with its ligand giving evidence of a native and active protein conformation. In contrast, addition of CF produced protein present in an unfolded or inactive conformation should result in the separation of the two compounds.

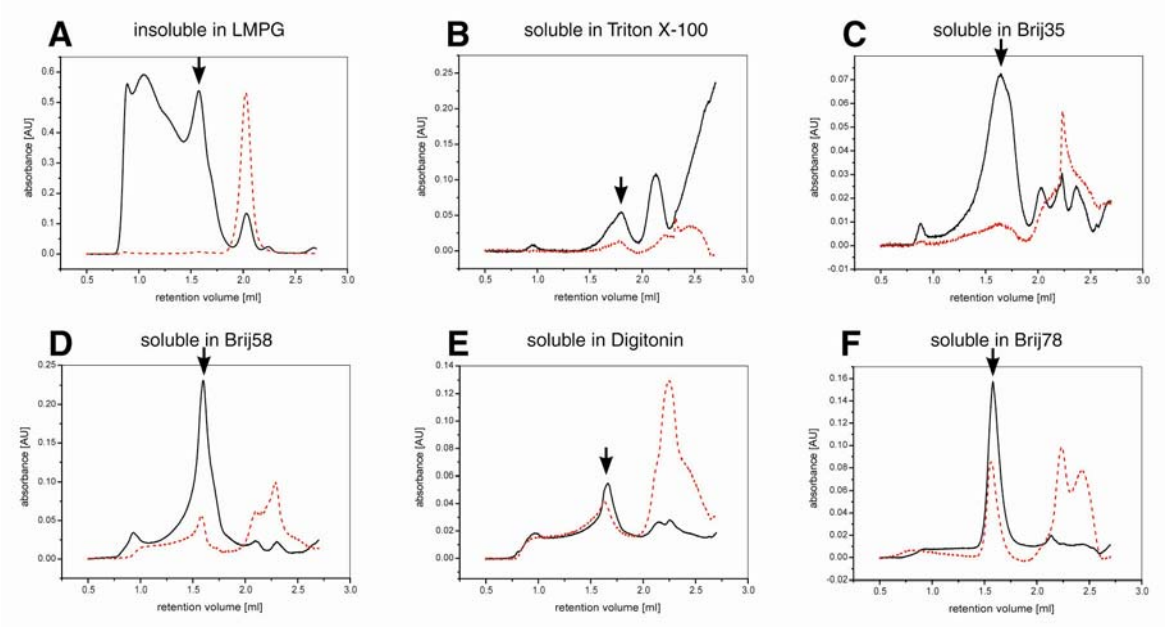


Fig. 23 Activity correlation of ETB with the soluble CF expression in different detergents.
ETB was CF expressed in presence of different detergents, purified by Ni-chelat chromatography and the binding of cET-1, after 3 hrs incubation at RT was analyzed on a Superose 6 PC 3.2/30 column. The elution chromatograms show the total protein absorption at 280 nm (black solid line) and the ligand absorption of cET-1 at 550 nm (red dotted line) The retention volume of ETB and bound ligand is indicated by arrows. **A**, ETB was CF expressed as precipitate without the presence of detergents and resolubilized in 1% LMPG; **B**, ETB was soluble expressed in presence of 0.1% Triton X-100; **C**, ETB was soluble expressed in presence of 0.2 % Brij35; **D**, ETB was soluble expressed in presence of 1.5% Brij58; **E**, ETB was soluble expressed in presence of 0.4% Digitonin; **F**, ETB was soluble expressed in presence of 1% Brij78. (The gel-filtration analysis was done in collaboration with Ankita Srivastava).

The ligand cET-1 was completely separated from ETB_{cHx} samples that were CF produced as precipitate and solubilized in LMPG, indicating that despite solubilization, the receptor might not have adopted its native conformation (Fig. 23 A). In contrast, significant amounts of cET-1 co-eluted with ETB_{cHx} synthesized in the soluble mode of CF expression in presence of the detergents digitonin, Brij58 and Brij78 whereas little

coelution was obtained in ETB soluble expressed in the presence of Triton X-100 and Brij35. The highest apparent binding of cET-1 was obtained with protein CF expressed in presence of Brij78 (Fig. 23 F). This expression condition was therefore chosen for further sample preparations of ETB_{cHx} and its derivatives.

The percentage of ligand binding receptor present in the purified ETB_{cHx} sample obtained after soluble CF expression in presence of 1 % Brij78 was determined by correlation of the molar ratio of complexed cET-1 with the known amount of supplied ETB_{cHx}. The amount of bound cET-1 in the gelfiltration elution profile was first estimated by standardizing the corresponding peak area with a calibration curve obtained from the peak areas of varying amounts of cET-1 subjected to gelfiltration at identical conditions. After subtracting non-specifically bound cET-1, the amount of ligand binding ETB_{cHx} present in samples was estimated with the described conditions at approx. 50 %.

In order to confine the ETB region essential for ligand binding, all eight ETB truncations were tested for their ligand binding ability. Ligand binding activity was assessed by analyzing co-elution profiles of ETB truncations mixed with the linear fluorescein labeled ligand $fET\text{-}1_{A(1,3,11,15)}$ (Fig. 24). The $fET\text{-}1_{A(1,3,11,15)}$ derivative has a co-elution profile with full-length ETB_{cHx} (Fig. 24 A)comparable to that with cET-1 (Fig 23 A). All truncated ETB fragments containing at least TMS1 like ETB_{131} (N-terminal domain (ND)-TMS1) showed ligand binding activity. Interestingly, also fragment ETB_{93a} (TMS1-TMS3) that is deleted for almost the complete ND still is able to bind $fET\text{-}1_{A(1,3,11,15)}$. However, co-elution with $fET\text{-}1_{A(1,3,11,15)}$ was largely reduced with truncations deleted of TMS1 like the 39 kDa ETB_{132} covering the ETB regions from TMS2 to C-terminal domain (CDO) (Fig. 24 D).

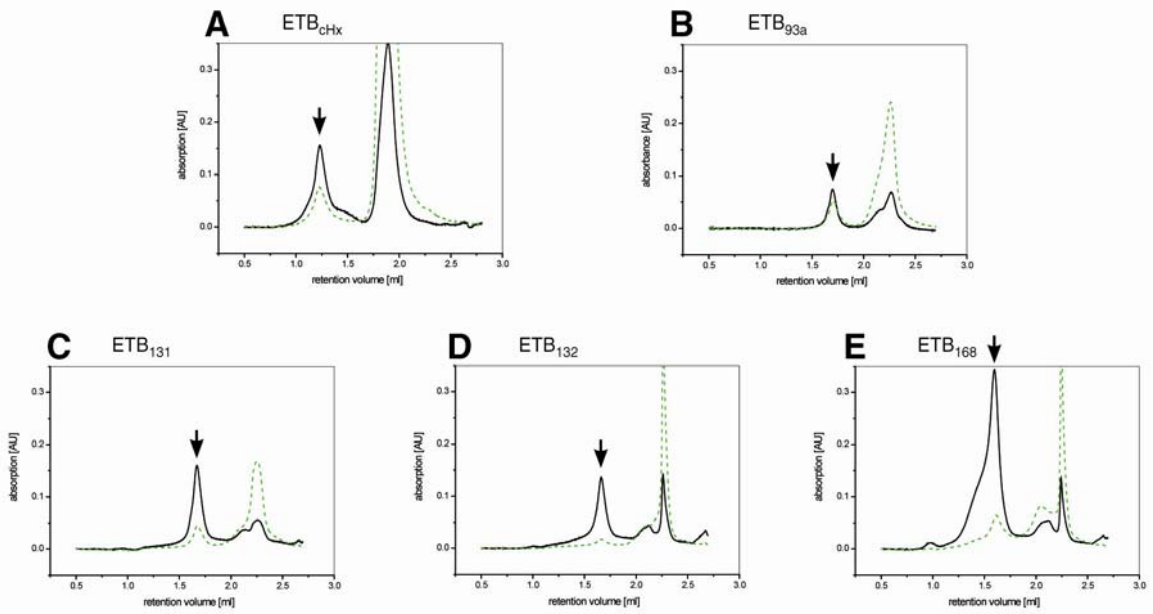


Fig. 24 Gelfiltration co-elution chromatograms of ETB and ETB fragments with fET-1$_{A(1,3,11,15)}$. ETB$_{cHx}$ and ETB fragments were CF expressed in the presence of Brij78, purified by Ni-chelat chromatography and incubated with fET-1$_{A(1,3,11,15)}$ for 3 hrs at 21°C and subsequently analyzed on a Superose 12 PC 3.2/30 (A) or Superose 6 PC 3.2/30 (B-E) column. The elution chromatograms show the total protein absorption at 280 nm (black solid line) and the ligand absorption of of linear-fluorescein-ET-1 at 495 nm (green dotted line). The retention volume of ETB and bound ligand is indicated by arrows. **A**, ETB$_{cHx}$ analyzed on a Superose 12 PC 3.2/30 column; **B**, ETB$_{93a}$ fragment; **C**, ETB$_{131}$ fragment; **D**, ETB$_{132}$ fragment; **E**, ETB$_{168}$ fragment. (The gel-filtration analysis was done in collaboration with Ankita Srivastava).

6.2.3.2. Biotinylated ligand binding

The ligand binding of ETB$_{cHx}$ and truncated derivatives was further characterized by pull-down assays of purified proteins with the immobilized bET-1 ligand as described in the methods section. Fractions containing complexes of bET-1 with ETB derivatives were separated by SDS-PAGE, blotted and the proteins were identified by immunodetection with the anti-T7-tag antibody (Fig. 25). In agreement with the results presented above, only fragments containing TMS1 like full-length ETB$_{cHx}$, ETB$_{131}$ (ND-TMS1), ETB$_{168}$ (ND-TMS2), ETB$_{203}$ (ND-TMS3) and also again the ND deleted fragment ETB$_{93}$ (TMS1-TMS3) were detected in the eluted fractions and complexed with bET-1. Accordingly, proteins devoid of TMS1 like ETB$_{132}$ (TMS2-CDO) and ETB$_{204}$ (TMS4-CDO) did not interact with bET-1.

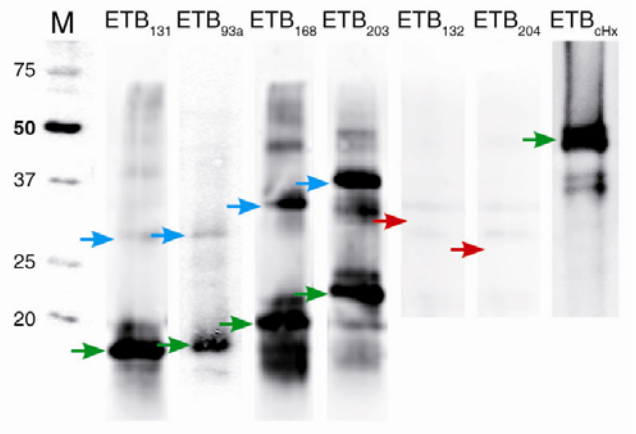


Fig. 25 Western blot of biotinylated ET-1 binding.
Elution fraction of pull-down experiments with bET-1 and ETB truncations analyzed by western blot using a HRP-coupled T7-tag antibody. Different ETB truncations are used as mentioned. Eluted ETB fragments are indicated by arrows for monomer (green), dimer (blue) and no elution (red); M, marker.

6.2.3.3. Analysis of ETB_{cHx}-ligand interaction by SPR

Total internal reflection of light at a surface-solution interface produces an electromagnetic field, or evanescent wave, that extends a short distance (~100-200 nm) into the solution. This so called evanescent-wave phenomenon occurs at certain metallic surfaces and describes surface plasmon resonance (SPR). SPR allows the detection of changes in the refractive index in the immediate vicinity of the surface layer of a sensor chip and it is observed as a sharp shadow in the reflected light from the surface at an angle that is dependent on the mass of material at the surface. In case biomolecules like ETB bind to the surface or to its ligand immobilized on the surface, the mass on the surface layer changes. This change in resonant angle can be monitored non-invasively in real time as a plot of resonance signal (proportional to mass change) versus time [125]. The recorded diagram further allows the determination of $k_{association}$ (k_a) and $k_{dissociation}$ (k_d) and further the rate constant ($K_D = 1/K_A = k_d/k_a$). Thus SPR allows the sensitive detection of molecular interactions in real time without labels. Biacore instruments (Biacore, Uppsala, Sweden) uses SPR for determining binding constants of interaction partners. The principle of biacore measurements is illustrated in Fig. 26.

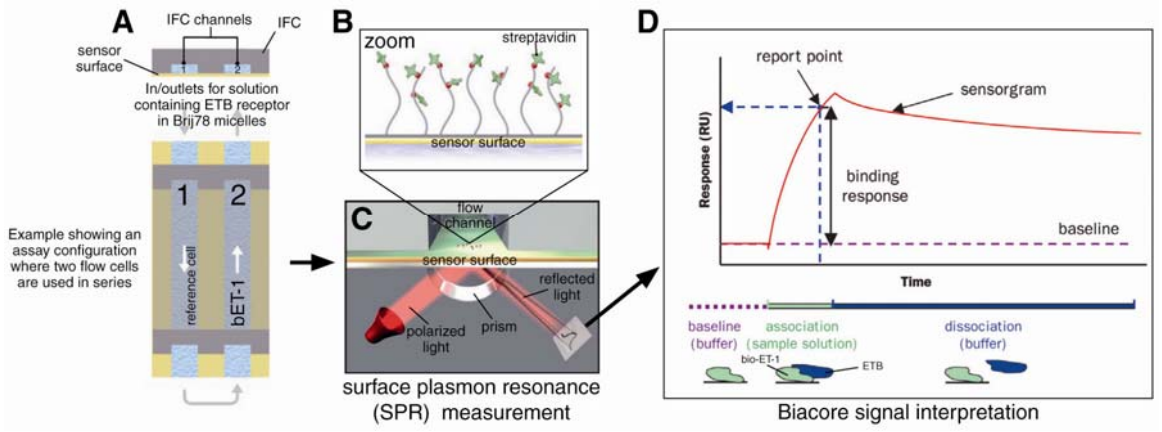


Fig. 26 Principle of biacore measurements.
A, Flow cell containing two cells for on-line referencing, allowing direct background subtraction for accurate kinetic analyses; **B**, zoom into sensor surface designed for immobilization of interacting partners via a streptavidin capture molecule; **C**, as molecules are immobilized on a sensor surface, the refractive index at the interface changes, altering the angle at which reduced intensity polarized light is reflected from the supporting glass plane. The change in angle caused by association or dissociation of molecules from the sensor surface is proportional to the mass of bound material and is recorded in a sensogram; **D**, the sensogram provides real-time information about an entire interaction, with binding responses measured in resonance units (RU). (Modified after Biacore technology note 23 from http://www.biacore.com).

Although the co-elution approach gives good evidence for a ligand binding activity of CF produced ETB samples, it is primarily not a quantitative assay. Here SPR was used to detect and quantify the molecular interactions of ETB_{cHx} and the ETB truncations ETB_{131} (ND-TMS1) and ETB_{93a} (TMS1-3) in real time. Therefore, the biotinylated ligand bET-1 was immobilized on the surface of the biosensor chip and the direct binding of functionally active ETB_{cHx} has been analyzed. ETB_{cHx} solutions with increasing concentrations from 10 nM to 250 nM were loaded on the bET-1 chip and binding kinetics were evaluated using the BIAevaluation software. The binding constant K_D of ETB_{cHx} to bET-1 was determined at $6.2 \pm 1.7 \times 10^{-9}$ (Fig. 27, Table 8). Similar assays with the C-terminal truncated derivatives ETB_{131} and ETB_{93a} revealed K_D values of $2.7 \pm 1.9 \times 10^{-8}$ and $1.7 \pm 0.5 \times 10^{-8}$, respectively.

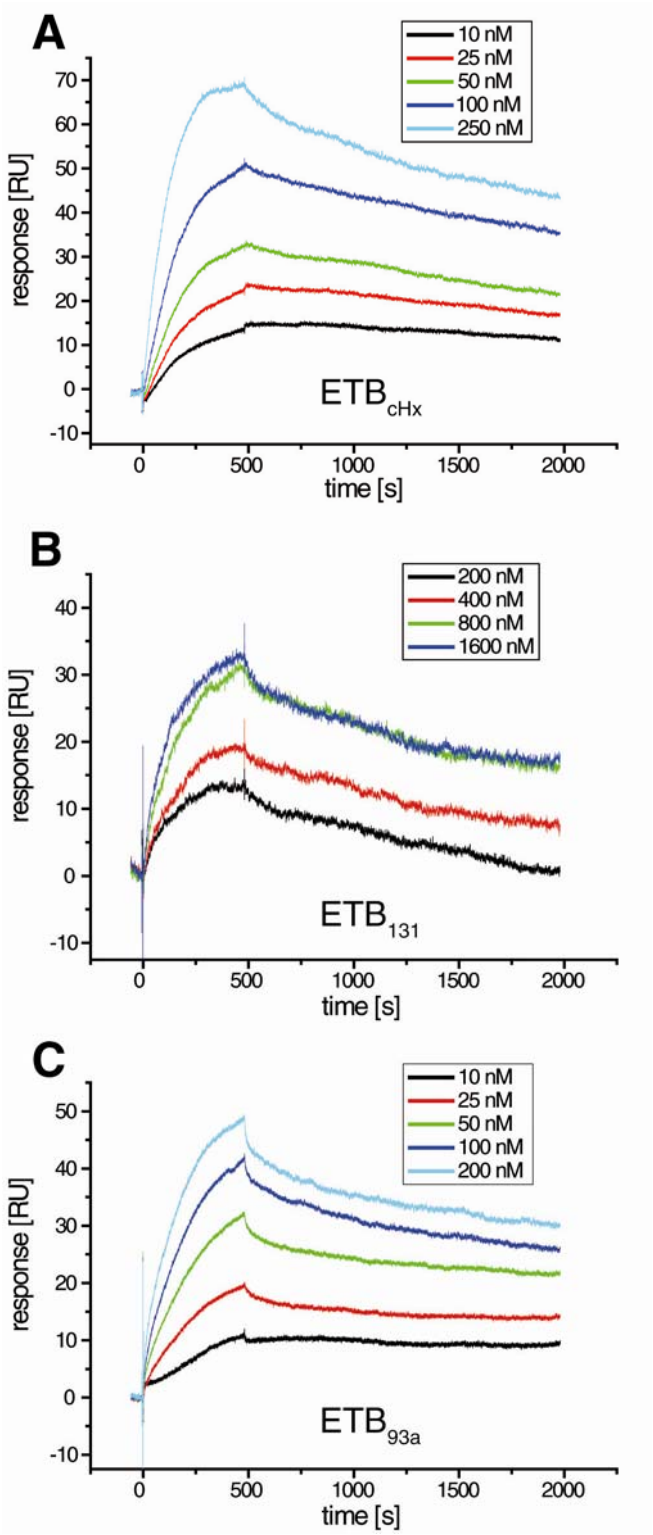


Fig. 27 Biacore data of ETB, ETB_{131} and ETB_{93a}.
SRP response curves for the bET-1/ETB_{cHx}, /ETB_{131} and /ETB_{93a} interaction using immobilized bET-1. 400 to 500 RU of bET-1 have been loaded on Biacore Sensor Chips SA and analyzed in a buffer containing 20 mM HEPES-NaOH (pH 7.4), 500 mM NaCl and 0.1% Brij78. The measurements were performed at 25°C and the response obtained for the blank Sensor Chip SA reference flow cell was subtracted from the curves, revealing negligible non-specific binding to the control peptide surface. **A**, ETB_{cHx} has been titrated between 10 and 250 nM; **B**, ETB_{131} has been analyzed between 200 and 1600 nM; **C**, ETB_{93a} has been titrated between 10 and 200 nM.

Table 8 SPR binding constants of ETB, ETB_{131} and ETB_{93a} and circular bET-1.

construct	k_a [1/Ms]	k_d [1/s]	K_A [1/M]	K_D [M]
ETB_{cHx}	**2.3E+05** ± 1.7E+05	**1.2E-03** ± 0.6E-03	**1.7E+08** ± 0.4E+08	**6.2E-09** ± 1.7E-09
ETB_{131}	**1.1E+05** ± 1.9E+05	**1.4E-03** ± 1.6E-03	**5.4E+07** ± 3.3E+07	**2.7E-08** ± 1.9E-08
ETB_{93a}	**3.6E+04** ± 1.3E+04	**5.7E-04** ± 0.2E-04	**6.3E+07** ± 2.1E+07	**1.7E-08** ± 0.5E-08

6.2.3.4. Analysis of ETB_{cHx}-ligand interaction by TIRFS

Like SPR, total internal reflection fluorescence spectroscopy (TIRFS) enables the determination of k_a and k_d rates and thus the dissociation constant K_D. In this study, ETB_{cHx} was immobilized after CF expression in Brij78 and purification on Tris-NTA Chips loaded with Nickel Ions as described by Lata and co-workers [126] and analyzed on a custom made TIRFS-reflectance interferometry (RIf) combined set-up [127]. The linear fluorescent labeled ligand fET-$1_{A(1,3,11,15)}$, once bound to immobilized ETB_{cHx} at the surface will be excited by the evanescent illumination, whereas other fET-$1_{A(1,3,11,15)}$ in solution will remain dark (Fig. 28). Thus k_a and k_d of fET-$1_{A(1,3,11,15)}$ to ETB_{cHx} can be determined.

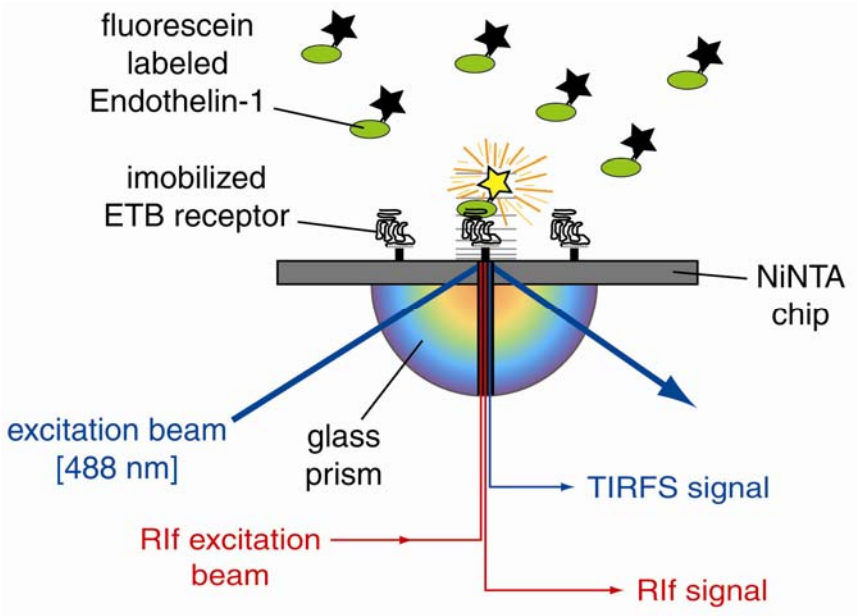


Fig. 28 Principle of total internal reflection fluorescence spectroscopy (TIRFS). The fluorescence labeled ligand, bound to the receptor immobilized by NiNTA on the chip is excited by a laser beam and the TIRFS and RIf signal is measured.

fET-$1_{A(1,3,11,15)}$ concentrations between 0.25 and 2 µM have been titrated before and after loading the surface with ETB_{cHx}. Whereas the RIf signal indicates the immobilization of ETB_{cHx} to the Ni-NTA surface, the TIRFS signal illustrates the amount of excited fET-$1_{A(1,3,11,15)}$ bound to the receptor (Fig. 29). Little background fluorescence is observed that has to be subtracted form the TIRFS signal after immobilizing ETB_{cHx} prior analysis. The binding constant K_D of fET-$1_{A(1,3,11,15)}$ to ETB_{cHx} was determined at $2.9 \pm 2.0 \times 10^{-8}$ M (Table 9) by using the BIAevaluation software.

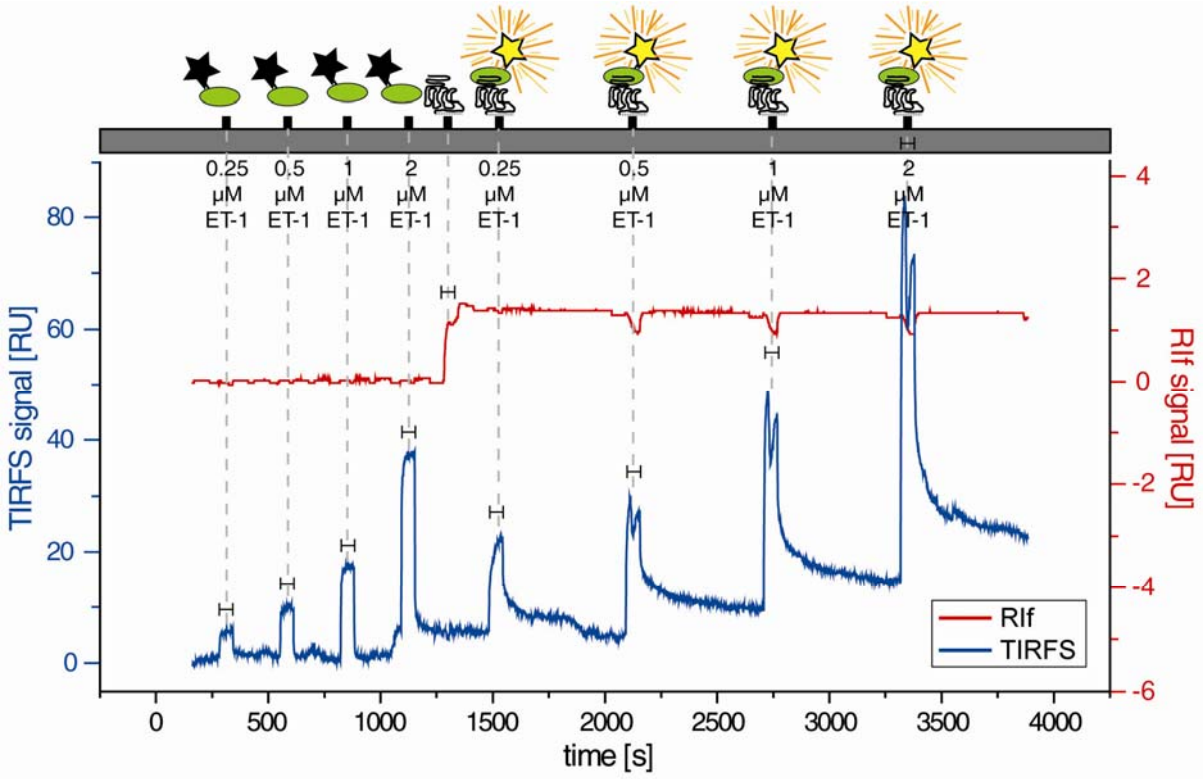


Fig. 29 TIRFS spectrum of ETB_{cHx} immobilized on Ni-NTA chip titrated with $fET\text{-}1_{A(1,3,11,15)}$. Concentrations of $fET\text{-}1_{A(1,3,11,15)}$ of 0.25 to 2 µM have been titrated to an empty chip as negative control first, subsequently ETB_{cHx} was loaded to 1.5 RU and titrated with the same concentrations of linear fluorescein-ET-1. The chromatogram show the RIf signal (red) and the TIRFS signal in blue. The length of $fET\text{-}1_{A(1,3,11,15)}$ and ETB_{cHx} injections are indicated by black bars.

Table 9 TIRFS binding constants for ETBcHx and $fET\text{-}1_{A(1,3,11,15)}$.

construct	k_a [1/Ms]	k_d [1/s]	K_A [1/M]	K_D [M]
ETB_{cHx}	**2.5E+05** ± 0.1E+05	**6.1E-03** ± 2.7E-03	**4.9E+07** ± 0.2E+07	**2.9E-08** ± 2.0E-08

6.2.4. Analysis of the oligomerization state

6.2.4.1. Pull down experiment with full-length protein and truncated constructs

Several GPCRs are known to form dimers that remain stable even after SDS-PAGE analysis. ETB is known to form homodimers in the membrane environment, as determined by fluorescence resonance energy transfer analysis [30]. Protein bands with MWs corresponding to dimers or even higher oligomers of full-length ETB_{cHx} and of most of the truncated derivatives are visible after separation of purified protein samples by SDS-PAGE (Fig. 21 B, C).

However, nothing is known about the regions where ETB oligomerization takes place. In order to identify structural elements responsible for ETB oligomerization, heterodimer formation between full-length ETB_{Strep} and the various truncated ETB fragments was analyzed in two different pull down assays. At first, purified ETB fragments and full-length ETB_{Strep} were incubated at equimolar concentrations and then loaded on Strep-Tactin® columns. In a second assay, the full-length ETB_{Strep} receptor was co-expressed with the various truncated fragments in CF reactions and the RMs were then loaded on Strep-Tactin® columns. In both assays, the interacting protein fragments were identified after washing, elution and SDS-PAGE separation by immunoblotting with the anti-T7-tag antibody (Fig. 30).

In the co-expression assays, the synthesis of full-length ETB_{Strep} and that of the corresponding ETB fragment was always visible by immunoblotting (Fig. 30 A). After loading of the RMs on Strep-Tactin® columns the fragments ETB_{93} (TMS1-TMS3), ETB_{131} (ND-TMS1), ETB_{168} (ND-TMS2), ETB_{203} (ND-TMS3) and ETB_{306} (ND-TMS5) were co-eluted together with ETB_{Strep} indicating an interaction of the proteins. However, fragment ETB_{132} (TMS2-CDO) lacking the TMS1 region was not detectable in the eluted fraction and therefore seems not to interact with ETB_{Strep}. Upon interaction of the purified proteins, again the fragments ETB_{131} (ND-TMS1), ETB_{168} (ND-TMS2), ETB_{203} (ND-TMS3) and ETB_{306} (ND-TMS5) were identified to interact with ETB_{Strep}, while fragments lacking TMS1 like ETB_{132} (TMS2-CDO) and ETB_{204} (TMS4-CDO) could not be co-eluted with full-length ETB and were localized only in the flow-through of the Strep-Tactin® column (Fig. 30 B).

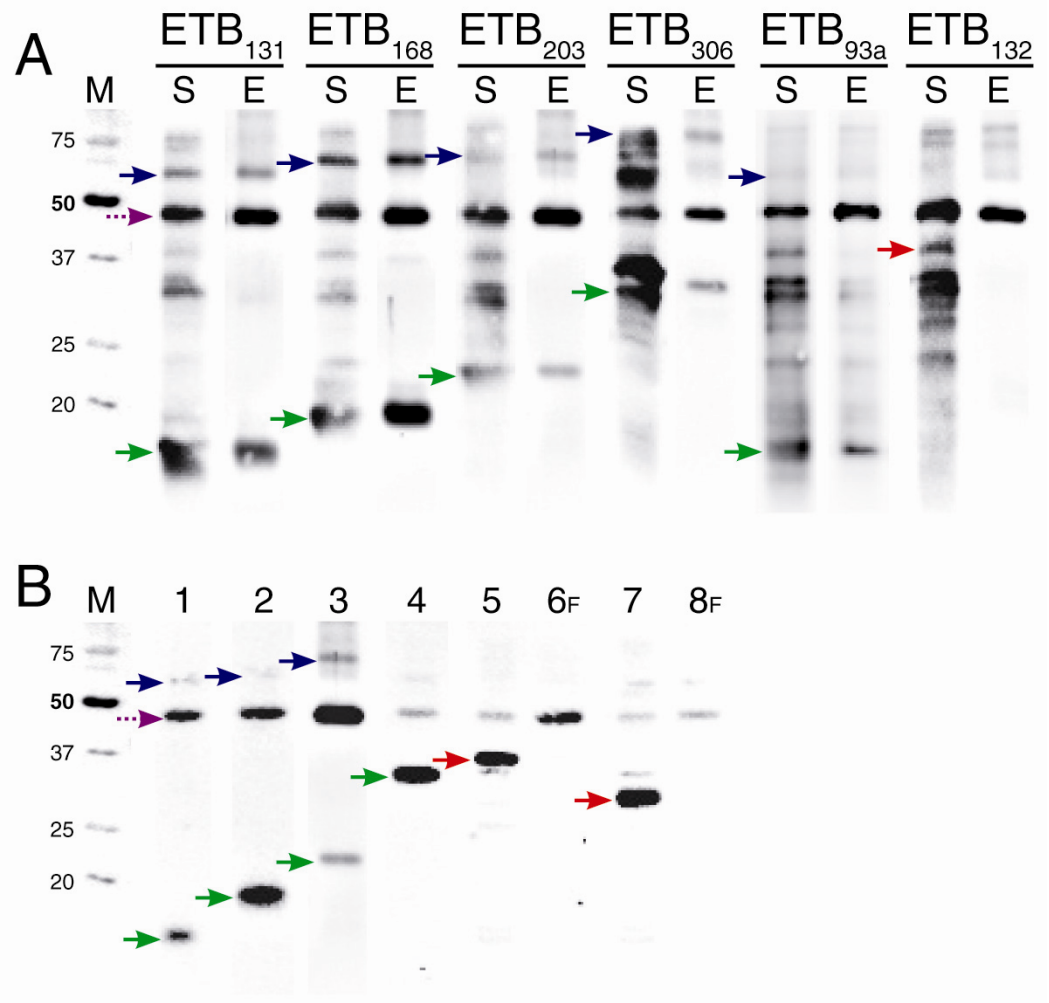


Fig. 30 Dimerization analysis of ETB fragments with ETB_{strep}.
ETB_{step} and ETB fragments containing a deca Histidine tag have been analyzed by pull-down experiments on Strep-Tactin® spin columns. **A**, ETB_{strep} and different fragments have been coexpressed in presence of Brij78 and the soluble part (S) before pull-down and after elution (E) from the Strep-Tactin® spin columns was analyzed by western blot against T7-tag. ETB fragment monomers are indicated by green arrows in case of pull-down or red arrows if not, ETB_{strep} by a purple dotted arrow and hetero-oligomers by dark blue arrows. M, marker; **B**, ETB fragments have been purified separately and incubated in equal amounts with purified ETB_{strep} construct. Flow-through (F) or elution of Strep-Tactin® spin columns are analyzed by western blot against T7-tag. M, marker; 1, elution of ETB_{131}; 2, elution of ETB_{168}; 3, elution of ETB_{203}; 4, elution of ETB_{306}; 5, flow-through of ETB_{132}; 6, elution of ETB_{132}; 7, flow-through of ETB_{204}; 8, elution of ETB_{204}. Monomer fragments are indicated by green arrows in case of pull-down or red if not, hetero-dimers by dark blue arrows and ETB_{strep} by a purple dotted arrow.

6.2.4.2. Single particle analysis of ETB

Besides validating the formation of ETB dimers in detergent micelles, the quality of CF expressed and purified ETB_{cHx} was analyzed by negative stain electron microscopy. In agreement with the observed inability to bind the ligand cET-1, ETB_{cHx} produced as precipitate and solubilized in LMPG was found to be mostly aggregated and is therefore

probably present in an non-native instable state (Fig. 31 A). In contrast, ETB_{cHx} protein that was CF synthesized in presence of the detergent Brij78 revealed mono-disperse particles displaying no detectable signs for aggregation (Fig. 31 B). ETB_{cHx} synthesized at these conditions appears to be predominantly dimeric and the good quality of the sample allowed further structural assessment using single particle analysis. 500 side views were reference-free aligned, classified and averaged within the classes (Fig. 31 B). ETB_{cHx} side view averages display a pair of rods with a length of 63 – 68 Å. The distance between the centers of the rods corresponds to 35 – 38 Å and the rods are closely associated at one end - potentially at TMS1. These values are in excellent agreement with the dimensions observed for the rhodopsin dimer [128, 129]. Single rods presumably representing ETB_{cHx} monomers are also present but they represent less than 10 % of all particles.

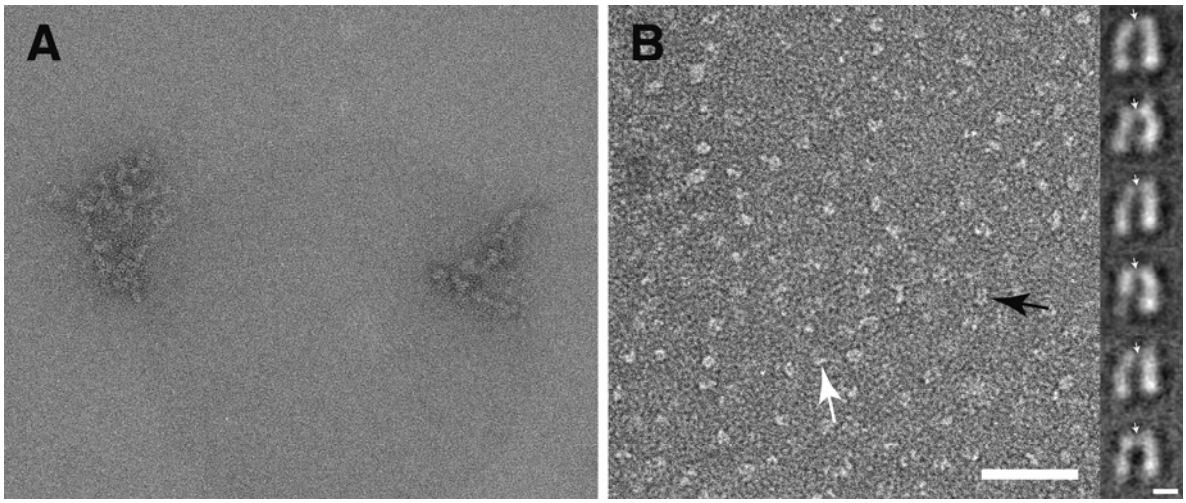


Fig. 31 Single particle analysis of ETB.
Electron-micrograph of negatively stained ETB_{cHx} construct purified by Nickel-chelate chromatography. **A**, ETB_{cHx} was CF expressed as precipitate and resolubilized in 1% LMPG. **B**, ETB_{cHx} was CF expressed in presence of Brij78. Here, ETB_{cHx} particles appear to be predominantly dimeric (black arrow); ETB_{cHx} monomers can be seen occasionally (white big arrow). The scale bar corresponds to 40 nm for A and B. Side view class averages of reference-free aligned ETB dimers are displayed in the gallery on the right. The scale bar corresponds to 3 nm and is valid for all averages. Regions where ETB_{cHx} dimerization takes place is indicated by small white arrows in side view class averages. (The single particle analysis was done in cooperation with Nora Eifler).

6.2.5. Solution NMR spectra of isotopically labeled ETB and its truncations

NMR samples are usually first analyzed by [^{15}N,^{1}H]-TROSY-HSQC measurements, providing a first impression of the sample behavior, displayed by the chemical shift distribution of the amide backbone resonances. However, GPCRs have so far almost been excluded from NMR analysis due to extreme difficulties in preparation of

isotopically labeled samples and the requirement for detergent solutions. The individual CF expression system overcome expression and labeling limitations and the use of LMPG micelles provides relatively good solution NMR spectra, considering the size of protein plus surrounding detergent micelle [23]. ETB_{cHx} and the C-terminal truncations of ETB have been uniformly isotopically labeled with ^{15}N by CF expression in the insoluble mode. ETB_{cHx}, containing 467 amino acids surrounded by LMPG micelles, seems way too large for solution NMR analysis (Fig 32 F). Interestingly, ETB_{131} (ND-TMS1) and ETB_{93a} (TMS1-3) provide reasonable good spectral quality (Fig. 32 A + B), promising for further investigations by solution NMR.

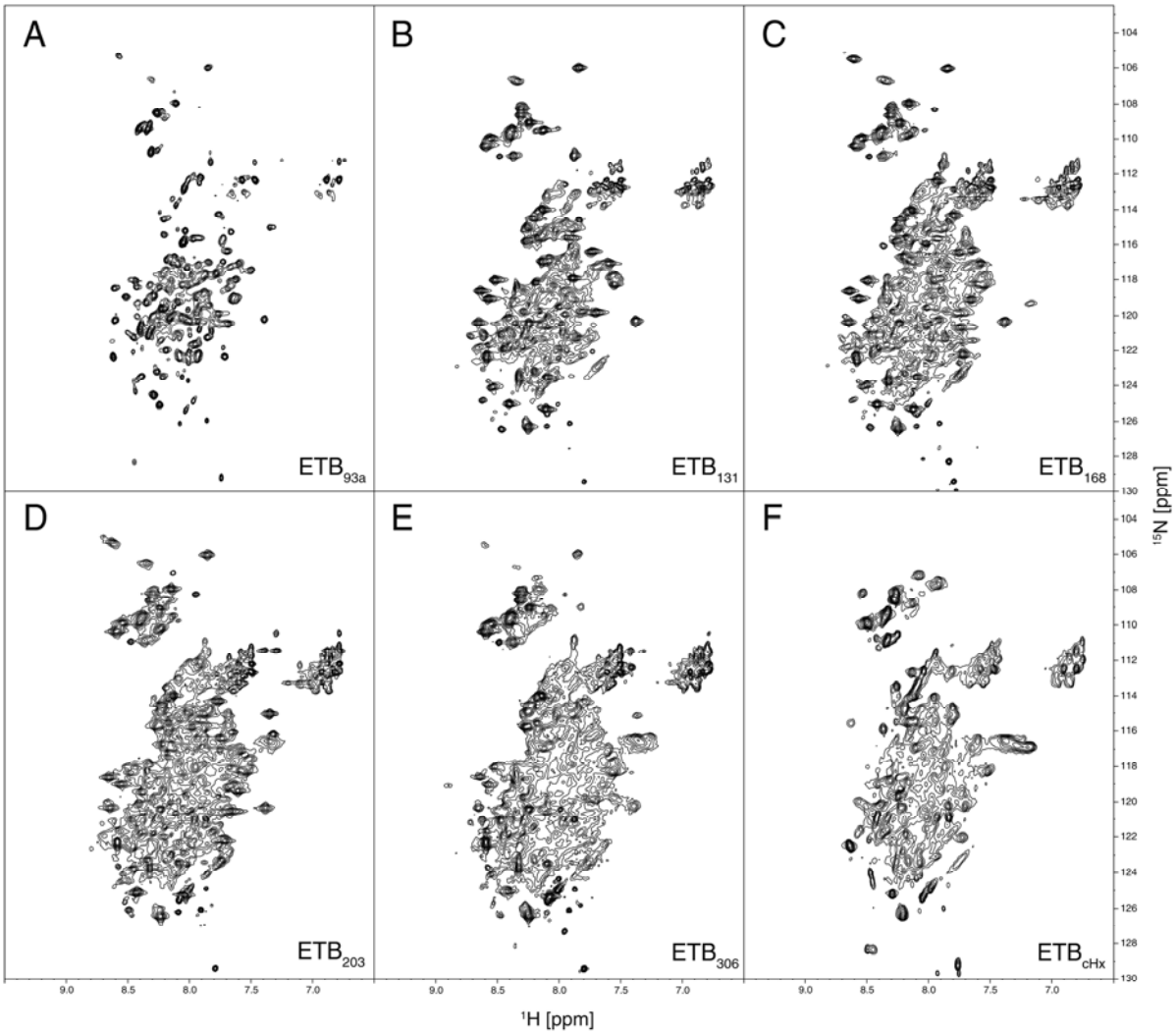


Fig. 32 [^{15}N-^{1}H]-TROSY-HSQC spectra of U-^{15}N labeled ETB_{cHx} and ETB truncations.
The proteins have been U-^{15}N labeled by CF expression in the insoluble mode and the precipitate was solubilized in 25 mM potassium phosphate buffer (pH 6.0), 2% TCEP, 1 % LMPG and concentrated 6 times for NMR analysis. **A**, 0.5 mM ETB_{93} purified by His-tag purification; **B**, 0.3 mM ETB_{131}; **C**, 0.5 mM ETB_{168}; **D**, ETB_{203}; **E**, ETB_{306}; **F**, ETB_{cHx}. The spectra where recorded on an Avance 900 MHz spectrometer (A, C-F) or on an drx 800 MHz spectrometer (B) both equipped with a cryogenic probe.

6.3. Discussion and conclusions

Structural as well as functional evaluation of GPCRs requires milligram quantities of purified and functional protein. GPCRs are typically expressed at very low levels *in vivo* and cell-based functional assays by using fluorescently labeled or radiolabeled ligands are difficult to interpret due to the presence of an undefined variety of other cellular components. Overexpression and native purification of GPCRs as a prerequisite for *in vitro* studies would therefore be essential for the understanding of their functional and structural properties. The high-level production of GPCRs in conventional *in vivo* systems like *E. coli* or *Pichia pastoris* cells is despite numerous efforts still very difficult. Expression of aggregated inclusion bodies or the construction of large fusion proteins are common strategies [130]. In addition, a variety of critical steps like kinetics of membrane insertion, saturation of the biosynthetic translocation machinery, control of proteolysis, growth conditions and solubilization out of the membrane deserve intensive optimizations.

I have established a fast and efficient protocol for the high level production of functionally folded human ETB that eliminates most critical steps of conventional expression systems as membranes are no longer involved. Furthermore, instability of synthesized membrane proteins due to proteolysis can easily be avoided by addition of protease inhibitors. N-terminal digestion of ETB that is frequently observed upon *in vivo* expression [131] was not detectable by CF expression. Terminal truncated derivatives of ETB that are often very difficult to express *in vivo* due to proteolysis or translocation problems [37] can be produced at high levels in the CF system. As a unique advantage of CF expression systems, the proteins can be directly inserted into detergent micelles upon translation. The solubilization efficiency in case of ETB was nearly 100 % as no residual precipitate was detectable and expression levels were similar to those obtained in absence of detergent. The protocol is most straightforward and purified ETB protein sufficient for structural analysis can now be obtained in less than two days. It should also be highlighted that the production of labeled membrane proteins, even with complicated label combinations, is easily feasible by CF expression without the need for extensive optimization screens and without any loss of productivity [18, 19, 101]. Thus ETBcHx and its C-terminal truncations were successfully labeled with ^{15}N isotopes and NMR sample preparation was performed within 24 hours.

As observed upon the production of the porcine vasopressin type 2 receptor [9], only the steroid detergent digitonin and several long chain polyoxyethylene derivatives like Bri35, Brij58 and Brij78 were suitable for the CF synthesis of soluble ETB in milligram amounts. However, the production of functionally folded receptor can vary dramatically

with the type of supplied detergent and the highest apparent ET-1 binding was obtained with Brij78 with some lower activity in digitonin and some other Brij derivatives. Accordingly, only digitonin was found to be suitable to isolate fully active ligand free ETB out of cell membranes [37]. It is known that the binding activity and structural integrity of GPCRs can be sensitive to the supplied detergents during solubilization [132]. Also the functional folding of other membrane proteins like the nucleoside transporter Tsx was found to require the presence of specific detergents during CF expression [9]. CF produced precipitates of Tsx as well as of ETB did not adopt a functional conformation upon solubilization after expression. Expression in the soluble mode and the initial screen for suitable detergents allowing the functional folding of the target proteins are therefore most important for the CF production of functionally folded membrane proteins. ETB is known to become posttranslationally modified by palmitoylation, phosphorylation and glycosylation. However, these modifications do not play a role in the ligand binding capacities of ETB [133] and they are most likely absent in CF produced ETB resulting in more homogenous sample preparations which might be even better suitable for crystallization studies.

SPR studies of GPCRs are generally difficult due to the intrinsic properties of these proteins. Hydrophobic environments are necessary and the SPR sensitivity level requires high receptor concentrations on the biosensor surface in order to detect the binding of low molecular weight ligands. Therefore, only few SPR measurements with GPCRs have been successful so far [132, 134], but these reports already demonstrate that ligands can bind solubilized GPCRs even in lipid-free environments and without the need for membrane reconstitution. Most recently, a modified assay by using the detergent solubilized neurotensin receptor (NTR) as the analyte has been described [135] and I could successfully apply this approach to the characterization of ETB. Interestingly, for both GPCRs the amplitude of the observed response is lower than might be expected if the relatively high mass of the receptor used as analyte is considered. Ligand occlusion by immobilization on the sensor chip surface as well as limited access to the ligand binding site of the receptor due to the presence of detergent molecules might account for this effect. While some potential for the optimization of this technique, e.g. by systematic evaluation of sensor chip surfaces or of linker structures, still might remain, the nice correlation of the presented study with the published results obtained with NTR indicate that the SPR system could be utilized for the optimization of GPCR expression conditions, for the localization of ligand binding sites and for the identification of compounds with new properties important for the pharmaceutical industry. TIRFS has so far not been reported for investigating GPCRs. Here the immobilization of ETB on Ni-NTA chips allowed the determination of the dissociation constant for fET-$1_{A(1,3,11,15)}$. Whereas some artifacts in the spectrum are

observed while association of fET-$1_{A(1,3,11,15)}$ to ETB_{cHx}, probably due to detergent effects, the spectra could successfully interpreted for determining the binding constant, that is about 5 times lower than obtained with SPR. This is in great agreement with the literature. Independent groups found exactly a five fold difference in ETB ligand binding by comparing the dissociation constants of wild type circular ET-1 to the linear ET-$1_{A(1,3,11,15)}$ derivative [136, 137].

Human ETB forms a very tight complex with its ligand ET-1 that remains stable even in 2 % SDS [138]. ET-1 binds with high affinity to purified ETB in Brij78 micelles which is indicated by the determined k_D of 6 nM for bET-1 as well as 29 nM for fET-$1_{A(1,3,11,15)}$. Nevertheless, dissociation constants between 40 – 50 pM of ligand/ETB complexes in various cell environments have been reported [139, 140]. It is known that the ligand binding kinetics of ETB in intact cells is already different from that in the corresponding membrane preparations [141]. In this work, I first determined the dissociation constant of ETB in the environment of detergent micelles and it is also the first analysis of ETB by SPR and TIRFS measurements. The different assay conditions in addition to the use of a modified biotinylated ET-1 or fluorescein labeled linear ET-1 derivative as a ligand most likely have therefore resulted in the modified binding kinetics.

The direct *in vitro* analysis of purified N- and C-terminally truncated ETB derivatives confined the ET-1 binding site to a 39 amino acid area between P_{93} in the N-terminal domain (ND) and C_{131} in the first cytoplasmic loop C1. A previous estimation of the ETB ligand binding site with radiolabeled ET-1 followed by chemical crosslinking and trypsin-digest analysis defined the ET-1 binding domain between residues I_{85} in ND and Y_{200} in the second cytoplasmic loop C2 [142]. The observed data is therefore in agreement with these findings and I further show that primarily TMS1 alone is the central determinant for ET-1 binding. In addition, deletions, mutations and the lack of glycosylation in ND did not impair ET-1 binding to ETB [37]. It is thus interesting to note that rather the presumably membrane embedded TMS1 than the relatively large extracellular domain ND is necessary for ET-1 binding. Based on further indirect analysis by using chimeric ETB derivatives and binding of antagonists, Wada and coworkers proposed a 60 amino acid area spanning I_{138} to I_{197} and thus covering TMS2 and TMS3 as ET-1 binding site. This result might have been caused by side effects of the crosslink approach, different binding sites of the supplied antagonists or by conformational changes of the chimeric ETB derivatives. I could clearly show that ETB truncations devoid of TMS1 but still retaining TMS2 and TMS3 are not able to bind ET-1 in detectable amounts. Nevertheless, the affinity of ETB_{93} and ETB_{131} to ET-1 was reduced approx. one order of magnitude, indicating that other regions of ETB still

might contribute to the ligand binding. Evidence for several and partially overlapping binding sites of ETB for different ligands has been documented [143].

Homooligomerization of rhodopsin–like GPCRs is observed with increasing frequency in recent times and it might represent an important platform for the modulation of GPCR activities like ligand binding, signaling or trafficking properties [128, 144-146]. Even SDS-resistant dimerization of ß2-adrenergic receptor and V2 vasopressin receptor has been reported [146] and SDS resistant dimers of the CF produced porcine vasopressin type 2 receptor could also be detected [9]. The ETB dimer bands observed during SDS PAGE analysis indicate a similar stable association also of this GPCR type. First evidence of homodimer formation of ETB and also its homologue ETA *in vivo* was recently obtained by fluorescence resonance energy transfer analysis in HEK293 cells [30]. Interestingly, ETB dimer formation *in vivo* did not depend on the presence of the ligand ET-1. This is in accordance with the observed oligomerization of CF produced ETB in absence of any ligand. The ETB dimer formation is furthermore strongly supported by single particle analysis and the bi-lobed structures are almost identical to that of rhodopsin [128]. By analyzing CF produced truncated ETB fragments, I could confine the site essential for dimer formation to a 39 amino acid region from P_{93} to C_{131} and covering TMS1 and part of the adjacent loops. The two fragments ETB_{131} and ETB_{93} overlapping in that region did still form homodimers as well as heterodimers with full-length ETB. These results therefore indicate that TMS1 is a key area for two main functions of ETB, the binding of ET-1 as one of the main natural peptide ligands and for providing the interface for ETB dimerization. This close co-localization might raise speculations whether dimer formation could modulate the ligand binding of ETB.

In summary, the presented work might provide a new alternative approach for the functional and structural characterization of ETB and similar GPCRs. Furthermore, the identified ETB_{131} and ETB_{93} fragments retaining ligand binding activity and dimerization properties are even small enough for structural studies by high-resolution NMR analysis.

6.4. Material and methods

6.4.1. Cell-free expression

Proteins were produced in the CECF system based essentially on the previously described protocol [9, 11]. CF extracts were prepared from *E. coli* strain A19. Analytical scale reactions for the optimization of reaction conditions were performed in Microdialysers® (Spectrum Laboratories, Rancho Dominguez, USA) with a molecular weight cut-off (MWCO) of 25 kDa in a RM volume of 70 µl with a RM : feeding mixture (FM) ratio of 1 : 14. Preparative scale reactions were carried out in Dispodialysers® (Spectrum Laboratories, Rancho Dominguez, USA) in a RM volume of 1 ml with a RM : FM ratio of 1 : 17. Dispodialysers were placed in suitable plastic tubes holding the FM and incubated overnight for approx. 20 h. with gentle shaking at 30°C. The reaction was optimized for the concentrations of the ions Mg^{2+} (15 mM) and K^{+} (290 mM). DTT was replaced by 2 mM Tris(2-carboxyethyl)phosphine hydrochloride (Pierce, Rockford, USA). For soluble expression, detergent was supplied during the reaction at the following final concentrations: TX-100 (0.1%), Brij-35 (0.1 %), Brij-58 (1.5 %), Brij-78 (1 %), digitonin (0.4 %).

6.4.2. Cloning procedures and protein analysis

Coding regions of full-length ETB and of truncated derivatives were amplified from cDNA by standard polymerase chain reaction (PCR) techniques using VentDNA-polymerase (NewEnglandBiolabs, Frankfurt, Germany). Purified PCR fragments were cleaved with the restriction enzymes *XhoI* and either *BamHI* or *NdeI* and inserted into the corresponding cloning sites of the expression vector pET21a(+) (Merck Biosciences, Darmstadt, Germany). Suitable restriction sites were added to the DNA fragments by PCR with suitable oligonucleotide primers (Table 10). Additional codons for an extended poly$(His)_{10}$ tag or for the StrepII-tag were inserted by the Quickchange procedure (Stratagene, La Jolla, USA). Plasmid DNA used as templates for CF expression were isolated by using commercial kits (Qiagen, Hilden, Germany).

For SDS-gel analysis, protein samples supplemented with SDS sample buffer were loaded on 12 % or 16.5 % (w/v) Tris/glycine/SDS gels and stained with coomassie blue. For western blot analysis, the gels were blotted on a 0.45 µm Immobilon-P poly(vinylidene difluoride) membrane (Millipore, Eschborn, Germany) in a Hoefer TE22 (GE Healthcare, Freiburg, Germany) wet western blot apparatus for 1 h at 400 mA. The membrane was then blocked for 1 h in blocking-buffer containing 1 x Tris

buffered saline, 7 % skim milk powder (Fluka, Bucks, Switzerland), 0.1 % (w/v) Triton X-100. The horseradish peroxidase-conjugated T7-tag antibody (Merck Biosciences, Darmstadt, Germany) were used at a dilution of 1 : 5000 and incubated for 1 h with the membrane. After extensive washing, the blots were analyzed by chemiluminescence in a Lumi-imager F1™ (Roche Diagnostics, Penzberg, Germany). Protein concentrations were routinely determined by the BCA assay (Sigma, Taufkirchen, Germany).

Table 10 Oligonucleotide primers for the construction of ETB derivatives[1].

Construct		Primer
pET21$_{cHx}$	ETBh10_up	**CTC GAG** CAC CAC CAC CAC CAC CAC CAT CAT CAT CAT TGA GAT CCG GCT
	ETBh10_low	AGC CGG ATC TCA ATG ATG ATG ATG GTG GTG GTG GTG GTG GTG **CTC GAG**
ETB$_{\Delta T7}$	ETB-up	CGG **CATATG** CAG CCG CCT CCA AGT CTG TGC GGA CGC
	ETB-low	CGG **CTC GAG** AGA TGA GCT GTA TTT ATT ACT GGA ACG
ETB$_{cHx}$	ETB_upB	CGG **GGA TCC** CAG CCG CCT CCA AGT CTG TGC GGA CGC
	ETB_low	CGG **CTC GAG** AGA TGA GCT GTA TTT ATT ACT GGA ACG
ETB$_{strep}$	ETB1-bam_up	CGG **GGA TCC** GAG GAG AGA GGC TTC CCG CCT GAC AGG
	ETBstrep_low1	CTG CGG GTG GCT CCA **CCA TGG** AGA TGA GCT GTA TTT ATT ACT GGA
	ETBstrep_low2	GGC **CTC GAG** TTA TTT TTC GAA CTG CGG GTG GCT CCA **CCA TGG**
ETB$_{131}$	ETB1-bam_up	CGG **GGA TCC** GAG GAG AGA GGC TTC CCG CCT GAC AGG
	ETB131_low	CGG **CTC GAG** GCA CTT GTT CTT GTA GAT AAT TCT CAG AAG
ETB$_{168}$	ETB1-bam_up	CGG **GGA TCC** GAG GAG AGA GGC TTC CCG CCT GAC AGG
	ETB168_low	CGG **CTC GAG** TGG CCA GTC CTC TGC CAG CAG CTT GTA GAC
ETB$_{203}$	ETB1-bam_up	CGG **GGA TCC** GAG GAG AGA GGC TTC CCG CCT GAC AGG
	ETBn3Txho_low	CGG **CTC GAG** AGC TCG ATA TCT GTC AAT ACT CAG AGC
ETB$_{306}$	ETB1-bam_up	CGG **GGA TCC** GAG GAG AGA GGC TTC CCG CCT GAC AGG
	ETB1-xho_low	CGG **CTC GAG** GCC ACT TTT CTT TCT CAA CAT TTC ACA
ETB$_{132}$	ETB132_up	CGG **GGA TCC** ATG CGA AAC GGT CCC AAT ATC TTG ATC
	ETB_low	CGG **CTC GAG** AGA TGA GCT GTA TTT ATT ACT GGA ACG
ETB$_{204}$	ETBc4Tbam_up	CGG **GGA TCC** GTT GCT TCT TGG AGT AGA ATT AAA GGA ATT
	ETB_low	CGG **CTC GAG** AGA TGA GCT GTA TTT ATT ACT GGA ACG
ETB$_{307}$	ETB2-bam_up	CGG **GGA TCC** ATG CAG ATT GCT TTA AAT GAT CAC CTA AAG
	ETB_low	CGG **CTC GAG** AGA TGA GCT GTA TTT ATT ACT GGA ACG
ETB$_{93}$	ETB93_up	CGG **GGA TCC** CCC ATC GAG ATC AAG GAG ACT TTC AAA
	ETBn3Txho_low	CGG **CTC GAG** AGC TCG ATA TCT GTC AAT ACT CAG AGC

[1], restriction linkers are in bold, inserted sequences encoding for tags are underlined.

6.4.3. Protein purification, solubilization and reconstitution

After CF expression, precipitated material was separated from the RM by centrifugation at 20000 x g for 10 min. After expression in the soluble mode in presence of detergent, the RM was diluted 1:10 in 20 mM Tris, pH 8.0, 500 mM NaCl in order to reduce the detergent concentration from 1 % to 0.1 % in case of Brij78 and to ensure a better binding to the column in the subsequent purification step. The diluted RMs were then applied on 1 ml HistrapHP columns (GE Healthcare, Freiburg, Germany) equilibrated in 20 mM Tris, pH 8.0, 500 mM NaCl, 0.1 % Brij78. Chromatography was performed with 1 ml/min flow-rate with washing steps of 6 column volumes of column buffer supplemented with 10 mM, 20 mM and 50 mM imidazol, respectively. The bound protein was finally eluted with 375 mM imidazol in column buffer and 0.5 ml fractions were collected. For buffer exchange, protein containing fractions of 0.5 ml were applied on 5 ml desalting columns (GE-Healthcare, Freiburg, Germany) equilibrated with suitable buffer at a flow-rate of 3 ml/min. StrepII-tagged full-length ETB was purified by Strep-Tactin® Spin columns (IBA GmbH, Göttingen, Germany) according to manufacturers recommendations. 1 ml RM was diluted 1 : 15 in 20 mM Tris-HCl, pH 8.0, 150 mM NaCl and divided onto equilibrated 6 columns and supplemented with 0.05 % Brij78. ETB and all its derivatives were further analyzed without removal of the small terminal peptide tags.

CF produced precipitate containing recombinant protein expressed in absence of detergent was suspended in 1 % LMPG in 20 mM phosphate buffer, pH 7.0, in a volume equal to the RM volume. The suspension was incubated for 1 h at RT with gentle shaking followed by centrifugation for 10 min at 20000 x g in order to remove residual precipitate. For reconstitution, a volume of 400 µl purified ETB at a concentration of approx. 0.3 mg/ml in 20 mM Tris, pH 8.0, 500 mM NaCl was mixed with *E. coli* lipid mixture at a molar protein : lipid ratio of 1 : 750. The mixture was incubated at 30^{o}C for 1 h and the detergent was subsequently removed with washed biobeads SM-2 (BioRad, Hercules, USA) by gentle shaking at 25^{o}C for another 36 h with three exchanges of the biobeads.

6.4.4. Ligand binding analysis

The Cy3-dye was attached at position Lys9 of the 21-mer peptide ET-1 (cET-1). Biotin was covalently attached to Cys1 and Lys9 of ET-1 resulting in bET-1. The N-terminal fluorescein labeled linear derivative fET-$1_{A(1,3,11,15)}$ was commercially synthesized (Biosyntan, Berlin, Germany).

For co-elution studies of ET-1 with ETB and its derivatives, 10-30 µg of the purified receptor in 20 mM Hepes, pH 7.4, 150 mM NaCl, 0.02 % DDM was mixed with cET-1 dissolved in 20 % DMSO or fET-$1_{A(1,3,11,15)}$ dissolved in water at a molar ratio of 1 : 5. The mixtures were incubated for 2 hrs at 21°C, filtered and then separated on a Superose 6 PC column (3.2 mm/30 cm) (GE Healthcare, Freiburg, Germany) equilibrated with 20 mM Hepes, pH 7.4, 150 mM NaCl, 0.02 % DDM with a flow-rate of 0.05 ml/min on a SMART chromatography station (GE Healthcare, Freiburg, Germany). The cET-1 ligand was detected by its specific absorption at 550 nm. The linear fluorescein labeled ligand fET-$1_{A(1,3,11,15)}$ was monitored by absorption at 495 nm. Peak area values were calculated by the SMART Manager software. Peak area values obtained for the cET-1 calibration curve were plotted with the Kaleidagraph 3.52 software. Non-specific binding of cET-1 was monitored by saturation of the ETB sample with unlabelled ET-1 ligand for 2 hrs at 21°C followed by incubation with supplied cET-1 for additional 2 hrs and subsequent gelfiltration.

For pull-down assays, the biotinylated ligand bET-1 was mixed with purified ETB or its derivatives in 20 mM Tris, pH 8.0, 500 mM NaCl, 0.1 % Brij78 in a molar ratio of 5 : 1 and incubated for 1 h at 25°C. The mixture was then added to 100 µl of monomeric avidin matrix (Pierce, Rockford, USA) and incubated for 1 h at 4°C with gentle mixing. The irreversible covalent binding sites of the avidin matrix had already been blocked by 2 mM biotin. The matrix was subsequently packed in 2 ml gravity flow columns and washed with approx. 5 column volumes of 20 mM Hepes, pH 7.4, 150 mM NaCl, 0.02 % DDM. The bound bET-1/protein complex was finally eluted with 20 mM Hepes, pH 7.4, 150 mM NaCl, 0.02 % DDM and 2 mM biotin in a total volume of 1 ml. To 1 ml of the elution fraction 25 µl of 1 % sodium deoxycholate was added, mixed and incubated for 15 min at 25°C. Then 1 ml 12 % ice-cold trichloric acid was added, mixed and centrifuged at 10.000 x g for 20 min. at 4°C. The resulting pellet was dried, suspended in 50 µl 0.1 % SDS and analyzed by SDS-PAGE.

6.4.5. Biacore measurements

Kinetic measurements were done with a Biacore T100 (Biacore, Uppsala, Sweden) in 20 mM Hepes-NaOH (pH 7.4), 500 mM NaCl and 0.1 % Brij78 at 25°C. The ligand bET-1 was loaded at 400 to 450 resonance unites (RU) on Biacore Sensor Chips SA with 60 s contact time, a flow rate of 10 µl/min and 1.500 s stabilization time. Responses obtained for the reference flow cell were directly subtracted from the curves, revealing negligible non-specific binding to the control surface. ETB binding was

analyzed at a flow rate of 30 µl/min. with 480 s contact time and 1.500 s dissociation time. Data were processed with the BIAevaluation 3.1 software.

6.4.6. TIRFS measurements

TIRFS measurements were done with a custom made TIRFS-RIf combined set-up [127] using home made Tris-NTA chips loaded with Nickel Ions as described elsewhere [126] in 20 M Hepes-NaOH (pH 7.4), 500 mM NaCl and 0.1% Brij78 at 25°C. As negative control the ligand fET-$1_{A(1,3,11,15)}$ have been titrated at concentrations between 0.25 and 2 µM on an empty chip first with for 36 s at a flow rate of 0.42 mL/min and dissociation times of 500 s with a flow rate of 0,6 mL/min. ETB_{cHx}, CF expressed in presence of Brij78 and purified by Nickel-chelat chromatography has been loaded to 1.5 RU on the Nickel-NTA surface (36s at 0.42 mL/min) and subsequently fET-$1_{A(1,3,11,15)}$ was titrated as before. The obtained TIRFS signals have been processed with the BIAevalution 3.1 software and the calculated dissociation constant (K_D) is the mean of 6 measurements.

6.4.7. Single particle analysis

Different concentrations of purified ETB particles were adsorbed to glow discharged 400 mesh carbon coated Parlodion grids and negatively stained with 2 % (w/v) uranyl acetate. Images were recorded at 50.000 x magnification on Kodak SO163 film using a Hitachi H-8.000 microscope operating at an acceleration voltage of 200 kV. Negatives were digitized on a Heidelberg PrimescanD 7100 drum scanner at a resolution of 2 Å/pixel at the specimen level. The EMAN boxer program [147] was used to select a total of approx. 500 particles from electron micrographs. Particle projections were subjected to reference-free alignment [148] and classification by multivariate statistical analysis [149] employing the SPIDER package [150].

6.4.8. Protein interaction studies

For binding of ETB_{Strep} the Strep-Tactin® Spin Column Kit (IBA GmbH, Göttingen, Germany) was used. For interaction analysis of soluble co-expressed proteins in Brij78, the soluble fraction of an analytical CF reaction obtained after centrifugation of the RM at 20000 x g for 10 min containing ETB_{Strep} and the various His-tagged ETB fragments was 20 times diluted 1 : 20 with 100 mM Tris-HCl, pH 8.0, 150 mM NaCl to a final

volume of 1.4 ml. The solution was split into 2 x 0.7 ml and loaded on pre-equilibrated Strep-Tactin® Spin columns. Washing and elution was essentially followed according to the manufacturers' recommendations with the exception that all buffers were adjusted to 0.05 % Brij78. For interaction of purified proteins, approx. equimolar concentrations of ETB_{Strep} and ETB fragments were combined in 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.05 % Brij78, incubated for 1 h at 25^{o}C and purified by Strep-Tactin® columns as described above.

7. References

1. Toyoshima C, Nakasako M, Nomura H & Ogawa H (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 A resolution. *Nature* **405**, 647-55.

2. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M & Miyano M (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739-45.

3. Murata K, Mitsuoka K, Hirai T, Walz T, Agre P, Heymann JB, Engel A & Fujiyoshi Y (2000) Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599-605.

4. Kim DM & Choi CY (1996) A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. *Biotechnol Prog* **12**, 645-9.

5. Alakov YB, Baranov VI, Ovodov SJ, Ryabova LA, Spirin AS & Morozov IJ. (1995) Method of preparing polypeptides in cell-free translation system in, U.S. Patent 5,478,730,

6. Ryabova LA, Desplancq D, Spirin AS & Pluckthun A (1997) Functional antibody production using cell-free translation: effects of protein disulfide isomerase and chaperones. *Nat Biotechnol* **15**, 79-84.

7. Yin G & Swartz JR (2004) Enhancing multiple disulfide bonded protein folding in a cell-free system. *Biotechnol Bioeng* **86**, 188-95.

8. Pornillos O, Chen YJ, Chen AP & Chang G (2005) X-ray structure of the EmrE multidrug transporter in complex with a substrate. *Science* **310**, 1950-3.

9. Klammt C, Schwarz, D., Fendler K, Haase W, Dötsch V & Bernhard F (2005) Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *Febs J* **272**, 6024-38.

10. Elbaz Y, Steiner-Mordoch S, Danieli T & Schuldiner S (2004) In vitro synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. *Proc Natl Acad Sci U S A* **101**, 1519-24.

11. Klammt C, Löhr F, Schafer B, Haase W, Dötsch V, Rüterjans H, Glaubitz C & Bernhard F (2004) High level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem* **271**, 568-80.

12. Lorch M, Lehner I, Siarheyeva A, Basting D, Pfleger N, Manolikas T & Glaubitz C (2005) NMR and fluorescence spectroscopy approaches to secondary and primary active multidrug efflux pumps. *Biochem Soc Trans* **33**, 873-7.

13. Klammt C, Schwarz D, Löhr F, Schneider B, Dötsch V & Bernhard F (2006) Cell-free expression as an emerging technique for the large scale production of integral membrane protein. *Febs J* **273**, 4141-53.

14. Ishihara G, Goto M, Saeki M, Ito K, Hori T, Kigawa T, Shirouzu M & Yokoyama S (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expr Purif* **41**, 27-37.

15. Berrier C, Park KH, Abes S, Bibonne A, Betton JM & Ghazi A (2004) Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry* **43**, 12585-91.

16. Ozawa K, Jergic S, Crowther JA, Thompson PR, Wijffels G, Otting G & Dixon NA (2005) Cell-free protein synthesis in an autoinduction system for NMR studies of protein-protein interactions. *J Biomol NMR* **32**, 235-41.

17. Ozawa K, Dixon NE & Otting G (2005) Cell-free synthesis of 15N-labeled proteins for NMR studies. *IUBMB Life* **57**, 615-22.

18. Torizawa T, Shimizu M, Taoka M, Miyano H & Kainosho M (2004) Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. *J Biomol NMR* **30**, 311-25.

19. Ozawa K, Headlam MJ, Schaeffer PM, Henderson BR, Dixon NE & Otting G (2004) Optimization of an Escherichia coli system for cell-free synthesis of selectively N-labelled proteins for rapid analysis by NMR spectroscopy. *Eur J Biochem* **271**, 4084-93.

20. Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T & Yokoyama S (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett* **442**, 15-9.

21. Yabuki T, Kigawa T, Dohmae N, Takio K, Terada T, Ito Y, Laue ED, Cooper JA, Kainosho M & Yokoyama S (1998) Dual amino acid-selective and site-directed stable-isotope labeling of the human c-Ha-Ras protein by cell-free synthesis. *J Biomol NMR* **11**, 295-306.

22. Kigawa T, Muto Y & Yokoyama S (1995) Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. *J Biomol NMR* **6**, 129-34.

23. Krueger-Koplin RD, Sorgen PL, Krueger-Koplin ST, Rivera-Torres IO, Cahill SM, Hicks DB, Grinius L, Krulwich TA & Girvin ME (2004) An evaluation of detergents for NMR structural studies of membrane proteins. *J Biomol NMR* **28**, 43-57.

24. Turner RJ, Taylor DE & Weiner JH (1997) Expression of Escherichia coli TehA gives resistance to antiseptics and disinfectants similar to that conferred by multidrug resistance efflux pumps. *Antimicrob Agents Chemother* **41**, 440-4.

25. Parker MJ, Aulton-Jones M, Hounslow AM & Craven CJ (2004) A combinatorial selective labeling method for the assignment of backbone amide NMR resonances. *J Am Chem Soc* **126**, 5020-1.

26. Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R & Choe S (2005) NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* **307**, 1317-21.

27. Battiste JL & Wagner G (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **39**, 5355-65.

28. Howard AD, McAllister G, Feighner SD, Liu Q, Nargund RP, Van der Ploeg LH & Patchett AA (2001) Orphan G-protein-coupled receptors and natural ligand discovery. *Trends Pharmacol Sci* **22**, 132-40.

29. Klabunde T & Hessler G (2002) Drug design strategies for targeting G-protein-coupled receptors. *Chembiochem* **3**, 928-44.

30. Gregan B, Schaefer M, Rosenthal W & Oksche A (2004) Fluorescence Resonance Energy Transfer Analysis Reveals the Existence of Endothelin-A and Endothelin-B Receptor Homodimers. *J Cardiovasc Pharmacol* **44**, S30-S33.

31. Grisshammer R (2006) Understanding recombinant expression of membrane proteins. *Curr Opin Biotechnol* **17**, 337-40.

32. Tate CG (2001) Overexpression of mammalian integral membrane proteins for structural studies. *FEBS Lett* **504**, 94-8.

33. Lundstrom K (2003) Semliki Forest virus vectors for rapid and high-level expression of integral membrane proteins. *Biochim Biophys Acta* **1610**, 90-6.

34. Cramer H, Schmenger K, Heinrich K, Horstmeyer A, Boning H, Breit A, Piiper A, Lundstrom K, Muller-Esterl W & Schroeder C (2001) Coupling of endothelin receptors to the ERK/MAP kinase pathway. Roles of palmitoylation and G(alpha)q. *Eur J Biochem* **268**, 5449-59.

35. Björklöf K, Lundstrom K, Abuin L, Greasley PJ & Cotecchia S (2002) Co- and posttranslational modification of the alpha(1B)-adrenergic receptor: effects on receptor expression and function. *Biochemistry* **41**, 4281-91.

36. Massotte D (2003) G protein-coupled receptor overexpression with the baculovirus-insect cell system: a tool for structural and functional studies. *Biochim Biophys Acta* **1610**, 77-89.

37. Doi T, Hiroaki Y, Arimoto I, Fujiyoshi Y, Okamoto T, Satoh M & Furuichi Y (1997) Characterization of human endothelin B receptor and mutant receptors expressed in insect cells. *Eur J Biochem* **248**, 139-48.

38. Tate CG, Haase J, Baker C, Boorsma M, Magnani F, Vallis Y & Williams DC (2003) Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter. *Biochim Biophys Acta* **1610**, 141-53.

39. Sarramegna V, Talmont F, Demange P & Milon A (2003) Heterologous expression of G-protein-coupled receptors: comparison of expression systems fron the standpoint of large-scale production and purification. *Cell Mol Life Sci* **60**, 1529-46.

40. Lundstrom K (2005) Structural biology of G protein-coupled receptors. *Bioorg Med Chem Lett* **15**, 3654-7.

41. White JF, Trinh LB, Shiloach J & Grisshammer R (2004) Automated large-scale purification of a G protein-coupled receptor for neurotensin. *FEBS Lett* **564**, 289-93.

42. Chrunyk BA, Evans J, Lillquist J, Young P & Wetzel R (1993) Inclusion body formation and protein stability in sequence variants of interleukin-1 beta. *J Biol Chem* **268**, 18053-61.

43. Charles IG, Chubb A, Gill R, Clare J, Lowe PN, Holmes LS, Page M, Keeling JG, Moncada S & Riveros-Moreno V (1993) Cloning and expression of a rat neuronal nitric oxide synthase coding sequence in a baculovirus/insect cell system. *Biochem Biophys Res Commun* **196**, 1481-9.

44. Maurizi MR (1987) Degradation in vitro of bacteriophage lambda N protein by Lon protease from Escherichia coli. *J Biol Chem* **262**, 2696-703.

45. Goff SA & Goldberg AL (1987) An increased content of protease La, the lon gene product, increases protein degradation and blocks growth in Escherichia coli. *J Biol Chem* **262**, 4508-15.

46. Blum P, Ory J, Bauernfeind J & Krska J (1992) Physiological consequences of DnaK and DnaJ overproduction in Escherichia coli. *J Bacteriol* **174**, 7436-44.

47. Tsalkova T, Zardeneta G, Kudlicki W, Kramer G, Horowitz PM & Hardesty B (1993) GroEL and GroES increase the specific enzymatic activity of newly-synthesized rhodanese if present during in vitro transcription/translation. *Biochemistry* **32**, 3377-80.

48. Jermutus L, Ryabova LA & Pluckthun A (1998) Recent advances in producing and selecting functional proteins by using cell-free translation. *Curr Opin Biotechnol* **9**, 534-48.

49. Zubay G (1973) In vitro synthesis of protein in microbial systems. *Annu Rev Genet* **7**, 267-87.

50. Gold LM & Schweiger M (1971) Synthesis of bacteriophage-specific enzymes directed by DNA in vitro. *Methods Enzymol* **20**, 537-542.

51. Gold LM & Schweiger M (1969) Synthesis of phage-specific alpha- and beta-glucosyl transferases directed by T-even DNA in vitro. *Proc Natl Acad Sci U S A* **62**, 892-8.

52. Baranov VI & Spirin AS (1993) Gene expression in cell-free system on preparative scale. *Methods Enzymol* **217**, 123-42.

53. Kawarasaki Y, Nakano H & Yamane T (1994) Prolonged cell-free protein synthesis in a batch system using wheat germ extract. *Biosci Biotechnol Biochem* **58**, 1911-3.

54. Anderson CW, Straus JW & Dudock BS (1983) Preparation of a cell-free protein-synthesizing system from wheat germ. *Methods Enzymol* **101**, 635-44.

55. Tulin EE, Tsutsumi KI & Ejiri SI (1995) Continuously coupled transcription-translation system for the production of rice cytoplasmic aldolase. *Biotechnol Bioeng* **45**, 511-516.

56. Woodward WR, Ivey JL & Herbert E (1974) Protein synthesis with rabbit reticulocyte preparations. *Methods Enzymol* **30**, 724-31.

57. DeVries JK & Zubay G (1967) DNA-directed peptide synthesis. II. The synthesis of the alpha-fragment of the enzyme beta-galactosidase. *Proc Natl Acad Sci U S A* **57**, 1010-2.

58. Lederman M & Zubay G (1967) DNA-directed peptide synthesis. 1. A comparison of T2 and Escherichi coli DNA-directed peptide synthesis in two cell-free systems. *Biochim Biophys Acta* **149**, 253-8.

59. Craig D, Howell MT, Gibbs CL, Hunt T & Jackson RJ (1992) Plasmid cDNA-directed protein synthesis in a coupled eukaryotic in vitro transcription-translation system. *Nucleic Acids Res* **20**, 4987-95.

60. Spirin AS. (1991) *Cell-free protein synthesis bioreactor*, American Chemical Society, Washington, DC.

61. Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K & Ueda T (2001) Cell-free translation reconstituted with purified components. *Nat Biotechnol* **19**, 751-5.

62. Madin K, Sawasaki T, Ogasawara T & Endo Y (2000) A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. *Proc Natl Acad Sci U S A* **97**, 559-64.

63. Chen HZ & Zubay G (1983) Prokaryotic coupled transcription-translation. *Methods Enzymol* **101**, 674-90.

64. Nevin DE & Pratt JM (1991) A coupled in vitro transcription-translation system for the exclusive synthesis of polypeptides expressed from the T7 promoter. *FEBS Lett* **291**, 259-63.

65. Martemyanov KA, Spirin AS & Gudkov AT (1997) Direct expression of PCR products in a cell-free transcription/translation system: synthesis of antibacterial peptide cecropin. *FEBS Lett* **414**, 268-70.

66. Burks EA, Chen G, Georgiou G & Iverson BL (1997) In vitro scanning saturation mutagenesis of an antibody binding pocket. *Proc Natl Acad Sci U S A* **94**, 412-7.

67. Lesley SA, Brow MA & Burgess RR (1991) Use of in vitro protein synthesis from polymerase chain reaction-generated templates to study interaction of Escherichia coli transcription factors with core RNA polymerase and for epitope mapping of monoclonal antibodies. *J Biol Chem* **266**, 2632-8.

68. Ahn JH, Chu HS, Kim TW, Oh IS, Choi CY, Hahn GH, Park CG & Kim DM (2005) Cell-free synthesis of recombinant proteins from PCR-amplified genes at a comparable productivity to that of plasmid-based reactions. *Biochem Biophys Res Commun* **338**, 1346-52.

69. Michel-Reydellet N, Woodrow K & Swartz J (2005) Increasing PCR fragment stability and protein yields in a cell-free system with genetically modified Escherichia coli extracts. *J Mol Microbiol Biotechnol* **9**, 26-34.

70. Ryabova LA, Vinokurov LM, Shekhovtsova EA, Alakhov YB & Spirin AS (1995) Acetyl phosphate as an energy source for bacterial cell-free translation systems. *Anal Biochem* **226**, 184-6.

71. Yao SL, Shen XC & Suzuki E (1997) Biochemical energy consumption by wheat germ extract during cell-free protein synthesis. *Journal of Fermentation and Bioengineering* **84**, 7-13.

72. Kitaoka Y, Nishimura N & Niwano M (1996) Cooperativity of stabilized mRNA and enhanced translation activity in the cell-free system. *J Biotechnol* **48**, 1-8.

73. Baranov VI, Morozov I, Ortlepp SA & Spirin AS (1989) Gene expression in a cell-free system on the preparative scale. *Gene* **84**, 463-6.

74. Spirin AS, Baranov VI, Ryabova LA, Ovodov SY & Alakhov YB (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* **242**, 1162-4.

75. Biryukov SV, Simonenko PN, Shirokov VA, Marjorov SG & Spirin AS. (1999) Method of preparing polypeptides in cell-free system and devise for its realization in, PCT WO 00 50436,

76. Shirokov VA, Simonenko PN, Biryukov SV & Spirin AS. (2002) *Continuous-flow and continuous-exchange cell-free translation systems and reactors*, Springer, Berlin, Germany.

77. Kigawa T & Yokoyama S (1991) A continuous cell-free protein synthesis system for coupled transcription-translation. *J Biochem (Tokyo)* **110**, 166-8.

78. Stiege W & Erdmann VA (1995) The potentials of the in vitro protein biosynthesis system. *J Biotechnol* **41**, 81-90.

79. Kim DM & Swartz JR (2004) Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of Escherichia coli. *Biotechnol Bioeng* **85**, 122-9.

80. Kawasaki T, Gouda MD, Sawasaki T, Takai K & Endo Y (2003) Efficient synthesis of a disulfide-containing protein through a batch cell-free system from wheat germ. *Eur J Biochem* **270**, 4780-6.

81. Ohuchi S, Nakano H & Yamane T (1998) In vitro method for the generation of protein libraries using PCR amplification of a single DNA molecule and coupled transcription/translation. *Nucleic Acids Res* **26**, 4339-46.

82. Rungpragayphan S, Nakano H & Yamane T (2003) PCR-linked in vitro expression: a novel system for high-throughput construction and screening of protein libraries. *FEBS Lett* **540**, 147-50.

83. Sawasaki T, Ogasawara T, Morishita R & Endo Y (2002) A cell-free protein synthesis system for high-throughput proteomics. *Proc Natl Acad Sci U S A* **99**, 14652-7.

84. Busso D, Kim R & Kim SH (2003) Expression of soluble recombinant proteins in a cell-free system using a 96-well format. *J Biochem Biophys Methods* **55**, 233-40.

85. Angenendt P, Nyarsik L, Szaflarski W, Glokler J, Nierhaus KH, Lehrach H, Cahill DJ & Lueking A (2004) Cell-free protein expression and functional assay in nanowell chip format. *Anal Chem* **76**, 1844-9.

86. Lee K & Cohen SN (2001) Effects of 3' terminus modifications on mRNA functional decay during in vitro protein synthesis. *J Biol Chem* **276**, 23268-74.

87. Hirao I, Yoshizawa S & Miura K (1993) Stabilization of mRNA in an Escherichia coli cell-free translation system. *FEBS Lett* **321**, 169-72.

88. Vinarov DA, Newman CL & Markley JL (2006) Wheat germ cell-free platform for eukaryotic protein production. *Febs J* **273**, 4160-9.

89. Vinarov DA & Markley JL (2005) High-throughput automated platform for nuclear magnetic resonance-based structural proteomics. *Expert Rev Proteomics* **2**, 49-55.

90. Vinarov DA, Lytle BL, Peterson FC, Tyler EM, Volkman BF & Markley JL (2004) Cell-free protein production and labeling protocol for NMR-based structural proteomics. *Nat Methods* **1**, 149-53.

91. Sengupta K, Klammt C, Bernhard F & Rüterjans H. (2003) Incorporation of fluorescent labels into cell-free produced proteins in *Cell-Free Protein Expression* (Swartz, JR, ed) pp. 81-88, Springer, Berlin Heidelberg New York.

92. Kiga D, Sakamoto K, Kodama K, Kigawa T, Matsuda T, Yabuki T, Shirouzu M, Harada Y, Nakayama H, Takio K, Hasegawa Y, Endo Y, Hirao I & Yokoyama S (2002) An engineered Escherichia coli tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. *Proc Natl Acad Sci U S A* **99**, 9715-20.

93. Kanda T, Takai K, Hohsaka T, Sisido M & Takaku H (2000) Sense codon-dependent introduction of unnatural amino acids into multiple sites of a protein. *Biochem Biophys Res Commun* **270**, 1136-9.

94. Ye J & van den Berg B (2004) Crystal structure of the bacterial nucleoside transporter Tsx. *Embo J* **23**, 3187-95.

95. Shimada Y, Wang ZY, Mochizuki Y, Kobayashi M & Nozawa T (2004) Functional expression and characterization of a bacterial light-harvesting membrane protein in Escherichia coli and cell-free synthesis systems. *Biosci Biotechnol Biochem* **68**, 1942-8.

96. Yerushalmi H, Lebendiker M & Schuldiner S (1995) EmrE, an Escherichia coli 12-kDa multidrug transporter, exchanges toxic cations and H+ and is soluble in organic solvents. *J Biol Chem* **270**, 6856-63.

97. Schwaiger M, Lebendiker M, Yerushalmi H, Coles M, Groger A, Schwarz C, Schuldiner S & Kessler H (1998) NMR investigation of the multidrug transporter EmrE, an integral membrane protein. *Eur J Biochem* **254**, 610-9.

98. Kelly AE, Ou HD, Withers R & Dötsch V (2002) Low-conductivity buffers for high-sensitivity NMR measurements. *J Am Chem Soc* **124**, 12013-9.

99. Tian C, Breyer RM, Kim HJ, Karra MD, Friedman DB, Karpay A & Sanders CR (2005) Solution NMR spectroscopy of the human vasopressin V2 receptor, a G protein-coupled receptor. *J Am Chem Soc* **127**, 8010-1.

100. Oxenoid K, Kim HJ, Jacob J, Sonnichsen FD & Sanders CR (2004) NMR assignments for a helical 40 kDa membrane protein. *J Am Chem Soc* **126**, 5048-9.

101. Trbovic N, Klammt C, Koglin A, Löhr F, Bernhard F & Dötsch V (2005) Efficient strategy for the rapid backbone assignment of membrane proteins. *J Am Chem Soc* **127**, 13504-5.

102. Liang B, Bushweller JH & Tamm LK (2006) Site-directed parallel spin-labeling and paramagnetic relaxation enhancement in structure determination of membrane proteins by solution NMR spectroscopy. *J Am Chem Soc* **128**, 4389-97.

103. Andre N, Cherouati N, Prual C, Steffan T, Zeder-Lutz G, Magnin T, Pattus F, Michel H, Wagner R & Reinhart C (2006) Enhancing functional production of G protein-coupled receptors in Pichia pastoris to levels required for structural studies via a single expression screen. *Protein Sci* **15**, 1115-26.

104. Trumpp-Kallmeyer S, Chini B, Mouillac B, Barberis C, Hoflack J & Hibert M (1995) Towards understanding the role of the first extracellular loop for the binding of peptide hormones to G-protein coupled receptors. *Pharm Acta Helv* **70**, 255-62.

105. Pantaloni C, Brabet P, Bilanges B, Dumuis A, Houssami S, Spengler D, Bockaert J & Journot L (1996) Alternative splicing in the N-terminal extracellular domain of the pituitary adenylate cyclase-activating polypeptide (PACAP) receptor modulates receptor selectivity and relative potencies of PACAP-27 and PACAP-38 in phospholipase C activation. *J Biol Chem* **271**, 22146-51.

106. George SR, O'Dowd BF & Lee SP (2002) G-protein-coupled receptor oligomerization and its potential for drug discovery. *Nat Rev Drug Discov* **1**, 808-20.

107. Bockaert J & Pin JP (1999) Molecular tinkering of G protein-coupled receptors: an evolutionary success. *Embo J* **18**, 1723-9.

108. Baldwin JM (1993) The probable arrangement of the helices in G protein-coupled receptors. *Embo J* **12**, 1693-703.

109. Yanagisawa M, Kurihara H, Kimura S, Goto K & Masaki T (1988) A novel peptide vasoconstrictor, endothelin, is produced by vascular endothelium and modulates smooth muscle Ca2+ channels. *J Hypertens Suppl* **6**, S188-91.

110. Inoue A, Yanagisawa M, Kimura S, Kasuya Y, Miyauchi T, Goto K & Masaki T (1989) The human endothelin family: three structurally and pharmacologically distinct isopeptides predicted by three separate genes. *Proc Natl Acad Sci U S A* **86**, 2863-7.

111. Baynash AG, Hosoda K, Giaid A, Richardson JA, Emoto N, Hammer RE & Yanagisawa M (1994) Interaction of endothelin-3 with endothelin-B receptor is essential for development of epidermal melanocytes and enteric neurons. *Cell* **79**, 1277-85.

112. Rubanyi GM & Polokoff MA (1994) Endothelins: molecular biology, biochemistry, pharmacology, physiology, and pathophysiology. *Pharmacol Rev* **46**, 325-415.

113. Yoshizawa T, Kimura S, Kanazawa I, Uchiyama Y, Yanagisawa M & Masaki T (1989) Endothelin localizes in the dorsal horn and acts on the spinal neurones: possible involvement of dihydropyridine-sensitive calcium channels and substance P release. *Neurosci Lett* **102**, 179-84.

114. Yanagisawa M & Masaki T (1989) Molecular biology and biochemistry of the endothelins. *Trends Pharmacol Sci* **10**, 374-8.

115. Marasciulo FL, Montagnani M & Potenza MA (2006) Endothelin-1: the yin and yang on vascular function. *Curr Med Chem* **13**, 1655-65.

116. Sakurai T, Yanagisawa M, Takuwa Y, Miyazaki H, Kimura S, Goto K & Masaki T (1990) Cloning of a cDNA encoding a non-isopeptide-selective subtype of the endothelin receptor. *Nature* **348**, 732-5.

117. Arai H, Hori S, Aramori I, Ohkubo H & Nakanishi S (1990) Cloning and expression of a cDNA encoding an endothelin receptor. *Nature* **348**, 730-2.

118. Seo B & Luscher TF (1995) ETA and ETB receptors mediate contraction to endothelin-1 in renal artery of aging SHR. Effects of FR139317 and bosentan. *Hypertension* **25**, 501-6.

119. de Nucci G, Thomas R, D'Orleans-Juste P, Antunes E, Walder C, Warner TD & Vane JR (1988) Pressor effects of circulating endothelin are limited by its removal in the pulmonary circulation and by the release of prostacyclin and endothelium-derived relaxing factor. *Proc Natl Acad Sci U S A* **85**, 9797-800.

120. Verhaar MC, Strachan FE, Newby DE, Cruden NL, Koomans HA, Rabelink TJ & Webb DJ (1998) Endothelin-A receptor antagonist-mediated vasodilatation is attenuated by inhibition of nitric oxide synthesis and by endothelin-B receptor blockade. *Circulation* **97**, 752-6.

121. Davenport AP (2002) International Union of Pharmacology. XXIX. Update on endothelin receptor nomenclature. *Pharmacol Rev* **54**, 219-26.

122. D'Orleans-Juste P, Labonte J, Bkaily G, Choufani S, Plante M & Honore JC (2002) Function of the endothelin(B) receptor in cardiovascular physiology and pathophysiology. *Pharmacol Ther* **95**, 221-38.

123. Miyauchi T & Masaki T (1999) Pathophysiology of endothelin in the cardiovascular system. *Annu Rev Physiol* **61**, 391-415.

124. Sokolovsky M (1995) Endothelin receptor subtypes and their role in transmembrane signaling mechanisms. *Pharmacol Ther* **68**, 435-71.

125. Cooper MA (2002) Optical biosensors in drug discovery. *Nat Rev Drug Discov* **1**, 515-28.

126. Lata S & Piehler J (2005) Stable and functional immobilization of histidine-tagged proteins via multivalent chelator headgroups on a molecular poly(ethylene glycol) brush. *Anal Chem* **77**, 1096-105.

127. Gavutis M, Lata S, Lamken P, Muller P & Piehler J (2005) Lateral ligand-receptor interactions on membranes probed by simultaneous fluorescence-interference detection. *Biophys J* **88**, 4289-302.

128. Fotiadis D, Jastrzebska B, Philippsen A, Muller DJ, Palczewski K & Engel A (2006) Structure of the rhodopsin dimer: a working model for G-protein-coupled receptors. *Curr Opin Struct Biol* **16**, 252-9.

129. Fotiadis D, Liang Y, Filipek S, Saperstein DA, Engel A & Palczewski K (2003) Atomic-force microscopy: Rhodopsin dimers in native disc membranes. *Nature* **421**, 127-8.

130. Grisshammer R, White JF, Trinh LB & Shiloach J (2005) Large-scale expression and purification of a G-protein-coupled receptor for structure determination -- an overview. *J Struct Funct Genomics* **6**, 159-63.

131. Satoh M, Miyamoto C, Terashima H, Tachibana Y, Wada K, Watanabe T, Hayes AE, Gentz R & Furuichi Y (1997) Human endothelin receptors ET(A) and ET(B) expressed in baculovirus-infected insect cells--direct application for signal transduction analysis. *Eur J Biochem* **249**, 803-11.

132. Stenlund P, Babcock GJ, Sodroski J & Myszka DG (2003) Capture and reconstitution of G protein-coupled receptors on a biosensor surface. *Anal Biochem* **316**, 243-50.

133. Okamoto Y, Ninomiya H & Masaki T (1998) Posttranslational modifications of endothelin receptor type B. *Trends Cardiovasc Med* **8**, 327-9.

134. Sen S, Jaakola VP, Pirila P, Finel M & Goldman A (2005) Functional studies with membrane-bound and detergent-solubilized alpha2-adrenergic receptors expressed in Sf9 cells. *Biochim Biophys Acta* **1712**, 62-70.

135. Harding PJ, Hadingham TC, McDonnell JM & Watts A (2006) Direct analysis of a GPCR-agonist interaction by surface plasmon resonance. *Eur Biophys J online*

136. De Leon H & Garcia R (1995) Characterization of endothelin receptor subtypes in isolated rat renal preglomerular microvessels. *Regul Pept* **60**, 1-8.

137. Nakamichi K, Ihara M, Kobayashi M, Saeki T, Ishikawa K & Yano M (1992) Different distribution of endothelin receptor subtypes in pulmonary tissues revealed by the novel selective ligands BQ-123 and [Ala1,3,11,15]ET-1. *Biochem Biophys Res Commun* **182**, 144-50.

138. Takasuka T, Sakurai T, Goto K, Furuichi Y & Watanabe T (1994) Human endothelin receptor ETB. Amino acid sequence requirements for super stable complex formation with its ligand. *J Biol Chem* **269**, 7509-13.

139. Schiller H, Haase W, Molsberger E, Janssen P, Michel H & Reilander H (2000) The human ET(B) endothelin receptor heterologously produced in the methylotrophic yeast Pichia pastoris shows high-affinity binding and induction of stacked membranes. *Receptors Channels* **7**, 93-107.

140. Elshourbagy NA, Korman DR, Wu HL, Sylvester DR, Lee JA, Nuthalaganti P, Bergsma DJ, Kumar CS & Nambi P (1993) Molecular characterization and regulation of the human endothelin receptors. *J Biol Chem* **268**, 3873-9.

141. Hara M, Tozawa F, Itazaki K, Mihara S & Fujimoto M (1998) Endothelin ET(B) receptors show different binding profiles in intact cells and cell membrane preparations. *Eur J Pharmacol* **345**, 339-42.

142. Wada K, Hashido K, Terashima H, Adachi M, Fujii Y, Hiraoka O, Furuichi Y & Miyamoto C (1995) Ligand binding domain of the human endothelin-B subtype receptor. *Protein Expr Purif* **6**, 228-36.

143. Lee JA, Brinkmann JA, Longton ED, Peishoff CE, Lago MA, Leber JD, Cousins RD, Gao A, Stadel JM, Kumar CS & et al. (1994) Lysine 182 of endothelin B receptor modulates agonist selectivity and antagonist affinity: evidence for the overlap of peptide and non-peptide ligand binding sites. *Biochemistry* **33**, 14543-9.

144. Rios CD, Jordan BA, Gomes I & Devi LA (2001) G-protein-coupled receptor dimerization: modulation of receptor function. *Pharmacol Ther* **92**, 71-87.

145. Schoneberg T, Schultz G & Gudermann T (1999) Structural basis of G protein-coupled receptor function. *Mol Cell Endocrinol* **151**, 181-93.

146. Hebert TE & Bouvier M (1998) Structural and functional aspects of G protein-coupled receptor oligomerization. *Biochem Cell Biol* **76**, 1-11.

147. Ludtke SJ, Baldwin PR & Chiu W (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* **128**, 82-97.

148. Penczek P, Radermacher M & Frank J (1992) Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy* **40**, 33-53.

149. Frank J, Verschoor A & Boublik M (1982) Multivariate statistical analysis of ribosome electron micrographs. L and R lateral views of the 40 S subunit from HeLa cells. *J Mol Biol* **161**, 107-33.

150. Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M & Leith A (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**, 190-9.

8. Appendix

List of Figures

List of Tables

8.1. Abbreviations

AcP	acetyl phosphate
ARSases	aminoacyl-tRNA synthetases
ATP	adenosine triphosphate
bET-1	biotinylated Endothelin-1
BHK	baby hamster kidney
Brij35	polyoxyethylene-(23)-lauryl-ether, ($C_{12/23}$)
Brij56	polyoxyethylene-(10)-cetyl-ether, ($C_{16/10}$)
Brij58	polyoxyethylene-(20)-cetyl-ether, ($C_{16/20}$)
Brij72	polyoxyethylene-(2)-stearyl-ether, ($C_{18/2}$)
Brij78	polyoxyethylene-(20)-stearyl-ether, ($C_{18/20}$)
Brij97	polyoxyethylene-(10)-oleyl-ether, ($C_{18\text{-}1/10}$)
Brij98	polyoxyethylene-(20)-oleyl-ether, ($C_{18\text{-}1/20}$)
CD	circular dichroism
CD_3OH	deuterated Methanol
$CDCl_3$	deuterated Chlorophorm
CDO	C-terminal domain
CECF	continuous exchange cell-free
cET-1	Cy3 labeled Endothelin-1
CF	cell-free
CFCF	continuous flow cell-free
CHAPS	3-[(3-cholamidopropyl)dimethly- ammonio]-1-propansulfonat
CHO	Chinese hamster ovary
CMC	critical micellar concentration
COS	simian fibroblast (CV-1 cells) transformed by SV40
CRF	corticotropin releasing factor receptor
CrP	creatine phosphate
CTP	cytidine triphosphate

DDM	n-dodecyl-β-D-maltoside
DDs	Dispodialysers®
DHPC	1,2-diheptanoyl-sn-glycero-3-phosphocholine
DiC_6PC	1,2-dihexanoyl-sn-glycero-3-phosphocholine
diC_8PC	1,2-dioctanoyl-sn-glycero-3-phosphocholine
DM	n-decyl-β-D-maltoside
DNA	desoxy-ribonucleic acid
DPC	Dodecyl-phosphocholine
DTT	dithiotreitol
E. coli	*Escherichia coli*
EFs	elongation factors
EM	electron microscopy
ET-1	Endothelin-1
ETA	endothelin A receptor
ETB	endothelin B receptor
FEBS	Federation of European Biochemical Societies
$fET\text{-}1_{A(1,3,11,15)}$	fluorescein labeled linear Endothelin-1
FM	feeding mixture
Genapol C-100	polyoxyethylene-(10)-dodecyl-ether, ($C_{12/10}$)
GFP	green fluorescent protein
GPCR	G-protein coupled receptor
GTP	guanosine triphosphate
HEK	human embryonic kidney
His	histidine
HNCO	[^{15}N,^{1}H]-TROSY-HN(CO)
HRP	horseradish peroxidase
HSQC	heteronuclear single quantum correlation
HT	high throughput
IFs	initiation factors
IMPs	integral membrane proteins
k_a	association constant
k_d	dissociation constant
K_D	dissociation rate constant (= $1/K_A = k_d/k_a$)
kDa	1000 Dalton
LMPG	1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)]
LMW	low molecular weight
LPPG	1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)]
MDs	Microdialysers®

MHz	1.000.000. Hertz
mRNA	messenger ribonucleic acid
MTSL	1-oxy(-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate
MWCO	molecular weight cut-off
ND	N-terminal domain
NMR	nuclear magnetic resonance
NO	nitric oxide
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser enhanced and exchange spectroscopy
NP40	nonylphenyl-polyethylene-glycol
NTPs	nucleotide triphosphates
NTR	neurotensin receptor
OD	optical density
OMP	outer membrane protein
PCR	polymerase chain reaction
PDI	protein disulfide isomerase
PEP	phosphoenol pyruvate
PRE	paramagnetic relaxation enhancement
PURE	protein synthesis using recombinant elements
RFs	release factors
RIf	reflectance interferometry
RM	reaction mixture
Rr	rabbit reticulocyte
RT	room temperature
RU	resonance unit
SFCF	semi-continuous-flow cell-free
SMR	small multi drug resistance
Strep	streptavidine
Thesit	polyethylene-glycol 400 dedecyl-ether
TIRFS	total internal reflection fluorescence spectroscopy
TMS	transmembrane segment
Triton X-100	polyethylene-glycol P-1,1,3,3-tetramethyl- butylphenyl-ether
tRNA	transport ribonucleic acid
TROSY	transverse relaxation optimized spectroscopy
Tween 20	polyoxyethylene-sorbitan-monolaurate 20
UTP	uracil triphosphate
V2R	vasopressin type 2 receptor
β-OG	n-octyl-β-D-glucopyranoside

9. Publications

(P 01)

High level cell-free expression and specific labeling of integral membrane proteins, C. Klammt, F. Löhr, B. Schäfer, W. Haase, V. Dötsch, H. Rüterjans, C. Glaubitz and F. Bernhard. *Eur. J. Biochem.*, **271**, 568-580. (2004)

Eur. J. Biochem. **271**, 568–580 (2004) © FEBS 2004 doi:10.1111/j.1432-1033.2003.03959.x

High level cell-free expression and specific labeling of integral membrane proteins

Christian Klammt[1], Frank Löhr[1], Birgit Schäfer[1], Winfried Haase[2], Volker Dötsch[1], Heinz Rüterjans[1], Clemens Glaubitz[1] and Frank Bernhard[1]

[1]*Centre for Biomolecular Magnetic Resonance, University of Frankfurt/Main, Institute for Biophysical Chemistry;*
[2]*Max-Planck-Institute for Biophysics, Department for Structural Biology, Frankfurt/Main, Germany*

We demonstrate the high level expression of integral membrane proteins (IMPs) in a cell-free coupled transcription/translation system using a modified *Escherichia coli* S30 extract preparation and an optimized protocol. The expression of the *E. coli* small multidrug transporters EmrE and SugE containing four transmembrane segments (TMS), the multidrug transporter TehA with 10 putative TMS, and the cysteine transporter YfiK with six putative TMS, were analysed. All IMPs were produced at high levels yielding up to 2.7 mg of protein per mL of reaction volume. Whilst the vast majority of the synthesized IMPs were precipitated in the reaction mixture, the expression of a fluorescent EmrE-sgGFP fusion construct showed evidence that a small part of the synthesized protein 'remained soluble and this amount could be significantly increased by the addition of *E. coli* lipids into the cell-free reaction. Alternatively, the majority of the precipitated IMPs could be solubilized in detergent micelles, and modifications to the solubilization procedures yielded proteins that were almost pure. The folding induced by formation of the proposed α-helical secondary structures of the IMPs after solubilization in various micelles was monitored by CD spectroscopy. Furthermore, the reconstitution of EmrE, SugE and TehA into proteoliposomes was demonstrated by freeze-fracture electron microscopy, and the function of EmrE was additionally analysed by the specific transport of ethidium. The cell-free expression technique allowed efficient amino acid specific labeling of the IMPs with ^{15}N isotopes, and the recording of solution NMR spectra of the solubilized EmrE, SugE and YfiK proteins further indicated a correctly folded conformation of the proteins.

Keywords: amino acid specific labeling; cell-free expression; integral membrane proteins; multidrug transporter; solution NMR.

Integral membrane proteins (IMPs) account for 20–25% of all open reading frames in fully sequenced genomes, and in bacteria half of all IMPs are estimated to function as transporters. The active efflux of antibiotics caused by multidrug transporter proteins results in the development of clinical resistance to antimicrobial agents and represents an increasing problem in the treatment of bacterial infections. Despite their importance, no high-resolution structure has been determined thus far from any secondary transporter, from either eukaryotic sources or from the bacterial inner membrane. This is due mainly to the tremendous difficulties generally encountered during the preparation of these multispan integral IMPs to the required purity and amounts [1]. Only some 20 IMPs have been overexpressed in *Escherichia coli* at a level of at least 1 $mg \cdot L^{-1}$ of culture [2,3]. Problems encountered by using conventional *in vivo* systems, such as toxicity of the overproduced protein upon insertion into the cytoplasmic membrane, poor growth of overexpressing strains and the proteolytic degradation of the proteins, could easily be eliminated by cell-free expression. Our primary goal was therefore to analyse whether these restrictions could be solved by the production of IMPs in a cell-free expression system. We have analyzed the efficiency of IMP production in a T7 based cell-free approach using an *E. coli* S30 cell extract in a coupled transcription/translation system [4,5]. During incubation the reaction mixture, containing all enzymes and high molecular mass compounds necessary for gene expression, was dialyzed against a low molecular mass substrate solution providing precursors to extend the protein synthesis for

Correspondence to F. Bernhard, Centre for Biomolecular Magnetic Resonance, University of Frankfurt/Main, Institute for Biophysical Chemistry, Marie-Curie-Str. 9, D-60439 Frankfurt/Main, Germany. Fax: + 49 69 798 29632, Tel.: + 49 69 798 29620, E-mail: fbern@bpc.uni-frankfurt.de

Abbreviations: β-OG, n-octyl-β-glucopyranoside; CMC, critical micellar concentrations; DDM, n-dodecyl-β-D-maltoside; DMPC, 1,2-dimyristoyl-sn-glycero-3-phosphocholine; DPC, dodecylphosphocholine; FID, free induction decay; FM, feeding mixture; GFP, green fluorescent protein; HSQC, heteronuclear single quantum correlation; IMP, integral membrane protein; LPC, L-α-phosphatidylcholine; MAS-NMR, magic angle spinning nuclear magnetic resonance; MHPG, 1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)]; NDSB, nondetergent sulfobetaines; NM, n-nonyl-β-maltoside; POGP, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; RM, reaction mixture; sgGFP, super-glow green fluorescent protein; TMS, transmembrane segment; TPP^+, tetraphenylphosphonium; TROSY, transverse relaxation optimized spectroscopy.

(Received 28 October 2003, revised 28 November 2003, accepted 8 December 2003)

more than 10 h [6,7]. Essential components of the cell-free system such as the bacterial S30 extract preparation, the energy system, the concentrations of precursors and of beneficial additives, have been optimized to yield up to 5 mg of recombinant protein per mL of reaction during a 12 h incubation.

For our expression studies we have chosen secondary transporter proteins from *E. coli* belonging to the families; small multidrug resistance (EmrE, SugE), TDT (TehA) and RhtB (YfiK) [8,9]. The small multidrug resistance (SMR) transporters are typically 110 amino acids in length and they are supposed to consist of four transmembrane segments (TMS) forming a tightly packed four-helix bundle [8–10]. EmrE is a polyspecific antiporter that exchanges hydrogen ions with aromatic toxic cations [11]. Its molecular transport mechanism, and probably also that of the homologous protein SugE, is an electrogenic drug/proton antiport. EmrE is thought to form homooligomeric complexes and specifically transports aromatic dyes, quaternary amines and tetraphenylphosphonium (TPP^+) derivatives [8,11], whilst SugE is presumably only specific for quaternary ammonium compounds [12]. The 36 kDa transporter TehA contains 10 TMS and is responsible for potassium tellurite efflux [13]. Overexpression of TehA further increases the resistance against monovalent cations such as tetraphenylarsonium and ethidium bromide and it decreases the resistance against divalent cations like dequalinium and methyl viologen [13]. A region including TMS 2 to 5, and homologous to proteins of the SMR family, might be primarily responsible for the activity of TehA. YfiK is a 22 kDa transporter with six putative TMS and part of a putative cysteine efflux system [14,15].

Large amounts of pure detergent solubilized IMPs are needed for biochemical characterization or even structural analysis by X-ray crystallography and NMR spectroscopy. This work is the first report of the fast cell-free production of milligram amounts of four different integral transporter proteins, three of which have been amino acid specifically labeled. Whilst a small part of the overproduced proteins could be stabilized post-translationally by the addition of lipids into the cell-free reaction, the precipitated major part of the IMPs could be folded efficiently and solubilized by various detergents. The structural reconstitution of EmrE, SugE, YfiK and TehA was demonstrated by CD spectroscopy, freeze fracturing electron microscopy, NMR spectroscopy and by functional assays.

Experimental procedures

Strains, plasmids, oligonucleotides and DNA techniques

Strains and plasmids used in this study are listed in Table 1. Standard DNA techniques were performed as described elsewhere [17]. The coding sequences for the *E. coli* EmrE, SugE, TehA and YfiK proteins were amplified by standard PCR using the corresponding oligonucleotide primers from MWG-Biotech (Ebersberg, Germany) (Table 2), Vent polymerase (New England Biolabs, Frankfurt/Main, Germany) and chromosomal DNA from strain C600 as a template. The purified amplified DNA fragments were cloned with the enzymes *Nde*I and *Hin*dIII (New England Biolabs) into the expression vector pET21a(+) resulting in the plasmids pET-emrE, pET-sugE, pET-tehA and pET-yfiK. Expression from these plasmids produced the wild type proteins without any modifications or additional tags.

In vitro expression of proteins

Bacterial cell-free extracts were prepared from the *E. coli* strain A19 (*E. coli* Genetic Stock Center CGSC) in a procedure modified after Zubay [18]. The cells were washed in washing buffer [10 mM Tris-acetate, pH 8.2, 14 mM $Mg(OAc)_2$], with 6 mM 2-mercaptoethanol and 0.6 mM KCl. The lysis buffer was the washing buffer supplemented with 1 mM dithiothreitol and 0.1 mM phenylmethanesulfonyl fluoride. The extract was dialysed in washing buffer supplemented with 0.5 mM dithiothreitol and 0.6 mM KOAc. Endogenous mRNA was removed from the ribosomes by incubation of the extract with 400 mM NaCl at 42 °C for 45 min. Aliquots of the cell-free extract were frozen in liquid nitrogen and stored at −80 °C. The cell-free expression was performed in the continuous exchange mode using a membrane with a cutoff of 15 kDa to separate the reaction mixture (RM) containing ribosomes and all enzymes, from the feeding mixture (FM) providing the low molecular mass precursors. The ratio of RM/FM was 1 : 17 (v/v). Reactions in the analytical scale of 70 μL RM

Table 1. Bacterial strains and plasmids used in this study.

Strains and plasmids	Relevant genotype	Reference
BL21 (DE3) Star	*E. coli* B *ompT rne131*	Novagen
C600	*thr-1 leuB6 thi-1 lacY1 glnV44 rfbD1*	CGSC[a]
XL1-Blue	*recA1 lac*[F'Tn10 (Tet^r) $lacI^q$ lacZM15]	[16]
A19	*rna19 gdhA2 his95 relA1 spoT1 metB1*	CGSC[a]
pET21a(+)	T7 promoter Ap^r	Novagen
pQB1-T7-gfp	super glow gfp, Ap^r	QBiogene
pQB1-emrE-gfp	*emrE NheI* in pQB1	this study
pET-gfp	Ap^r, gfp	Roche
pET-emrE	*emrE NdeI-HindIII* in pET21a(+)	this study
pET-sugE	*sugE NdeI-HindIII* in pET21a(+)	this study
pET-tehA	*tehA NdeI-HindIII* in pET21a(+)	this study
pET-yfiK	*yfiK NdeI-HindIII* in pET21a(+)	this study

[a] *E. coli* Genetic Stock Center.

Table 2. Oligonucleotides used in this study.

Oligonucleotide	Sequence
SugE-upNd	cgg cat atg tcc tgg att atc tta gtt att gc
SugE-low	gga aag ctt tta gtg agt gct gag ttt cag acc
EmrE-upNd	cgg cat atg aac cct tat att tat ctt ggt ggt gc
EmrE-low	cgg aag ctt tta atg tgg tgt gct tcg tga c
TehA-up	cgg cat atg cag agc gat aaa gtg ctc aat ttg
TehA-low	cgg aag ctt tta ttc ttt gtc ctc tgc ttt cat taa aac
YfiK-up	cgg cat atg aca ccg acc ctt tta agt gct ttt tgg
YfiK-low	cgg aag ctt tta ata gaa aat gcg tac cgc gca ata gac
EmrE-upNh	cgg gct agc aac cct tat att tat ctt ggt gg
EmrE-lowNh	cgg gct agc atg tgg tgt gct tcg tga c
SugE-upNh	cgg gct agc tcc tgg att atc tta gtt att gc
SugE-lowNh	gga gct agc gtg agt gct gag ttt cag acc

Table 3. Protocol for cell-free protein expression. Amino acids were adjusted according to the composition of the expressed protein. RM, reaction mixture; FM, feeding mixture.

Component	Final concentration in RM	Final concentration in FM
S30-extract	35%	–
Tris-acetate, pH 8.2	3.5 mM	3.5 mM
plasmid DNA	15 μg·mL	–
RNasin[a]	0.3 U·μL^{-1}	–
T7-RNA polymerase	3 U·μL^{-1}	–
E. coli tRNA[b]	500 μg·mL	–
pyruvate kinase	40 μg·mL	–
amino acids	0.5–1 mM	1–1.5 mM
acetyl phosphate	20 mM	20 mM
phosphoenol pyruvate	20 mM	20 mM
ATP	1.2 mM	1.2 mM
CTP	0.8 mM	0.8 mM
GTP	0.8 mM	0.8 mM
UTP	0.8 mM	0.8 mM
1.4-dithiothreitol	2 mM	2 mM
folinic acid	0.2 mM	0.2 mM
complete protease inhibitor[b]	1 tablet per 10 mL	1 tablet per 10 mL
Hepes-KOH pH 8.0	100 mM	100 mM
EDTA	2.8 mM	2.8 mM
magnesium acetate	13 mM	13 mM
potassium acetate	290 mM	290 mM
polyethylenglycol 8000	2%	2%
sodium azide	0.05%	0.05%

[a] Amersham Biosciences. [b] Roche Diagnostics.

were performed in microdialysers (Spectrum Laboratories Inc., Breda, the Netherlands), and larger dispodialysers (Spectrum Laboratories Inc.) were used for preparative scale reactions with RM volumes of 500 μL to 1 mL. The reactions were incubated at 30 °C in a suitable shaker for 20 h. The protocol for the cell-free reaction mixtures is given in Table 3. Amino acid concentrations were adjusted with regard to the amino acid composition of the overproduced proteins. The least abundant amino acids (present at ≤3% in the protein) were added at 1.25 mM, medium abundant (between 3 and ≤8%) at 1.8 mM and highly abundant (more than 8%) at 2.5 mM final concentration. Amino acid specific labeling was achieved by replacing the corresponding amino acids by their isotopically labeled derivatives.

Detergent solubilization of precipitated IMPs

The pellets of cell-free reaction containing the IMPs were suspended in three volumes of washing buffer (15 mM sodium phosphate, pH 6.8, 10 mM dithiothreitol) and centrifuged for 5 min at 5000 ***g***. The washing step was repeated twice. For the reconstitution of proteoliposomes, EmrE was dissolved in one volume of 2% *n*-dodecyl-β-D-maltoside (DDM) in 15 mM Tris/HCl, pH 6.5, and 2 mM dithiothreitol. The mixture was sonified for 1 min in a water bath and then incubated for 1 h at 75 °C. Non dissolved protein was removed by centrifugation at 20 000 ***g*** at 15 °C for 5 min. TehA and SugE were additionally washed in 3% *n*-octyl-β-glucopyranoside (β-OG) in 15 mM sodium phosphate, pH 6.8, 2 mM dithiothreitol for 1 h at 40 °C. YfiK was first washed in 1% *n*-nonyl-β-maltoside (NM) in 25 mM sodium phosphate, pH 7.0, 5 mM dithiothreitol for 1 h at 40 °C. Impurities were removed by centrifugation and the pellet was further washed with 1% dodecyl-phosphocholine (DPC) at the previous conditions. Dissolved impurities were removed by centrifugation at 20 000 ***g*** for 5 min. The pellets were then dissolved with various concentrations of DDM, DPC, 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (MHPG) or SDS if appropriate. β-OG and SDS were from Sigma, DDM, DPC, NM and MHPG were from Avanti Polar Lipids (Alabaster, AL).

Protein analysis

Protein production was analyzed by SDS/PAGE in 17.5% (v/v) Tricine gels [19]. The proteins were silver stained or visualized with Coomassie-Blue (Sigma) as described [17]. Dissolved proteins were quantified according to their specific molar extinction coefficient by measuring the UV absorbance at 280 nm in 6 M guanidine hydrochloride, pH 6.5.

Circular dichroism spectroscopy

Circular dichroism (CD) spectrometry of IMPs dissolved in 15 mM sodium phosphate, pH 6.8, 2 mM dithiothreitol, and containing the appropriate detergents was performed with a Jasco J-810 spectropolarimeter (Jasco Labortechnik, Gross-Umstadt, Germany). Assays were carried out at standard sensitivity with a band width of 3 nm and a response of 1 s. The data pitch was 0.2 nm and the scanning rate 50 nm·min^{-1}. The spectra were recorded from 188 to 260 nm. The presented data are the average of three scans and smoothed by means-movement with a convolution width of 15. The α-helical content of the analyzed proteins was then calculated by the Jasco SECONDARY STRUCTURE ESTIMATION software. In addition, the α-helical content of proteins was calculated according to their primary structure with the PREDICT PROTEIN server at http://cubic.bioc.columbia.edu/pp/ [20].

Reconstitution of proteoliposomes

The protein concentration of membrane proteins solubilized in 1% DDM was determined by UV measurement at 280 nm in 6 M guanidine hydrochloride, pH 6.5, according to their molar extinction coefficients. Approximately 200 μM of the individual protein samples were used for the reconstitution,

and *E. coli* lipids were added at a molar ratio of protein : lipid of 1 : 500. The solutions were then adjusted to 150 mM NH_4Cl and incubated at 40 °C for 30 min. Washed biobeads SM-2 (Bio-Rad), presaturated with *E. coli* lipids were then added in 10-fold excess to the detergent, and the mixture was incubated overnight at 30 °C on a shaker. The biobeads were exchanged twice. The supernatant was then removed, sonified for 1 min in a water bath sonicator, and assayed immediately or stored in liquid nitrogen.

Freeze-fracture electron microscopy

Droplets of the vesicle suspension were placed between two copper blades used as sample holders and then frozen by plunging into liquid ethane cooled to −180 °C by liquid nitrogen. Freeze-fracturing was performed in a Balzers 400T freeze-fracture apparatus (Balzers, Lichtenstein) with the specimen stage at −160 °C. Platinum/carbon shadowing was at 45° (with respect to the specimen stage) whereas pure carbon was evaporated at 90° onto the sample. After thoroughly cleaning the metal replicas in chromosulfuric acid, they were placed on copper grids and analyzed in an EM208S electron microscope (Philips, Eindhoven, the Netherlands).

Ethidium transport by EmrE proteoliposomes

Transport of ethidium bromide into reconstituted EmrE proteoliposomes was carried out as described [11]. Unilamelar vesicles were prepared by extrusion using 400 nm micropore filters. Fluorescence was measured at excitation and emission wavelengths of 545 and 610 nm, respectively, with a band width of 2.5 nm and a data pitch of 0.1 s. Ten microliters of proteoliposomes (approximately 140 nM EmrE) in 15 mM Tris/HCl, pH 6.5; 2 mM dithiothreitol, 150 mM NH_4Cl and 20 $\mu g \cdot mL^{-1}$ circular plasmid DNA (pUC18) were suspended in 980 μL of outside buffer (15 mM Tris/HCl, pH 8.5; 2 mM dithiothreitol; 150 mM KCl) and measured immediately. If appropriate, ligands were added at the following final concentrations: tetraphenylphosphonium (TPP; 50 μM), ethidium bromide (2.5 μM) and nigericine (5 $\mu g \cdot mL^{-1}$) (Sigma). Green fluorescent protein (GFP) fluorescence was measured at excitation and emission wavelengths of 395 and 509 nm, and at 474 and 509 nm for the red shifted mutant superglow (sgGFP).

NMR spectroscopy

Two dimensional ^1H,^{15}N correlated spectra of [98% ^{15}N]Gly,[98% ^{15}N]Ala labeled samples of 0.1 mM EmrE and 0.5 mM SugE in $CDCl_3/CD_3OH/H_2O$ (6 : 6 : 1, v/v/v) with 200 mM ammonium acetate (pH 6.2) and 10 mM dithiothreitol, and of 0.3 mM YfiK in 4% MHPG (v/v) in 25 mM sodium phosphate (pH 7.0) and 5 mM dithiothreitol were obtained with a gradient-sensitivity enhanced [^{15}N,^1H]-transverse relaxation optimized spectroscopy (TROSY) pulse sequence [21,22]. The spectra of EmrE (T = 15 °C) and YfiK (T = 30 °C) were recorded on a Bruker DRX600 spectrometer (Bruker BioSpin GmbH, Karlsruhe, Germany) equipped with a $^1H\{^{13}C,^{15}N\}$ triple-resonance cryoprobe with z-gradient accessory. Acquisition times were adjusted to 140 ms in both dimensions for EmrE. Accumulation of four scans per free induction decay (FID) resulted in a measurement time of 1 h. The spectrum of YfiK resulted from 200 × 768 time-domain data points corresponding to acquisition times of 55 and 53 ms in the ^{15}N and ^1H dimensions, respectively. The total recording time was 16 h using 128 scans per FID. The spectrum of SugE was taken at a Bruker DMX500 spectrometer using a xyz-gradient $^1H\{^{13}C,^{15}N\}$ triple-resonance probe at 15 °C. Acquisition times were 102 ms in both dimensions. Thirty-two transients were recorded for each FID, giving rise to a measurement time of 6 h.

Results

Cell-free expression of integral transporter proteins

The cell-free reaction conditions were first optimized in order to obtain high yields of protein production by titration of each component and by using the expression of green fluorescent protein (GFP) as a monitor. The most critical parameters appeared to be the concentrations of potassium, magnesium and amino acids, and the quality of the prepared S30 extract. The energy regenerating system was most efficient if a combination of phosphoenol pyruvate, acetyl phosphate and pyruvate kinase was used. With the final protocol (Table 3) we received approximately 3 mg of soluble and fluorescent GFP per mL of reaction mixture and almost 80% of the protein was synthesized during the first 7 h of incubation (Fig. 1). Identical reaction conditions were then subsequently used for the expression of the selected IMPs with the only modification being that the amino acid concentrations of the reaction mixtures were specifically adjusted according to the composition of each target protein. The coding sequences of the genes *emrE*, *sugE*, *tehA* and *yfiK* were amplified from the *E. coli* genome by PCR and cloned into the expression vector pET21a(+) containing the T7 regulatory sequences. All four proteins were expressed without any modifications and in each case we obtained a high level production in our cell-free system (Fig. 2). In contrast, the conventional *in vivo* expression

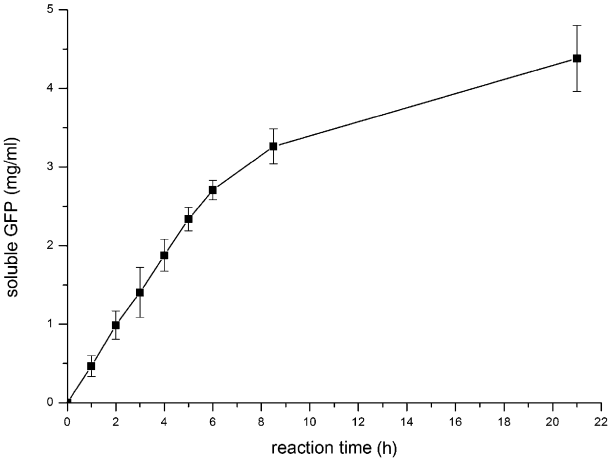


Fig. 1. Protein production kinetics in the cell-free system. Soluble GFP production in a standard cell-free reaction with a membrane cut-off of 25 kDa and an RM/FM ratio of 1 : 17 was monitored by fluorescence at an emission at 509 nm and after excitation at 395 nm. Data are averages of at least three determinations.

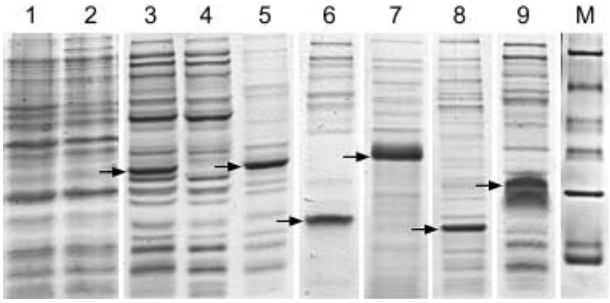


Fig. 2. Cell-free production of membrane proteins. Lanes 1 and 2, *in vivo* expression. Samples of total cell extracts containing 10 μg of protein were analysed by SDS/PAGE in 17.5% (v/v) tricine gels. Lane 1, total protein of BL21 (DE3) Star × pET21-tehA before induction; lane 2, total protein of BL21 (DE3) Star × pET21-tehA 4 h after induction with 1 mM IPTG. Lanes 3–9, cell-free reactions, samples of 1 μL of the reaction mixtures were analysed. Lane 3, pET21-tehA total protein; lane 4, pET21-tehA soluble protein; lane 5, pET21-tehA pellet; lane 6, pET21-emrE pellet; lane 7, pET21-emrE-GFP pellet; lane 8, pET21-sugE pellet; lane 9, pET21-yfiK pellet. M, marker from top to bottom: 116, 66, 45, 35, 25, 18 and 14 kDa. Arrows indicate the overproduced proteins.

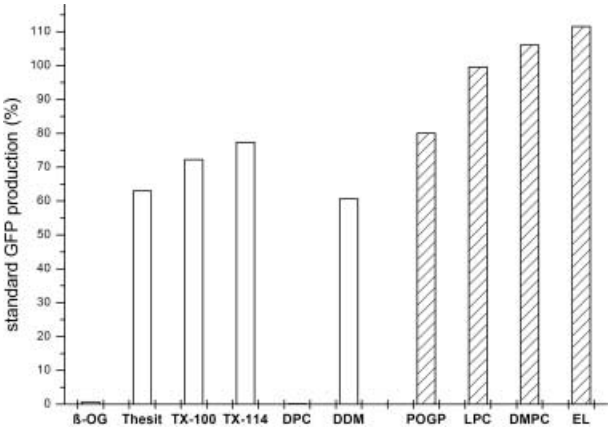


Fig. 3. Effect of selected lipids and detergents on the efficiency of cell-free GFP expression. The reactions were incubated for 7 h at 30 °C. The fluorescence of GFP in a standard cell-free reaction corresponding to an average concentration of 2.6 mg·mL^{-1} was set as 100%. Blank bars, detergents; hatched bars, lipids. Detergent concentrations were 1.5-fold CMC. Lipid concentrations were 4 mg·mL^{-1}. DDM, *n*-dodecyl-β-D-maltoside; DPC, dodecyl phosphocholine; β-OG, *n*-octyl-β-glucopyranoside; TX-100, Triton X-100; TX-114, Triton X-114; LPC, L-α-phosphatidylcholine; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; POGP, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; EL, *E. coli* lipid mixture.

using BL21 (DE3) star cells transformed with the same plasmids yielded no expression detectable by SDS/PAGE analysis. The production rate of all four proteins in the cell-free system was estimated to be at least 1 mg IMP per mL of reaction mixture. However, most of the synthesized IMPs precipitated during the cell-free expression remained insoluble. In order to detect whether a small part of the overproduced proteins might stay soluble, we constructed a fusion of *emrE* to the 5′ end of the gene of the reporter protein sgGFP, resulting in the expression of an EmrE-sgGFP fusion protein. Soluble and correctly folded sgGFP protein can be monitored by its fluorescence at 509 nm and in addition to the more than 1 mg of insoluble fusion protein we could calculate an average of approximately 6 μg of soluble EmrE-sgGFP protein per mL of reaction mixture after standard cell-free expressions.

Modification of the cell-free expression system by addition of detergents and lipids

The results obtained with the EmrE-sgGFP fusion gave evidence that a cell-free expression of IMPs in a soluble condition might be feasible and a major reason for the observed precipitation of the vast majority of the IMPs might be the lack of any hydrophobic environment in the cell-free reaction. We therefore analysed whether the addition of detergents or lipids could increase the solubility of overproduced IMPs. As the addition of those substances might impact the general efficiency of the cell-free reaction, we first tested the production of GFP in the presence of various detergents which have been known to support the functional reconstitution of certain IMPs. DDM, DPC, β-OG, Thesit (Avanti Polar Lipids), Triton X-100 and Triton X-114 (Sigma) were added to the reaction mixtures in concentrations starting from the specific critical micellar concentrations (CMC) up to 1.5-fold CMC. With the highest concentrations tested, all detergents showed a negative effect on the GFP expression, and with DPC and β-OG no synthesized GFP was detectable even at the CMC concentrations (Fig. 3). The detergents DDM, Thesit, Triton X-110 and Triton X-114 showed less drastic effects on the GFP expression and even at the highest concentration analysed, only reductions of ≈ 60–80% of that of the control were observed. A slight increase in amount of soluble EmrE-sgGFP expression was only detectable after addition of Triton X-100 at 1.5-fold CMC (Fig. 4). As expected, DPC and β-OG also completely inhibited the EmrE-sgGFP production when at the CMC (data not shown).

We next analysed the effect of lipids on the cell-free GFP expression. L-α-phosphatidylcholine (LPC), 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POGP) and an *E. coli* lipid mixture were added in increasing concentration only to the RM. POGP resulted in a slight reduction of GFP expression down to approximately 80%, while no negative effects even at the highest analysed concentration of 4 mg lipid per mL RM was noticed with the other three lipids (Fig. 3). The addition of POGP, DMPC and *E. coli* lipids to the cell-free reaction proved to be beneficial for the soluble expression of EmrE-sgGFP protein. An increase in fluorescent EmrE-sgGFP of up to > threefold could be obtained upon addition of *E. coli* lipids (Fig. 4), resulting in a concentration of soluble fusion protein of 20 μg·mL.

Detergent solubilization of EmrE, SugE, YfiK and TehA

As the vast majority of the IMPs still remained insoluble we next approached the solubilization of the precipitated proteins using membrane mimicking detergent micelles.

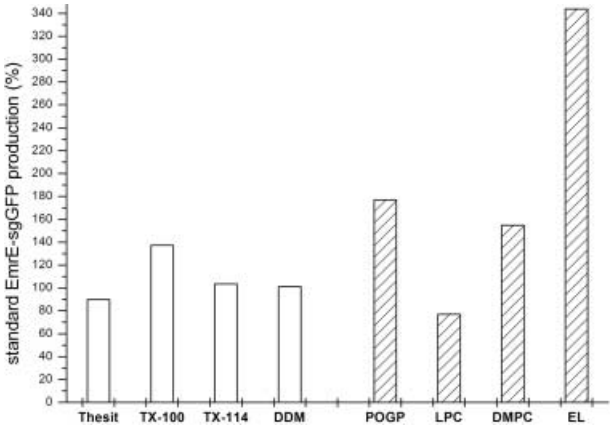


Fig. 4. Increase of soluble EmrE-sgGFP expression in presence of selected lipids and detergents. The fluorescence was measured at 509 nm. The reactions were incubated for 7 h at 30 °C. The fluorescence of EmrE-sgGFP in a standard cell-free reaction corresponding to an average concentration of 5.8 $\mu g \cdot mL^{-1}$ was set as 100%. Blank bars, detergents; hatched bars, lipids. Detergent concentrations were 1.5-fold CMC (TX-110, TX-114, DDM) and twofold CMC (Thesit). Lipid concentrations were 4 $mg \cdot mL^{-1}$. DDM, *n*-dodecyl-β-D-maltoside; TX-100, Triton X-100; TX-114, Triton X-114; LPC, L-α-phosphatidylcholine; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; POGP, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; EL, *E. coli* lipid mixture.

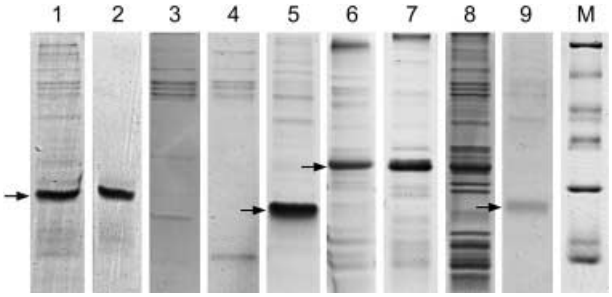


Fig. 5. Purification of cell-free expressed membrane proteins by selective solubilization. Pellets containing the precipitated membrane proteins were dissolved in various detergents in a volume corresponding to the volumes of the original reaction mixtures, and nonsolubilized proteins were removed by centrifugation. 5 μL samples of the soluble fractions were analysed by SDS/PAGE in 17.5% (v/v) tricine gels. Lane 1, EmrE in 2% (v/v) DDM after 1 h at 45 °C; lane 2, EmrE in 2% (v/v) DDM after 1 h at 75 °C; lane 3, SugE in 3% (v/v) β-OG after 2 h at 40 °C; lane 4, SugE in 20% (v/v) NDSB-201 after 2 h at 40 °C; lane 5, SugE in 1% (v/v) DDM after washing with 20% (v/v) NDSB-201; lane 6, TehA in 1% (v/v) DDM after washing with 3% (v/v) β-OG; lane 7, TehA in 1% (v/v) SDS after washing with 3% (v/v) β-OG; lane 8, TehA in 3% (v/v) DPC; lane 9, YfiK in 1% (v/v) DDM after washing with 25% (v/v) NDSB-256. M, marker from top to bottom: 116, 66, 45, 35, 25, 18 and 14 kDa. Arrows indicate the overproduced membrane proteins.

First, the solubility of the IMPs in different detergents dissolved in 15 mM sodium phosphate, pH 6.8, and 2 mM dithiothreitol was analysed, and impurities present in the insoluble pellets of the cell-free reactions were removed where possible. The detergents tested for their ability to solubilize the IMPs were β-OG, DDM, DPC, MHPG, NM, nondetergent sulfobetaines (NDSB-195, -201 and -256), SDS, Thesit, Triton X-100 and Triton X-114. The protein pellets containing the overproduced IMPs and other impurities were first washed twice with 15 mM sodium phosphate, pH 6.8, and 10 mM dithiothreitol. EmrE could then be almost quantitatively dissolved in a buffered 2% (v/v) DDM solution. Co-solubilized impurities could be removed easily by heating the solution to 75 °C for 1 h and apparently pure EmrE remained in solution (Fig. 5). The precipitated SugE and TehA proteins could be further purified by washing the pellets first with 3% (v/v) β-OG or with 20% (v/v) NDSBs. These IMPs dissolved only barely in β-OG or NDSB derivatives, and could be harvested by centrifugation, while most impurities remained β-OG or NDSB soluble (Fig. 5). SugE could then be solubilized in 2% (v/v) DPC, 0.1% (v/v) SDS or 1% (v/v) DDM and TehA solubilized best in 3%(v/v) DPC, 1% (v/v) DDM, or 1% (v/v) SDS. YfiK was washed with 1% (v/v) NM and with 1% (v/v) DPC and then solubilized in 3% (v/v) MHPG. For an efficient solubilization, the proteins were incubated on a shaker at 40 °C for 1 h. In addition, the presence of dithiothreitol was important and a higher molecular mass of the proteins observed after SDS/PAGE analysis without reducing agents indicated the formation of disulfide bridges in the protein precipitates (data not shown).

Structural analysis of solublized EmrE, SugE and TehA by CD spectroscopy

The solubilization of precipitated IMPs into detergent micelles might result in the refolding of the proteins. We therefore analysed the formation of secondary structures of the solubilized IMPs. SugE (15 μM) and TehA (10 μM) were measured in 15 mM sodium phosphate buffer, pH 6.8, 2 mM dithiothreitol, and supplemented with DPC, DDM and SDS, respectively. EmrE was measured in 10 mM sodium phosphate, pH 7.4, 2 mM dithiothreitol and with 2% (v/v) DDM. The spectra measured in the various detergent micelles at 25 °C, showing minima at 208 and 222 nm and a large peak of positive ellipticity centered at 193 nm, were characteristic of α-helical proteins (Fig. 6). The analysis of the spectra yielded an estimate of 55 ± 4% α-helical content for EmrE, 72 ± 11% (DPC), 60 ± 11% (SDS) and 84 ± 10% (DDM) for SugE and 78 ± 8% (DDM), 49 ± 3% (DPC) and 40 ± 15% (SDS) for TehA. The predicted α-helical contents, after primary stuctural analysis, were 69% for EmrE, 67% for SugE and 70% for TehA. According to these data, the adoption of the mostly folded conformation of SugE might be favoured upon solubilization with DPC, and with DDM for TehA, respectively.

Reconstitution of solubilized EmrE, SugE and TehA into proteoliposomes

The precipitated proteins produced by cell-free reactions were solubilized in a 1% (v/v) DDM solution in 15 mM sodium phosphate, pH 6.8, and 2 mM dithiothreitol.

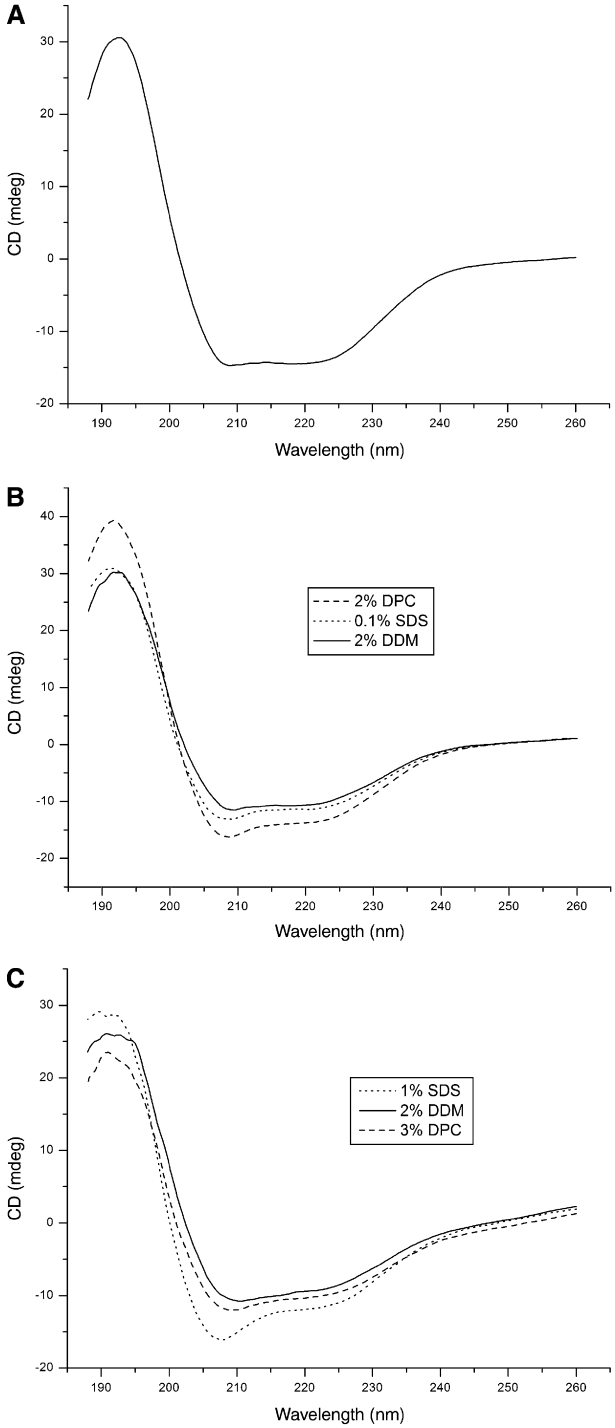


Fig. 6. CD spectroscopy of solubilized multidrug transporter in detergent micelles. Far-UV spectra were taken at 25 °C in buffered detergent solutions. (A) 24 μM EmrE in 2% (v/v) DDM in 10 mM sodium phosphate, pH 7.4. (B) 15 μM SugE in 15 mM sodium phosphate, pH 6.8, 2 mM dithiothreitol with various detergents. (C) 15 μM TehA in 15 mM sodium phosphate, pH 6.8, 2 mM dithiothreitol with various detergents. SDS, sodium dodecylsulfate; DDM, *n*-dodecyl-β-D-maltoside; DPC, dodecyl phosphocholine.

Reconstitution into proteoliposomes with *E. coli* lipids was carried out at a molar protein/lipid ratio of 1 : 500. The insertion of EmrE, SugE and TehA into the lipid membranes was monitored by freeze-fracture electron microscopy (Fig. 7). As would be expected by a functional reconstitution, all three proteins inserted as homogenously dispersed particles into the vesicles. The efficiency of insertion of SugE and EmrE was comparable and an estimated 80% of the vesicles contained inserted proteins. In the case of TehA, the efficiency of proteoliposome generation was less, and ≈10% of the vesicles contained proteins.

Ethidium/H^+ antiport in reconstituted EmrE proteoliposomes

The functional reconstitution of EmrE into proteoliposomes was tested with an established transport assay using ethidium bromide as a ligand [11]. Intercalation of ethidium into DNA causes an effect on the quantum yield of its fluorescence. Active EmrE protein should therefore generate a significant increase in the fluorescence intensity, by pumping ethidium into the proteoliposomes where it is accumulated in the DNA molecules. Approximately 140 nM EmrE embedded in *E. coli* lipids were assayed in a total volume of 1 mL. After establishing the baseline, proteoliposomes were added, followed by ethidium bromide after 10 s to a final concentration of 2.5 μM. An immediate large biphasic increase in the fluorescence was monitored (Fig. 8). The first phase of the increase can be attributed to the binding of ethidium to residual DNA in the extraliposomal space [11], while the second phase represents the accumulation of ethidium inside the liposomes due to the transport activity of EmrE. Preincubation of the proteoliposomes with an excess of 50 μM of the high affinity substrate TPP^+ completely eliminated the second phase, probably through competition with the ethidium binding site at EmrE. In addition, the collapse of the pH gradient upon addition of nigericine also prevented the accumulation of ethidium in the proteoliposomes, resulting only in the single phase increase of fluorescence after addition of ethidium bromide. The results clearly demonstrate that the ethidium/H^+ antiport was responsible for the observed increase in fluorescence, indicating the functional reconstitution of EmrE in *E. coli* lipids.

Structural analysis of selectively labeled EmrE, SugE and YfiK by NMR spectroscopy

One advantage of the cell-free expression technique is the rapid and efficient uniform or amino acid specific labeling of the overproduced proteins. Selected amino acids can be replaced by their labeled derivatives and provided in the reaction mixtures. We selected the relatively abundant amino acids glycine and alanine for a specific labeling approach of EmrE, SugE and YfiK and for the generation of samples suitable for NMR spectroscopy. The quality and dispersion of recorded two dimentional $^1H,^{15}N$ correlation spectra could provide information on whether the solubilized IMPs are either aggregated or present in a folded conformation. However, in addition to

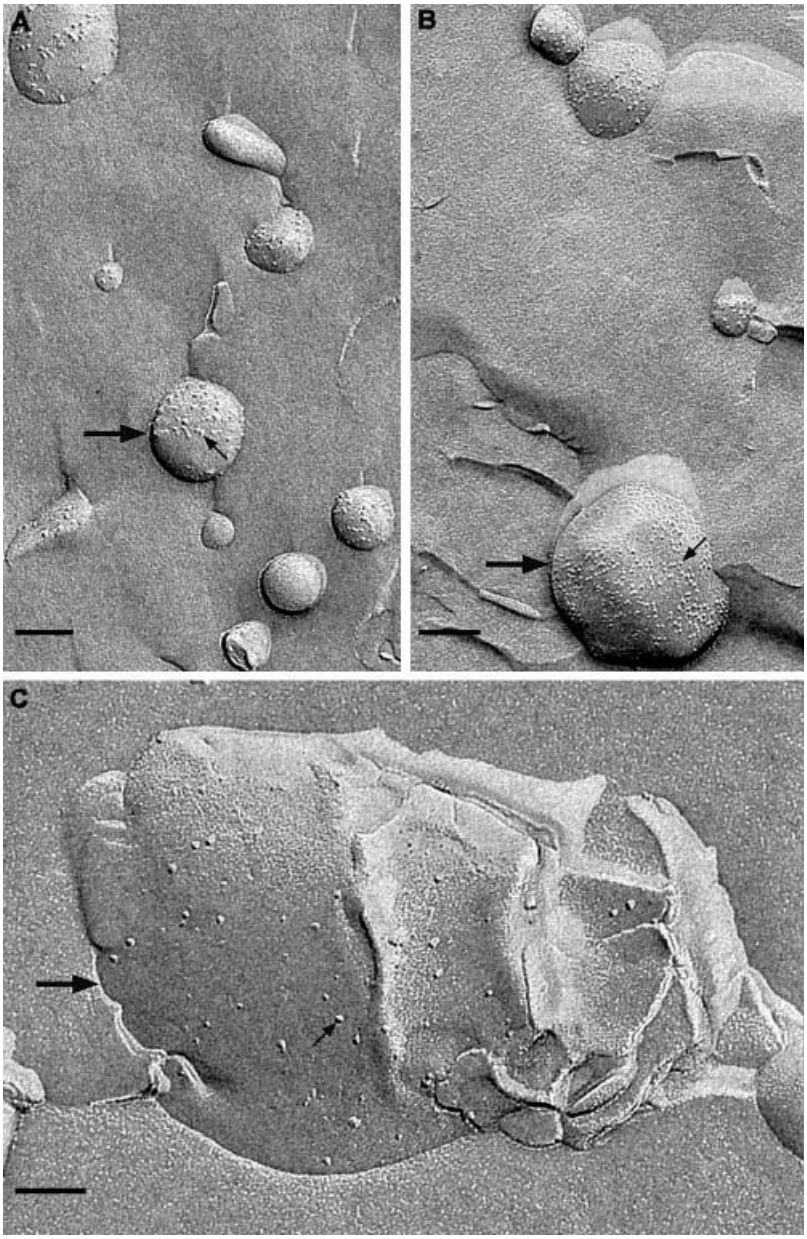


Fig. 7. Freeze-fracture electron microscopical analysis of reconstituted proteoliposomes. The membrane proteins EmrE (A), SugE (B) and TehA (C) were solubilized in 1% (v/v) DDM and reconstituted in *E. coli* lipid vesicles (bold arrows). Randomly distributed particles (small arrows) in the fracture faces indicate incorporation of proteins into vesicular membranes. Scale bar = 100 nm.

the size of the proteins, a major problem for the solution NMR analysis of IMPs, is the size of the detergent micelles necessary for the solubilization. We therefore took advantage of the reported high stability of EmrE in the organic solvent mixture $CDCl_3/CD_3OH/H_2O$ (6 : 6 : 1, v/v/v) with 200 mM ammonium acetate, pH 6.2, and 10 mM dithiothreitol [11,23]. The pellets of preparative scale cell-free reactions with a total of 2 mL RM were washed twice with 15 mM sodium phosphate, pH 6.8, and 2 mM dithiothreitol and then suspended in the chloroform mixture in a volume corresponding to one fourth of the volume of the RM. The suspension was incubated on a shaker for 2 h at 40 °C and then centrifuged at 20 000 ***g*** for 5 min at 15 °C. The supernatant was then used directly for NMR analysis. Interestingly, the SugE protein shared this stability in the chloroform mixture with its homologue EmrE and could be dissolved by using identical procedures. Both proteins were apparently pure in the chloroform mixture as judged by SDS/PAGE analysis and the impurities obviously remained insoluble during this treatment.

The YfiK protein did not dissolve in the chloroform mixture but it showed good solubility in buffered MHPG solutions. The pellets of six preparative reactions with 0.5 mL RM, each containing the YfiK protein, were combined, washed in 1% (v/v) NM and in 1% (v/v) DPC and dissolved in 2 mL of 1% (v/v) MHPG in 25 mM sodium phosphate, pH 6.0, with 5 mM dithiothreitol. After removal of insoluble protein by centrifugation, the sample was concentrated fourfold and measured by NMR. The final protein concentration of YfiK in the sample was calculated at approximately 6 $mg{\cdot}mL^{-1}$, indicating a yield of solubilized labeled YfiK of approximately 1 mg per ml of cell-free RM.

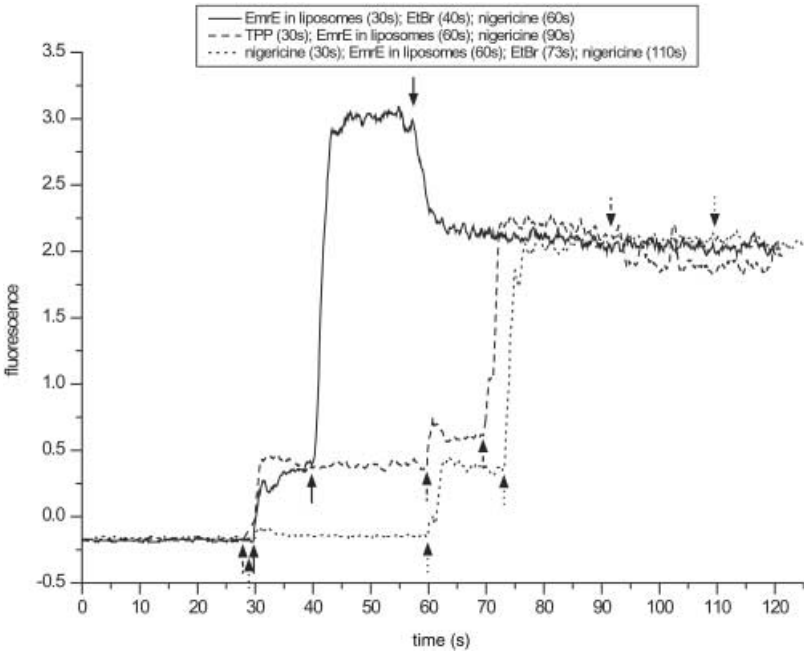


Fig. 8. Ethidium transport assay of EmrE proteoliposomes. Transport of ethidium into reconstituted EmrE proteoliposomes in 15 mM Tris/Cl, pH 8.5, 2 mM dithiothreitol, 150 mM KCl was measured by an increase in fluorescence at excitation and emission wavelengths of 545 and 610 nm, respectively. Ten microliters of proteoliposomes (approximately 140 nM EmrE) were added after 30 or 60 s. If appropriate, substances were added at the following final concentrations: TPP (50 μM), ethidium bromide (2.5 μM) and nigericine (5 $\mu g \cdot mL^{-1}$). Arrows indicate the time points of addition.

The selectively labeled proteins were subsequently analysed by heteronuclear [^{15}N,^{1}H]-TROSY experiments at 500 or 600 MHz ^{1}H frequency. In the EmrE spectrum, all nine alanine residues and 12 glycine residues are visible and well resolved, spanning an area between 7.5 and 9 p.p.m and indicating a specific folded conformation of the solubilized EmrE protein (Fig. 9A). The spectrum could be nicely aligned with a previously published [^{15}N,^{1}H]-HSQC spectrum of uniformly labeled EmrE, prepared by conventional *in vivo* expression and labeling in *E. coli* [23], and all signals of the specifically labeled residues could be assigned accordingly. The dispersion of the amide proton signals also indicated a monomeric conformation of EmrE. The [^{15}N,^{1}H]-TROSY spectrum of the SugE protein also showed a good resolution, and signals of all the 14 alanine and 11 glycine residues were detectable, spanning an area between 7.5 and 8.9 p.p.m, and indicating again a folded conformation of the solubilized protein (Fig. 9B). Despite the size of the 21.3 kDa YfiK protein, the dispersion of its [^{15}N,^{1}H]-TROSY spectrum in MHPG micelles showed a reasonable resolution, and signals of most of the 24 alanine and 13 glycine residues were visible (Fig. 9C).

Discussion

We describe a new and versatile approach for the rapid production, purification and reconstitution of large amounts of structurally folded IMPs, and for the generation of amino acid specific labeled samples suitable for NMR spectroscopy. The production of sufficient amounts of protein is the major bottleneck for the structural and functional analysis of membrane proteins *in vitro*. In addition, if a protein is produced it has to be isolated from complex cellular membranes by time consuming procedures that frequently involve considerable losses. The small multidrug transporter EmrE is one of the few exceptions of IMPs which can also be produced in relatively high amounts by *in vivo* expression. Yields of up to 1 $mg \cdot L^{-1}$ after intensive optimizations in *E. coli* systems have been reported [24] and a hemagglutinin epitope-tagged functional EmrE derivative was expressed in the yeast *Saccharomyces cerevisiae* at levels of approximately 0.5 $mg \cdot L^{-1}$ [25]. For SugE, TehA and YfiK are no quantitative data available for *in vivo* expression, and this is the first report of preparative expression of these proteins. We have been able to demonstrate the cell-free production of at least 1 $mg \cdot mL^{-1}$ of reaction mixture of all of our four target proteins. In the case of SugE and TehA, the production rates were considerably higher. After purification and solubilization into detergent micelles, we could calculate a yield of resolubilizable protein of 1 $mg \cdot mL^{-1}$ RM for YfiK, 1.5 $mg \cdot mL^{-1}$ RM for SugE and of 2.7 $mg \cdot mL^{-1}$ RM for TehA. These calculations did not take into account the amount of proteins which remained insoluble. The obtained production rates of membrane proteins by cell-free expression are therefore comparable to that of other proteins [7,26,27].

The structural reconstitution of EmrE, SugE, YfiK and TehA was monitored by different techniques. EmrE represents one of the best characterized model systems of an integral membrane transporter and its reconstitution is a very well established technique. We included a simple incubation step at 75 °C for the rapid purification of EmrE as it was previously reported that the exposure of EmrE to 80 °C did not affect its transport activity after reconstitution [28]. EmrE is tightly packed without any hydrophilic cytoplasmatic domains [29] and this conformation might cause its somewhat unique solubility and stability in organic solvents [11], and might also favour the observed rapid reconstitution in micelles or liposomes. Homologous proteins of EmrE such as SugE and probably also YfiK and TehA, seem to share these properties and the presented strategy of a cell-free production as precipitate might therefore be advantageous even for this class of IMPs, in

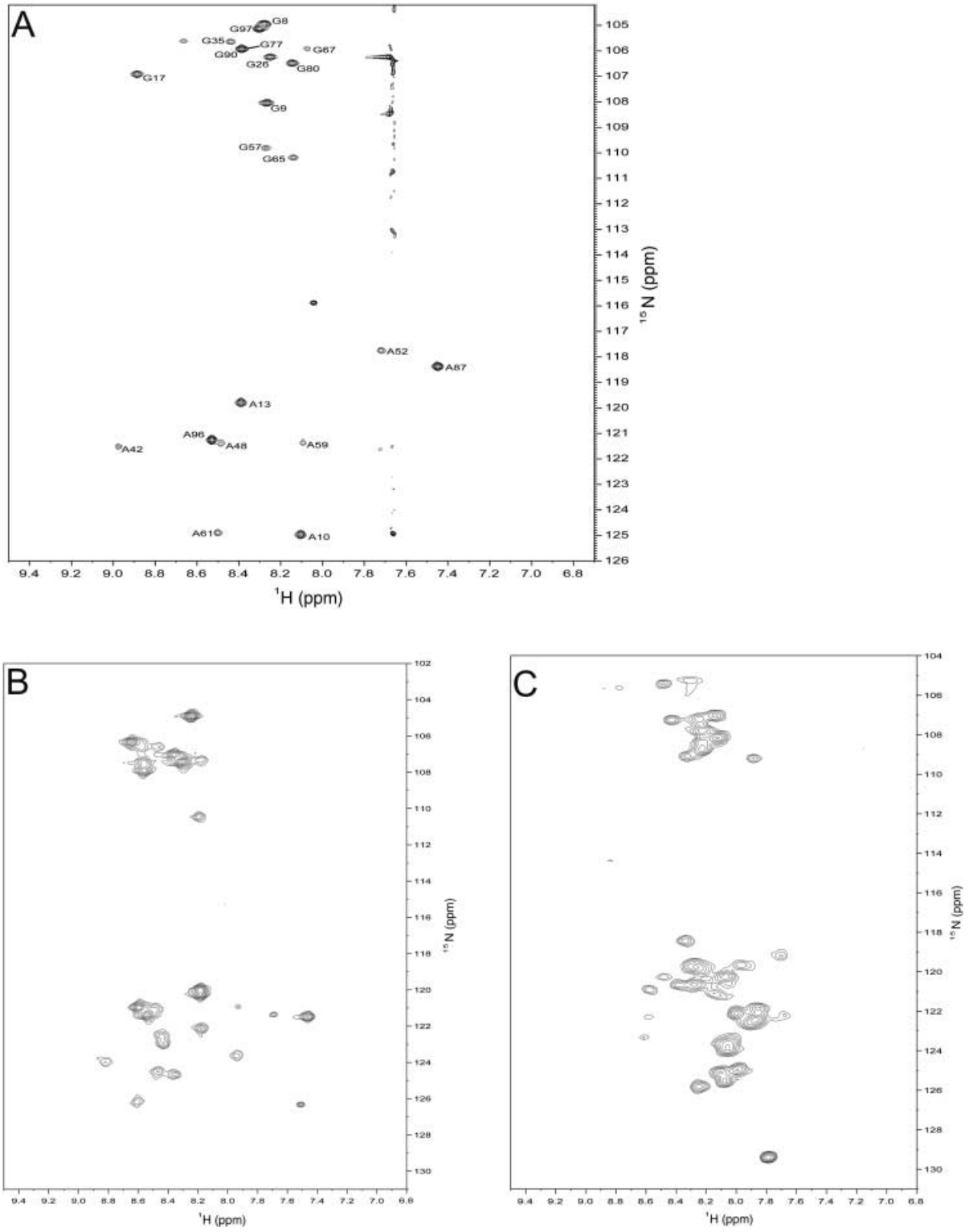


Fig. 9. [^{15}N,^{1}H]-TROSY spectra of solubilized membrane proteins. The proteins were specifically labeled with [^{15}N]alanine and [^{15}N]glycine by cell-free expression. (A) 0.1 mM EmrE dissolved in $CDCl_3/CD_3OH/H_2O$ (6 : 6 : 1, v/v/v) with 200 mM ammonium acetate (pH 6.2) and 10 mM dithiothreitol. The assignments for the amide proton-nitrogen pairs according to Schwaiger *et al.* [23] are indicated. The spectrum was taken at 15 °C with a 600 MHz spectrometer. (B) 0.5 mM SugE dissolved in $CDCl_3/CD_3OH/H_2O$ (6 : 6 : 1, v/v/v) with 200 mM ammonium acetate (pH 6.2) and 10 mM dithiothreitol. The spectrum was taken at 15 °C with a 500 MHz spectrometer. (C) YfiK (0.3 mM) solubilized with 4% (v/v) MHPG in 25 mM sodium phosphate (pH 7.0) and 5 mM dithiothreitol. The spectrum was taken at 30 °C with a 600 MHz spectrometer.

order to obtain pure samples of the nonmodified proteins just by using selective resolubilization protocols in suitable detergents. We could demonstrate for the first time that SugE has a high stability in organic solvents comparable to that of EmrE and that it was able to refold into a structural conformation in the identical chloroform mixture. SugE, like EmrE, appears to be monomeric in chloroform as judged by the dispersion of its [^{15}N,^{1}H]-TROSY spectrum. The spectra of both proteins were well resolved, and the [^{15}N,^{1}H]-TROSY spectrum of the cell-free produced and reconstituted EmrE protein is comparable to that of EmrE prepared after *in vivo* expression [23].

Far-UV CD spectroscopy of solubilized EmrE, SugE and TehA in various detergents revealed spectra typical for predominantly α-helical proteins [30]. EmrE has α-helical estimates of 78% and 80% in chloroform/methanol/water and DMPC, respectively [29,31]. Accordingly, the predicted predominantly α-helical secondary structures of SugE and TehA were in good agreement with the data obtained from CD spectroscopy of the solubilized proteins. The observed differences in α-helicity, in combination with the various detergents, might reflect variations in the protein conformations depending on the type of micelles [32]. An extensive analysis of the effects of different membrane mimetic environments on the conformation of EmrE has recently been published and remarkably, differences in the conformational dynamics, were monitored [33]. The largest amount of α-helical content of EmrE was observed in DDM and the authors assumed that the protein is in a slightly more denatured state in other environments. Their data are in full agreement with our results. Additionally, SugE and TehA also showed the highest α-helicity in DDM.

In MHPG micelles, the YfiK protein showed a reasonable resolution in the [^{15}N]-TROSY spectrum, as would be expected from a protein with a mass in the range ≈ 50–100 kDa. Classical multidrug transporters contain 12 TMS per monomer or functional unit. The EmrE monomer would therefore be three times smaller than this 12 TMS consensus, and it is speculated that functional EmrE might be composed of three subunits [10,34]. It could therefore be possible that the six TMS containing YfiK monomers might reconstitute as oligomers. Considering the estimated micellar size of DPC of ≥ 25 kDa, even as a monomer the analysed molecules would have a minimum size of 47 kDa, which is then in agreement with the observed data.

For the functional analysis of the multidrug transporter EmrE, we could take advantage of a previously established activity assay [11], and the functional reconstitution of the cell-free produced and solubilized protein into proteoliposomes could be clearly demonstrated. The ethidium transport could be specifically competed against the high affinity substrate TPP^+ [34], and it was eliminated by affecting the membrane proton gradient with nigericine. Unfortunately, ethidium is not a substrate for SugE and as only nonfluorescent quarternary ammonium compounds have been reported as potential ligands [12], analoguous assays have not be established to date. Ethidium is a potential substrate of TehA but we have not been able to detect any transport activity with proteoliposomes of TehA solubilized either in DPC, DDM or SDS and reconstituted with an *E. coli* lipid mixture (data not shown). However, the analysis of proteoliposomes by freeze-fracture electron microscopy gave evidence of a structural reconstitution of SugE, EmrE and TehA in *E. coli* lipid vesicles, and no differences between SugE and EmrE proteoliposomes could be observed. It should also be noted that the function of TehA is not very well analysed yet, and it is not clear so far whether the transport activity requires TehA alone or in a complex with other proteins or cofactors [13].

GFP has been shown to be a sensitive folding indicator for the study of globular and membrane protein overexpression in *E. coli* [35,36], and it is most likely to become correctly folded as a C-terminal fusion that is not translocated through the membrane into the periplasm. Therefore, at least the C-terminus of the target protein should remain in the cytoplasm. Approximately 70% of all predicted membrane proteins are believed to have this topology. For EmrE, the cytoplasmic localization of the N- and C-terminal ends has been shown [29], and the C-terminal fusion of GFP should therefore not prevent its reconstitution into membranes. In addition, a fully functional chimera between EmrE and GFP was expressed in *S. cerevisiae* and it conferred resistance against TPP^+, acriflavine and ethidium [25]. It can therefore be assumed that the observed fluorescent part of the cell-free produced EmrE-sgGFP fusion also contains a functionally folded EmrE protein. Despite optimized conditions upon addition of *E. coli* lipids, only an estimate of approximately 1% of the total overproduced protein stayed soluble. While this could already be sufficient for certain analytical assays, higher yields of soluble membrane proteins might be possible by increasing the added amounts of lipids or by providing alternative hydrophobic environments. Dog pancreas microsomes have, for example, been used to produce analytical amounts of completely assembled human T-cell receptor by *in vitro* expression [37]. The cell-free expression principally offers the opportunity to insert the translated protein directly into the desired membrane of choice. Tedious efforts of delipidation and reinsertion of the overproduced membrane proteins into artificial membranes could therefore be avoided, and the possibility of soluble cell-free membrane protein expression might be considered if the reconstitution of a protein is not possible or if only analytical amounts of protein are needed.

Membrane proteins are difficult to analyse by solution NMR techniques, and the main problems are caused by the sizes of the detergent micelles needed for solubilization. Spectra are frequently very crowded and the low dispersion of signals prevents the effective assignment of residues. A valuable tool to approach this problem is the amino acid specific labeling of membrane proteins by cell-free expression. Whilst the selective labeling of proteins for NMR studies in both individual and commercial cell-free expression systems has already been demonstrated [26,38–41], this report shows the first application of this technique to membrane proteins. The selective labeling of proteins by cell-free expression is highly efficient and advantageous compared with the *in vivo* labeling. No auxotrophic strains and minimal media are needed, and commonly encountered problems with reduced expression rates are thus eliminated. In addition, due to the lack of any metabolism during cell-free expression, cross-labeling problems usually do not occur. The presented [^{15}N,^{1}H]-TROSY spectra of EmrE, SugE and YfiK nicely demonstrate the highly efficient

amino acid selective labeling of membrane proteins without any losses in the protein yields. Together with the fast generation of samples, the selective cell-free labeling of membrane proteins could considerably accelerate the assignment of proteins showing a reasonable resolution. The approach presented here might become especially valuable for solid-state NMR studies. The possibility of producing mg quantites of membrane proteins, with the option of using a range of different isotope labeling schemes, enables structural studies of IMPs reconstituted into lipid membranes. So far, only ligand studies by MAS-NMR have been feasible for these protein families [42], but solid-state NMR studies on some of the presented proteins are already in progress.

Cell-free expression has a high potential to become a valuable tool for the rapid generation of samples suitable for structural analysis [43], and commercially available systems have been developed for the efficient production of proteins on a preparative scale [39,44]. In addition, cell-free expression might also yield more homogenous protein samples more readily suitable for crystallization. The main advantages of the cell-free expression of IMPs were the high level production of insoluble protein and the efficient selective labeling. This is the first report of the solubilization of SugE, YfiK and TehA in micelles and of their reconstitution into membranes. The dissolving of the proteins in suitable detergents obviously resulted in the refolding of the proteins, and renaturation procedures with strong denaturants such as guanidine hydrochloride could be omitted. At least for the family of small multidrug transporters, the cell-free expression technique seems to be a highly appealing way to generate samples suitable for NMR spectroscopy. The production rate of membrane proteins in the cell-free system was not related to the number of transmembrane domains, and the cell-free expression of even larger membrane proteins might therefore be possible. Regardless of this, the cell-free expression could be a suitable tool for the rapid screening of the general likelihood of expression of membrane proteins.

Acknowledgements

We are grateful to Vladimir Shirokov and Alexander Spirin for valuable advice in establishing the cell-free expression system, Marc Lorch for his help in proteoliposome preparation and Klaus Fendler for helpful discussions. The work was financially supported by DFG (grant GL307-1/3) and BMBF project ProAMP.

References

1. Arora, A. & Tamm, L.K. (2001) Biophysical approaches to membrane protein structure determination. *Curr. Opin. Struct. Biol.* **11**, 540–547.
2. Tate, C.G. (2001) Overexpression of mammalian integral membrane proteins for structural studies. *FEBS Lett.* **504**, 94–98.
3. Wang, D.N., Safferling, M., Lemieux, M.J., Griffith, H., Chen, Y. & Li, X.D. (2003) Practical aspects of overexpressing bacterial secondary membrane transporters for structural studies. *Biochim. Biophys. Acta* **1610**, 23–365.
4. Baranov, V.I. & Spirin, A.S. (1993) Gene expression in cell-free system on preparative scale. *Methods Enzymol.* **217**, 123–142.
5. Spirin, A.S., Baranov, V.I., Ryabova, L.A., Ovodov, S.Y. & Alakho, Y.B. (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* **242**, 1162–1164.
6. Alakov, Y.B., Baranov, V.I., Ovodov, S.J., Ryabova, L.A., Spirin, A.S. & Morozov, I.J. (1995) United States Patent No. 5,478,730.
7. Kim, D.M. & Choi, C.Y. (1996) A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. *Biotechnol. Prog.* **12**, 645–649.
8. Paulsen, I.T., Skurry, R.A., Tam, R., Saier, M.H., Turner, R.J., Weiner, J.H., Goldberg, E.B. & Grinius, L.L. (1996) The SMR family: a novel family of multidrug efflux proteins involved with the efflux of lipophilic drugs. *Mol. Microbiol.* **19**, 1167–1175.
9. Putman, M., van Veen, H.W. & Konings, W.N. (2000) Molecular properties of bacterial multidrug transporters. *Microbiol. Mol. Biol. Rev.* **64**, 672–693.
10. Yerushalmi, H., Lebendiker, M. & Schuldiner, S. (1996) Negative dominance studies demonstrate the oligomeric structure of EmrE, a multidrug antiporter from *Escherichia coli*. *J. Biol. Chem.* **271**, 31044–31048.
11. Yerushalmi, H., Lebendiker, M. & Schuldiner, S. (1995) Negative dominance studies demonstrate the oligomeric structure of EmrE, a multidrug antiporter from *Escherichia coli*. *J. Biol. Chem.* **270**, 6856–6863.
12. Chung, Y.J. & Saier, M.H. (2002) Overexpression of the *Escherichia coli* sugE gene confers resistance to a narrow range of quaternary ammonium compounds. *J. Bacteriol.* **184**, 2543–2545.
13. Turner, R.J., Taylor, D.E. & Weiner, J.H. (1997) Expression of *Escherichia coli* TehA gives resistance to antiseptics and disinfectants similar to that conferred by multidrug resistance efflux pumps. *Antimicrob. Agents Chemother.* **41**, 440–444.
14. Aleshin, V.V., Zakataeva, N.P. & Livshits, V.A. (1999) A new family of amino-acid-efflux proteins. *Trends Biochem. Sci.* **24**, 133–135.
15. Franke, I., Resch, A., Dassler, T., Maier, T. & Böck, A. (2003) YfiK from *Escherichia coli* promotes export of *O*-acetylserine and cysteine. *J. Bacteriol.* **185**, 1161–1166.
16. Bullock, W.O., Fernandez, J.M. & Stuart, J.M. (1987) XL1-Blue: a high efficiency plasmid transforming recA *Escherichia coli* strain with beta-galactosidase selection. *Biotechniques* **5**, 376–379.
17. Sambrock, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual* (Ford, N., Nolan, C. & Ferguson, M., eds), 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
18. Zubay, G. (1973) *In vitro* synthesis of protein in microbial systems. *Annu. Rev. Genet.* **7**, 267–287.
19. Schagger, H. & Jagow, G.V. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368–379.
20. Rost, B. (1996) PHD: predicting one-dimensional protein structure by profile based neural networks. *Methods Enzymol.* **266**, 525–539.
21. Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl Acad. Sci. USA* **94**, 12366–12371.
22. Czisch, M. & Boelens, R. (1998) Sensitivity enhancement in the TROSY experiment. *J. Magn. Reson.* **134**, 158–160.
23. Schwaiger, M., Lebendiker, M., Yerushalmi, H., Coles, M., Gröger, A., Schwarz, C., Schuldiner, S. & Kessler, H. (1998) NMR investigation of the multidrug transporter EmrE, an integral membrane protein. *Eur. J. Biochem.* **254**, 610–619.
24. Winstone, T.L., Duncalf, K.A. & Turner, R.J. (2002) Optimization of expression and the purification by organic extraction of the integral membrane protein EmrE. *Protein Expr. Purif.* **26**, 111–121.

25. Yelin, R., Rotem, D. & Schuldiner, S. (1999) EmrE, a small *Escherichia coli* multidrug transporter, protects *Saccharomyces cerevisiae* from toxins by sequestration in the vacuole. *J. Bacteriol.* **181**, 949–956.
26. Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T. & Yokoyama, S. (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett.* **442**, 15–19.
27. Madin, K., Sawasaki, T., Ogasawara, T. & Endo, Y. (2000) A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. *Proc. Natl Acad. Sci. USA* **97**, 559–564.
28. Rotem, D., Sal-man, N. & Schuldiner, S. (2001) *In vitro* monomer swapping in EmrE, a multidrug transporter from *Escherichia coli*, reveals that the oligomer is the functional unit. *J. Biol. Chem.* **276**, 48243–48249.
29. Steiner-Mordoch, S., Granot, D., Lebendiker, M. & Schuldiner, S. (1999) Scanning cysteine accessibility of EmrE, an H^+-coupled multidrug transporter from *Escherichia coli*, reveals a hydrophobic pathway for solutes. *J. Biol. Chem.* **274**, 19480–19486.
30. Venyaminov, S.Y. & Yang, J.T. (1996) Determination of protein secondary structure. In *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G.D., eds), pp. 69–107. Plenum, New York.
31. Arkin, I.T., Russ, W.P., Lebendiker, M. & Schuldiner, S. (1996) Determining the secondary structure and orientation of EmrE, a multi-drug transporter, indicates a transmembrane four-helix bundle. *Biochemistry* **35**, 7233–7238.
32. Kuroda, Y., Maeda, Y., Sawa, S., Shibata, K., Miyamoto, K. & Nakagawa, T. (2003) Effects of detergents on the secondary structures of prion protein peptides as studied by CD spectroscopy. *J. Pep. Sci.* **9**, 212–220.
33. Federkeil, S.L., Winstone, T.L., Jickling, G. & Turner, R.J. (2003) Examination of EmrE conformational differences in various membrane mimetic environments. *Biochem. Cell Biol.* **81**, 61–70.
34. Muth, T.R. & Schuldiner, S. (2000) A membrane-embedded glutamate is required for ligand binding to the multidrug transporter EmrE. *EMBO J.* **19**, 234–240.
35. Waldo, G.S., Standish, B.M., Berendzen, J. & Terwilliger, T.C. (1999) Rapid protein-folding assay using green fluorescent protein. *Nat. Biotechnol.* **17**, 691–695.
36. Drew, D.E., von Heijne, G., Nordlund, P. & de Gier, J.W.L. (2001) Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*. *FEBS Lett.* **507**, 220–224.
37. Huppa, J.B. & Ploegh, H.L. (1997) *In vitro* translation and assembly of a complete T cell receptor-CD3 complex. *J. Exp. Med.* **4**, 393–403.
38. Guignard, L., Ozawa, K., Pusglove, S.E., Otting, G. & Dixon, N.E. (2002) NMR analysis of *in vitro*-synthesized proteins without purification: a high-throughput approach. *FEBS Lett.* **524**, 159–162.
39. Betton, J.M., Palibroda, N., Namane, A., Metzler, T. & Bârzu, O. (2002) Selective labelling of proteins in the RTS 500 System. In *Cell-Free Translation Systems* (Spirin, A.S., ed), pp. 219–225. Springer, Berlin.
40. Kigawa, T., Muto, Y. & Yokoyama, S. (1995) Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. *J. Biomol. NMR* **6**, 129–134.
41. Yabuki, T., Kigawa, T., Dohmae, N., Takio, K., Terada, T., Ito, Y., Laue, E.D., Cooper, J.A., Kainosho, M. & Yokoyama, S. (1998) Dual amino acid-selective and site-directed stable-isotope labeling of the human c-Ha-Ras protein by cell-free synthesis. *J. Biomol. NMR* **11**, 295–306.
42. Glaubitz, G., Groger, A., Gottschalk, K., Spooner, P., Watts, A., Schuldiner, S. & Kessler, H. (2000) 31P-CP-MAS NMR studies on TPP+ bound to the ion-coupled multidrug transport protein EmrE. *FEBS Lett.* **480**, 127–131.
43. Yokoyama, S. (2003) Protein expression systems for structural genomics and proteomics. *Curr. Opin. Chem. Biol.* **7**, 39–43.
44. Betton, J.M. (2003) Rapid translation system (RTS): a promising alternative for recombinant protein production. *Curr. Protein Pept. Sci.* **4**, 73–80.

(P 02)

Evaluation of detergents for the soluble expression of α-helical and β-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. C. Klammt, D. Schwarz, K. Fendler, W. Haase, V. Dötsch and F. Bernhard. *FEBS Journal*, **272**, 6024-6038. (2005)

the FEBS Journal

Evaluation of detergents for the soluble expression of α-helical and β-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system

Christian Klammt[1,*], Daniel Schwarz[1,*], Klaus Fendler[2], Winfried Haase[3], Volker Dötsch[1] and Frank Bernhard[1]

1 Centre for Biomolecular Magnetic Resonance, University of Frankfurt/Main, Institute for Biophysical Chemistry, Frankfurt/Main, Germany
2 Max-Planck-Institute for Biophysics, Department for Biophysical Chemistry, Frankfurt/Main, Germany
3 Max-Planck-Institute for Biophysics, Department for Structural Biology, Frankfurt/Main, Germany

Keywords
cell free membrane protein synthesis; detergent micelles; functional β-barrel membrane protein; G-protein coupled receptor; transporter

Correspondence
F. Bernhard, Centre for Biomolecular Magnetic Resonance, University of Frankfurt/Main, Institute for Biophysical Chemistry, Marie-Curie-Str. 9, D-60439 Frankfurt/Main, Germany
Fax: +49 69 798 29632
Tel: +49 69 798 29620
E-mail: fbern@bpc.uni-frankfurt.de

*These authors contributed equally to this manuscript.

(Received 1 July 2005, revised 26 September 2005, accepted 3 October 2005)

doi:10.1111/j.1742-4658.2005.05002.x

Cell-free expression has become a highly promising tool for the fast and efficient production of integral membrane proteins. The proteins can be produced as precipitates that solubilize in mild detergents usually without any prior denaturation steps. Alternatively, membrane proteins can be synthesized in a soluble form by adding detergents to the cell-free system. However, the effects of a representative variety of detergents on the production, solubility and activity of a wider range of membrane proteins upon cell-free expression are currently unknown. We therefore analyzed the cell-free expression of three structurally very different membrane proteins, namely the bacterial α-helical multidrug transporter, EmrE, the β-barrel nucleoside transporter, Tsx, and the porcine vasopressin receptor of the eukaryotic superfamily of G-protein coupled receptors. All three membrane proteins could be produced in amounts of several mg per one ml of reaction mixture. In general, the detergent 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-RAC-(1-glycerol)] was found to be most effective for the resolubilization of membrane protein precipitates, while long chain polyoxyethylene-alkyl-ethers proved to be most suitable for the soluble expression of all three types of membrane proteins. The yield of soluble expressed membrane protein remained relatively stable above a certain threshold concentration of the detergents. We report, for the first time, the high-level cell-free expression of a β-barrel type membrane protein in a functional form. Structural and functional variations of the analyzed membrane proteins are evident that correspond with the mode of expression and that depend on the supplied detergent.

Abbreviations
BLM, black lipid membrane; β-OG, *n*-octy-b-glucopyranoside; Brij-35, polyoxyethylene-(23)-lauryl-ether; Brij-56, polyoxyethylene-(10)-cetyl-ether; Brij-58, polyoxyethylene-(20)-cetyl-ether; Brij-72, polyoxyethylene-(2)-stearyl-ether; Brij-78, polyoxyethylene-(20)-stearyl-ether; Brij-97, polyoxyethylene-(10)-oleyl-ether; Brij-98, polyoxyethylene-(20)-oleyl-ether; CE, continuous exchange; CF, cell-free; CMC, critical micellar concentration; C_{mic}, micellar concentration; DDM, *N*-dodecyl-β-D-maltoside; DHPC, 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine; diC_6PC, 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine; diC_8PC, 1,2-dioctanoyl-*sn*-glycero-3-phosphocholine; DM, *n*-decyl-β-maltoside; DPC, dodecylphosphocholine; FM, feeding mixture; Genapol C-100, polyoxyethylene-(10)-dodecyl-ether; GPCR, G-protein coupled receptor; LMPG, 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)]; LPPG, 1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)]; MP, integral membrane protein; NP40, nonylphenyl polyethylene glycol; RM, reaction mixture; Thesit, polyethylene glycol 400 dedecyl ether; TMS, transmembrane segment; Triton X-100, polyethylene glycol P-1,1,3,3-tetramethyl-butylphenyl ether; Tween 20, polyoxyethylene sorbitan monolaurate 20.

Recent modifications of cell-free (CF) transcription/translation systems resulted in the design of the continuous-exchange cell-free (CECF) expression system [1]. A semipermeable membrane separates the reaction mixture (RM), holding all high molecular mass compounds, from the feeding mixture (FM), providing low molecular mass precursors and removing undesirable breakdown products. CECF systems based on *Escherichia coli* CF extracts have been shown to be especially promising for the high-level production of α-helical integral membrane proteins (MPs) [2–6]. Potential toxic effects caused by the overproduced MPs *in vivo* are eliminated by using CF expression systems, and even proteins with no detectable expression in *E. coli* cells can be produced in milligram quantities in 1 mL of RM. In addition, specific amino acids of the CF expressed MPs can be efficiently labeled with stable isotopes, without any reduction in productivity [4]. Samples of MPs that are ready for structural analysis by NMR spectroscopy can therefore now be generated in less than 2 days. The fact that nearly 40% of all identified proteins are MPs is in strong contrast with their highly underrepresented structural information deposited in data banks. This discrepancy is mainly determined by the tremendous difficulties encountered when sufficient amounts of MPs for a structural analysis have to be produced by conventional *in vivo* expression techniques. The observed high-level production of MPs by CF expression might therefore have important potential for the fast and efficient generation of protein samples suitable for structural analysis by NMR techniques or by crystallization.

Owing to the lack of any membranes or other hydrophobic compartments, the MPs expressed in a standard CF system are primarily insoluble and can be isolated only as a precipitate. Interestingly, the resolubilization of CF produced MP precipitates is faster and more efficient if compared with the resolubilization of *in vivo* produced inclusion bodies, indicating that significant structural differences between MPs obtained by the two different modes of expression might exist [4]. However, the precipitation of the MPs in the RM could cause their partial or complete denaturation, and enzymatic activities or other functions could be affected. The CF expression of MPs in the presence of detergents, resulting in the production of soluble MPs inserted into micelles, could therefore represent an interesting alternative to the CF expression of protein precipitates. Initial data suggest that CF expression systems might be tolerant for some detergents, and this would open the unique opportunity to insert the target protein, directly after translation, into the detergent micelles of choice [2,3,5]. It is evident that the appropriate choice of detergent, and also its concentration, should have a major impact on its tolerance by the expression system as well as on the expression efficiency and on the solubility of a specific MP. In this report, we therefore intended to evaluate the suitability of a representative range of detergents for the soluble expression of MPs in a defined individual CF system.

To date, only α-helical MPs have been produced in CF systems based on S30 extracts of *E. coli* cells [2–6]. It was unclear whether the preparative CF production is applicable to a wider range of structurally diverse MPs. We therefore included, in our work, the expression of three very different MPs (a) the α-helical multidrug transporter, EmrE, of *E. coli*, with four transmembrane segments (TMS), (b) the β-barrel type nucleoside transporter, Tsx, of the outer membrane of *E. coli* and (c) the α-helical eukaryotic porcine vasopressin type 2 receptor of the family of G-protein coupled receptors (GPCRs) with seven TMS. The production of the three MPs in the presence of different detergents was quantified and resulted in the identification of several detergents that are generally useful for the high-level soluble expression of structurally very diverse MPs. Representative samples of CF expressed MPs have been structurally and functionally analyzed and we present evidence for variations in the functional folding of MPs that depend on their mode of CF expression. This report extends the application of the high-level CF production of MPs to β-barrel-type proteins and will provide a comprehensive support for the general CF production of soluble and functional complex integral MPs with multiple TMS in the presence of detergents.

Results

Cell-free expression of EmrE, Tsx and V2R as precipitates

All three proteins were expressed with an N-terminal T7 tag and with a C-terminal poly(His)$_6$ tag, by using the plasmids pET21-emrE2, pET21-tsx and pET21-v2R, in order to accelerate their purification and to enable a fast quantification. The addition of the two tags did not affect the protein expression, and the obtained yields per 1 mL of RM were $\approx$2 mg for EmrE, 4 mg for Tsx and 2 mg for V2R, by using the standard protocol of our individual CF system [4] without the addition of any detergents. All of the detectable recombinant protein was produced as precipitate (Fig. 1). Addition of the T7 tag proved to be highly beneficial for expression of the V2R protein as almost no expression was detectable without the tag

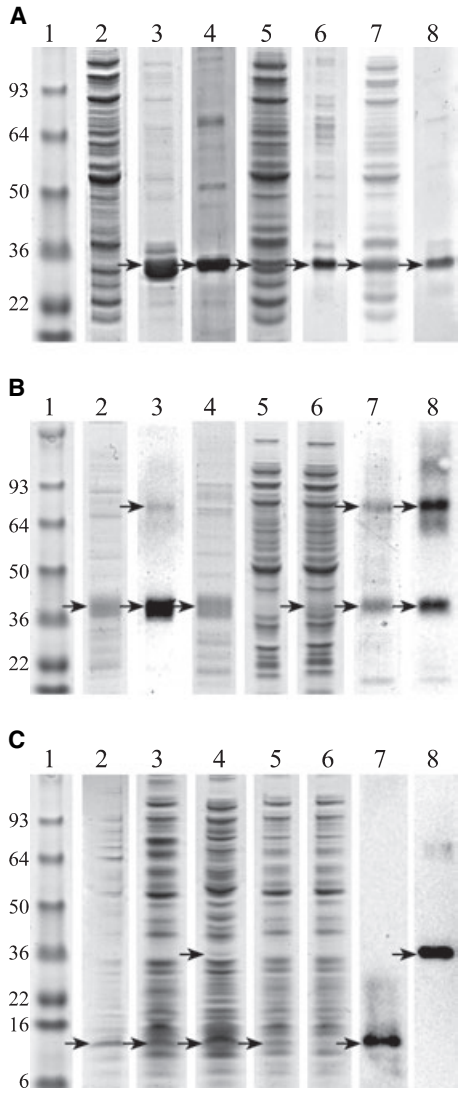


Fig. 1. Cell free (CF) production of the membrane proteins (MPs) Tsx, V2R and EmrE, in different modes. The molecular mass (*m*) values of the marker proteins are shown. The arrows indicate the expressed MPs. The protein concentration in 1 μL of reaction mix (RM) is ≈ 10 μg. The nonsoluble parts of the RM were resolubilized in suitable detergents in volumes equal to the corresponding RM volume. Nonsolubilized protein was removed by centrifugation. (A) CF expression of the β-barrel-type nucleoside transporter, Tsx, after 20 h of incubation. Samples of 0.8 μL were analyzed on a 12% (w/v) SDS/Tris/glycine gel. Lane 1, marker; lane 2, supernatant without added detergent; lane 3, suspended pellet without added detergent; lane 4, pellet without added detergent resolubilized in 2% (w/v) 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LMPG); lane 5, supernatant with 0.1% (w/v) polyethylene glycol P-1,1,3,3 tetramethyl butylphenyl ether (Triton X-100); lane 6, 8 μg of purified Tsx in 0.1% (w/v) Triton X-100; lane 7, supernatant with 0.2% (w/v) polyoxyethylene-(20)-cetyl-ether (Brij-58); lane 8, 8 μg of purified Tsx in 0.2% (w/v) Brij-58. (B) CF expression of the G-protein coupled receptor (GPCR) protein, V2R, after 20 h of incubation. If not otherwise stated, samples of 1 μL were analyzed on a 12% SDS/Tris/glycine gel. Lane 1, marker; lane 2, suspended pellet without added detergent; lane 3, suspended pellet without added detergent stained with the Invision staining kit (Invitrogen) directed against the poly(His)$_6$ tag; lane 4, pellet without added detergent resolubilized in 1% (w/v) LMPG; lane 5, supernatant with 1.5% (w/v) Brij-58 at time = 0 h; lane 6, supernatant with 1.5% (w/v) Brij-58 at time = 20 h; lane 7, 5 μg of purified V2R in 0.5% (w/v) Brij-58; lane 8, 0.01 μL of supernatant with 0.5% (w/v) Brij-58 stained with an anti-(T7-tag) immunoglobulin coupled to horseradish peroxidase. (C) CF expression of the small multidrug transporter, EmrE, after 20 h of incubation. Samples of 0.5 μL were analyzed on a 4–12% gradient gel (lanes 1–6) or on a 10% tricine gel (lanes 7 and 8). Lane 1, marker; lane 2, suspended pellet without added detergent; lane 3, supernatant with 1.5% (w/v) Brij-58; lane 4, supernatant with 0.2% 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine (DHPC); lane 5, supernatant with 0.1% (w/v) *n*-dodecyl-β-D-maltoside (DDM); lane 6, supernatant with 0.75% (w/v) Chaps; lane 7, 0.2 μL of supernatant with 0.2% (w/v) Digitonin stained with an anti-(T7-tag) immunoglobulin coupled to horseradish peroxidase; lane 8, 0.2 μL of supernatant with 0.1% (w/v) dodecyl-phosphocholine (DPC) stained with an anti-(T7-tag) immunoglobulin coupled with horseradish peroxidase.

(data not shown). A nonoptimal initiation of translation in the *E. coli* system, caused by the eukaryotic codon sequence of the V2R gene, might account for that observation. The CF expressed V2R protein has an apparent molecular mass (*m*) of 36 kDa after analysis by SDS/PAGE (Fig. 1), whereas the calculated *m* should be 43 kDa. Incomplete denaturation of MPs upon SDS/PAGE analysis is often considered to be the reason for variations of the apparent and calculated *m* values. However, in order to exclude the possibility of premature termination during the CF translation caused, for example, by rare codon usage of the *v2R* gene, we verified the full-length expression of V2R by detecting the C-terminal poly(His)$_6$ tag with the commercial Invision™ in a gel staining procedure (Invitrogen, Carlsbad, CA, USA) (Fig. 1). The additional larger band detected, with an *m*-value of ≈ 70 kDa, can probably be attributed to an oligomeric form of V2R.

Selected detergents from the types of phosphocholines, long-chain phosphoglycerols, alkyl-glucosides, steroids, polyoxyethylene-alkyl-ethers and polyethyleneglycol derivatives were tested for their efficiency in the resolubilization of the CF produced precipitates at a final concentration of 2% (Table 1). The strong denaturating anionic detergent, sodium dodecyl sulfate (SDS), was included into the analysis as a control for the complete solubilization of the proteins. The washed precipitates of CF produced MPs were suspended in various detergents in 15 mM sodium phosphate, pH 8.0, and incubated for 90 min on a shaker at 30 °C. The

Table 1. Solubilization of cell free (CF) produced protein precipitates. Solubilization assays of CF produced membrane protein precipitates were carried out with ≈30 μg of protein in 2% (w/v) detergent solution at 30 °C for 90 min. Nonsoluble protein was removed by centrifugation. Brij-35, polyoxyethylene-(23)-lauryl-ether; β-OG, *n*-octy-b-glucopyranoside; CHAPS, 3-[(3-cholamidopropyl) dimethylammonio]-1-propansulfonate; DDM, *n*-dodecyl-β-D-maltoside; DHPC, 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine; DPC, dodecyl-phosphocholine; LMPG, 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)]; SDS, sodium dodecylsulfate; Thesit, polyethylene glycol 400 dedecyl ether; Tween 20, polyoxyethylene sorbitan monolaurate 20.

Detergent	Solubilized Tsx (%)	Solubilized V2R (%)	Solubilized EmrE (%)
SDS	100	100	100
DPC	90	10	90
DHPC	40	2	80
LMPG	90	90	95
β-OG	10	2	20
DDM	40	15	90
Digitonin	25	2	0
Chaps	20	0	90
Brij-35	10	10	10
Triton X-100	20	5	90
Thesit	5	0	50
Tween 20	0	0	60

suspension was centrifuged (20 000 ***g***, 20 min, 20–24 °C) and the supernatant was analyzed on appropriate SDS gels. The small multidrug transporter, EmrE, was readily resolubilized in high quantities in most of the detergents analyzed (Table 1). However, this is obviously not a general attribute for all CF produced MP precipitates. The quantitative resolubilization of the GPCR protein, V2R, was only possible with the long-chain phosphocholine, 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LMPG), with the other detergents having no, or only minor, resolubilization effects. In addition, only LMPG and dodecyl phosphocholine (DPC) were highly effective for the resolubilization of the β-barrel type transporter, Tsx. LMPG and probably similar long chain-phosphoglycerols appear therefore to be the primary choice for a general and efficient resolubilization of even structurally diverse types of MPs in mild detergents.

Cell-free expression of EmrE, Tsx and V2R in the presence of detergents

CF expression provides, in general, an open system without any membrane barriers, as in living cells, and many additives that could potentially become beneficial for the recombinant protein can be added. The supply of detergents to the CF reaction could thus provide an important alternative approach in order to produce directly soluble MPs. We therefore started a systematic evaluation of a comprehensive selection of detergents with respect (a) to their impact on the productivity of a CF expression system, (b) to their efficiencies in the direct solubilization of CF produced MPs and (c) to their effects on the activity of the solubilized MPs. In order to obtain relatively representative results, we analyzed the synthesis of the three structurally very different MPs: EmrE, Tsx and V2R. A total of 24 detergents, selected from the major types commonly used for the solubilization of MPs, were tested for their suitability as an additive for the CF expression of the three MPs; some important physicochemical characteristics of the selected detergents are summarized in Table 2. All detergents were supplemented above their specific critical micellar concentrations (CMCs), as specified in Table 2, and most were added initially in concentration ranges from 0.1% to 0.2% (w/v). The production of MPs was analyzed in analytical scale CF reactions after incubation for 20 h, and the proteins in the soluble and unsoluble fractions were quantified separately by western blot analysis using specific antibodies directed against the terminal T7- or poly$(His)_6$ tags.

Few detergents turned out to become inhibitory to the CF system at low concentrations and resulted in an at least considerably reduced production of all three MPs (Table 3). In particular, some of the phosphocholine derivatives seemed to be problematic as additives in CF systems, while the mono-chain phosphocholines were clearly more critical than the bi-chain phosphocholines. DPC almost completely inhibited general protein synthesis. However, the bi-chain phosphocholine, 1,2-dioctanoyl-*sn*-glycero-3-phosphocholine (diC_8PC), still resulted in the production of up to 0.6 mg of soluble EmrE per mL of RM. The production of soluble Tsx was not as effective and remained below 0.5 $mg{\cdot}mL^{-1}$ in the RM, while soluble V2R was only spuriously detectable. Reducing the chain length in diC_7PC [1,2-diheptanoyl-*sn*-glycero-3-phosphocholine (DHPC)] and 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (diC_6PC) decreased the soluble expression of EmrE to 0.47 $mg{\cdot}mL^{-1}$ and to 2 $\mu g{\cdot}mL^{-1}$ in the RM, respectively. The detergent, diC_6PC, clearly affected the general CF protein synthesis, while DHPC and diC_8PC were still tolerated at even higher CMCs and both were also relatively effective in the solubilization of the Tsx protein. However, none of the phosphocholine derivatives tested was suitable for the high-level CF expression of soluble V2R protein. Negative effects on the CF system were further observed with the

Table 2. Physicochemical properties of the analyzed detergents

Detergent	Short name	Charge[a]	*m* (Da)	CMC[b] (mM)
Alkyl glucosides				
n-Dodecyl-β-D-maltoside	DDM	N	511	0.12
n-Decyl-β-D-maltoside	DM	N	483	1.8
n-Octyl-β-D-glucopyranoside	β-OG	N	292	19
Steroid derivatives				
Digitonin	Digitonin	N	1229	0.73
3-[(3-Cholamidopropyl)dimethylammonio]-1-propansulfonat	Chaps	Z	615	2.4–8.6
Long chain-phosphoglycerols:				
1-Myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)]	LMPG	A	479	0.05
1-Palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)]	LPPG	A	507	NA
Mono-/bi-chain phosphocholines				
1,2-Dioctanoyl-*sn*-glycero-3-phosphocholine	diC_8PC	Z	510	0.22
1,2-Diheptanoyl-*sn*-glycero-3-phosphocholine	DHPC	Z	482	1.4
1,2-Dihexanoyl-*sn*-glycero-3-phosphocholine	diC_6PC	Z	454	14
Dodecyl-phosphocholine	DPC	Z	352	0.9–1.5
Polyoxyethylene alkyl-ether				
Polyoxyethylene-(23)-lauryl-ether ($C_{12/23}$)	Brij-35	N	1200	0.08
Polyoxyethylene-(10)-dodecyl-ether ($C_{12/10}$)	Genapol C-100	N	627	0.1
Polyoxyethylene-(10)-cetyl-ether ($C_{16/10}$)	Brij-56	N	682	0.035
Polyoxyethylene-(20)-cetyl-ether ($C_{16/20}$)	Brij-58	N	1123	0.075
Polyoxyethylene-(2)-stearyl-ether ($C_{18/2}$)	Brij-72	N	359	NA
Polyoxyethylene-(20)-stearyl-ether ($C_{18/20}$)	Brij-78	N	1152	0.046
Polyoxyethylene-(10)-oleyl-ether ($C_{18\text{-}1/10}$)	Brij-97	N	709	0.217
Polyoxyethylene-(20)-oleyl-ether ($C_{18\text{-}1/20}$)	Brij-98	N	1150	0.025
Polyoxyethylene-sorbitan-monolaurate 20	Tween 20	N	1228	0.059
Polyethylene glycol derivatives				
Polyethylene-glycol P-1,1,3,3-tetramethyl-butylphenyl-ether	Triton X-100	N	647	0.23
Polyethylene-glycol 400 dedecyl-ether	Thesit	N	583	0.1
Nonylphenyl-polyethylene-glycol	NP40	N	603	0.05–0.3

[a] A, anionic; N, nonionic; Z, zwitterionic. [b] Estimated critical micellar concentrations (CMCs). Averaged values are given if different CMCs for a detergent have been published; NA, not available.

nonionic alkyl-glucoside, *n*-octy-b-glucopyranoside (β-OG), the steroid, Chaps, and the long-chain phosphoglycerol, LMPG. Chaps and *β*-OG are generally considered to be relatively mild detergents, but their high CMC, resulting in relatively high final detergent concentrations in the RM, might be a reason for the observed inhibitory effect on protein expression. LMPG, which was found to be highly suitable for the resolubilization of CF produced precipitates, resulted, only with Tsx, in the production of some 100 μg of soluble protein, while the total yield of protein was considerably affected for all three MPs.

Approximately one-third of the tested detergents did not have any, or only a marginal, effect on the solubility of any of the expressed MPs, but they did not cause marked decreases of the total protein production in the range of the supplied concentrations. This group includes polyethylene glycol 400 dedecyl ether (Thesit), nonylphenyl polyethylene glycol (NP40), 1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LPPG) and the polyoxyethylene derivatives polyoxyethylene-(10)-cetyl-ether (Brij-56), polyoxyethylene-(2)-stearyl-ether (Brij-72), polyoxyethylene-(10)-oleyl-ether (Brij-97), polyoxyethylene-(10)-dodecyl-ether (Genapol C-100) and polyoxyethylene sorbitan monolaurate 20 (Tween 20). Even with concentrations as high as 209× CMC in the case of Brij-56, only precipitated Tsx protein was obtained in high yields. These detergents are obviously tolerated by the CF system, even at higher CMC concentrations, and they might still be beneficial for the soluble expression of other MPs. However, they are clearly not suitable for general usage.

Several detergents from almost all analyzed groups were found to be suitable for the soluble CF expression of MPs, yielding production rates of more than 0.5 mg of soluble MP per mL of RM (Table 3). With a total of 10 detergents, we obtained preparative scale soluble expression yields for at least one of the analyzed MPs, while the individual proteins behaved quite differently. Again, EmrE and also Tsx could be solubilized by numerous detergents, while the V2R protein was much more selective. The most outstanding group of effective

Table 3. Cell free expression of membrane proteins in the presence of detergents. CMC, critical micellar concentration; S, soluble protein production; P, protein produced as precipitate; NA, not analyzed; u, unknown. [a] Detergents that result at least in one instance of the production of > 500 μg·mL^{-1} soluble membrane proteins are given in bold. [b] The final concentrations of the detergents in the cell-free system for the expression of the α-helical proteins EmrE and V2R are given. The numbers in parentheses indicate instances where different concentrations were used for expression of the Tsx protein.

	Concentration		EmrE[a]	Tsx[a]	V2R[a]
Detergent	(%)	(x CMC)	S/P	S/P	S/P
None		–	0, *0*/4, NA	0/4	0, *0*/4, NA
DDM	0.1	15.0	3, *1*/1, *0*	4/4	1, *2*/2, *3*
DM	0.2	2.3	4, *1*/3, *0*	3/2	1, *0*/1, *0*
β-OG	0.75	1.3	1, *2*/0, *0*	0/0	0, *0*/2, *1*
Digitonin	0.4	4.5	3, *0*/0, *0*	4/4	3, *3*/2, *2*
Chaps	0.75	1.5	1, *0*/3, *0*	2/3	0, *0*/0, *1*
LMPG	0.01	4.2	1, *0*/3, *0*	3/3	1, *1*/3, *3*
LPPG	0.025	U	0, *0*/3, *1*	0/4	2, *1*/4, *4*
diC$_8$PC	0.1	8.9	4, *2*/4, *2*	3/4	1, *2*/3, *3*
DHPC	0.2	3.0	3, *3*/2, *0*	4/3	1, *1*/3, *3*
diC$_6$PC	0.75	1.2	1, *2*/2, *2*	3/3	1, *1*/3, *2*
DPC	0.1	1.9	0, *1*/0, *0*	0/0	2, *0*/0, *1*
Brij-35	0.1	10.4	4, *0*/0, *2*	4/0	3, *3*/2, *2*
Brij-58	1.5	178.1	4, *0*/2, *0*	4/3	4, *4*/2, *2*
Brij-78	1.0	189.0	4, *3*/1, *0*	4/2	4, *4*/2, *2*
Brij-98	0.2	69.6	4, *0*/0, *0*	3/4	4, *4*/2, *2*
Brij-56	0.01 (0.5)	4.2 (209)	1, *0*/2, *2*	0/4	2, *2*/1, *1*
Brij-72	0.2 (0.33)	U	2, *0*/3, *0*	0/3	2, *2*/2, *2*
Brij-97	0.2 (0.5)	13.0 (17)	0, *0*/4, *4*	0/4	2, *1*/0, *0*
Genapol C-100	0.1	21.3	2, 0/3, *2*	0/4	2, *2*/1, *2*
Tween 20	0.075	10.4	0, *0*/4, *0*	0/4	0, *0*/2, *2*
Triton X-100	0.1 (0.2)	6.7 (13.4)	1, *0*/4, *0*	4/4	2, *2*/2, *2*
Thesit	0.1	17.2	0, *0*/3, *0*	3/4	1, *2*/2, *2*
NP40	0.1 (0.4)	9.8 (39)	2, *0*/4, *3*	2/4	1, *1*/1, *0*

[a] The production of recombinant proteins, in μg·mL^{-1} of reaction mixture, was classified into five groups: 0, no detectable expression; 1, spurious expression (< 10 μg·mL^{-1}); 2, 10–100 μg·mL^{-1}; 3, 101–500 μg·mL^{-1}; and 4, > 500 μg·mL^{-1}. The estimated amounts of the detected putatively oligomeric forms of the EmrE and V2R proteins are shown separately with italic numbers.

detergents covered the long chain polyoxyethylene-alkyl-ethers, polyoxyethylene-(23)-lauryl-ether (Brij-35), polyoxyethylene-(20)-cetyl-ether (Brij-58), polyoxyethylene-(20)-stearyl-ether (Brij-78) and polyoxyethylene-(20)-oleyl-ether (Brij-98), and they resulted in the high-level soluble expression of all three MPs analyzed. In particular, in the case of V2R, production rates of more than 1 mg·mL^{-1} in the RM were only possible with Brij derivatives. Both the number of polyoxyethylene groups, as well as the length of the alkyl group, were identified as important parameters for the solubilization effect. The reduction in the number of polyoxyethylene moieties to ≤ 10 almost completely abolished the solubilization of all three MPs produced (Table 3). In addition, the chain length of the alkyl moiety specifically affected the efficiency of soluble MP production. The chain length of the alkyl moiety had a clear optimum at C16 and C18 for the soluble expression of V2R, yielding up to 3 mg of protein per mL of RM (Fig. 2). The double bond in the oleyl-ether of Brij-98 and the short chain lauryl-ether in Brij-35 considerably reduced the yield of soluble V2R. The solubilization of the Tsx protein was less sensitive to the chain length, but was

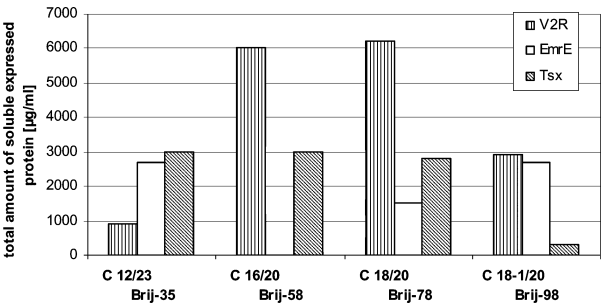


Fig. 2. Soluble expression of membrane proteins (MPs) in the presence of polyoxyethylene derivatives. The total amount (monomers and oligomers) of soluble expressed protein was quantified after 20 h of CF expression. Values represent the means of at least two determinations, and the variations were in the range of not more than 15%.

also significantly affected by a double bond in the alkyl moiety. In contrast, the highest amounts of the soluble small multidrug transporter, EmrE, were produced with Brij-35, containing the shorter lauryl-ether, and with Brij-98, containing a double bond in the alkyl chain.

The steroid derivative, Digitonin, was identified as a second relatively general solubilizing detergent (Table 3). However, the yields obtained in the case of EmrE and V2R remained mostly below 0.5 mg·mL^{-1} in the RM and were thus lower by comparison with the above-mentioned Brij derivatives. The alkyl glucosides, *n*-dodecyl-β-D-maltoside (DDM) and *n*-decyl-β-maltoside (DM) resulted in the soluble expression of only EmrE and Tsx in the range between 0.5 mg and 1.5 mg per mL of RM. The beneficial effect of polyethylene glycol P-1,1,3,3-tetramethyl-butylphenyl ether (Triton X-100) was specific for the Tsx protein and yielded soluble protein at a concentration of nearly 1 mg·mL^{-1} in the RM.

High yields of soluble MPs can be obtained within a wide range of detergent concentrations

Besides the detergent structure, the supplied final detergent concentration in the RM is an important parameter for the production of soluble MP. As a critical factor, the micellar concentration (C_{mic}) of the detergent in the RM should be considered. The C_{mic} is calculated from the molar detergent concentration minus the specific CMC and divided by the specific aggregation number (number of detergent molecules per micelle). In the optimal case, all protein can become solubilized as soon the C_{mic} equals the molar concentration of the synthesized MP. A further increase in detergent concentration would then only result in the production of additional and empty micelles. Likewise, with a C_{mic} below the molar protein concentration, either heterogeneous micelles with protein aggregates or residual protein precipitates in addition to the solubilized protein should be formed. The titration of soluble expression of the Tsx protein with different detergent concentrations completely agreed with this assumption (Fig. 3). With Triton X-100, a concentration of 7× CMC resulted in maximal production of soluble protein and the yield remained relatively constant upon further increases of the detergent concentration. Similar results were obtained with the detergents Brij-58 and Brij-78, with threshold concentrations of ≈47× CMC and 76× CMC, respectively (Fig. 3). However, it should be noted that some residual precipitate of the Tsx protein remained, even if the optimal detergent concentration was exceeded several times (Table 3, data not shown). Some protein obviously escapes; therefore, the solubilization

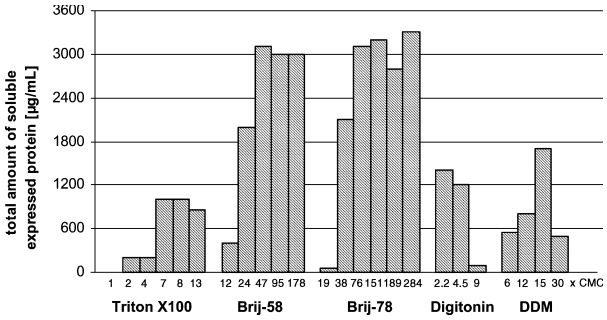


Fig. 3. Effect of detergent concentration on the soluble expression of the transporter, Tsx. The total amount of soluble expressed protein was quantified after 20 h of cell free (CF) expression. Values represent the means of at least two determinations and variations were in the range of not more than 15%.

process in the CF reaction. With Digitonin and DDM, concentration optima for the soluble Tsx production were obtained with 2.2× CMC and 15× CMC, and higher detergent concentrations resulted in a drastic decrease of soluble protein. However, this reduction did not favor an increase of nonsoluble precipitate and could thus be attributed to an inhibitory effect of these detergents at higher concentrations that generally affects the efficiency of the CF expression system.

Secondary structure analysis of the V2R and Tsx proteins and reconstitution into proteoliposomes

The functional folding and the structural conformation of MPs are likely to depend on their mode of expression. Variations in the specific activity of a MP could be obtained regardless of whether it has been produced primarily as a precipitate or whether it could be inserted in a hydrophobic environment, such as detergent micelles, directly after translation. The transport activity of CF expressed EmrE protein, either as soluble protein or resolubilized from precipitates, has already been published [3,4]. CD spectroscopy is a first and fast technique for the analysis of secondary structures, and V2R and Tsx samples produced by different CF expression modes were therefore analyzed using CD spectroscopy (Fig. 4). The β-sheet content of the Tsx protein varied between 45% and 58%, depending on whether the protein was produced as a precipitate with subsequent resolubilization in LMPG, or whether it was synthesized directly as soluble protein in the presence of various detergents. Accordingly, the α-helical content of the V2R protein varied considerably, between 32% and almost 77%, depending on its mode of expression. The soluble expression in the detergent, Brij-58, should therefore result in the most native-like conformation if a proposed, mostly α-helical, composition is

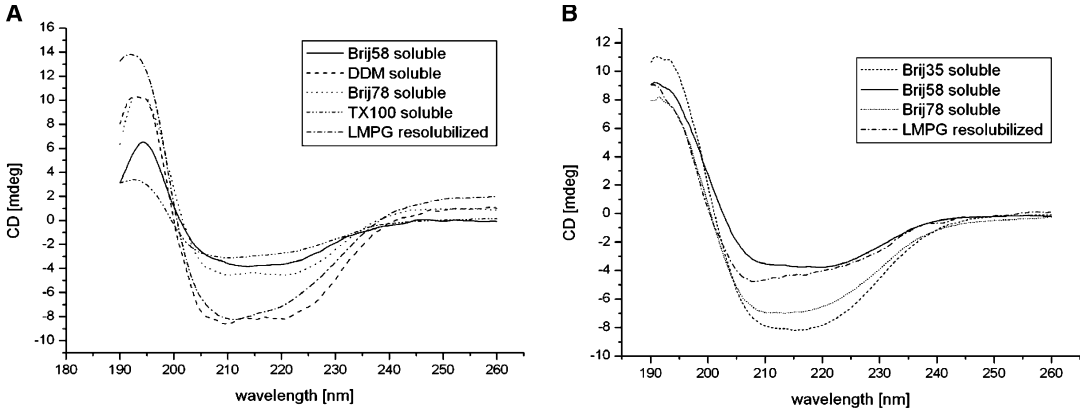


Fig. 4. Secondary structure analysis of cell free (CF) produced membrane proteins (MPs) by CD spectroscopy. Approximately 15 μM protein solubilized in detergent was analyzed at 25 °C in 10 mM sodium phosphate, pH 8.0, containing 0.5 mM dithiothreitol. The proteins were expressed either in the soluble form with Brij derivatives, *n*-dodecyl-β-D-maltoside (DDM) and Triton X-100 or resolubilized from CF produced precipitates in 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LMPG). The proteins were purified by Ni-chelate chromatography prior to analysis. The secondary structure composition was calculated by JASCO secondary structure estimation software. (A) Tsx; β-sheet: Brij-58 (0.2% w/v), 58.5%; Brij-78 (0.4% w/v), 51.2%; DDM (0.1% w/v), 43.6%; Triton X-100 (0.1% w/v), 47.5%; LMPG (2% w/v), 45.4%. (B) V2R; α-helix: Brij-35 (0.1% w/v), 31.8%; Brij-58 (0.5% w/v), 76.6%; Brij-78 (0.5% w/v), 27.3%; LMPG (1% w/v), 35.4%.

considered. A structural rearrangement of a MP into a native like conformation might also become induced upon its transfer from detergent micelles into a lipid membrane environment. LMPG-resolubilized V2R protein, with a relatively low calculated α-helical content of ≈35%, was therefore reconstituted into membranes composed of an *E. coli* lipid mixture, and the resulting proteoliposomes were analyzed by freeze-fracture electron microscopy (Fig. 5). The evenly and homogeneously distributed V2R particles in the membranes gave evidence for a structural folding of the protein.

Functional folding of the CF expressed nucleoside transporter, Tsx, depends on its mode of expression

The activity of the nucleoside transporter, Tsx, synthesized in different CF expression modes, was

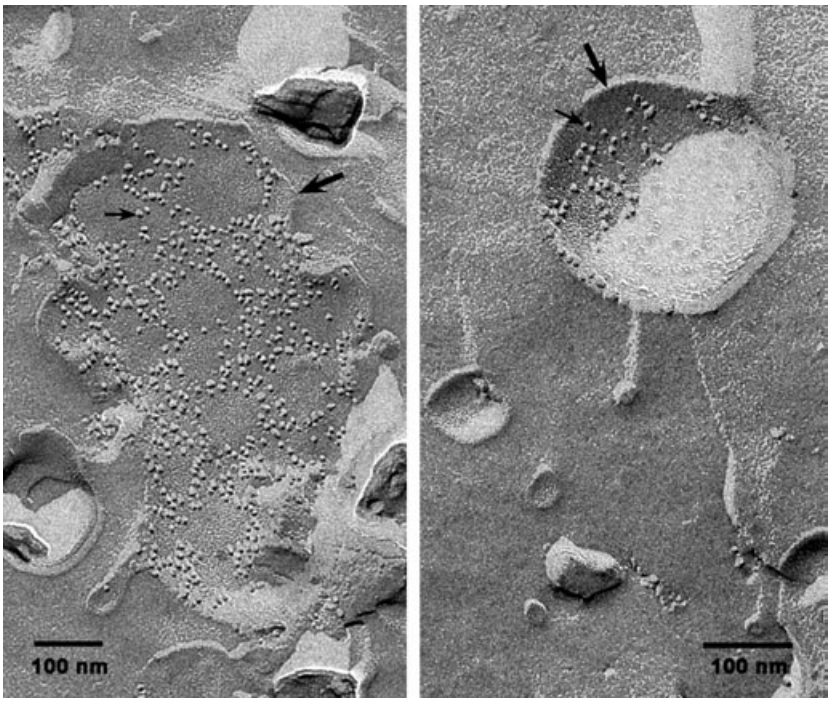


Fig. 5. Freeze-fracture electron microscopy of reconstituted V2R protein. Cell free (CF) produced precipitate of V2R was resolubilized in 1% (w/v) 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LMPG) in 15 mM sodium phosphate, pH 7.5, and reconstituted into *Escherichia coli* lipid mixture at a molar protein/lipid ratio of 1 : 2500. The large arrow indicates liposomes and the small arrow indicates reconstituted protein.

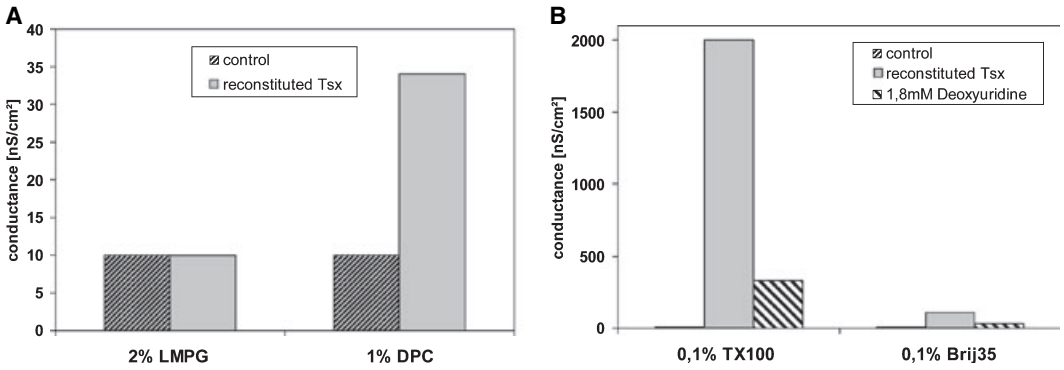


Fig. 6. Functional analysis of the cell free (CF) expressed nucleoside transporter, Tsx, by the black lipid membrane (BLM) assay. Approximately 100 ng of purified Tsx protein solubilized in different detergents was reconstituted into black lipid membranes. A pore-forming activity of Tsx was measured by an increase in conductance after 20 min of incubation. The Tsx pores were specifically blocked by addition of the substrate, deoxyuridine. The background conductance of the membrane in the control was ≈ 10 nS·cm^{-1}. (A) Resolubilized Tsx in 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LMPG) and dodecyl-phosphocholine (DPC). (B) soluble expressed Tsx in Triton X-100 and Brij-35.

determined using the black lipid membrane (BLM) assay. Tsx precipitates produced by standard CF reactions without any supplemented detergent were resolubilized in the detergents LMPG and DPC. In addition, Tsx was soluble expressed in the presence of the detergents Brij-35 and Triton X-100. All soluble Tsx samples were purified by Ni-chelate chromatography and aliquots of ≈ 100 ng of protein, dissolved in a small volume (5 μL maximum), were injected into preformed BLMs separating two buffer chambers. An electric current was measured between the two chambers, and the functional reconstitution of the Tsx protein was determined by its pore-forming activity in the BLM, resulting in an increase of the conductance. The specificity of the pore-forming activity for Tsx was further demonstrated by inhibition with the substrate deoxyuridine. A high activity, with a 200-fold increase in conductance that specifically could be inhibited by the nucleoside substrate, was observed with the Tsx sample directly produced in a soluble form in the presence of Triton X-100 (Fig. 6). The Tsx sample, soluble expressed in the presence of Brij-35, showed a considerably lower, but still significant, 15-fold increase in conductance if compared with the control. In contrast, with the two resolubilized Tsx samples, only a threefold increase of the protein in DPC, and no increase at all of the sample in LMPG, was detected.

Discussion

The choice of type and concentration of a detergent can have an enormous impact on the yield and solubility of an individual MP, and the influence of detergents on the CF reaction may further limit their application. Few reports on the soluble expression of integral MPs are available [2,3,5], but the relevance of the nature of a detergent is not clear as extensive systematic studies have not been carried out and different CF systems have been used. We therefore analyzed the soluble synthesis of representatives of prokaryotic and eukaryotic α-helical MPs with multiple TMS, as well as a β-barrel type outer membrane protein, in an individual and highly productive coupled transcription/translation CF system in the continuous exchange mode. In order to avoid the unfolding of the solubilized MPs or of essential components of the CF system, we focused, in our work, on relatively mild detergents. A representative cross-section of biological and industrial detergents, including nonionic detergents [alkyl-glycosides (β-OG, DDM, DM)], polyoxyethylene-alkyl-ethers (Brij derivatives, Thesit, Genapol), polyethylene-glycol derivatives (Triton, Tween, NP40), steroid derivatives (Chaps, Digitonin), zwitterionic detergents (DPC, DHPC, diC_6PC, diC_8PC), and long-chain ionic detergents (LMPG, LPPG), were analyzed for their suitability for the production of soluble MPs. The addition of the alkyl-glycoside, β-OG, and of the phosphocholine derivatives, DPC and diC_6PC, abolished any significant protein expression or reduced it to levels of ≤ 100 μg·mL^{-1} in the RM at relatively low concentrations of 1–2× CMC. Common features of this group are the short chain lengths and relatively high CMCs of > 1 mM (and up to 19 mM in the case of β-OG). The relatively high concentration required of those detergents in the CF system, in order to obtain the required CMC levels for the MP solubilization, could therefore be one reason for the observed marked reduction in protein expression.

A surprising result was that a single group of detergents, the polyoxyethylene-alkyl-ethers, with a higher polymerization number of the polyoxyethylene moieties, is especially suitable for the soluble CF expression of even structurally diverse MPs. Polyoxyethylenes with a shorter alkyl-chain length have a higher aggregation number, and extensive fluctuations in the micellar shape could affect the stability of an inserted protein and reduce its solubility [7]. The influence of detergents on protein conformations becomes stronger with an increase of their own flexibility and with a decrease in the size of their hydrophilic head [8,9]. This might explain why especially long chain polyoxyethylene-alkyl-ethers have this generally high potential of MP solubilization. It is noteworthy that Brij derivatives, together with Digitonin, have been found to be the only detergents suitable for the high-level soluble expression of the V2R protein and also for other GPCR proteins [5]. In contrast, the smaller MPs – EmrE and Tsx – could also be effectively solubilized by a much wider range of detergents, such as Triton X-100, DDM, DM, DHPC and diC_8PC. Accordingly, Triton X-100, Brij 58, DDM and Chaps were found to be compatible with a commercial CF system upon expression of the α-helical bacterial mechanosensitive channel, MscL [2]. Most of those detergents were also optimal for the crystallization of several MPs [10].

The solubilization properties of a detergent depend on its structure as well as on its concentration. In particular, the actual micellar concentration, C_{mic}, which is determined by the specific aggregation number of a detergent, has to be considered. The CMC is not a constant parameter and is difficult to measure with high accuracy as it depends on experimental conditions such as pH, salt concentrations and temperature. Micelles do not have static structures, and even within one type of detergent, the polydispersity of micelles can vary considerably with the chain length [11,12]. Furthermore, the CMC and the effective *m* of protein–detergent aggregates are believed to be determined more by the structural properties of the protein than by the properties of the detergent [13]. However, it is generally assumed that the amount of detergent binding to the hydrophobic surface of an integral MP is relatively constant above CMC [12] and an increase in detergent concentration above C_{mic} barely influences the protein solubility [14]. This is consistent with our observation that the detergent concentrations for the optimal solubilization of MPs follow a plateau-like kinetics, and similar observations have been made upon the expression of several GPCR proteins in a batch mode CF system [5]. Further increased concentrations do not affect the protein expression as long as the detergent is tolerated by the CF system. The specific threshold concentration for an MP–detergent combination might therefore be considered as the point of molar equilibrium between C_{mic} and MP.

Several detergents are tolerated to some extent by the CF system, but have no, or only minor, effects on the solubilization of the analyzed MPs. Those detergents comprise the long-chain phosphoglycerols, LMPG and LPPG, the bile acid derivative, Chaps, and the nonionic detergents, Tween 20, Thesit, Genapol C-100, NP40 alternative and Brij-72. In this report, we focused on the solubilization effects of pure detergent micelles. Micelles mostly do not have regular spherical shapes but are more of rather disorganized and compact structures [12]. The inability of certain detergents to solubilize or to stabilize an MP could thus arise from an unstable or unfavorable packing of detergent monomers on the surface of the protein and they still might become beneficial for the production of soluble MPs in combination with other amphipathic solutes [12]. The addition of lipids or other detergents could help to eliminate packing defects upon the formation of mixed micelles or bicelles. It is well known that specific lipids are often tightly associated with MPs [15] and they can become important elements for the stabilization of MPs and sometimes they can even be essential for a functional folding. In addition, the refolding of precipitated MPs in detergents might benefit from added lipids [16].

Inclusion body formation is a well-known phenomenon upon conventional cellular expression of proteins, especially in *E. coli*. The overproduced proteins precipitate in a nonstructured and inactive form, and their structural reconstitution usually requires solubilization in a strong detergent, a chaotrope or in organic solvent [17,18]. It should therefore be noted that the CF produced precipitates of MPs could be solubilized by much milder conditions. In particular, LMPG was found to have outstanding properties in the efficient solubilization of structurally different MPs. Lysolipids, such as LMPG and the related LPPG, have been used in the past to isolate large MPs in native and functional form and to analyze small MPs by NMR spectroscopy [19,20]. A survey of 25 different detergents revealed lyso-phosphatidyl-glycerols as superior in stabilizing protein conformations and in preventing the aggregation of monotopic and polytopic MPs at high concentrations [20]. This type of detergent might therefore become highly interesting for the structural analysis of CF produced MPs, especially when analyzed by NMR techniques. However, solubilized precipitate of the Tsx transporter in LMPG did not show any pore-forming activity in the BLM assay, in contrast to the directly

soluble expressed protein in the presence of Triton X-100. The functionality, but also the efficiency, of reconstitution of a MP can be clearly affected by the type of detergent [21]. Despite the use of equal amounts of protein in the different BLM assays, the definite amount of effectively reconstituted protein into the membrane cannot be assessed. It can therefore not be differentiated whether LMPG has inactivated Tsx by adoption of a nonfunctional conformation or whether LMPG-solubilized Tsx is just unable to effectively reconstitute into the membrane of the BLM assay.

In general, the CF expression of MPs in the presence, as well as in the absence, of detergent can result in functionally active protein. The EmrE protein was highly active after resolubilization from a CF produced precipitate in DPC [4], as well as after soluble expression in the presence of DDM [3]. The bacterial MscL protein was functional after soluble expression in the presence of Triton X-100 [2], and, moreover, CF produced precipitate of a bacterial light-harvesting MP that has been solubilized in Triton X-100 was found to be functionally folded [6]. The type of detergent is clearly crucial in the solubilization process for the recovery of MP activity, and high variations in the specific activity can be obtained [11,22]. The secondary structure analysis of the CF expressed V2R and Tsx proteins revealed considerable variations that were dependent on the mode of expression or on the added type of detergent. EmrE, and also V2R, build multimers in their native state, whereas the nucleoside transporter, Tsx, only forms monomers [23–25]. The detection of putative multimeric forms of EmrE in the presence of, for example, Brij-78, DHPC or DM, while only monomers could be observed in the presence of Digitonin or other Brij derivatives, were further evidence of structural variations that correlate with the provided detergent. Furthermore, the marked differences in the specific activity of Tsx after soluble expression in the presence of either Triton X-100 or Brij-35 might also be attributed to structural variations. However, unfavourable detergents could be exchanged after CF expression by gel filtration with another detergent that is more likely to support the functional conformation of an MP [26]. Detergents that have even reported to be deleterious for the activity of a solubilized MP could furthermore still be effective in reconstitution trials because the addition of lipids and the removal of detergent often protects the MP against denaturation [27].

Integral outer membrane proteins of Gram-negative bacteria are characteristic in using amphipathic β-strands to traverse the membrane. All known structures show a typical meander topology of a closed β-barrel in which the last β-strand is hydrogen bonded to the first. However, a considerable versatility of β-barrels is illustrated by the variety of solved structures, such as those of general or specific porins [16,28]. Owing to the limitations on membrane-targeted expression, the majority of structurally analyzed outer membrane proteins have been produced by the refolding of inclusion bodies [16]. The 34 kDa outer membrane protein, Tsx, of *E. coli* is functionally and structurally highly conserved within enteric bacteria, with an overall identity of 78% [29]. Three different functions could be assigned to Tsx: first, the transport of desoxynucleosides and nucleosides; second, the uptake of colicine-like antibiotics; and third, the binding of specific bacteriophages. According to its recently solved 3D crystal structure, Tsx forms a monomeric, 12-stranded β-barrel with six surface-exposed loops and a long and narrow channel spanning the outer membrane and containing several distinct nucleoside-binding sites [25]. The production of Tsx in concentrations of up to 4 $mg \cdot mL^{-1}$ in the RM is the first example of the high-level expression of a β-barrel protein in CF systems. Moreover, the demonstrated functionality of the CF produced Tsx opens a new avenue for the fast and efficient generation of correctly folded outer membrane proteins for functional and structural analysis, and conventional denaturation and tedious refolding steps could be completely avoided.

GPCR proteins transmit a remarkable diversity of endogenous signals into cellular responses by their ligand-induced association with hetero-trimeric G-protein complexes that initiate downstream signalling cascades involving effector enzymes, such as adenylate cyclase and a variety of second messengers. The V2R receptor belongs to the rhodopsin or A-type family of GPCRs, having relatively short loop regions and the proposed classical structure with seven TMS. GPCRs are linked to a number of human hereditary diseases and it is estimated that over 50% of all modern drugs are targeted at GPCRs. The high-level CF expression of GPCRs could provide an interesting alternative tool for structural approaches. Besides the vasopressin type 2 receptor, the CF expression of the β2 adrenergic receptor, the muscarinic acetylcholine receptor and the neurotensin receptor in a batch mode CF system, in concentrations of up to 150 μg per mL of RM, has been reported [5]. The ligand-binding activity of the CF-produced and refolded β2 adrenergic receptor could be demonstrated, and also the V2R protein appears to become folded in the presence of certain detergents. It should be noted that the GPCRs mentioned had to be produced as a fusion with the relatively large thioredoxin and were not expressed with

their native N-terminal ends [5]. The authors concluded that thioredoxin might confer a certain stabilization effect to the MPs. Accordingly, the production of the V2R protein was not detectable with a DNA construct that generated the protein with its native N-terminal end. However, we obtained a high-level expression of the V2R protein with an N-terminal extension of only 14 amino acids, comprising the small T7 tag to the N-terminal end. In addition, the same effect was observed with three further GPCR proteins, (a) the endothelin B receptor, (b) the corticotropin-releasing factor receptor and (c) the human vasopressin type 2 receptor (data not shown). We therefore speculate that not a stabilization effect of an N-terminal fusion protein, but rather an efficient initiation of translation provided by the codon region of the T7 tag is effective for the high-level CF production of the analyzed GPCR proteins.

There are marked differences in the solubilization of MPs, during or after synthesis in a CF system, compared with conventional solubilization from biological membranes. The most obvious difference is that no membranes have to be destabilized that might cause conformational changes or even denaturation of the inserted MPs [30]. The sudden loss of lateral pressure, exposure of hydrophobic surfaces to water, and transient contacts with different solution conditions during extensive purification steps, often cause the structural perturbation of MPs. During CF expression, the large hydrophobic sector of integral MPs could become surrounded by a detergent–micellar-like structure immediately after translation and/or folding, and the exposure to aqueous solution conditions is extremely minimized or even prevented. In addition, the elimination of purification steps further avoids the structural perturbation of MPs. The absence of any labelling background enables the analysis of labelled MP samples by NMR spectroscopy directly in the CF reaction mixture or after simple purification steps [4,31]. Upon solubilization of membranes, often a balance must be found between the mildness of a detergent and its efficiency in covering the hydrophobic surfaces of the membrane lipids. Suitable for the solubilization of MPs out of membranes in a functional form are several detergents that have also been involved in our study, such as derivatives of Brij, Triton, bile-acid salts and short-chain phospholipids, like DHPC [32]. The latter, in particular, seem to be superior to most other detergents as they better resemble the typical membrane phospholipids as the natural environment of MPs [11].

While similar results cannot be guaranteed for every MP, our results strongly suggest several detergents that should be among the first of choice when testing the soluble CF expression of a new protein, and that can generally be recommended for the production of a broader range of structurally different MPs. However, optimized methods must certainly be elaborated for each protein individually. The high-level CF production of integral MPs, in combination with the significantly decreased concentration of proteins needed for the production and X-ray analysis of microcrystals, could enable new perspectives for combined structural approaches.

Experimental procedures

DNA techniques

The *tsx* coding sequence was amplified from plasmid pTX5 [33] by standard PCR using the primers Tsx-up7 (CGGGGATCCGCTGAAAACGACAAACCGCAGTATC) and Tsx-low7 (CGGCTCGAGGAAGTTGTAACCTACTA CCAGGTAAC). The PCR product was restricted with *Bam*HI and *Xho*I, and ligated into the vector pET21a (Merck Biosciences, Darmstadt, Germany), resulting in the plasmid pET-tsx. The Tsx protein was produced without signal peptide and, in addition to the native 272 amino acids, it was modified at its N terminus with an additional 14 amino acids (MASMTGGQQMGRGS) containing a T7 tag and with the C-terminal poly(His)$_6$ tag-containing sequence, LE-HHHHHH. The coding sequence for the porcine V2R protein was amplified from cDNA [34] using standard PCR and the primers V2R-upB (CGGGGATCCCTCAGAGCC ACCACCTCGGCTGTG) and V2R-low (CGGCTCGA GGGACGAGGTGTCCCTGGCCGAGAAGG). The PCR product was ligated into the vector pET21a using *Bam*HI and *Xho*I restriction sites, resulting in plasmid pET-v2r. The recombinant V2R protein contained, in addition to its native 370 amino acids, an N-terminal T7 tag and a C-terminal poly(His)$_6$ tag identical to that of the Tsx protein. The *emrE* gene was amplified from chromosomal DNA of *E. coli* using the primers EmrE-upB (CGGGGATCCATGAACCCTTA TATTTATCTTGGTGG) and EmrE-lowX (CGGCTCG AGATGTGGTGTGCTTCGTGACAA). The PCR product was inserted into the *Bam*HI and *Xho*I restriction sites of the vector pET21a, and the resulting plasmid, pET21-emrE2, was used for the expression of EmrE with an N-terminal T7 tag and a C-terminal poly(His)$_6$ tag.

Black lipid transport assay

The transport assay of the Tsx protein with black lipid bilayer experiments was carried out as described in previous publications [35,36] with minor modifications. The device consists of a Teflon chamber with two aqueous compartments filled with 100 mM potassium phosphate, pH 7.4,

and 1 M KCl. Circular holes in the wall separating the two compartments had an area of ≈ 1 mm^2. Membranes were formed across the holes by applying a 1.25% diphytanoyl phosphatidylcholine and 0.025% octadecylamin solution in *n*-decane. Inactivation of Tsx by high-salt concentrations was prevented by adding ≈ 100 ng of protein to the aqueous phase after the membranes had turned completely black. The membrane current was measured at different voltages by using a pair of matched calomel electrodes, containing salt bridges, inserted into the aqueous solutions on both sides of the membrane. The macroscopic conductance measurements were performed with an ITHACO electrometer (ITHACO, NY, USA) and monitored using an oscilloscope (Nicolet 310, Nicolet, WI, USA). The experiments were carried out at 20–24 °C, and an aqueous solution of 200 mM deoxyuridine was used for the specific inhibition of Tsx.

Cell-free expression

The bacterial CF extracts were prepared from the *E. coli* strain, A19 (*E. coli* Genetic Stock Center CGSC, Department of Molecular, Cellular and Development Biology, Yale University, New Haven, CT, USA), as described previously [4]. The reaction was performed in the CECF mode using a membrane with a cut-off of 15 kDa, either in the analytical scale of 70 μL of RM with an RM/FM ratio of of 1 : 14 (v/v) in microdialysers (Spectrum Laboratories Inc., Breda, the Netherlands) or in a preparative scale of 1 mL of RM using a ratio of RM/FM of 1 : 17 (v/v) in suitable dispodialysers (Spectrum Laboratories Inc.). The reactions were incubated for 20 h at 30 °C in a shaker. The reaction conditions for the CF reaction were as follows. RM and FM: 290 mM potassium acetate; 15 mM magnesium acetate; 100 mM Hepes-KOH, pH 8.0; 3.5 mM Tris-acetate, pH 8.2; 1.8 mM EDTA; 0.2 mM folinic acid; 0.05% sodium azide; 2% polyethylenglycol 8000; 2 mM 1,4-dithiothreitol; 1.2 mM ATP; 0.8 mM each of CTP, UTP, and GTP; 20 mM acetyl phosphate; 20 mM phosphoenol pyruvate; 1 tablet per 10 mL of complete protease inhibitor (Roche Diagnostics, Mannheim, Germany); 2 mM of each of the amino acids R, C, W, M, D and E; 1 mM of each of the other amino acids; RM: 40 μg·mL^{-1} pyruvate kinase (Roche Diagnostics); 500 μg·mL^{-1} *E. coli* tRNA (Roche Diagnostics); 0.3 U·μL^{-1} RNasin (Amersham Biosciences, Freiburg, Germany); 6 U·μL^{-1} T7 RNA polymerase; 35% S30 extract and 15 μg·mL^{-1} plasmid DNA.

Detergents were obtained from Avanti Polar Lipids (Alabaster, AL, USA) (DHPC, DPC, LMPG, LPPG), Merck Biosciences [Genapol C-100, Karlsruhe, Germany NP-40], Carl Rothe (Chaps), Glycon (Luckenwalde, Germany) (DDM, DM) and Sigma (Taufkirchen, Germany) (Brij-35, Brij-56, Brij-58, Brij-72, Brij-78, Brij-97, Brij-98, diC$_6$PC, diC$_8$PC, Digitonin, β-OG, Thesit, Triton X-100, Tween 20).

CD spectroscopy

CD spectroscopy was performed with a Jasco J-810 spectropolarimeter (Jasco Labortechnik, Gross-Umstadt, Germany) in 10 mM sodium phosphate, pH 8.0, 0.5 mM 1,4-dithiothreitol, and with the appropriate detergent. Assays were carried out at standard sensitivity with a band width of 2 nm and a response of 4 s. The data pitch was 0.2 nm and the scanning rate 100 nm·min^{-1}. The spectra were recorded from 190 to 260 nm at 25 °C in a cuvette of 1 mm cell length. The presented data are the average of five scans and smoothed if necessary by means-movement with a convolution width of 25. The α-helical and β-barrel content of proteins was calculated according to their secondary structure with the Jasco secondary structure estimation program.

Preparation and analysis of proteoliposomes

The protein concentration of the GPCR V2R solubilized in 1% (w/v) LMPG was determined by UV measurement at 280 nm in 6 M guanidinum hydrochloride, pH 6.5, according to its molar extinction coefficient. Approximately 10 μM of the protein sample in 15 mM sodium phosphate, pH 7.5, was used for the reconstitution, and an *E. coli* lipid mixture was added at a molar ratio of protein/lipid of 1 : 2500. The solution was incubated at 30 °C for 60 min. The detergent was removed with washed biobeads SM-2 (Bio-Rad, Hercules, CA, USA), presaturated with *E. coli* lipids that were added in 10-fold excess to the detergent. The mixture was incubated for 18 h at 30 °C on a shaker while the biobeads were exchanged twice. The supernatant was then removed, sonified for 1 min in a water bath sonicator and analyzed immediately or stored in liquid nitrogen. The proteoliposomes were analyzed by freeze-fracture electron microscopy, as described previously [4].

Analysis and purification of proteins

For MP purification, the RM was first buffer exchanged on a HiTrap-desalting column of an Äkta Basic system (Amersham Bioscience), at 4 °C, in 20 mM Tris/HCl, pH 8.0, containing the desired detergent. The proteins were further purified by affinity chromatography using a 1 mL HisTrap-HP column at 4 °C with a stepwise elution at 10, 20, 50, 250 and 500 mM imidazole in 20 mM Tris/HCl, pH 8.0, 500 mM sodium chloride and detergent. The Tsx and V2R proteins eluted at 50 mM and 250 mM imidazole, respectively. Suitable fractions were pooled and, if necessary, the buffer was exchanged with HiTrap-desalting columns and concentrated to ≈ 10 μM using a Vivaspin 0.5 mL concentrator (Vivascience AG, Hannover, Germany) with an *m* cut-off of 10 kDa.

For western blot analysis, the proteins were separated on a 10% (w/v) Tricine/SDS gel (EmrE) or on 12% (w/v) Tris/glycine/SDS gels (Tsx and V2R). The gels were

blotted on a 0.45 μm poly(vinylidene difluoride) membrane (Immobilon-P; Millipore, Eschborn, Germany) in a Hoefer TE22 (Amersham Bioscience) wet western blot apparatus for 1 h at 400 mA. The membrane was then blocked for 1 h in blocking-buffer containing 1 × NaCl/Tris (TBS), 7% (w/v) skim milk powder (Fluka, Buchs, Switzerland), 0.1% (w/v) sodium azide and 0.1% (w/v) Triton X-100. The horseradish peroxidase-conjugated T7 tag antibody (Merck Biosciences) and the primary mouse anti-His immunoglobulin (Merck Biosciences) were used at a dilution of 1 : 5000. After extensive washing, the blots were analyzed by chemiluminescence in a Lumi-imager F1^{TM} (Roche Diagnostics).

Proteins were quantified either according to a T7-tagged standard protein (Merck Biosciences) on western blots, or with SDS gels and appropriate mass standards by densitometry using the BIODOCANALYZE software (Biometra, Göttingen, Germany).

Acknowledgements

We are grateful to Vladimir Shirokov for valuable help and Erhard Bremer for providing us with plasmid pTX5. The work was financially supported by the DFG and the SFB 628 'Functional membrane proteomics'.

References

1 Spirin AS, Baranov VI, Ryabova LA, Ovodov SY & Alakov YB (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* **242**, 1162–1164.
2 Berrier C, Park KH, Abes S, Bibonne A, Betton JM & Ghazi A (2004) Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry* **43**, 12585–12591.
3 Elbaz Y, Steiner-Mordoch S, Danieli T & Schuldiner S (2004) *In vitro* synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. *Proc Natl Acad Sci USA* **101**, 1519–1524.
4 Klammt C, Löhr F, Schäfer B, Haase W, Dötsch V, Rüterjans H, Glaubitz C & Bernhard F (2004) High level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem* **271**, 568–580.
5 Ishihara G, Goto M, Saeki M, Ito K, Hori T, Kigawa T, Shirouzu M & Yokoyama S (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Prot Express Purif* **41**, 27–37.
6 Shimada Y, Wang ZY, Mochizuki Y, Kobayashi M & Nozawa T (2004) Functional expression and characterization of a bacterial light-harvesting membrane protein in *Escherichia coli* and cell-free synthesis systems. *Biosci Biotechnol Biochem* **68**, 1942–1948.
7 Le Maire M, Champeil P & Moller JV (2000) Interaction of membrane proteins and lipids with solubilizing detergents Biochim. *Biophys Acta* **1508**, 86–111.
8 De Grip W (1982) Thermal stability of rhodopsin and opsin in some novel detergents. *Methods Enzymol* **81**, 256–265.
9 Kühlbrandt W (1988) Three-dimensional crystallization of membrane proteins. *Q Rev Biophys* **21**, 429–477.
10 Rigaud JL, Chami M, Lambert O, Levy D & Ranck JL (2000) Use of detergents in two-dimensional crystallization of membrane proteins. *Biochim Biophys Acta* **1508**, 112–128.
11 Hauser H (2000) Short-chain phospholipids as detergents. *Biochim Biophys Acta* **1508**, 164–181.
12 Garavito RM & Ferguson-Miller S (2001) Detergents as tools in membrane biochemistry. *J Biol Chem* **276**, 32403–32406.
13 Vinogradova O, Sönnichsen F & Sanders CR (1998) On choosing a detergent for solution NMR studies of membrane proteins. *J Biomol NMR* **4**, 381–386.
14 Odahara T (2004) Stability and solubility of integral membrane proteins from photosynthetic bacteria solubilized in different detergents. *Biochim Biophys Acta* **1660**, 80–92.
15 Palsdottir H & Hunte C (2004) Lipids in membrane protein structures. *Biochim Biophys Acta* **1666**, 2–18.
16 Buchanan SK (1999) β-barrel proteins from bacterial outer membranes: structure, function and refolding. *Curr Opin Struct Biol* **9**, 455–461.
17 Kiefer H (2003) *In vitro* folding of alpha-helical membrane proteins. *Biochim Biophys Acta* **1610**, 57–62.
18 Vallejo LF & Rinas U (2004) Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb Cell Fact* **3**, 11.
19 Williams KA, Farrow NA, Deber CM & Kay LE (1996) Structure and dynamics of bacteriophage IKe major coat protein in MPG micelles by solution NMR. *Biochemistry* **35**, 5145–5157.
20 Krueger-Koplin RD, Sorgen PL, Krueger-Koplin ST, Rivera-Torres IO, Cahill SM, Hicks DB, Grinius L, Krulwich TA & Girvin ME (2004) An evaluation of detergents for NMR structural studies of membrane proteins. *J Biomol NMR* **17**, 43–57.
21 Czerski L & Sanders CR (2000) Functionality of a membrane protein in bicelles. *Anal Biochem* **284**, 327–333.
22 Shivanna BD & Rowe ES (1997) Preservation of the native structure and function of Ca^{2+}-ATPase from sarcoplasmic reticulum: solubilization and reconstitution by new short-chain phospholipid detergent 1,2-diheptanoyl-*sn*-phosphatidylcholine. *Biochem J* **325**, 533–542.

23 Terrillon S, Durroux T, Mouillac B, Breit A, Ayoub MA, Taulan M, Jockers R, Barberis C & Bouvier M (2003) Oxytocin and vasopressin V1a and V2 receptors form constitutive homo- and heterodimers during biosynthesis. *Mol Endocrinol* **17**, 677–691.
24 Ubarretxena-Belandia I, Baldwin JM, Schuldiner S & Tate CG (2003) Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. *EMBO J* **22**, 6175–6181.
25 Ye J & van den Berg B (2004) Crystal structure of the bacterial nucleoside transporter Tsx. *EMBO J* **23**, 3187–3195.
26 Casey JR & Reithmeier AFR (1993) Detergent interaction with band 3, a model polytopic membrane protein. *Biochemistry* **32**, 1172–1179.
27 Lévy D, Gulik A, Bluzat A & Rigaud JL (1992) Reconstitution of the sarcoplasmic reticulum Ca-ATPase: mechanism of membrane insertion into liposomes during reconstitution procedures involving the use of detergents. *Biochim Biophys Acta* **1107**, 283–298.
28 Schirmer T (1998) General and specific porins from bacterial outer membranes. *J Struct Biol* **121**, 101–109.
29 Nieweg A & Bremer E (1997) The nucleoside-specific Tsx channel from the outer membrane of *Salmonella typhimurium*, *Klebsiella pneumoniae* and *Enterobacter aerogenmes*: functional characterization and DNA sequence of the tsx genes. *Microbiology* **143**, 603–615.
30 Rosenbusch JP (2001) Stability of membrane proteins: Relevance for the selection of appropriate methods for high-resolution structure determinations. *J Struct Biol* **136**, 144–157.
31 Guignard L, Ozawa K, Pursglove SE, Otting G & Dixon NE (2002) NMR analysis of *in vitro*-synthesized proteins without purification: a high-throughput approach. *FEBS Lett* **524**, 159–162.
32 Kessi J, Poirée JC, Wehrli E, Bachofen R, Semenza G & Hauser H (1994) Short-chain phosphatidylcholines as superior detergents in solubilizing membrane proteins and preserving biological activity. *Biochemistry* **33**, 10825–10836.
33 Bremer E, Middendorf A, Martinussen J & Valentin-Hansen P (1990) Analysis of the tsx gene, which encodes a nucleoside-specific channel-forming protein (Tsx) in the outer membrane of *Escherichia coli*. Gene **96**, 59–65.
34 Kojro E, Postina R, Gilbert S, Bender F, Krause G & Fahrenholz F (1999) Structural requirements for V2 vasopressin receptor proteolytic cleavage. *Eur J Biochem* **266**, 538–548.
35 Benz R, Janko K, Boos W & Lauger. P (1978) Formation of large, ion-permeable membrane channels by the matrix protein (porin) of *Escherichia coli*. *Biochim Biophys Acta* **511**, 205–319.
36 Maier C, Bremer E, Schmid A & Benz R (1988) Pore-forming activity of the Tsx protein from the outer membrane of *Escherichia coli*. Demonstration of a nucleoside-specific binding site. *J Biol Chem* **263**, 2493–2499.

(P 03)

Cell-free production of integral membrane proteins on a preparative scale. C. Klammt, D. Schwarz, V. Dötsch and F. Bernhard. *Meth. Mol. Biol.*, accepted. (2006)

Running title: Cell-free membrane proteins

Cell-free production of integral membrane proteins on a preparative scale

*Christian Klammt, Daniel Schwarz, Volker Dötsch, and Frank Bernhard**

Centre for Biomolecular Magnetic Resonance, University of Frankfurt/Main, Institute for Biophysical Chemistry, Max-von-Laue-Str.9, D-60438 Frankfurt/Main

* Corresponding author

Phone: +49-69-798-29620

FAX: +49-69-798-29632

Email: fbern@bpc.uni-frankfurt.de

Key words: integral membrane proteins, structural analysis, solubilization, reconstitution, detergent, NMR spectroscopy, amino acid transporter, G-protein coupled receptors, ß-barrel proteins, multidrug resistance.

Abstract

The chapter will focus on the high level cell-free production of integral membrane proteins having multiple transmembrane segments by using an individual coupled transcription/translation system based on an *Escherichia coli* S30-extract. We describe in detail the set-up and optimization of the cell-free expression technique in order to obtain the maximum yield of recombinant proteins. The protocol can be used for the expression of soluble membrane proteins as well as for their production as a precipitate. In addition, we will provide protocols for the efficient solubilization and reconstitution of membrane proteins directly from the cell-free produced precipitates.

1. Introduction

Although membrane proteins (MP's) spanning lipid bilayers with multiple transmembrane segments (TMS) are highly abundant in fully sequenced genomes, only few high-resolution structures have been determined so far. This obvious discrepancy can mainly be attributed to the tremendous difficulties that generally emerge when MP's in the required amounts for a structural analysis by X-ray crystallography and NMR spectroscopy need to be prepared (***1, 2***). Unfortunately, only a very limited number of MP's could so far be produced in conventional *Escherichia coli in-vivo* expression systems at a level of at least 1 mg/liter of culture (***2, 3***).

In contrast, *in-vitro* cell-free (CF) expression systems have recently been shown to be considerably suitable for the high level expression of MP's (***4-6***). This finding might represent a breakthrough for the structural and functional analysis of MP's whenever high amounts of protein are required. The advantage of CF expression systems in the production of MP's might be due to the elimination of some principal problems occurring in conventional *in-vivo* systems, like toxicity of the overproduced MP's upon insertion into the cytoplasmic membranes, poor growth of overexpressing strains, proteolytic degradation of the expressed MP's, or generally unfavourable impacts on cellular metabolisms. In this chapter, we describe the efficient production of MP's in a CF coupled transcription/translation system using *Escherichia coli* S30 cell extract and phage T7 RNA-polymerase for transcription. The set up of the reaction is in the continuous exchange mode, where two compartments, holding the reaction mixture (RM) and the feeding mixture (FM), are separated by a semipermeable membrane (***7-9***). Key elements of the CF system like the bacterial S30-extract preparation (***10***), the energy system and the concentrations of precursors and of beneficial additives have been optimized in order to yield several mg of recombinant MP per one single ml of RM.

It should be highlighted that MP's can be produced either as a precipitate or as soluble proteins. The CF expression system is considerably tolerant upon relatively high

concentrations of a variety of detergents and lipids (***4-6***), and we compile suitable substances that can be provided as a hydrophobic environment in order to stabilize the synthesized MP's immediately after translation. Alternatively, the vast majority of MP's produced without supplemented detergents most likely will end up as precipitates (***4, 5***). However, those precipitates usually can easily be solubilized with suitable detergents without the necessity to apply extensive denaturation and renaturation steps as known from refolding protocols of inclusion bodies. The organization of the CF produced MP precipitates might therefore be different from that of inclusion bodies formed during *in-vivo* expressions.

2. Materials

2.1. S30-extract preparation

1. 10 liter fermenter.
2. French Press cell disruption device.
3. Liquid N_2.
4. S30-A buffer: 10 mM Tris-acetate, pH 8.2, 14 mM $Mg(OAc)_2$, 6 mM ß-mercaptoethanol, 0.6 mM KCl. Prepared as 50x stock solution and stored at 4°C.
5. S30-B buffer: 10 mM Tris-acetate, pH 8.2, 14 mM $Mg(OAc)_2$, 0.6 mM KCl, 1 mM DTT, 0.1 mM phenylmethanesulfonyl fluoride. Prepared as 50x stock solution and stored at 4°C.
6. S30-C buffer: 10 mM Tris-acetate, pH 8.2, 14 mM $Mg(OAc)_2$, 0.5 mM DTT, 0.6 mM KOAc. Prepared as 50x stock solution and stored at 4°C.
7. NaCl, 4 M. Store at 4°C.
8. Terrific Broth (TB) medium (per liter): 24 g yeast extract, 12 g tryptone, 4 ml 100 % glycerol, 100 mM potassium phosphate buffer.
9. Dialysis tubes, type 27/32 MWCO 14 kDa (Roth, Karlsruhe, Germany; cat. no. 1784.1).

10. Bacterial strain A19 [*rna19 gdh A2 his95 relA1 spoT1 metB1*] *E. coli* Genetic Stock Center (*E. coli* Genetic Stock Center, New Haven, USA, CGSC no. 5997).
11. Centriprep devices YM-10 (Amicon, Witten, Germany; cat. no. 4305).

2.2. Reaction for cell-free protein expression

1. Total tRNA from *E. coli* (Roche Diagnostics GmbH, Mannheim, Germany; cat. no. 109550), 40 mg/ml in distilled water. Store at -20°C.
2. Pyruvate kinase, 10 mg/ml (Roche Diagnostics GmbH, Mannheim, Germany; cat. no. 109045). Store at -20°C.
3. T7-RNA-polymerase, 40 U/ul (Roche Diagnostics GmbH, Mannheim, Germany; cat. no. 881775), (*see* **Note 1**). Store at -20°C.
4. NTP-Mix, 75-fold, 90 mM ATP, 60 mM each CTP, GTP, UTP, pH 7.0 with NaOH. Store at -20°C.
5. Dithiothreitol, 0.5 M. Store at -20°C.
6. Folinic acid, 20 mM. Store at -20°C.
7. PEG 8.000, 40 %. Store at -20°C.
8. NaN_3, 10 %. Store at 4°C.
9. Amino acid mixture, 4 mM of each of the 20 amino acids (*see* **Note 2**). Store at -20°C.
10. Amino acid mixture R, C, W, M, D, E, 16.7 mM (*see* **Note 3**). Store at -20°C.
11. RNAguard porcine RNAse inhibitor, 40 U/µl (Amersham; cat. no. AP 27-0816-01). Store at -20°C.
12. Complete mini protease inhibitor (Roche Diagnostics GmbH, Mannheim, Germany; cat. no. 1836153), 50-fold concentrated in water. Store at -20°C.
13. Acetyl phosphate, potassium salt, 1 M in distilled water, pH 7.0 with KOH. Store at -20°C.

14. Phosphoenol-pyruvic-acid, mono-potassium salt, 1 M in distilled water, pH 7.0 with KOH. Store at -20°C.
15. HEPES buffer, 2.5 M, pH 8.0 with KOH. Store at 4°C.
16. KOAc, 4 M. Store at 4°C.
17. $Mg(OAc)_2$, 1 M. Store at 4°C.
18. KOH, 10 M.
19. NaOH, 5 M.
20. Reaction container for the cell-free expression: Microdialysers (cutoff 15 kDa, Roth, Karlsruhe, Germany; cat. no. T698.1), and dispodialysers (reaction volume 1ml or higher, Roth, Karlsruhe, Germany; cat. no. O359.1). Devices may be re-used several times (*see* **Note 4**).
21. Incubator for the cell-free reaction: e.g. standard shaker with temperature control or rollers like the Universal Turning Device (Vivascience, Göttingen, Germany; cat. no. IV-76001061) placed in an incubator (*see* **Note 5**).
22. Plasmid vectors containing the T7 promoter regulatory region: e.g. pET21 (Merck Biosciences, Darmstadt, Germany).
23. Glass vials, 50 ml (Roth, Karlsruhe, Germany, cat. no.: X663.1).

2.3 Preparation of solubilized or reconstituted membrane proteins

1. Detergents: α-[4-(1,1,3,3-tetramethylbutyl)-phenyl]-ω-hydroxy-poly(oxy-1,2-ethandiyl) (Triton X-100), Digitonin, Dodecylpoly(ethyleneglycolether)$_9$ (Thesit), Sodium dodecylsulfate (SDS), Polyethylenglycododecylether (Brij-35), n-octyl-ß-D-glucopyranosid (ß-OG); (Sigma-Aldrich, Taufkirchen, Germany).
   3-[N-(cholanamidopropyl)-dimethyl-ammonio]-1-propanesulfonate (CHAPS), Polyoxy-ethylene (20) sorbitane monolaurate (Tween 20); (Roth, Karlsruhe, Germany).

   n-dodecylphosphocholine (DPC), 1-myristoyl-2-hydroxy-sn-glycero-3-[phosphor-rac-(1-glycerol)] (LMPG); (Avanti-Lipids, Alabaster, USA).
   n-dodecyl-ß-D-maltoside (DDM); (Glycon Biochemicals GmbH Luckenwalde, Germany).
2. *E. coli* lipid mixture (Avanti-Lipids, Alabaster, USA) stock suspension, 50 mg/ml in distilled water.
3. BT Bio-beads SM-2 (Biorad, München, Germany).

3. Methods

3.1. Construction of expression plasmids

The transcription of the target gene depends on the T7 RNA-polymerase and therefore the essential T7 regulatory sequences (promoter, ε-enhancer or g10 sequence, ribosomal binding site (RBS), terminator) need to flank the coding region (*see* **Fig. 1**). Standard expression vectors like that of the pET series can be used for cell-free expression (*see* **Note 6**).

3.2. Preparation of the S30 cell-free extract

The preparation of the S30-extract includes fermentation and harvesting of the cells, the cell disruption, run-off procedure and dialysis. While we recommend to finish the complete process consecutively, the protocol could optionally be interrupted after the harvesting step (*see* **Note 7**). The quality of individual batches of S30-extracts can vary considerably and at least an optimization for the Mg^{2+} and K^+ ion concentrations should be made for each new preparation.

1. Take a fresh overnight culture from the *E. coli* strain A19 in TB medium and inoculate 10 liters of TB medium in a fermenter at a ratio of 1:10 (*see* **Note 8**).
2. Incubate the cells at 37°C with vigorous stirring and good aeration.

3. Control the bacterial growths by measuring the optical density at 595 nm until the cells reach the mid-log phase corresponding to an OD_{595} of appr. 3.5 (*see* **Note 9**).
4. Switch off the heating of the fermenter and chill down the broth to ≤ 10°C as quickly as possible (≤ 45 min.) in order to stall the growth of the bacteria (see **Note 10**). The final OD_{595} of the broth after cooling should be ≤ 4.5.
5. Harvest the cells by centrifugation at 4°C for 10 min. at 7.000 x g in pre-cooled beakers.
6. Resuspend the pellet with a glass-rod or similar in 100 ml of S30-A buffer pre-cooled to 4°C. Centrifuge at 8.000 x g at 4°C for 10 min. Repeat the washing step two more times while the final centrifugation step should last 30 min. Out of a 10 liter fermenter with TB medium, you should now yield appr. 60 g -70 g wet-weight of bacterial cells.
7. Suspend the cell pellet in an equal volume (e.g. for 65 g cells use 65 ml of buffer) of S30-B buffer pre-cooled at 4°C.
8. Disrupt the cells by passing through a French-Press cell pre-cooled at 4°C with one pass at 1.200 psi (*see* **Note 11**).
9. Pellet cell debris by centrifugation at 30.000 x g at 4°C for 30 min. The pellet can be quite smooth at this stage and a turbid solution immediately above the pellet, if present, should not be transferred.
10. Transfer the upper 2/3 of the supernatant in a fresh vial and repeat the centrifugation step.
11. Remove the upper 2/3 of the supernatant carefully, adjust to a final concentration of 400 mM NaCl and incubate at 42°C for 45 min. in a water bath (*see* **Note 12**). The solution will become turbid.
12. Fill the turbid solution in a dialysis tubing (cut-off 14 kDa) and dialyse at 4°C against 60 volumes of S30-C buffer with gentle stirring. Exchange the dialysis buffer after 2 h and continue to dialyse overnight.
13. Centrifuge the extract at 30.000 x g at 4°C for 30 min.

14. Remove the clear supernatant, fill suitable aliquots (*see* **Note 13**) in plastic tubes and freeze in liquid nitrogen. The frozen extract can be stored at -80°C for months. One 10 liter fermenter yields appr. 50 ml – 60 ml of extract.
15. Optional step: Before aliquoting, the extract could be concentrated to appr. half of the original volume by using microconcentrator devices with a cut-off of 10 kDa (*see* **Note 14**).

3.3. Set-up of a cell-free reaction

Generally, the reaction can be set up in analytical modes for optimization and screening reactions as well as in preparative modes for the production of mg amounts of recombinant protein. As a device for analytical scale reactions with a RM volume of 70 µl, microdialysers that are commercially available with different MWCO's can be used (*see* **Fig. 2A** and **Note 15**). For preparative reactions with RM volumes of 500 µl and larger dispodialysers (or simple dialysis tubes) can be used (*see* **Fig. 2B**). The ratio of RM/FM should be 1:14 (microdialysers) or 1:17 (dispodialysers) (v/v) (*see* **Note 16**). The final concentrations of the individual components are listed in **Table 1**. An example of a pipetting protocol is given for a preparative reaction with 1 ml RM and 17 ml of FM in **Table 2**.

1. Thaw aliquoted stock solutions and mix carefully. The enzymes, tRNA and the S30 extract should be kept on ice after thawing.
2. First make the master-mix FRM by pipetting the components common to FM and RM (*see* **Table 2**).
3. Take the appropriate aliquots from the master-mix FRM (*see* **Table 2**) and first complete the FM.
4. Pre-incubate the FM at 30°C in a water bath.

5. Start completing the RM (*see* **Table 2**). All steps should be carried out on ice. Mix but do not vortex. Keep the solution on ice.
6. Fill 17 ml of the FM into the FM-compartment. (*see* **Note 17**).
7. Fill 1 ml of RM into the RM-compartment. Avoid any air-bubbles at the dialysis membrane that might restrict an efficient exchange between the two compartments. All dialysers should be thoroughly rinsed with distilled water before use.
8. Incubate the reaction on a suitable shaking or rolling device, e.g. the Universal Turning Device with at least 40 rpm at 30°C for 15 h to 20 h.
9. Harvest the RM containing the recombinant protein. Depending on the cut-off of the dialysis membrane, the recombinant protein might also be present in the FM.

3.4. Optimization of the expression yields

This step is especially important in order to achieve the highest yields possible, i.e. incorporation efficiencies of the added amino acids of up to 20 %. We recommend to run an optimization experiment at least for the highly critical concentrations of Mg^{2+} and K^+ ions with each newly prepared batch of S30-extract. The final Mg^{2+} and K^+ concentrations should be varied by addition of suitable amounts of $Mg(OAc)_2$ or KOAc in the range between 12 - 17 mM and 270 - 330 mM, respectively (*see* **Note 18**). The Mg^{2+} concentrations should be titrated in 1 mM steps and the K^+ concentrations in 10 mM steps. Consider that the optima of the two ions depend on each other.

3.5. Cell-free expression of membrane proteins as a precipitate

MP's can be expressed in the cell-free system either as a precipitate or as soluble proteins by the addition of suitable detergents or lipids directly into the reaction (*see* **Fig. 3**). While it might be more likely to obtain higher yields if a MP is produced as a precipitate, preparative amounts of protein can generally be synthesized by both ways. The precipitates are

predominantly formed by the target protein but they still can contain a considerable variety of co-precipitated proteins originating from the S30-extract (*see* **Fig. 4**). The cell-free production of MP's in form of a precipitate follows exactly the standard protocol (***see* 3.3**). The synthesis of the MP can be followed by the increasing turbidity of the RM. After incubation, the MP can be harvested by centrifugation of the RM at 10.000 x g for 10 min.

3.6. Solubilization of CF-expressed MP's

Interestingly, the precipitated MP's do not seem to be completely unfolded if compared with the inclusion body formation by conventional *in-vivo* expression. The solubilization of the cell-free produced MP precipitates usually does not require extensive unfolding and refolding steps like the treatment with guanidinium hydrochloride or urea. Most MP precipitates start to solubilize more or less immediately after addition of a suitable detergent. The choice of the detergent certainly depends on the nature of the recombinant MP. A list of detergents that have been useful for solubilization is given in **Table 3**. It is important to mention that some detergents like e.g. LMPG, that are only fairly suitable for the soluble expression of MP's, are on the other hand highly efficient in the solubilization of MP-precipitates. However, a 100% efficiency of solubilization might never been achieved.

1. Centrifuge the RM containing the suspended MP-precipitate at 20.000 x g for 10 min. at room temperature.
2. Discard supernatant and wash the precipitate in an adequate solubilization buffer (e.g. 50 mM Na_2HPO_4, pH 7.8, 1 mM DTT) of a volume equal to the RM (*see* **Note 19**). The pellet should be carefully suspended by pipetting.
3. Repeat centrifugation step.
4. Suspend the pellet in solubilization buffer in a volume equal to the RM volume and containing detergent (e.g. 2% DPC (w/v)). Incubate at 30°C - 37°C up to several hours (*see* **Note 20**).

5. Remove unsoluble protein by centrifugation at 20.000 x g for 10 min. at room temperature and remove the solubilized MP in the supernatant for further analysis.

3.7. Reconstitution of solubilized MP's

The successful reconstitution strongly depends on the right choice of a suitable lipid or lipid mixture and those should be primarily selected according to the origin of the recombinant MP. *E. coli* lipids are predominantly composed of phosphatidylethanolamine (PE), phosphatidylglycerol and cardiolipin, while eukaryotic lipids are mainly mixtures of phosphatidylcholine, cholesterol, PE and phosphatidylserine. A huge variety of defined lipids or lipid mixtures as well as crude lipid isolates of different origin are commercially available for selection. As a rapid control of an effective and homogeneous reconstitution of MP's we recommend the analysis of the proteoliposomes by freeze-fracture electron microscopy. The homogeneous insertion of protein particles into the membrane as shown in **Fig. 5** for the cysteine transporter YfiK provides first evidence of a structural and functional reconstitution of the MP. As an example, we describe the reconstitution of a CF produced MP in an *E. coli* lipid mixture.

1. Add *E. coli* lipids (50 mg/ml in water) in a molar ratio of 2000 : 1 (lipid : MP) to the solubilized protein.
2. Incubate at 30°C for one hour.
3. Wash 750 mg biobeads in 10 ml of 100 % methanol, let the biobeads settle down and discard the supernatant.
4. Wash the biobeads three-times with 10 ml of distilled water
5. Finally wash the biobeads two-times in the buffer used for the detergent solution
6. Presaturation with lipids: Add 100 ul of the lipid suspension to 1 ml of washed biobeads in 10 ml of detergent buffer (e.g. 25 mM HEPES, pH 7.4, 150 mM NaCl) and incubate for 30 min. at RT.

7. Add lipid-presaturated biobeads in a ratio biobeads : detergent of 100 : 1 (w/w) to a vial containing the micelle suspension and let the detergent adsorb to the biobeads by incubation on a shaker at 30°C.
8. Incubate over-night on a shaker at 30°C.
9. Next morning, remove the vial from the shaker, let the biobeads settle down and transfer the supernatant into a new vial. Then incubate with a second aliquot of lipid-presaturated biobeads for 6 hours at 30°C on a shaker.
10. Harvest the supernatant containing the proteoliposomes. The proteoliposomes can be stored at 4°C until or frozen in liquid nitrogen until further analysis.

3.8. Cell-free expression of soluble membrane proteins

1. The expression of soluble MP's in the preparative scale can require first some optimization steps in order to evaluate the optimal reaction conditions. At the beginning, only some µg of protein might be produced. We therefore recommend the addition of an antigen epitope tag like the T7-tag at the N-terminal end of the recombinant protein in order to facilitate the detection of the expressed MP in western-blots. Furthermore, an additional poly(His)$_6$-tag at the C-terminal end of the recombinant protein would enable the fast purification out of the RM by standard metal-chelate-chromatography .
2. The cell-free expression system is considerably tolerant against a variety of detergents even if supplied in relatively high concentrations (***4-6***). Detergents that are compatible with the CF synthesis can be either ionic or non-ionic, but in most cases they have a relatively low critical micellar concentration (CMC). First, the most appropriate detergents have to be identified as their effects on the expression and solubility of specific MP's can be very different (**see Fig. 6**). Overall, Brij-35 has so far turned out to be one of the most effective detergents for the production of soluble MP's (*see* **Table 3** and **Fig. 6**). A detergent screen including a variety of detergents (*see* **Table 3**) should be set up in the

analytical scale mode. An example for the CF expression of the nucleoside transporter Tsx in presence of different detergents is shown in **Fig. 6**. In that case, Brij-35, Triton X-100 and DDM would clearly be the best choices in order to produce soluble protein. Also the addition of mixed micelles composed of two or more detergents might be considered. The production of MP's can be monitored by Coomassie-Blue staining or western-blotting after separation of the RM by SDS-polyacrylamide gel-electrophoresis. In order to differentiate between soluble MP's and precipitate, the RM should be fractionated by centrifugation at 20.000 x g for 30 min. at RT prior to analysis.

3. After the most suitable detergents for your specific protein have been identified, run a second series of analytical scale reactions in order to define the optimal concentration of the preferred detergent. E.g. the highest yields of soluble Tsx protein in presence of Triton X-100 were only obtained at concentrations of 4x CMC and above (*see* **Fig. 6**).
4. The solubilized MP's in micelles might not necessarily adopt a structured and functional conformation. A rapid first analysis of the MP structure should therefore be carried out in order to verify the structured folding of the CF expressed MP. Suitable techniques might be Circular Dichroism spectroscopy or the recording of heteronuclear single quantum correlation spectra by solution NMR spectroscopy. Highly valuable would be certainly the availability of any activity or binding assay.

4. Notes

1. T7 RNA-polymerase will be the most expensive component of the reaction and in addition, commercial enzymes are often too low concentrated. At least 3 units final concentration of T7 RNA-polymerase per µl of RM should be added but up to 40 units per µl can be used. We therefore highly recommend the overproduction of the enzyme in *E. coli* (***11***). We isolate the T7-RNA-polymerase from strain BL21 (DE3) by a single-step purification with a Q-sepharose column. Out of a 4 liter fermentation, the yield can be

appr. 5 x 10^5 units. The isolated T7 RNA-polymerase can be stored in glycerol at -80^{o}C for many months.

2. L-tyrosine is prepared as 20 mM, the other 19 amino acids are made as 100 mM stock solutions dissolved in water. The solubility of L-aspartic acid, L-cysteine, L-glutamic acid and L-methionine could be improved by using 100 mM HEPES pH 7.4. L-tryptophan is dissolved in 100 mM HEPES, pH 8.0, upon sonification in a water-bath. Appropriate aliquots of the individual stocks are then combined in the amino acid mixture. The final concentration of amino acids in the reaction is 1 mM. The yields of recombinant protein might be improved by using higher concentrations, and by adjusting the amino acid concentrations according to the composition of the recombinant protein. Least abundant amino acids (present ≤ 3 % in the protein) should then be added at 1.25 mM, medium abundant (between 3 and ≤ 8 %) at 1.8 mM and highly abundant (more than 8 %) at 2.5 mM final concentration.
3. Some amino acids tend to be unstable and increasing their concentration in the reaction significantly improves the synthesis of recombinant proteins (***12, 13***).
4. Dialysers containing membranes made of regenerated cellulose should be preferred as they enable a better exchange of compounds if compared with cellulose-ester. Dialysers might be reused after washing with distilled water and they can be stored in distilled water supplemented with 0.01 % sodium azide at 4^{o}C.
5. Depending on the design of the reaction container, quite a variety of incubators can be used. The only consideration is that temperature control and an efficient agitation of the reaction device, either by rolling, shaking or stirring, can be provided.
6. RNAse in the lysis buffer during plasmid isolation principally could cause problems in the cell-free reaction and it might be avoided during DNA preparation. Distilled water should be used to finally dissolve the DNA in order to prevent the addition of undesired ions into the cell-free reaction. The plasmid DNA should be concentrated in a speedvac to a

final concentration of at least 0.1 µg/µl. The final concentration of DNA in the cell-free reaction should be at least 15 µg per one ml of RM.

7. At this stage, the cells could be frozen at -80°C as a thin plate wrapped in a sheet of aluminum foil and stored until further usage. However, in our hands the quality of the S30 extract was always considerably improved if the preparation was consecutively completed without interruption.
8. A variety of bacterial strains including *E. coli* A19, D10 or BL21 has been described as source for cell-free S30-extracts. In our hands, A19 yielded the most efficient extracts with the best reproducible quality. Using a fermenter to grow the bacteria might be essential because of the better supply of oxygen.
9. Harvesting the cells in the mid-log phase is essential. For new modifications of the fermentation conditions, an initial pilot experiment should always be carried out in order to define the exact optical density of the culture at mid-log phase.
10. It is critical to cool down the broth as quickly as possible. As the cells rapidly divide, too slow cooling might result in entering the late-log or even stationary phase of growth and that could considerably reduce the quality of the extract. For the described conditions, the reduction of temperature from 37°C to 10°C should be optimally completed in not more than 45 minutes. An external cooling-unit could be connected or the broth in the fermenter could be cooled down by adding blocks of frozen TB medium.
11. Using a French Press device for cell disruption is important. Disruption by ultra-sonification would result in bad quality of the extract, probably by disintegration of the ribosomes.
12. This step causes the dissoziation of endogenous mRNA from the ribosomes. Other run-off procedures by adding substrates in order to terminate transcriptions have been described (***10, 14***) and are similar effective, but more expensive. The incubation at the relatively high heat shock temperature of 42°C has been proven to be essential. Using lower

temperatures like 37°C are by far not as effective. A lot of proteins in the extract precipitate during this step while the proteins necessary for the translation process remain stable.

13. Make aliquots of different appropriate sizes according to your requirements (e.g. 100 µl, 500 µl, 1ml) as the individual aliquots must not be re-frozen.
14. Condensing the extract two-fold can increase the efficiency to appr. 1.5 times. The final protein concentration in the extract should be between 20 and 30 mg/ml.
15. Relatively large cut-offs up to 50 kDa can be used without having significant leaching of the proteins relevant for the transcription/translation process. It is therefore assumed that those proteins form a large macromolecular complex together with the mRNA. The only protein that is released from that complex is the synthesized recombinant target protein. The selection of the MWCO of the reaction device should therefore depend on the MW of the target protein and whether it should remain in the RM. Generally, MWCO's < 10 kDa tend to reduce the expression yields probably due to a more restricted exchange of low molecular weight substances and we recommend to routinely use a MWCO of 15 kDa.
16. The ratio RM/FM is certainly important for an efficient production of recombinant protein. The higher the volume of the FM, the more protein can be produced. However, the FM contains relatively expensive precursors like PEP, acetyl phosphate, nucleotides or even labelled amino acids. We found the given ratio to be an optimal compromise in order to yield high levels of protein at reasonable costs. The exchange of the FM after several hours of incubation would additionally increase the yield of the recombinant protein. However, as the increase usually is only some 30 %, it would be more economically to run two identical reactions separately.
17. For preparative scale reactions, e.g. in dispodialysers, standard glass vials of suitable sizes can be used as FM compartment. For analytical scale reactions, e.g. in microdialysers,

suitable plastic containers like the lower part of 50 ml polypropylene test-tubes (Greiner, Solingen, Germany, cat. no.: 210261) that can be fixed to the microdialysers with a tape (e.g. Nescofilm, Azwell, Osaka, Japan) can be used.

18. For the calculation of the final concentrations, it has to be considered that 5 mM of Mg^{2+} and 140 mM of K^{+} are already present in the reaction from other components.
19. It could be beneficial to include detergents that are not suitable for dissolving the recombinant protein into the washing buffer. Co-precipitated contaminants might then be dissolved, resulting in a purer sample of the MP.
20. Considerably higher detergent concentrations can be used for the re-solubilization of precipitated MP's if compared with the direct addition of detergent into the CF reaction for the production of soluble MP's.

Acknowledgements

We are grateful to Vladimir Shirokov, Alexander Spirin and Heinz Rüterjans for their help in setting up the cell-free expression system. We further thank Winfried Haase for the electron microscopical analysis of proteoliposomes.

References

1. Arora, A., and Tamm, L.K. (2001) Biophysical approaches to membrane protein structure determination. *Curr. Opin. Struct. Biol.* **11**, 540-547.
2. Wang, D.N., Safferling, M., Lemieux, M.J., Griffith, H., Chen, Y., and Li, X.D. (2003) Practical aspects of overexpressing bacterial secondary membrane transporters for structural studies. *Biochim. Biophys. Acta* **1610**, 23-365.
3. Tate, C.G. (2001) Overexpression of mammalian integral membrane proteins for structural studies. *FEBS Lett.* **504**, 94-98.

4. Klammt, C., Löhr, F., Schäfer, B., Haase, W., Dötsch, V., Rüterjans, H., Glaubitz, C., and Bernhard, F. (2004) High level cell-free expression and specific labeling of integral membrane proteins. *Eur. J. Biochem.* **271**, 568-580.
5. Elbaz, Y., Steiner-Mordoch, S., Danieli, T., and Schuldiner, S. (2004). In vitro synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. *Proc. Natl. Acad. Sci.* USA. **101**, 1519-1524.
6. Berrier, C., Park, K.H., Abes, S., Bibonne, A., Betton, J.M., and Ghazi, A. (2004). Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry* **43**, 12585-12591.
7. Baranov, V.I., and Spirin, A.S. (1993) Gene expression in cell-free system on preparative scale. *Methods Enzymol* **217,** 123-42.
8. Spirin, A.S., Baranov, V.I., Ryabova, L.A., Ovodov, S.Y., and Alakov,Y.B. (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* **242**, 1162-1164.
9. Kim, D.M., and Choi, C.Y. (1996) A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. *Biotechnol. Prog.* **12**, 645-649.
10. Zubay, G. (1973) *In vitro* synthesis of protein in microbial systems. *Annu. Rev. Genet.* **7**, 267-287.
11. Li, Y., Wang, E., and Wang, Y. (1999). A modified procedure for fast purification of T7 RNA-polymerase. *Protein Expr. Purif.* **16**, 355-358.
12. Kim, D.M., Swartz, J.R .(1999) Prolonging cell-free protein synthesis with a novel ATP regeneration system. *Biotech. Bioengineering* **66,** 180-188.
13. Kim, D.M., Swartz, J.R. (2000) Prolonging cell-free protein synthesis by selective reagent additions. *Biotech. Prog.* **16**, 385-390.

14. Kigawa, T., Yabuki, T., Matsuda, N., Matsuda, T., Nakajima, R., Tanaka, A., and Yokoyama, S. (2004) Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. *J. Struct. Funct. Genomic* **5,** 63–68.

Legends to the figures

Fig. 1. Essential T7 regulatory sequences for optimal CF protein production.

Fig. 2. Design of cell-free reaction containers. A, microdialyser for analytical scale reactions. The microdialyser holding the approx. 70 µl of RM is placed into a suitable vial (e.g. bottom part of a 50 ml Falcon plastic tube) holding the 1 ml of FM. The vial is fixed to the microdialyser by a sealing tape (e.g. nescofilm). The set-up is than incubated on a shaker at 30°C. B, Dispodialyser for preparative scale reactions. A dispodialyser holding a suitable volume of RM is placed into a plastic vial containing the corresponding volume of FM. The set-up is then incubated on a roller at 30°C.

Fig. 3. Different modes of MP production by CF expression. A: Production of MP precipitates followed by re-solubilization after addition of detergents. B: Production of soluble MP's by addition of detergents directly into the CF system. The MP's solubilized in micelles by the two different modes can be directly analyzed or reconstituted into liposomes for further analysis.

Fig. 4. MP's from different families that have been produced as a precipitate by cell-free expression. The precipitate of an analytical reaction (70 µl of RM) was suspended in 70 µl of buffer and the proteins in 1 µl of each suspension were separated by electrophoresis in a 12 % SDS-polyacrylamide-gel. The arrows indicate the overproduced recombinant MP's. The right lane shows the molecular weight marker. GPCR: G-protein coupled receptor; OMP: outer membrane protein; SMR: small multidrug transporter.

Fig. 5. Freeze-fracture electron microscopy of the cell-free expressed *E. coli* cysteine transporter YfiK after reconstitution into an *E. coli* lipid mixture. The YfiK protein was first produced as a precipitate in a standard CF reaction, then solubilized in 1 % (v/v) LMPG and subsequently reconstituted into *E. coli* lipid vesicles by the above described procedure. The randomly distributed particles (small arrow) indicate the homogenous *in-vitro* incorporation of the MP into the vesicular membranes (bold arrow) (figure courtesy of W. Haase).

Fig. 6. Soluble CF expression of the ß-barrel-like nucleoside transporter Tsx in presence of detergents. Samples were from analytical scale CF reactions with 70 µl RM. The concentrations of the added detergents are given in CMC. After the reaction, 0.8 µl of the RM, or of the precipitated proteins suspended in a volume equal to the RM volume, were loaded on a 12 % SDS-polyacrylamide-gel. The arrows indicate the soluble Tsx protein in the reaction mixture. S, soluble protein; P, precipitated protein; M, marker proteins (kDa); TX-100, Triton X-100.

Fig. 1.

promoter ε-enhancer RBS

TAATACGACTCACTATA - N_{38} - TTAACTTTA – N_2 - AAGGAG – N_7 -

coding sequence terminator

ATG – N_x – TAA – N_x – AACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTT

Fig. 2.

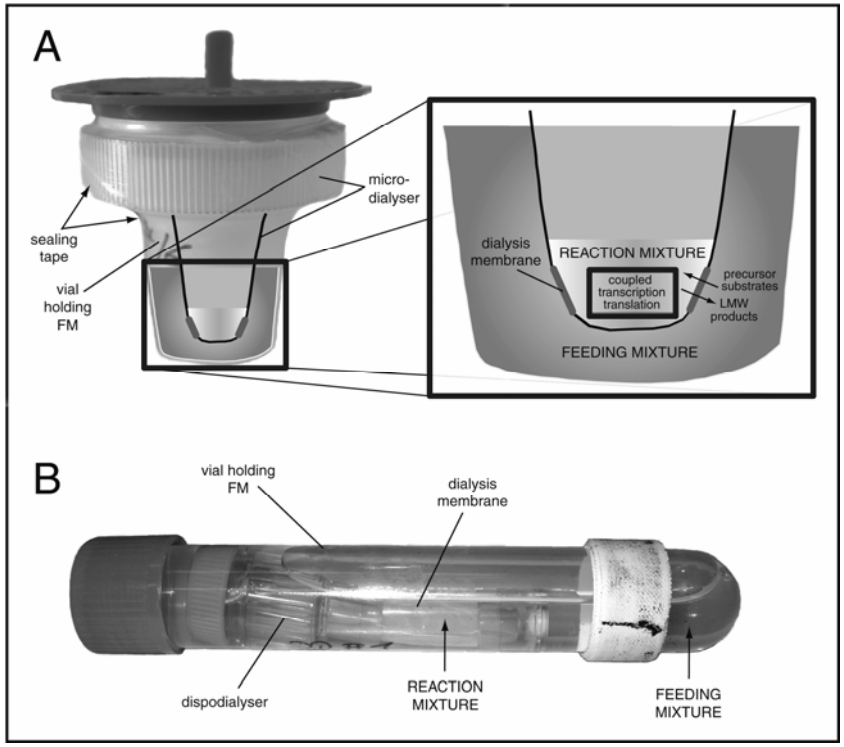


Fig. 3.

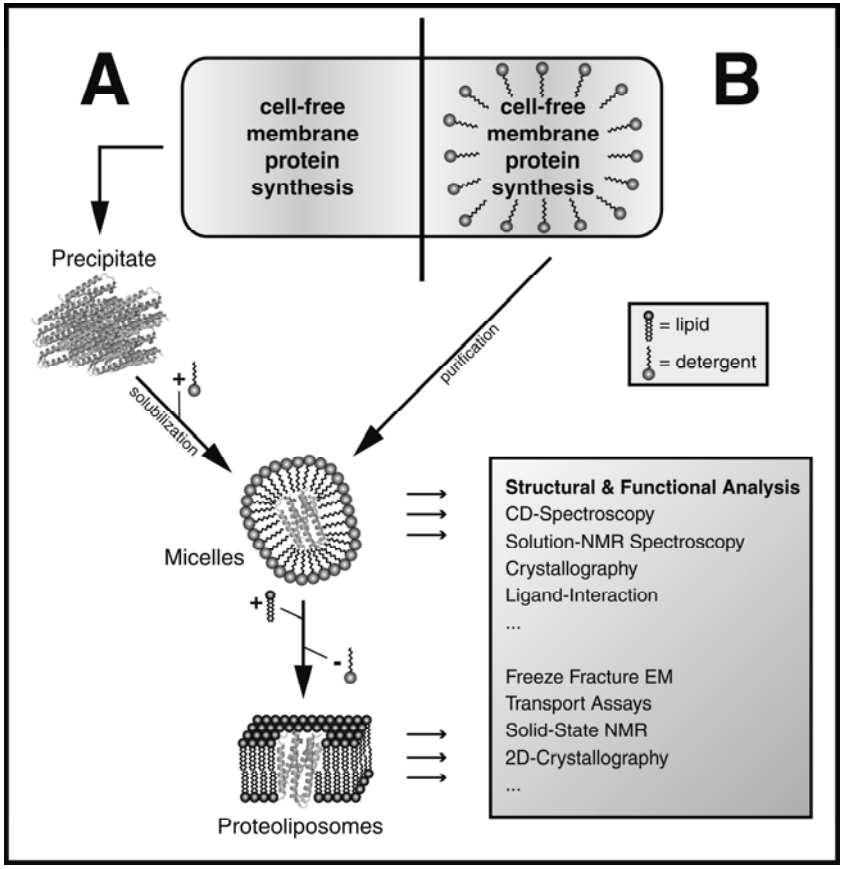

A
B
cell-free membrane protein synthesis
cell-free membrane protein synthesis
Precipitate
= lipid
= detergent
solubilization
purification
Micelles
Proteoliposomes
Structural & Functional Analysis
CD-Spectroscopy
Solution-NMR Spectroscopy
Crystallography
Ligand-Interaction
...
Freeze Fracture EM
Transport Assays
Solid-State NMR
2D-Crystallography
...

Fig. 4.

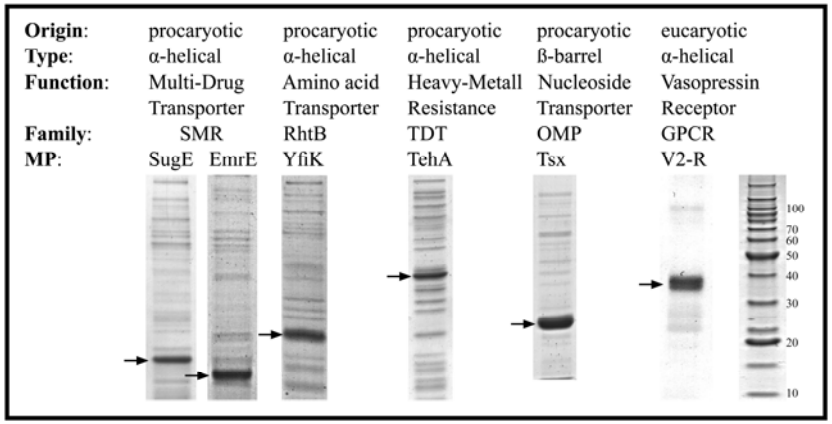

Origin: procaryotic procaryotic procaryotic procaryotic eucaryotic
Type: α-helical α-helical α-helical ß-barrel α-helical
Function: Multi-Drug Transporter Amino acid Transporter Heavy-Metall Resistance Nucleoside Transporter Vasopressin Receptor
Family: SMR RhtB TDT OMP GPCR
MP: SugE EmrE YfiK TehA Tsx V2-R
100
70
60
50
40
30
20
10

Fig. 5.

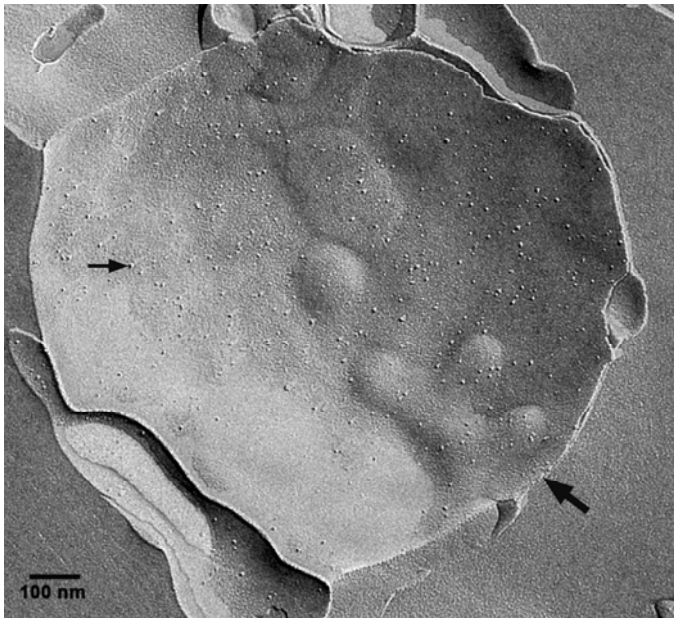

100 nm

Fig. 6.

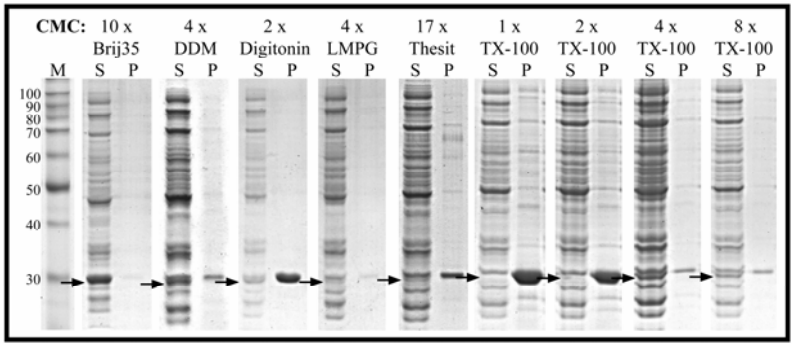

CMC:
10 x Brij35
4 x DDM
2 x Digitonin
4 x LMPG
17 x Thesit
1 x TX-100
2 x TX-100
4 x TX-100
8 x TX-100
M
S P
100
90
80
70
60
50
40
30

Table 1

Protocol for cell-free protein expression

Component	**stock concentration**	**final concentration in RM**
S30-extract / S30-C-buffer	100%	35 %
plasmid DNA	0.3 mg/ml	≥ 15 µg/ml
RNAguard	39.8 U/µl	0.3 U/µl
T7 RNA-polymerase	40 U/µl	≥ 3 U/µl
E. coli tRNA	40 mg/ml	500 µg/ml
pyruvate kinase	10 mg/ml	40 µg/ml
		in RM + FM
amino acids	4 mM / 16.7 mM	≈ 1 mM
acetyl phosphate	1 M	20 mM
phosphoenol-pyruvic-acid	1 M	20 mM
ATP	360 mM	1.2 mM
CTP, GTP, UTP	240 mM	0.8 mM each
1.4-dithiothreitol (DTT)	500 mM	2 mM
folinic acid	20 mM	0.2 mM
complete protease inhibitor	50 x (1 tablet /0.2 ml)	1 tablet /10 ml
HEPES-KOH pH 8.0	2.5 M	100 mM
magnesium acetate	1 M	≈ 14 mM
potassium acetate	4 M	≈ 290 mM
polyethylenglycol 8000	40 %	2 %
sodium azide	10 %	0.05 %

Note: Concentrations of Mg^{2+} and K^{+} are highly critical and should be subject of optimization. 9.1 mM magnesium acetate and 150.8 mM potassium acetate are added, 4.9 mM Mg^{2+} results from the S30-extract and 139.2 mM K^{+} results from other reaction components. Amino acids and T7-RNA-polymerase are limiting compounds and only minimal concentrations are given. The standard amino acid concentrations according to the protocol will be 0.5 mM in the RM and 1 mM in the FM.

Table 2

Pipetting protocol for a 1 ml cell-free reaction

Stock solution	master-mix FRM
10% NaN_3	92 µl
40% PEG8000	918 µl
4 M KOAc	692 µl
1 M $Mg(OAc)_2$	167 µl
25x Buffer	646 µl
50x Complete + EDTA	367 µl
20 mM folinic acid	184 µl
0.5 M DTT	73 µl
75x NTP-mix	245 µl
1 M PEP	367 µl
1 M AcP	367 µl
4 mM AA-mix	2295 µl
16.7 mM RCWMDE-mix	1099 µl
	7513 µl

	FM	RM
Master-mix FRM	6957 µl	409 µl
4 mM AA-Mix	2125 µl	-
S30-C-buffer	5950 µl	-
10 mg/ml pyruvate kinase	-	4 µl
40 mg/ml *E.coli* tRNA	-	13 µl
40 U/µl T7-RNA-Polym.	-	75 µl
39.8 U/µl RNAsin	-	8 µl
S30-Extract	-	350 µl
0.3 mg/ml Plasmid-DNA	-	50 µl
H_2O	1968 µl	91 µl
	17000 µl	1000 µl

Note: All volumes less then one µl have been rounded up.

Table 3

Detergents used for the solubilization of cell-free expressed MP's

Detergent	nature	mass (Da)	CMC (mM)	concentration (x CMC)	solubilization of precipitates	soluble MP's Em	Tsx	V2R
SDS	A	288.38	2.6	2	+++	0	0	0
Tween 20	N	1228	0.059	10	+	0	0	0
CHAPS	Z	614.9	8.0	1.5	±	0	I	0
ß-OG	N	292.4	19	2	±	I	I	0
DPC	Z	351.5	1.5	1	++	I	0	I
Thesit	N	583	0.1	17	+	I	I	I
LMPG	A	478.5	0.05	4	+++	I	I	I
Triton X-100	N	647	0.23	15	+	I	II	I
DDM	N	348.5	0.19	10	+	II	II	I
Digitonin	N	1229.31	0.73	2	+	II	II	II
Brij-35	N	1199.57	0.08	10	+	III	III	III

Note: The highest concentrations (x CMC) that have been added into the CF reaction are shown. For the solubilization of precipitates, considerably higher concentrations can be used. A: anionic; N: nonionic; Z: zwitterionic; CMC: critical micellar concentration; The production of soluble MP is compared with the yields obtained from the production as precipitates in standard CF reactions without detergent. 0: no soluble MP production; I: low level soluble MP production ≤ 20 % ; II: medium level soluble MP production between 20 % and 50 %; III: high level soluble MP production ≥ 50 %. Em: *E. coli* α-helical multidrug transporter EmrE; Tsx: *E. coli* ß-barrel transporter Tsx; V2R: porcine vasopressin receptor;

(P 04)

Cell-free expression as an emerging technique for the large scale production of integral membrane protein. C. Klammt, D. Schwarz, F. Löhr, B. Schneider, V. Dötsch and F. Bernhard. *FEBS Journal*, **273**, 4141-4153. (2006)

the FEBS Journal

MINIREVIEW

Cell-free expression as an emerging technique for the large scale production of integral membrane protein

Christian Klammt, Daniel Schwarz, Frank Löhr, Birgit Schneider, Volker Dötsch and Frank Bernhard

Centre for Biomolecular Magnetic Resonance, University of Frankfurt/Main, Institute for Biophysical Chemistry, Frankfurt/Main, Germany

Keywords
cell-free expression; G-protein coupled receptors; integral membrane proteins; NMR spectroscopy; small multidrug transporter; X-ray crystallography

Correspondence
F. Bernhard, Centre for Biomolecular Magnetic Resonance, Institute for Biophysical Chemistry, University of Frankfurt/Main, Max-von-Laue-Str. 9, D-60438 Frankfurt/Main, Germany
Fax: +49 69 798 29632
Tel: +49 69 798 29620
E-mail: fbern@bpc.uni-frankfurt.de

(Received 12 May 2006, revised 5 July 2006, accepted 11 July 2006)

doi:10.1111/j.1742-4658.2006.05432.x

Membrane proteins are highly underrepresented in structural data banks due to tremendous difficulties that occur upon approaching their structural analysis. Inefficient sample preparation from conventional cellular expression systems is in many cases the first major bottleneck. Preparative scale cell-free expression has now become an emerging alternative tool for the high level production of integral membrane proteins. Many toxic effects attributed to the overproduction of recombinant proteins are eliminated by cell-free expression as viable host cells are no longer required. A unique characteristic is the open nature of cell-free systems that offers a variety of options to manipulate the reaction conditions in order to protect or to stabilize the synthesized recombinant proteins. Detergents or lipids can easily be supplemented and membrane proteins can therefore be synthesized directly into a defined hydrophobic environment of choice that permits solubility and allows the functional folding of the proteins. Alternatively, cell-free produced precipitates of membrane proteins can efficiently be solubilized in mild detergents after expression. Highly valuable for structural approaches is the fast and efficient cell-free production of uniformly or specifically labeled proteins. A considerable number of membrane proteins from diverse families like prokaryotic small multidrug transporters or eukaryotic G-protein coupled receptors have been produced in cell-free systems in high amounts and in functionally active forms. We will give an overview about the current state of the art of this new approach with special emphasis on technical aspects as well as on the functional and structural characterization of cell-free produced membrane proteins.

Introduction

Integral membrane proteins (MPs) play key roles in numerous human diseases and they currently represent one of the most prevalent drug targets. However, low expression levels, inefficient purification protocols and problematic handling procedures have so far significantly hampered the functional and structural analysis

Abbreviations
β2AR, human β2 adrenergic receptor; Brij-35, polyoxyethylene-(23)-lauryl-ether; Brij-58, polyoxyethylene-(20)-cetyl-ether; Brij-78, polyoxyethylene-(20)-stearyl-ether; Brij-98, polyoxyethylene-(20)-oleyl-ether; β-OG, *n*-octyl-β-glucopyranoside; CECF, continuous exchange cell-free; CF, cell-free; CMC, critical micellar concentration; C_{mic}, micellar concentration; CRF, rat corticotropin releasing factor 1 receptor precursor; DHA, dihydroalprenolol; DHPC, 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine; diC_8PC, 1,2-dioctanoyl-*sn*-glycero-3-phosphocholine; DM, *n*-decyl-β-maltoside; DDM, *n*-dodecyl-β-D-maltoside; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; DPC, dodecyl-phosphocholine; ETB, human endothelin B receptor precursor; FM, feeding mixture; GPCR, G-protein coupled receptor; HT, high-throughput; LMPG, 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)]; LPPG, 1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)]; M2, human muscarinic acetylcholine receptor M2; MP, membrane protein; NTPs, nucleotide triphosphates; NTR, rat neurotensin receptor; RM, reaction mixture; SMR, small multidrug resistance family; TMS, transmembrane segment; TROSY, transverse relaxation optimized spectroscopy; TX-100, polyethylene glycol P-1,1,3,3-tetramethyl-butylphenyl-ether; V2R, vasopressin receptor type 2.

of this extremely difficult class of proteins. Membrane integration of recombinant MPs upon their production in conventional cell culture systems of bacterial, yeast, mammalian or insect origin often affects the integrity of the cellular membranes resulting in growth retardation or even lysis of the host cells. Blocking of cellular transport and post-translational processing systems by the overproduced heterologous MPs may cause further toxic effects. Examples where MPs have been produced in preparative amounts are therefore relatively rare and even frequently associated with the aggregation and inactivation of the recombinant MPs by inclusion body formation. As a consequence, only some 80 different MP structures are currently deposited in the protein data banks in contrast to several thousands of nonmembrane-associated proteins [1].

Protein production by cell-free (CF) expression techniques does not depend on cellular integrity and could therefore provide a general advantage for the synthesis of problematic MPs. CF expression has been used since the early 1950s as a single compartment batch system with only very low productivity of recombinant proteins in the nanogram or microgram scale [2–4]. The lifetime of batch reactions was limited mainly by the fast consumption of precursors and also by inhibition of the translation process through rapidly accumulating breakdown products. The modification of CF expression systems by separation of the reaction chamber into two compartments resulted in considerably higher production rates reaching now preparative scales [5–7]. All high molecular mass compounds of the transcription and translation machinery in the reaction mixture (RM) are separated by a semipermeable membrane from a feeding mixture (FM) that defines a fixed volume reservoir of low molecular mass precursors. The new design ensures the removal of undesired breakdown products from the RM concomitant with a continuous supply of fresh precursors and energy substrates from the FM. This continuous exchange setup of a CF system (continuous exchange cell-free; CECF) extends protein synthesis for several hours and is able to produce several mg of recombinant protein per mL of RM [8]. The final yield of recombinant protein is furthermore determined by the volume ratio of RM and FM that is usually between 1 : 10 and 1 : 30. Refreshing the FM during the reaction could give an additional increase in productivity.

Two main sources of CF extracts have been established; wheat germs [9–12] and *Escherichia coli* cells [13–16]. The stability of eukaryotic wheat germ extracts offers extended reaction times up to several days. The very low endogenous RNAse activity allows their use in translation systems with purified mRNA as a template. In contrast, prokaryotic *E. coli* extracts are generally used in coupled transcription/translation systems with double-stranded DNA as a template. Both systems are comparable with respect to their productivities [17].

This minireview summarizes the current knowledge on the preparative scale CF expression of MPs and emphasizes on their production in the absence as well as in the presence of detergents. We will give an overview on reported examples of CF produced MPs and we further discuss effects on protein activity that correlate with the mode of CF expression. Finally we will point out new applications that could become feasible by the CF expression of MPs.

Cell-free expression of membrane proteins

For the preparative scale expression of MPs, mostly CF systems based on *E. coli* extracts have been used so far. Individual [18–20] as well as commercially available CF systems like the RTS^{TM} system (Roche Diagnostics, Penzberg, Germany) [21–23] or the Expressway^{TM} Milligram system (Invitrogen, Carlsbad, CA, USA) [24] can give satisfying results. An important advantage of individual systems is the detailed knowledge of the reaction protocol which is an essential prerequisite for rationally designed optimization strategies. Basic parameters for optimization are ion concentrations, quality and concentration of the DNA template, reducing conditions, the composition of the energy regenerating system and concentrations of amino acids. Subject of target specific optimizations are supplemented detergents, lipids and other potentially beneficial compounds like putative cofactors, inhibitors or substrates. It should be realized that high level CF expression of a particular MP with yields of up to several mg of protein per mL RM often has to be preceded by intensive optimization screens of several individual parameters starting with only a few micrograms of recombinant protein per mL RM. Effective detection systems like immunoblotting with tag-specific antibodies or radioactive labeling might therefore be necessary for the evaluation of first expression screens.

Individual *E. coli* systems are based on crude cell lysates supplemented with optimized buffer and salt conditions and all additional components necessary for transcription and translation. Extensive protocols for the preparation of *E. coli* lysates have been published [15,20,25–28]. An efficient energy regenerating system based on high energy phosphate donors like phosphoenol pyruvate, acetyl phosphate or creatine phosphate

in combination with the corresponding kinases ensures the recycling of hydrolyzed nucleotide triphosphates (NTPs). Transcription in the *E. coli* lysate depends on the strong T7 promotor and is facilitated by addition of purified T7 RNA polymerase, NTPs and DNA templates, either in the form of plasmid DNA or as linear PCR products [29,30]. Furthermore, the translation process requires the addition of *E. coli* tRNA mixtures and of the 20 amino acids. Finally, the CF expression system has to be stabilized by addition of protease- and RNAse-inhibitors. Several detailed protocols for individual CF expression systems are available [13–15,20].

MPs can be produced in two different modes by CF expression: As precipitate or as soluble protein in presence of detergents (Fig. 1). Both modes can result in functionally active MPs after subsequent solubilization and reconstitution procedures and the optimal expression strategy should be analysed for each new MP target. Production as precipitate might give higher yields and it facilitates downstream purification processes while the soluble expression is obviously more likely to result instantly into functionally folded MPs. However, it should be considered that due to the relatively low number of analysed MPs still only preliminary conclusions can be made.

Expression of membrane proteins as precipitate

Several diverse MPs of prokaryotic or eukaryotic origin could be produced in CF systems supplemented with *E. coli* extracts, while their synthesis in *E. coli* cells was much lower or even absent [18,20]. We propose that the elimination of the described negative effects upon *in vivo* expression of MPs like membrane insertion, blocking of transport systems and others might for the most part contribute to this observation. Hydrophobic environments like lipids or detergents are essential for the solubility and proper folding of MPs. Standard CF systems lack such hydrophobic compartments and the synthesized recombinant MPs therefore almost quantitatively precipitate in the RM (Fig. 2) [18,20–23]. After the reaction, the precipitated MPs are harvested from the RM by centrifugation. CF produced precipitates of many MPs solubilize quickly and effectively in mild detergents during short incubation with gentle shaking for a few minutes. This relatively fast solubilization procedure distinguishes CF produced MP precipitates from inclusion bodies that can often be observed after *in vivo* expression of MPs [31]. Solubilization of inclusion bodies usually requires extensive denaturation steps involving SDS or high concentrations of urea. In practice, the CF generated MP pellets are suspended in a suitable buffer supplemented with $\approx$1–2% of an appropriate detergent and in a volume that corresponds to the volume of the RM. The suspension is incubated between 20 °C and 30 °C on a shaker for $\approx$1 h and residual precipitate is removed by centrifugation for 15 min at 20 000 ***g***. The efficiency of solubilization of a particular MP depends on the detergent type, and coprecipitated impurities can be removed by subsequent washing steps using detergents that do not dissolve the target MP. It can further be effective to increase the incubation temperature in order to selectively eliminate impurities [20].

Various mild detergents of different classes like alkyl glycosides, steroid derivatives, long chain phosphoglycerols, phosphocholines, polyoxyethylene alkylethers or polyethylene glycols have been tested for their suitability to solubilize CF produced precipitates of even structurally very different MPs [19]. The detergent 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LMPG) and closely related derivatives like 1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phosphor-rac-(1-glycerol)] (LPPG) were found to be most useful for general applications. Pellets of the bacterial α-helical multidrug transporter EmrE, the β-barrel type nucleoside transporter Tsx, and the porcine vasopressin receptor V2R of the eukaryotic family of G-protein coupled receptors (GPCRs) could almost be completely solubilized by LMPG, while other mild detergents like *n*-dodecyl-β-D-maltoside (DDM) or dodecyl-phosphocholine (DPC) are only efficient in the solubilization of some of the analysed MP pellets [19]. A detergent exchange after solubilization by standard procedures might be considered if particular and different detergents are required for the further downstream analysis of the MP.

The CF expression in the absence of any supplemented detergents or lipids is a reliable technique to direct almost all of the synthesized MP into precipitate. Certainly, the precipitation could cause the partial or complete unfolding of a MP and it should be considered that an efficient functional refolding into its native conformation upon solubilization might not always occur. We could reconstitute several CF produced and solubilized MPs like the bacterial multidrug transporters EmrE and SugE, the tellurite transporter TehA, the cysteine exporter Yfik and the GPCR member V2R into liposomes and freeze-fracture analysis revealed homogenously distributed particles giving evidence of a functional reconstitution [19,20]. The bacterial multidrug transporter EmrE showed specific transport activity with its cognate substrate ethidium bromide after CF production as a precipitate,

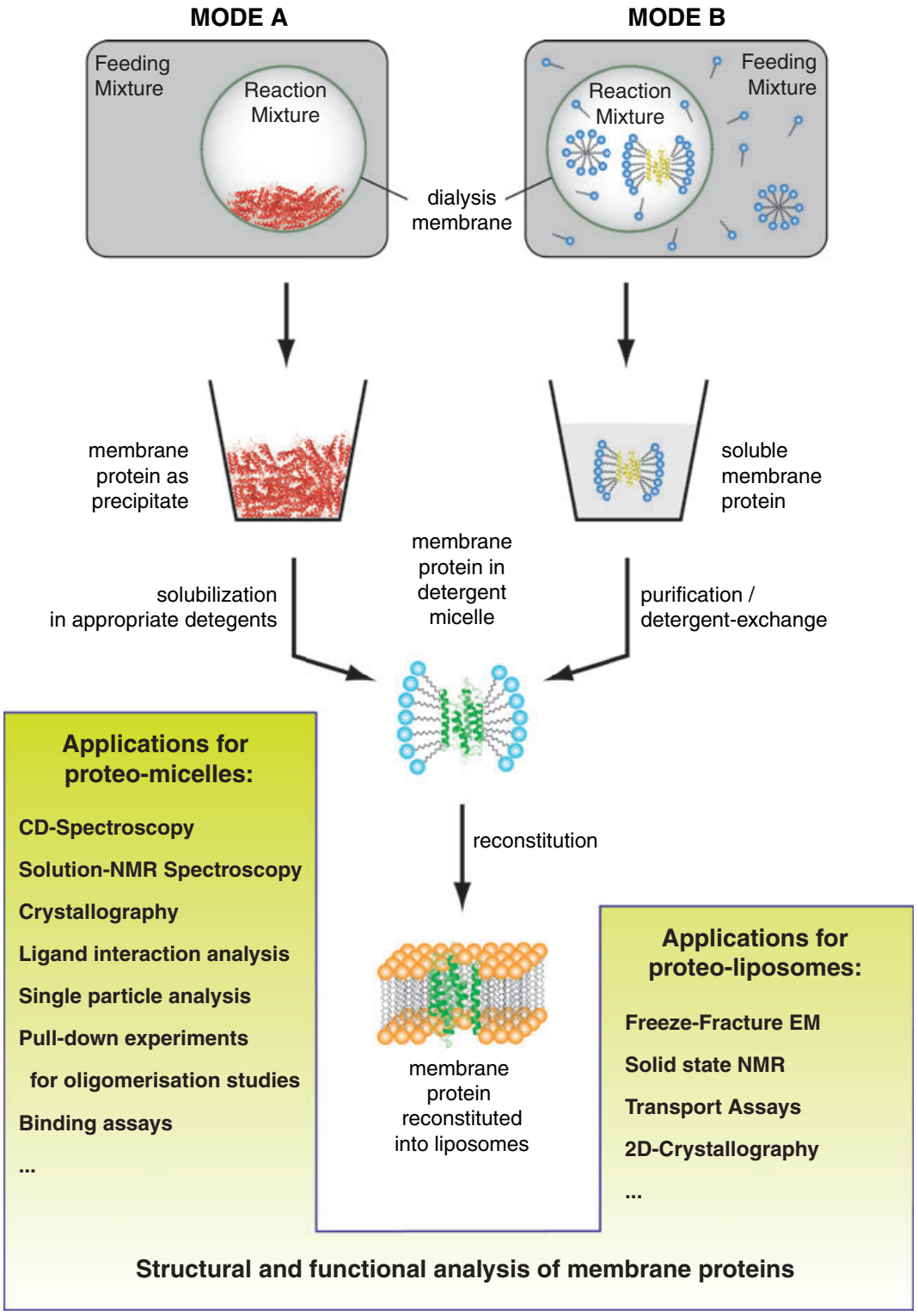


Fig. 1. Different modes for CF expression of MPs. MPs can be expressed as precipitate (Mode A), solubilized in an appropriate detergent and further reconstituted into proteoliposomes, or they can be soluble expressed in the presence of detergent (Mode B), purified and reconstituted into proteoliposomes. EM, electron microscopy.

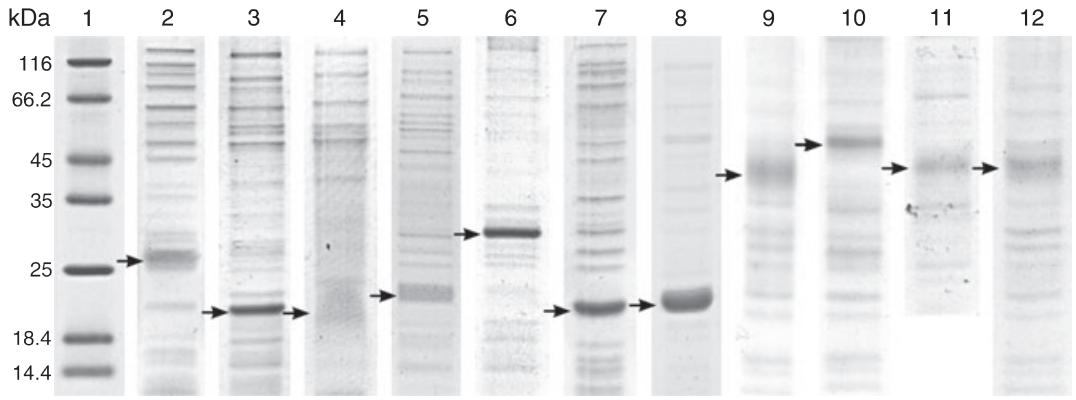


Fig. 2. CF production of different MPs as precipitates. The nonsoluble part of the RM was resuspended in appropriate buffer in a volume equal to the RM and 1 μL was analyzed on 10% tricine/SDS (lanes 2–6), 16.5% SDS (lane 7) or 12% SDS gels (lanes 8–12). The MPs are indicated by arrows. Lane 1, marker, 14.4, 18.4, 25, 35, 45, 66.2, 116 kDa; lane 2, EmrE; lane 3, SugE; lane 4, Tbsmr; lane 5, YfiK; lane 6, TehA; lane 7, ΔTehA; lane 8, Tsx; lane 9, V2R (porcine); lane 10, V2R (human); lane 11, CRF; lane 12, human endothelin B receptor precursor (ETB).

solubilization in DPC and reconstitution into *E. coli* lipids [20]. However, on the other hand, we have been unable to functionally reconstitute the β-barrel type *E. coli* outer membrane protein Tsx into membranes after its solubilization in DPC or LMPG from CF produced precipitates [19].

Cell-free expression of membrane proteins in soluble form

The intrinsic open nature of the CF system allows the direct supplementation of almost any desired compound into the RM. Selected detergents or lipids could therefore already be present at the beginning of the reaction in order to provide a suitable hydrophobic environment for the synthesized MPs. This gives nascent polypeptides the opportunity to become inserted into preformed micelles or liposomes during or shortly after translation. However, detergents are interfering nonspecifically with hydrophobic protein regions and it has to be expected that their addition to the RM might generally affect transcriptional and translational processes of the CF system by inactivation of essential components. A total of 24 frequently used detergents of a comprehensive variety have therefore been systematically evaluated with respect to (i) their impact on the general productivity of a CF expression system, (ii) their efficiency to keep synthesized MPs in solution, and (iii) their effects on the activity of the solubilized MPs [19]. Surprisingly, only a few of the analysed detergents like DPC, *n*-octyl-β-D-glucopyranoside (β-OG), the steroid Chaps and the members of the long chain phosphoglycerols LPPG and LMPG significantly reduced the CF protein production [19]. In addition, more aggressive detergents like sodium deoxycholate, sodium cholate and *N*-laurylsarcosine completely inhibit CF MP synthesis [18,23].

A considerable variety of mild detergents is suitable for the CF expression of soluble MPs in preparative scales. Most useful is the group of long chain polyoxyethylene-alkyl-ethers. Polyoxyethylene-(23)-lauryl-ether (Brij-35), polyoxyethylene-(20)-cetyl-ether (Brij-58), polyoxyethylene-(20)-stearyl-ether (Brij-78) and polyoxyethylene-(20)-oleyl-ether (Brij-98) resulted in high level soluble expression of all analysed MPs. Bacterial α-helical and β-barrel proteins as well as the larger eukaryotic GPCR could be synthesized with these detergents as soluble proteins in amounts of 3–6 mg per mL RM [19]. It is noteworthy that the corresponding short chain derivatives with less than 10 polyoxyethylene groups have not been effective at all. The number of polyoxyethylene groups as well as the length of the alkyl moiety is critical and both could influence the solubilization efficiency and the final yield of a specific MP. The optimal Brij-derivative for the solubilization of a specific MP should therefore initially be selected by an optimization screen. A second detergent with a completely different structure but also having a relatively general ability to solubilize MPs during CF expression is the steroid derivative digitonin [18,19].

Smaller MPs like the small multidrug resistance family (SMR) transporter EmrE obviously solubilize in a higher variety of detergents. The alkyl-glucosides *n*-dodecyl-β-D-maltoside (DDM) and *n*-decyl-β-D-maltoside (DM), the bi-chain-phosphocholines 1,2-dioctanoyl-*sn*-glycero-3-phosphocholine (diC_8PC), 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine (DHPC) and 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and the polyethylene-glycol derivative polyethylene-glycol

P-1,1,3,3-tetramethyl-butylphenyl-ether (TX-100) were, in the case of the CF soluble production of EmrE, as efficient as the above described long chain Brij-derivatives [19,24].

Detergents are known to form micellar structures above a specific threshold concentration, the critical micellar concentration (CMC). The CMC and also the aggregation number, i.e., the number of detergent molecules per micelle, is specific to each detergent type, but they also depend on buffer conditions, inserted proteins and other parameters. An important consideration for the optimal production of soluble MPs in a CF expression reaction is that sufficient micelles are provided. The micellar concentration (C_{mic}) should therefore be at least as high as the molar concentration of the MP in order to prevent aggregates or the formation of nonhomogenous proteomicelles. As the amount of detergent that binds to the hydrophobic surfaces of an integral MP remains relatively constant above the CMC, excessive detergent will form additional empty micelles but does not affect the solubility of a MP. Depending on its tolerance by the CF system, high detergent concentrations of several times CMC could be added in order to ensure the possibility of complete and homogenous MP solubilization. Polyoxyethylene-alkyl-ethers like the highly effective Brij-derivatives can be added to more than 100 times CMC [18,19]. However, the tolerated levels of other detergents like DDM or TX-100 are much lower.

Disadvantages of the soluble CF expression of MPs are sometimes lower final yields of recombinant protein if compared with a CF production as precipitate. It should furthermore be considered that the addition of even high concentrations of detergents frequently does not result in the complete soluble expression of MPs, and rather mixtures of soluble and precipitated MPs are produced. Finally it should be realized that the type of supplemented detergent has an impact not only on the yield of the synthesized MP but also on its functional conformation, the specific activity and perhaps even on its probability to form oligomers [19]. The choice of supplemented detergent should therefore be one of the first optimization parameters when the CF expression of a new MP target is anticipated.

Examples of cell-free expressed membrane proteins

Although only a limited number of CF expressed MPs has been analysed so far, they already comprise a considerable diverse pool of proteins of prokaryotic as well as eukaryotic origin (Table 1). A common characteristic is their almost completely integral nature, and the analysed MPs are embedded in cellular membranes with up to 10 transmembrane segments (TMSs). It still remains to be shown whether, e.g., integral MPs with large hydrophilic domains or peripheral MPs with only short membrane anchors can also be produced in a functional form by CF systems.

The vast majority of CF produced MPs of prokaryotic origin belongs to the family of SMR efflux transporters, having four TMSs and a molecular mass of $\approx$12 kDa. A well characterized prototype of this family is the *E. coli* protein EmrE that has been *in vitro* expressed by different groups in individual or commercial CF expression systems (Table 1). In all cases, preparative amounts of protein ranging from 0.2 mg up to several mgs per mL of RM could be achieved. Other CF produced members of the SMR family include SugE and YdgF from *E. coli*, Hsmr from *Halobacterium salinarium*, Tbsmr from *Mycobacterium tuberculosis*, Psmr from *Pseudomonas aeruginosa*, BPsmr from *Bacillus pertussis and* YfbW from *Salmonella typhimurium* [20,22,32]. All proteins could be obtained in preparative amounts suitable for their characterization by solution NMR [20], by functional assays [22], by solid state NMR [32] and even by crystallization [24].

Two small bacterial α-helical MPs, the mechanosensitive channel MscL of *E. coli* and the light harvesting protein α-LH1 of *Rhodospirillum rubrum* have been produced in a commercial CECF system [21,23]. The 14 kDa MscL protein consists of two TMSs while the 6 kDa α-LH1 subunit has only one TMS. The authors could document an activity of the two CF produced MPs that was similar to samples analysed after conventional *in vivo* expression in *E. coli* (Table 2). MscL has been produced soluble in TX-100 micelles [23] while α-LH1 was made as precipitate and solubilized in TX-100 after the reaction [21]. These results demonstrate nicely the versatility of CF expression and show that both expression modes can result in functionally folded protein. NMR spectra of CF produced samples of the 21 kDa cysteine exporter YfiK (Fig. 3), of the 36 kDa tellurite resistance transporter TehA and of its 24 kDa truncated version ΔTehA give evidence of structurally folded proteins [20,33]. The amide proton assignment of ΔTehA could be completed by using rationally designed combinatorial labeling schemes that are based on the CF production of labeled protein samples [33].

The family of eukaryotic GPCRs currently attracts considerable attention as potential drug targets for the pharmaceutical industry. GPCRs represent the largest single class of receptors and they play key roles in many signaling pathways, cellular recognition

Table 1. Membrane proteins synthesized in preparative scale CF expression systems. Type refers to main secondary structure element, either α (α-helical) or β (β-sheets). Mode of expression is either as precipitate (P) or in the presence of different detergents. Reaction is Batch (individual batch system set-up), CECF, RTS-100 [commercial batch system (Roche Diagnostics)], RTS-500 [commercial CECF system (Roche Diagnostics)] or Expressway [commercial batch system (Invitogen)]. TrxA-M2, thioredoxin fused human muscarinic acetylcholine receptor M2; TrxA-M2-Gi1α, TrxA-M2 fused with the $G_{i1}\alpha$ subunit of a G-protein; TrxA-NTR, thioredoxin fused rat neurotensin receptor; TrxA-β2AR, thioredoxin fused human β2-adrenergic receptor; TrxA-β2AR-$G_s\alpha$, TrxA-β2AR fused to the $G_s\alpha$ subunit of a G-protein. DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; TDT, tellurite-resistance/dicarboxylate transporter family; RhtB, resistance to homoserine/threonine family; n.p., not published; n.r., not reported.

Protein	TMS	*m* [kDa]	Family	Type	Reaction	Expression mode	Yield [mg·mL^{-1}]	Ref.
Eukaryotic								
CRF	7	48	GPCR	α	CECF	P	3	n. p.
ETB	7	52	GPCR	α	CECF	P	3	n. p.
TrxA-M2	7	45	GPCR	α	Batch	P	0.17	[18]
						Brij-35, digitonin	0.15	[18]
TrxA-M2-$G_{i1}\alpha$	7	95	GPCR	α	Batch	P	0.17	[18]
TrxA-NTR	7	52	GPCR	α	Batch	P	0.18	[18]
						Brij-35, digitonin	0.15	[18]
TrxA-β2AR	7	66	GPCR	α	Batch	P	0.16	[18]
						Brij-35, digitonin	0.16	[18]
TrxA-β2AR-$G_s\alpha$	7	105	GPCR	α	CECF	Brij-35, digitonin	1	[18]
					Batch	P	0.15	[18]
V2Rh	7	43	GPCR	α	CECF	P	3	n. p.
V2Rp	7	43	GPCR	α	CECF	P, Brij-58, -78	6	[19]
						Brij-98	2.9	[19]
						Brij-35	0.9	[19]
						Digitonin	0.38	[19]
Prokaryotic								
BPsmr	4	12	SMR	α	RTS-100	P	n. r.	[22]
EmrE	4	12	SMR	α	CECF	Brij-35, -98,	2.7	[19]
					CECF	P	3	[19]
					RTS-500	P, DDM	3	[22]
					CECF	P	1	[20]
					Expressway	DMPC	0.2	[24]
Hsmr	4	15	SMR	α	CECF	P	n. r.	[32]
Psmr	4	12	SMR	α	RTS-100	P	n. r.	[22]
SugE	4	11	SMR	α	CECF	P	1.5	[20]
Tbsmr	4	11	SMR	α	CECF	P	n. r.	[32]
					RTS-100	P	n. r.	[22]
YdgF	4		SMR	α	CECF	P	n. r.	[32]
Yfbw	4		SMR	α	CECF	P	n. r.	[32]
TehA	10	36	TDT	α	CECF	P	2.7	[20]
△TehA	7	24	TDT	α	CECF	P	3	[33]
α-LH1	1	6.1	LHP	α	RTS-500	P	0.7	[21]
MscL	2	14	MscL	α	RTS-500	TX-100	3.6	[23]
YfiK	6	21	RhtB	α	CECF	P	1	[20]
Tsx	–	31	OMP	β	CECF	P	4	[19]
						Brij-35, -58, -78	3	[19]
						DDM	1.7	[19]
						Digitonin, TX-100	4	[19]

processes and cell–cell communications. Despite numerous attempts, only the three-dimensional structure of bovine rhodopsin has been solved so far [34,35]. GPCRs have a predicted similar topology, with seven TMSs interconnected by loop regions. Several GPCRs of human and other mammalian origin like the human β2 adrenergic receptor (β2AR), the rat muscarinic acetylcholine receptor M2 (M2) and the human neurotensin receptor (NTR) could be synthesized in individual CF systems at relatively high yields [18] (Table 1). While the expression of the wildtype GPCRs was found to be inefficient, the construction of translational fusions to the C-terminal end of the 11.7 kDa *E. coli* thioredoxin (TrxA) tremendously

Table 2. Activity and yield of MPs synthesized in different expression systems. Comparison of activities and expression levels of CF and *in vivo* expressed MPs. Trx-β2AR-$G_s\alpha$, TrxA-β2AR fused to the $G_s\alpha$ subunit of a G-protein; [^{3}H]DHA, dihydroalprenolol; ND, not determined; EC_{50}, effective concentration to reach 50% of binding activity.

	Activity			Expression levels		
Protein	Assay conditions	CF	*in vivo*	CF [mg·mL^{-1}]	*in vivo* [mg·L^{-1}]	Ref.
EmrE	[^{3}H]-TPP binding in DDM	K_d = 2.3 nM	K_d = 2.8 nM	3	1	[19,22,48,49]
	[^{14}C]-Methylviologen transport in proteoliposomes	7.5 nmol·min^{-1} per μg EmrE	5.5 nmol·min^{-1} pe r μg EmrE			[22,49]
	Ethidium transport into proteoliposomes	Specific transport	Specific transport			[20,50]
α-LH1	Specific absorption spectrum	Maximum at 820 nm	Maximum at 820 nm	0.7	1.2	[21]
MscL	Channel opening in proteoliposomes	8.3 ± 1.8 open channels per patch	9.3 ± 5.5 open channels per patch	3.6	1	[23,51]
Trx-β2AR-$G_s\alpha$	[^{3}H]DHA binding affinity in proteoliposomes	K_d = 4.3 nM	K_d = 3.9 nM	1	ND	[18]
	Competition study with antagonist alprenolol [one-site (1), two-site (2) competition]	EC_{50} = 2.2 × 10^{-8} M (1) 1.0 × 10^{-5} M (2)	EC_{50} = 1.0 × 10^{-8} M (1) 1.0 × 10^{-8} M (2)			
Tsx	Black lipid membrane assay with proteoliposomes	Specific transport	Specific transport	4	0.25	[19,52]

enhanced the expression levels in all three cases. Final yields of several 100 μg·mL^{-1} in batch systems and more than 1 mg·mL^{-1} in a commercial CECF system were produced (Table 1). Those expression levels could even be obtained with a 108 kDa construct of β2AR that was fused at its N-terminus to TrxA and at its C-terminal to the $G_s\alpha$-protein [18]. The relatively large TrxA fusion was essential in order to increase the translation efficiency of GPCRs, but it can obviously be replaced by short N-terminal tags like the T7 tag. The porcine vasopressin receptor (V2R) was efficiently produced in an individual CECF containing only the short T7 tag with 14 amino acid residues [19]. Similar observations were made by our group upon the CF expression of the human V2R, the endothelin B receptor (ETB) and the corticotropin releasing factor 1 precursor receptor (CRF). All three GPCRs were synthesized in an individual CECF system with only an N-terminal T7 tag in amounts of ≈3 mg per mL RM (Table 1). In accordance with the results obtained upon the CF expression of the GPCRs β2AR, M2 and NTR this indicates that an efficient initiation of translation might not occur with the wildtype GPCR sequences. Only the steroid digitonin and long chain Brij-derivatives have been suitable for the soluble CF expression of mammalian GPCRs at elevated levels and after optimization of the detergent concentration up to 6 mg of soluble V2R could be obtained [19].

The 31 kDa nucleoside transporter Tsx of *E. coli* is so far the only example of a CF produced outer membrane protein (OMP) having a β-barrel type structure [19]. Tsx is synthesized in both CF expression modes, either as a precipitate or as soluble protein in the presence of detergents, in final amounts of up to 4 mg·mL^{-1} in an individual CECF system (Table 1). The soluble Tsx expression was quite tolerant for a wide range of detergents and besides digitonin and Brij-derivatives, also TX-100 and DDM were highly effective (Table 1). The Tsx transporter is an example of where the CF expression conditions might have a significant impact on the functional folding of the recombinant MP. CF produced Tsx precipitate that was solubilized in detergent micelles was not functionally active in specific transport assays, in contrast to Tsx protein produced in the soluble mode of CF expression (Table 2) [19]. Despite its efficient solubilization in mild detergents, the precipitated Tsx might have problems recovering its fully functional conformation. In addition, remarkable differences in the specific activity of soluble CF produced Tsx were noticed in dependency to the supplemented detergent. The functionally active conformation of Tsx might therefore be relatively susceptible for its hydrophobic environment, and a clear preference was

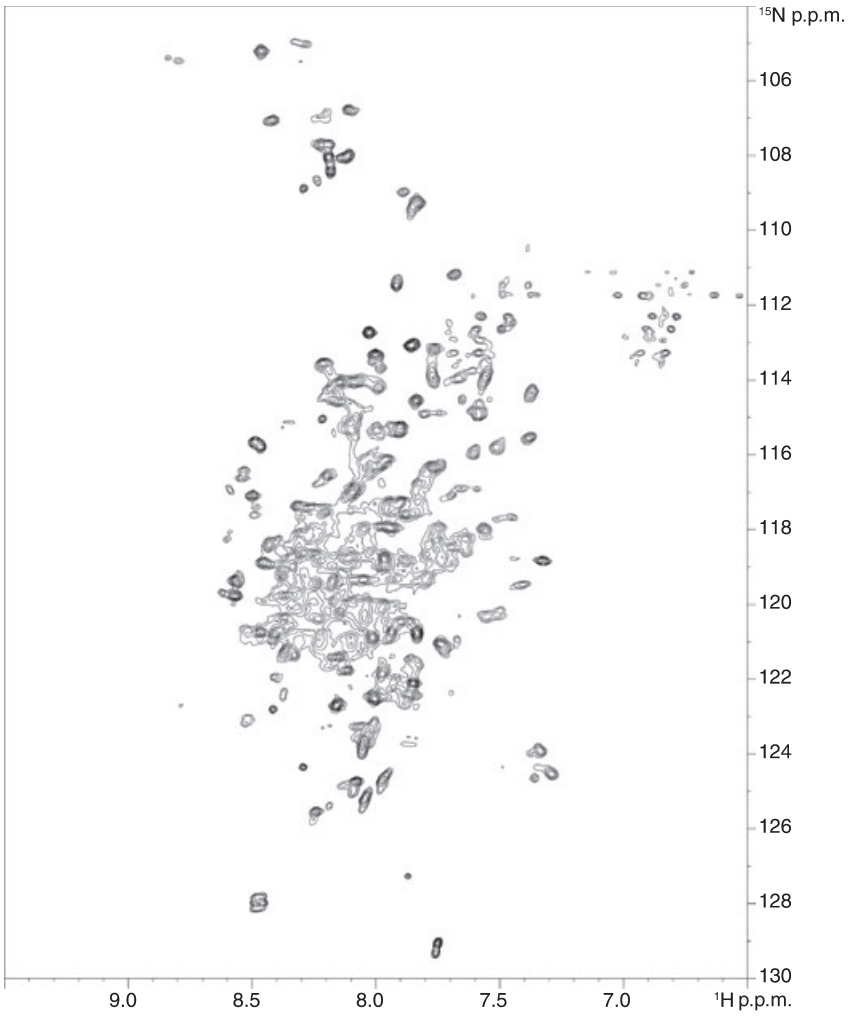


Fig. 3. ^{1}H-^{15}N TROSY spectrum of CF expressed U-^{2}H-^{15}N-labeled YfiK. The protein was expressed in the insoluble mode as precipitate and resolubilized in LMPG. The protein was dissolved in 25 mM potassium phosphate buffer (pH 6.0) containing 5% LMPG. The spectrum was measured with a protein concentration of 0.3 mM at 29 °C on a 900 Mhz NMR spectrometer (Bruker, Rheinstetten, Germany) equipped with a cryo probe with eight scans per free induction decay and 640 increments in the indirect dimension.

found for TX-100 [19]. Optimization strategies for CF expression should therefore focus not only on highest yields but also on conditions that ensure a high specific activity of the synthesized MPs.

Specific applications for cell-free produced membrane proteins

Structural analysis of CF synthesized MPs

CF expression systems could help to accelerate the production of MPs of diverse structure and origin, especially for the purpose of structural characterization. The two main directions of high resolution structural analysis, X-ray crystallography as well as solution NMR spectroscopy, could considerably benefit from the advantages of CF expression. Both techniques require protein samples that have been specifically labeled either with seleno-methionine or with stable isotopes. In conventional *in vivo* expression systems those labeling procedures often require auxotrophic mutants of host cells and fermentation protocols based on defined minimal media. Frequent consequences are growth retardation, significantly decreased productivities of the labeled proteins, incomplete label incorporation and scrambling problems due to metabolic conversion of the label precursors. Time-consuming and expensive optimization runs therefore have to be performed before final preparations can be started. Those problems have finally prevented structural approaches in many cases of MPs that are already difficult to produce at normal conditions.

The open nature of CF systems in combination with the complete control over the amino acid pool allows the production of labeled MP samples as fast and as efficiently as the production of nonlabeled proteins. Nonlabeled amino acids in the CF mixtures are simply exchanged against the labeled derivatives and no other changes of conditions are required. This ensures a constant yield of the recombinant MP concomitant with

100% label incorporation at the desired positions. This characteristic defines a precious advantage especially for the purpose of NMR analysis. Moreover, CF production provides significantly reduced scrambling or background labeling problems [8,13,14,26,36,37]. It is even possible to analyse the CF produced labeled protein directly in the RM by NMR without prior purification [38]. This is especially important for MPs where downstream purification processes can result in tremendous losses of protein. MP samples ready for NMR analysis can thus be prepared virtually in less than 24 h.

Solution NMR studies of MPs are still an exception due to the so far limited availability of labeled samples in combination with the required size restriction of the analysed proteins. Proteins exceeding molecular masses of $\approx$20 kDa increasingly generate signals of low dispersion with significant spectral overlaps that retard or even prevent a structural analysis by solution NMR. The α-helical nature of many MPs and the resulting concentration of most signals in the range between 7.5 p.p.m. and 8.5 p.p.m. often further reduce the size limits. The insertion of MPs into detergent micelles usually dramatically increases the MP–detergent complex size and thus considerably influences the quality of NMR spectra. The few available NMR structures of MPs have therefore been recorded mostly from β-barrel type outer membrane proteins. Fortunately, recent screens for NMR suitable detergents helped to identify the group of lyso-phosphoglycerols as highly promising detergents that do not restrict the spectral quality of solubilized MPs [39]. The detergent LMPG that was found to be superior for the solubilization of CF produced precipitates of MPs belongs to this group and NMR spectra of [^{15}N]-labeled samples of the relatively large *E. coli* α-helical MPs YfiK (Fig. 3) and TehA in LMPG micelles showed resolutions that are very promising for structural approaches by solution NMR [20,33].

A first example of the NMR analysis of a CF produced MP is the putative multidrug resistance and tellurite resistance transporter, TehA, of *E. coli*. The 36 kDa wildtype TehA with 10 TMSs as well as a C-terminal truncated functionally active version ΔTehA of 24 kDa and with seven TMSs can be produced in an individual CF system in amounts of 2–3 mg per mL RM [20,33]. The structural analysis of ΔTehA has been approached by solution NMR as a first model system of a CF produced relatively large α-helical integral MP. The almost complete backbone assignment of ΔTehA has been obtained using the explicit advantages of CF expression. While only $\approx$40% of the backbone amide protons could be assigned based on uniformly isotopically labeled ΔTehA measured in LMPG micelles, the final assignment of the residual amide protons was successful by using a combinatorial labeling scheme [33]. Selected sets of amino acid combinations have been labeled either with ^{15}N or ^{13}C by a rational designed strategy. This technique allowed specific amino acid pairs having a [^{15}N]-labeled amino acids N-terminally preceded by a [^{13}C]-labeled amino acid type to be identified, and it subsequently resulted in new anchor points for further sequence-specific assignments. Such combinatorial labeling schemes turned out to be superior in order to obtain unambiguous assignments if compared with single amino acid specific labeling approaches, but they are almost exclusively feasible by using CF synthesis.

Oligomerization studies

The conventional analysis of the oligomeric state of MPs is frequently performed by cross linking experiments of purified protein. CF expression could offer an alternative approach by providing two different DNA templates with one encoding for the wildtype MP and the other for a derivative that has been modified by a translational tag. Plasmids of even identical incompatibility groups could be used as DNA templates as no replication or selection is required. Even linear DNA fragments generated by standard PCR might be applicable. As an example, the CF produced wildtype mutidrug transporter EmrE was successfully pulled down by a poly(His)$_6$ tagged EmrE derivative that had been coexpressed in the same reaction, indicating the formation of EmrE oligomers [22]. Moreover, negative dominance studies with inactive coexpressed EmrE mutants also gave insights into the functionality of the oligomers. An advantage of CF expression is that the ratios of both protein species can easily be manipulated by titration of the corresponding DNA templates. It would further be possible to add already purified proteins that have been isolated from other sources in order to analyse their interaction with the synthesized target protein. This enables accurate quantitative experiments that are independent from the expression rates and proper folding conditions of the putative interaction partners.

High-throughput expression screening of MPs

CF expression has been optimized for direct PCR product expression and high-throughput (HT) applications of soluble proteins [40,41]. Expression screening in 96 well formats [42] as well as in even smaller nanowell chip formats is possible [43]. Instability problems

of linear DNA templates can be addressed by preparing CF extracts from nuclease deficient strains [29,44]. Alternatively, modifications of the mRNA, like the positioning of a stem-loop structure at the 3′ end, the addition of poly(G) tails or mini-hairpin sequences, can substantially help to increase the half-life of the transcripts [30,44,45].

MPs have usually been excluded from common HT proteome expression approaches as they are difficult to produce. However, the combined advances of CF expression systems for HT applications and MP production will open up new potentials for membrane proteomics. The possibility to use linear PCR products as expression templates will be highly valuable as a variety of translational tags could be screened in a relatively short time for potential benefits, without the need of time-consuming cloning procedures. Furthermore, the option to provide stabilizing agents like protease inhibitors, chaperones, ligands or others will additionally increase the likeability of a detectable protein production. It will also be interesting to analyse the potential of MP production in CF systems based on wheat germ extracts, as folding pathways and potential modification patterns of eukaryotic MPs might become improved in the eukaryotic background. Labeling protocols for NMR-based structural proteomics based on wheat germ extracts that can be extended to an automated platform have already been established [46,47].

Conclusion and future challenges

The reported number of CF produced MPs is still limited and also comprises several well characterized model proteins that can be produced in considerable amounts in *E. coli* cells. However, GPCRs for instance are known as problematic proteins and their high level CF production in functional form gives the first evidence of an immense potential of this new technology for the generation of difficult MP samples. A current restriction is that all analysed MPs so far have a more or less completely membrane-integrated topology with only relatively small hydrophilic loops. More comprehensive expression screenings are therefore needed in order to evaluate the application of CF expression to a wider range of MP families including those that contain larger nonmembrane integrated domains. It should be realized that besides the fast production of MPs that are otherwise very difficult to obtain, the CF expression provides further unique and highly valuable advantages. The destabilization of membranes by conventional purification protocols is often associated with the denaturation of the inserted MPs by transient exposure of hydrophobic surfaces to water. The direct synthesis of MPs into the desired detergents avoids such critical isolation steps. Most important is the fast and efficient labeling of MPs with modified amino acids that is an essential prerequisite for structural studies by NMR spectroscopy as well as by X-ray crystallography. The sole fact that labeled samples ready for analysis can now be obtained in less than 24 h will enable many new approaches for the structural analysis of MPs. Finally, it should be considered that CF expression protocols are currently the subject of a variety of optimization strategies. Usage of wheat germ extracts, efficient disulfide bridge formation, supplementation of lipids and others will help to continuously provide new options in order to improve the quality of CF produced MPs.

Acknowledgements

The work was financially supported by the DFG and the SFB 628 ‘Functional Membrane Proteomics’.

References

1 Raman P, Cherezov V & Caffrey M (2006) The membrane protein data bank. *Cell Mol Life Sci* **63**, 36–51.
2 Winnick T (1950) Studies on the mechanism of protein synthesis in embryonic and tumor tissues. I. Evidence relating to the incorporation of labeled amino acids into protein structure in homogenates. *Arch Biochem* **27**, 65–74.
3 Winnick T (1950) Studies on the mechanism of protein synthesis in embryonic and tumor tissues. II. Inactivation of fetal rat liver homogenates by dialysis, and reactivation by the adenylic acid system. *Arch Biochem* **28**, 338–347.
4 Borsook H (1950) Protein turnover and incorporation of labeled amino acids into tissue proteins *in vivo* and *in vitro*. *Physiol Rev* **30**, 206–219.
5 Kim DM & Choi CY (1996) A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. *Biotechnol Prog* **12**, 645–649.
6 Alakov YB, Baranov VI, Ovodov SJ, Ryabova LA, Spirin AS & Morozov IJ (1995) *Method of Preparing Polypeptides in Cell-Free Translation System In*. US Patent No. 5, 478,730.
7 Spirin AS, Baranov VI, Ryabova LA, Ovodov SY & Alakhov YB (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* **242**, 1162–1164.
8 Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T & Yokoyama S (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett* **442**, 15–19.

9 Sawasaki T & Endo Y (2004) *In vitro* protein synthesis system: cell-free protein synthesis system prepared from wheat germ]. *Tanpakushitsu Kakusan Koso* **49**, 1514–1519.

10 Madin K, Sawasaki T, Ogasawara T & Endo Y (2000) A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. *Proc Natl Acad Sci USA* **97**, 559–564.

11 Endo Y, Otsuzuki S, Ito K & Miura K (1992) Production of an enzymatic active protein using a continuous flow cell-free translation system. *J Biotechnol* **25**, 221–230.

12 Anderson CW, Straus JW & Dudock BS (1983) Preparation of a cell-free protein-synthesizing system from wheat germ. *Methods Enzymol* **101**, 635–644.

13 Ozawa K, Headlam MJ, Schaeffer PM, Henderson BR, Dixon NE & Otting G (2004) Optimization of an *Escherichia coli* system for cell-free synthesis of selectively N-labelled proteins for rapid analysis by NMR spectroscopy. *Eur J Biochem* **271**, 4084–4093.

14 Torizawa T, Shimizu M, Taoka M, Miyano H & Kainosho M (2004) Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. *J Biomol NMR* **30**, 311–325.

15 Kigawa T, Yabuki T, Matsuda N, Matsuda T, Nakajima R, Tanaka A & Yokoyama S (2004) Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. *J Struct Funct Genomics* **5**, 63–68.

16 Jewett MC & Swartz JR (2004) Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol Bioeng* **86**, 19–26.

17 Spirin AS (2004) High-throughput cell-free systems for synthesis of functionally active proteins. *Trends Biotechnol* **22**, 538–545.

18 Ishihara G, Goto M, Saeki M, Ito K, Hori T, Kigawa T, Shirouzu M & Yokoyama S (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expr Purif* **41**, 27–37.

19 Klammt C, Schwarz D, Fendler K, Haase W, Dötsch V & Bernhard F (2005) Evaluation of detergents for the soluble expression of α-helical and β-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *FEBS J* **272**, 6024–6038.

20 Klammt C, Löhr F, Schafer B, Haase W, Dötsch V, Rüterjans H, Glaubitz C & Bernhard F (2004) High level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem* **271**, 568–580.

21 Shimada Y, Wang ZY, Mochizuki Y, Kobayashi M & Nozawa T (2004) Functional expression and characterization of a bacterial light-harvesting membrane protein in *Escherichia coli* and cell-free synthesis systems. *Biosci Biotechnol Biochem* **68**, 1942–1948.

22 Elbaz Y, Steiner-Mordoch S, Danieli T & Schuldiner S (2004) *In vitro* synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. *Proc Natl Acad Sci USA* **101**, 1519–1524.

23 Berrier C, Park KH, Abes S, Bibonne A, Betton JM & Ghazi A (2004) Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry* **43**, 12585–12591.

24 Pornillos O, Chen YJ, Chen AP & Chang G (2005) X-ray structure of the EmrE multidrug transporter in complex with a substrate. *Science* **310**, 1950–1953.

25 Yamane T, Ikeda Y, Nagasaka T & Nakano H (2005) Enhanced cell-free protein synthesis using a S30 extract from *Escherichia coli* grown rapidly at 42 degrees C in an amino acid enriched medium. *Biotechnol Prog* **21**, 608–613.

26 Ozawa K, Dixon NE & Otting G (2005) Cell-free synthesis of 15N-labeled proteins for NMR studies. *IUBMB Life* **57**, 615–622.

27 Liu DV, Zawada JF & Swartz JR (2005) Streamlining *Escherichia coli* S30 extract preparation for economical cell-free protein synthesis. *Biotechnol Prog* **21**, 460–465.

28 Zubay G (1973) *In vitro* synthesis of protein in microbial systems. *Annu Rev Genet* **7**, 267–287.

29 Michel-Reydellet N, Woodrow K & Swartz J (2005) Increasing PCR fragment stability and protein yields in a cell-free system with genetically modified *Escherichia coli* extracts. *J Mol Microbiol Biotechnol* **9**, 26–34.

30 Ahn JH, Chu HS, Kim TW, Oh IS, Choi CY, Hahn GH, Park CG & Kim DM (2005) Cell-free synthesis of recombinant proteins from PCR-amplified genes at a comparable productivity to that of plasmid-based reactions. *Biochem Biophys Res Commun* **338**, 1346–1352.

31 Kiefer H (2003) *In vitro* folding of alpha-helical membrane proteins. *Biochim Biophys Acta* **1610**, 57–62.

32 Lorch M, Lehner I, Siarheyeva A, Basting D, Pfleger N, Manolikas T & Glaubitz C (2005) NMR and fluorescence spectroscopy approaches to secondary and primary active multidrug efflux pumps. *Biochem Soc Trans* **33**, 873–877.

33 Trbovic N, Klammt C, Koglin A, Löhr F, Bernhard F & Dötsch V (2005) Efficient strategy for the rapid backbone assignment of membrane proteins. *J Am Chem Soc* **127**, 13504–13505.

34 Okada T, Sugihara M, Bondar AN, Elstner M, Entel P & Buss V (2004) The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J Mol Biol* **342**, 571–583.

35 Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Trong IL, Teller DC, Okada T, Stenkamp RE & Miyano M (2000) Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **289**, 739–745.

36 Yabuki T, Kigawa T, Dohmae N, Takio K, Terada T, Ito Y, Laue ED, Cooper JA, Kainosho M & Yokoyama S

(1998) Dual amino acid-selective and site-directed stable-isotope labeling of the human c-Ha-Ras protein by cell-free synthesis. *J Biomol NMR* **11**, 295–306.
37 Kigawa T, Muto Y & Yokoyama S (1995) Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. *J Biomol NMR* **6**, 129–134.
38 Guignard L, Ozawa K, Pursglove SE, Otting G & Dixon NE (2002) NMR analysis of *in vitro*-synthesized proteins without purification: a high-throughput approach. *FEBS Lett* **524**, 159–162.
39 Krueger-Koplin RD, Sorgen PL, Krueger-Koplin ST, Rivera-Torres IO, Cahill SM, Hicks DB, Grinius L, Krulwich TA & Girvin ME (2004) An evaluation of detergents for NMR structural studies of membrane proteins. *J Biomol NMR* **28**, 43–57.
40 Rungpragayphan S, Nakano H & Yamane T (2003) PCR-linked *in vitro* expression: a novel system for high-throughput construction and screening of protein libraries. *FEBS Lett* **540**, 147–150.
41 Sawasaki T, Ogasawara T, Morishita R & Endo Y (2002) A cell-free protein synthesis system for high-throughput proteomics. *Proc Natl Acad Sci USA* **99**, 14652–14657.
42 Busso D, Kim R & Kim SH (2003) Expression of soluble recombinant proteins in a cell-free system using a 96-well format. *J Biochem Biophys Methods* **55**, 233–240.
43 Angenendt P, Nyarsik L, Szaflarski W, Glokler J, Nierhaus KH, Lehrach H, Cahill DJ & Lueking A (2004) Cell-free protein expression and functional assay in nanowell chip format. *Anal Chem* **76**, 1844–1849.
44 Lee K & Cohen SN (2001) Effects of 3′ terminus modifications on mRNA functional decay during *in vitro* protein synthesis. *J Biol Chem* **276**, 23268–23274.
45 Hirao I, Yoshizawa S & Miura K (1993) Stabilization of mRNA in an *Escherichia coli* cell-free translation system. *FEBS Lett* **321**, 169–172.
46 Vinarov DA & Markley JL (2005) High-throughput automated platform for nuclear magnetic resonance-based structural proteomics. *Expert Rev Proteomics* **2**, 49–55.
47 Vinarov DA, Lytle BL, Peterson FC, Tyler EM, Volkman BF & Markley JL (2004) Cell-free protein production and labeling protocol for NMR-based structural proteomics. *Nat Methods* **1**, 149–153.
48 Winstone TL, Duncalf KA & Turner RJ (2002) Optimization of expression and the purification by organic extraction of the integral membrane protein EmrE. *Protein Expr Purif* **26**, 111–121.
49 Ninio S, Rotem D & Schuldiner S (2001) Functional analysis of novel multidrug transporters from human pathogens. *J Biol Chem* **276**, 48250–48256.
50 Yerushalmi H, Lebendiker M & Schuldiner S (1995) EmrE, an *Escherichia coli* 12-kDa multidrug transporter, exchanges toxic cations and H^+ and is soluble in organic solvents. *J Biol Chem* **270**, 6856–6863.
51 Park KH, Berrier C, Martinac B & Ghazi A (2004) Purification and functional reconstitution of N- and C-halves of the MscL channel. *Biophys J* **86**, 2129–2136.
52 Benz R, Schmid A, Maier C & Bremer E (1988) Characterization of the nucleoside-binding site inside the Tsx channel of *Escherichia coli* outer membrane. Reconstitution experiments with lipid bilayer membranes. *Eur J Biochem* **176**, 699–705.

(P 05)

Cell-free expression of integral membrane proteins for structural studies. C. Klammt, D. Schwarz, B. Schneider, F. Löhr, I. Lehner, C. Glaubitz, V. Dötsch and F. Bernhard. In: *Cell-free expression techniques,* A. Spirin (ed.), Wiley-VCH, Weinheim, submitted. (2006)

Cell-free expression of integral membrane proteins for structural studies

Christian Klammt[1], Daniel Schwarz[1], Ines Lehner[1], Solmaz Sobhanifar[1], Frank Löhr[1], Johan Zeelen[2], Clemens Glaubitz[1], Volker Dötsch[1] and Frank Bernhard[1*]

[1] Centre for Biomolecular Magnetic Resonance, University of Frankfurt/Main, Institute for Biophysical Chemistry, Max-von-Laue-Str. 9, D-60438 Frankfurt/Main, Germany

[2] Max-Planck-Institute for Biophysics, Dept. for Structural Biology, Max-von-Laue-Str. 3, D-60438 Frankfurt/Main

* Corresponding author
Phone: +49-69-798-29620
FAX: +49-69-798-29632
Email: fbern@bpc.uni-frankfurt.de

Abstract

Cellular expression systems are often very inefficient for the high-level production of membrane protein. Toxic effects, instability or formation of inclusion bodies are frequent observed effects that prevent the synthesis of sufficient amounts of functional protein. The development of preparative scale cell-free expression systems has provided new alternative tools with several attractive benefits for the production of membrane proteins. Unique and fascinating properties are the possibilities to synthesize recombinant membrane proteins directly into detergent micelles or into liposomes of defined composition. Considerable success has already been made with the expression of structurally diverse membrane proteins in cell-free systems and this chapter will summarize the recent approaches. We will discuss distinct applications with a special focus on the cell-free production of functionally folded membrane proteins.

Abbreviations: Brij-35, polyoxyethylene-(23)-lauryl-ether; Brij-56, polyoxyethylene-(10)-cetyl-ether; Brij-58, polyoxyethylene-(20)-cetyl-ether; Brij-78, polyoxyethylene-(20)-stearyl-ether; Brij-97, polyoxyethylene-(10)-oleyl-ether; Brij-98, polyoxyethylene-(20)-oleyl-ether; ß-OG, n-octyl-ß-glucopyranoside; CD, Circular dichroism; CECF, continuous exchange cell-free; CF, cell-free; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propansulfonat; CMC, Critical micellar concentration; C_{mic}, micellar concentration; CP, cross polarisation; DHPC, 1,2-diheptanoyl-sn-glycero-3-phosphocholine; diC_8PC, 1,2-dioctanoyl-sn-glycero-3-phosphocholine; DM, n-decyl-ß-maltoside; DDM, n-dodecyl-ß-D-maltoside; DMPC, 1,2-dimyristoyl-sn-glycero-3-phosphocholine; DOPC, 1,2-dioleoyl-sn-glycero-3-phosphocholine; DPC, dodecyl-phosphocholine; FM, feeding mixture; GPCR, G-protein coupled receptor; HSQC, heteronuclear single quantum correlation; IMP, integral membrane protein; LMPG, 1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)]; LPPG, 1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)]; MAS, magic angle sample spinning; NMR, nuclear magnetic resonance; OMP, outer membrane protein; RM, reaction mixture; SDS, sodium dodecylsulfate; ssNMR, solid state NMR; Thesit, polyethylene-glycol 400 dodecyl-ether; TMS, transmembrane segment; TPP^+, tetraphenylphosphonium; TROSY, transverse relaxation optimized spectroscopy; TX-100, polyethylene glycol P-1,1,3,3-tetramethyl-butylphenyl ether (Triton X-100); Tween 20, polyoxyethylene-sorbitan-monolaurate 20.

Introduction

Integral membrane proteins (IMPs) are embedded into cellular membranes by multiple hydrophobic transmembrane segments (TMSs) and they control numerous essential functions like transport activities, energy generation, signal perception and communication of the cell with its environment. The topology of IMPs is generally dominated by α-helical structures, while typical ß-barrel arrangements are prevalent in IMPs inserted into the outer membrane of Gram-negative bacteria. An average whole cell proteome is supposed to consist to 20-40 % of IMPs. Many pharmaceutical studies focus currently on IMPs as they play key roles in a variety of global human diseases. IMPs provide thus an estimate of approx. 60 % of all modern drug targets. A basic prerequisite for a directed drug design as well as for the understanding of biological functions is the knowledge of the three-dimensional structure of a protein. Relatively high amounts of protein in the range of several 100 mg are often needed for structural approaches by either X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. With few exceptions where natural abundance of a protein offers the opportunity of high yield preparations, heterologous cellular expression systems of prokaryotic or eukaryotic origin have to be used in order to produce sufficient amounts of protein. Unfortunately, the hydrophobic nature of IMPs and their distinct localization in cellular membranes does often cause tremendous problems upon their synthesis in conventional expression systems. Blocking of cellular protein targeting systems, complex formation with cellular membrane proteins or disintegration of membranes are frequent effects that result in toxicity to living cells leading to low expression rates.

Cell-free (CF) expression systems offer in principle the possibility to eliminate toxic effects of recombinant proteins as no living and metabolic active environments are necessary. Only care has to be taken that the essential translation and transcription processes will not be affected by the overproduced protein. CF systems are therefore predestined for the synthesis of even strong toxins that cannot obtained with conventional *in-vivo* expression systems (1). Moreover, a considerable number of additional benefits are intrinsic properties of CF expression systems. Expression conditions like pH, redox potential or buffer systems could easily be modified. A highly valuable characteristic is the open nature of the reaction. This allows the addition of nearly any compound at any time-point directly into the reaction. In contrast to cellular expression systems, there is no selection of additives by specific transport mechanisms or risk of metabolic conversion or even breakdown of the added compound. The proteolytic degradation of synthesized proteins could be inhibited by the addition of protease inhibitors and their folding into functional conformations could be facilitated by

supplemented chaperones, cofactors or other helper proteins. Recombinant proteins could further be stabilized by providing ligands, substrates or inhibitors. A unique and fascinating option is the generation of artificial hydrophobic environments in CF expression reactions for the production of soluble IMPs.

Different designs of preparative scale CF expression systems with protein yields starting from few 100 µg up to several mg per one ml of reaction volume are possible. Batch systems with only one compartment can be carried out in various common labware containers like simple plastic reaction tubes. More advanced systems like the continuous exchange CF (CECF) reactions are composed of two compartments holding a reaction mixture (RM) and a feeding mixture (FM) with a volume ratio of RM/FM usually in between 1:10 and 1:20 (2-4). The two compartments are separated by a semipermeable membrane that ensures an efficient exchange of precursors from the FM into the RM. Inhibitory breakdown products are continuously removed from the translation machinery by diffusion into the FM. CECF systems are superior in yields if compared with batch systems while considerable efforts have recently been made in order to improve the productivity and longevity of batch CF reactions (5-10).

Reliable and detailed protocols for the set-up of CF reactions have been published by several groups (11-13). In addition to the individual CF designs, also commercial systems are available that can result in the high-level expression of IMPs (14-16). Key compound of CF expression systems is a cellular extract that can be derived from either bacterial origin, mostly *Escherichia coli* cells, or from eukaryotic origin, preferentially selected wheat germs. High quality bacterial extracts can be obtained from various common lab strains like BL21 derivatives (12) or RNAse deficient strains like A19 or D10 (17, 18). The preparation of bacterial extracts is a reliable and routine technique (11, 18-21) and an important step is the removal of the cellular mRNA that virtually eliminates any background expression in subsequent CF expression reactions. Nucleotides and amino acids can be added in order to release residual mRNA from the ribosomes in a “run-off” translation procedure (19, 20). The non-protected mRNA will then rapidly become degraded by endogenous RNAses. We prefer a modification of this step by adjusting the extract to 400 mM NaCl followed by incubation at 42°C for 45 min (11). This treatment causes the efficient dissociation of mRNA and ribosomes while no addition of precursors is necessary. So far only *E. coli* extracts have been used for the CF production of IMPs. This fact might be mostly attributed to the more difficult preparation procedure of wheat germ extracts (22). However, CF expression systems based on eukaryotic cell extracts might be considered in the future as they could provide several

advantages especially for the production of eukaryotic IMPs like improved folding pathways or the possibility of posttranslational modifications. CF reactions with *E. coli* extracts are operated as coupled transcription/translation systems with added circular plasmid DNA or linear DNA fragments as a template (3). Even several different templates could be transcribed simultaneously in a CF reaction, thus enabling the coexpression of different proteins, e.g. various subunits of heterooligomeric IMP complexes. The ratio of the recombinant proteins could furthermore easily be manipulated by varying the amounts of the supplied templates. An efficient transcription is provided by placing the target genes under control of a T7 promoter and by addition of T7 RNA polymerase into the reaction.

This chapter will summarize recent advancements in the CF production of IMPs in preparative scales, while results with non-preparative scale CF systems will not be covered. CF expression systems generally provide an attractive alternative technique for protein targets that cannot efficiently be produced in conventional *in-vivo* expression systems. The technique does not require special equipments and can be operated in most standard biochemical labs. The CF expression of IMPs is still an emerging application but the already obtained achievements are highly promising and make a rapidly increasing request for this technology very likely.

1. Specific characteristics for the cell-free expression of membrane proteins

Different modes are feasible for the expression of IMPs in CF systems (Fig. 1). Most cellular membranes have been removed during extract preparation and only spurious amounts of lipids might remain (14). Standard CF reactions therefore result consequently in the production of IMPs as precipitates as no hydrophobic compartments are present in the RM. The CF reaction protocol in that mode is virtually identical to the production of globular proteins (11, 14). Critical components are the ions Mg^{2+} and K^{+} because of their narrow optimal concentration ranges usually in between 12 - 17 mM and 270 - 330 mM, respectively. Optimization screens have therefore often to be employed before high levels of protein synthesis can be obtained. The supplementation of CF reactions with detergents or lipids generates preformed micelles or liposomes in the RM. This option is an attractive feature specific to CF expression and not possible with any other expression system. It enables the direct synthesis of IMPs into micelles composed of the desired detergents or into liposomes of defined compositions. Critical steps in common IMP preparation protocols like the disintegration of cellular membranes or repeated transfers of IMPs into micelles of different compositions can thus be reduced or even completely avoided. In the best case, the IMPs can

become inserted in the desired environment right after translation and the resulting proteomicelles or proteoliposomes could be directly used for further analysis. A high level production of IMPs is therefore generally possible with three different CF expression modes: (A) the production of IMP precipitates without any hydrophobic additives (11, 23), (B) the synthesis of IMPs into micelles in presence of detergents (13, 14, 23, 24) and (C) the synthesis of IMPs into liposomes in presence of lipids (16) (Fig. 1).

1.1. Cell-free expression of membrane proteins in presence of detergents or lipids

Detergents have to be added at concentrations above their specific critical micellar concentrations (CMC) in order to become effective for the solubilization of IMPs. However, detergents also interfere non-specifically with any hydrophobic protein regions and they may thus affect the productivity of the CF system. Several detergents like sodium deoxycholate, sodium cholate and N-laurylsarcosine as well as dodecyl-phosphocholine (DPC) and n-octyl-ß-glucopyranoside (ß-OG) severely inhibit CF systems already at low concentrations and they are not suitable for the soluble expression of IMPs (13, 14, 24). However, the majority of the commonly employed relatively mild detergents appear to be tolerated by CF systems at concentrations that exceed several times the proposed specific CMCs (13, 14, 24) (Table 1).

The evaluation of the most effective detergent for the CF production of a specific IMP should be one of the primary subjects of initial optimization screens. Most IMPs are probably likely to become soluble expressed in a variety of different detergents. The bi-chain-phosphocholines 1,2-dioctanoyl-sn-glycero-3-phosphocholine (diC_8PC), 1,2-diheptanoyl-sn-glycero-3-phosphocholine (DHPC), the alkyl-glucosides n-dodecyl–ß-D-maltoside (DDM) and n-decyl-ß-D-maltoside (DM), the alkyl-ether polyoxyethylene-sorbitan-monolaurate 20 (Tween 20) as well as the polyethylene-glycol derivatives polyethylene-glycol P-1,1,3,3-tetramethyl-butylphenyl-ether (TX-100) and polyethylene-glycol 400 dodecyl-ether (Thesit) are some detergents that resulted in the production of preparative amounts of protein (13, 14, 23, 24) (Table 1). In contrast, all these detergents have not been very effective for the CF expression of other IMPs like the porcine vasopressin type 2 receptor (V2R), a member of the family of G-protein coupled receptors (GPCRs).

A variety of polyoxyethylene-alkyl-ethers and the steroid derivative digitonin are exceptionally suitable for the CF expression of structurally different IMPs (13, 14, 24). EmrE, Tsx and even the V2R protein and other members of the GPCR family could be produced in amounts of up to several mg of soluble protein in one ml of RM. Some structural details of these detergents like the length of polyoxyethylene chains are important for the efficiency of

IMP solubilization. Long chain derivatives like polyoxyethylene-(23)-lauryl-ether (Brij-35), polyoxyethylene-(20)-cetyl-ether (Brij-58), polyoxyethylene-(20)-stearyl-ether (Brij-78) and polyoxyethylene-(20)-oleyl-ether (Brij-98) resulted in the high level soluble expression of the different IMPs. In presence of shorter chain length derivatives like polyoxyethylene-(10)-cetyl-ether (Brij-56) or polyoxyethylene-(10)-oleyl-ether (Brij-97), the IMPs have still been synthesized at high amounts but they remained almost quantitatively as precipitate.

The supplied final detergent concentration deserves special considerations. Most detergent will become inhibitory above certain concentrations and some guidelines for the supply of detergents into CF reactions have to be considered (13, 14, 24). The particular micellar concentration (C_{mic}) of a detergent in the RM should be at least equal to the estimated molar concentration of the synthesized IMP at the end of a CF reaction. An excess of synthesized IMP might form undesired precipitates or even heterogeneous micelles that could prevent further structural approaches. The C_{mic} is difficult to calculate as it is a result of the specific CMC as well as of the aggregation number, the proposed number of detergent molecules per micelle. These parameters, however, are highly variable as they depend on several environmental factors like pH and temperature as well as on the topology of the solubilized IMP (25). A good compromise of C_{mic} has thus to be found in order to ensure optimal solubilization of the synthesized amount of protein but to avoid inhibitory effects of too high detergent concentrations. Detergents with a low aggregation number resulting in a high C_{mic} at still relatively low molar concentrations appear generally to be better suitable for the efficient solubilization of CF produced IMPs. However, specific characteristics of detergents further contribute to their behaviour in CF reactions. DDM and Digitonin became already inhibitory above 15 x CMC (1.8 mM) and 4.5 x CMC (3.3 mM), respectively, while e.g. Brij-58 and Brij-78 are completely tolerated at even 170 x CMC (12.8 mM) and 280 x CMC (12.9 mM). The amount of soluble produced IMP in a CF reaction as a function of detergent concentration follows plateau-like kinetics and the yield remains constant above a certain concentration level. However, not all of the synthesized protein might become soluble and it is common that some residual precipitated IMP is still present at the end of the CF reaction. For the CF expression of soluble Tsx protein, optimal detergent concentrations have been determined for TX-100 (7 x CMC), Brij-58 (47 x CMC), Brij-78 (76 x CMC), Digitonin (2.2 x CMC) and DDM (15 x CMC) (24).

Increased yields of functionally folded mechanosensitive channel MscL were obtained in CF reactions with mixed micelles of TX-100 and *E. coli* lipids or in presence of liposomes prepared from *E. coli* lipid mixtures (14). In addition, seleno-methionine labelled EmrE was

synthesized in presence of 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) liposomes and the quality of the isolated protein was sufficient for structural analysis by X-ray crystallography (16). However, no further characterization of the proteoliposomes e.g. by freeze-fracture analysis has been done. Subsequent purification steps involved the solubilization of the IMPs with detergents and the efficiencies of IMP reconstitution into the provided liposomes during the CF reaction remained unclear so far. Nevertheless, these first data provide promising results and the expression of IMPs into artificial liposomes might have a high potential to become an important technique for membrane protein analysis.

1.2. Detergents for the efficient resolubilization of cell-free produced membrane proteins

For the CF production of IMP precipitates, standard reaction protocols could instantly be used and no time-consuming evaluation of detergents is needed. High yields of synthesized IMP are ensured as no inhibitory effects of supplied detergents are present. In addition, the IMPs could be obtained in highly purified form by only few steps (11). This expression mode mostly resembles conventional *in-vivo* approaches of the production of IMPs in form of inclusion bodies. However, structural differences might exist between cellular inclusion bodies and CF produced precipitates. IMP precipitates obtained by CF reactions usually solubilize rapidly upon addition of relatively mild detergents and they do not require intensive denaturation and refolding steps as it is known from the solubilization of inclusion bodies (14, 23, 24). Denaturating agents like Guanidinium HCl or excessive amounts of urea could thus be avoided and already gentle mixing of the IMP precipitate with a suitable detergent at a final concentration of 1-5 % at room temperature is often sufficient for a quantitative solubilization. In particular, 1-myristoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)] (LMPG) and related derivatives like 1-palmitoyl-2-hydroxy-sn-glycero-3-[phospho-rac-(1-glycerol)] (LPPG) were found to have outstanding properties in the efficient solubilization of CF produced precipitates of structurally different IMPs (24, 26). Other mild detergents like DDM, DPC, TX-100, DHPC or 3-[(3-cholamidopropyl)dimethylammonio]-1-propansulfonat (CHAPS) are further effective in the solubilization of specific IMPs like EmrE, Tsx or the mechanosensitive channel McsL, but they failed to solubilize the GPCR protein V2R (14, 24) (Table 1).

Most important is always the preparation of functionally folded proteins and the solubilization of precipitated IMPs might generally be more critical than the production as soluble proteins in presence of detergents. Some proteins like the transporter EmrE or the channel MscL can be produced in a functional form in both ways. Solubilization out of CF

produced precipitates as well as the direct soluble expression in presence of detergents resulted in active proteins (14, 23, 24). However, some IMPs like the nucleoside transporter Tsx could only produced in fully functional form with the soluble mode of CF expression (24).

2. Case studies for the high level cell-free expression of membrane proteins

The application of CF expression to the production of membrane proteins is an emerging technique and we will describe recent approaches of the preparative scale CF production of structurally different IMPs. The available data are still confined to a relatively limited number of proteins, but the considerable structural diversity of the synthesized IMPs and the recent success in the expression of GPCR proteins already indicate an enormous potential of the CF technique for the high-level production of difficult IMPs (Figs. 2 and 3). The functionality of the synthesized IMPs and also corresponding yields obtained with *in-vivo* expression systems will be discussed if data are available. The yields of recombinant protein obtained from CECF systems can easily be manipulated by varying the RM/FM ratio or by refreshing the complete FM during the reaction. The specific reaction conditions should therefore be considered when the efficiencies of different CF reactions are compared. It is even more difficult to compare yields of CF reactions with that of *in-vivo* reactions. If cell volumes should be compared, one could consider that one litre of an outgrown *E. coli* culture in standard Luria broth (10 g tryptone, 10 g yeast extract, 5 g NaCl) contains a cell volume equivalent to approx. 5 - 10 ml of RM. However, this calculation does neither consider the preparation of S30 extracts, enzymes or reaction mixtures nor fermentation and purification procedures. In general, CF expression is a relatively expensive technique and it might therefore represent the first choice primarily for specific labelling purposes or if a target protein cannot become expressed at all or only at low amounts in other systems.

2.1. α-helical transporters

Transporters control the conditions between in- and outside of their host-membrane like the uptake of substrates, the efflux of toxic substances or the regulation of ion concentrations. The 110 amino acid small multidrug transporter EmrE was subject of several approaches of high-level production by CF expression. Conventional *in-vivo* expression of EmrE routinely yields approx. one mg of protein per litre *E. coli* culture (27). In CF systems based on *E. coli* S30 extracts amounts of 2-3 mg per one ml of RM in preparative scale set-ups and 1-2 µg of protein per 20 µl of RM in analytical set-ups were reported (23). Supplied detergents did not alter the production rate of EmrE and directed most of the protein into the soluble fraction.

Pull down experiments as well as cross-linking studies with CF produced EmrE revealed a homodimer as the functional conformation and activity assays of the transporter in DDM micelles showed specific substrate binding of [^{3}H] tetraphenylphosphonium (TPP^+). The determined K_d value of 2.3 nM matches very nicely with the K_d of 2.8 nM obtained from *in-vivo* produced EmrE (28). The methyl viologen uptake of reconstituted EmrE was measured to verify the specific substrate transport across a membrane. An H^+ gradient dependent accumulation of the substrate against its concentration gradient could be determined to a rate of 7.5 nmol per min per µg of CF produced EmrE which is very similar to that of EmrE isolated from *E. coli* cells with 5.5 nmol per min and µg (23). CF produced EmrE precipitate could be solubilized in DDM and reconstituted into liposomes based on *E. coli* lipids. These proteoliposomes also showed specific transport of the substrate ethidium bromide (11). Recorded NMR-spectra indicated in addition an identical conformation of CF produced EmrE to that obtained by *in-vivo* expression. In further CF expression studies, the four homologous bacterial multidrug transporters TBsmr, BPsmr, Psmr and Hsmr could be produced (23) (Fig. 2). All four proteins can also be produced in an *in-vivo E. coli* system but at least the Psmr protein apparently showed higher expression levels in the CF system.

The high-level CF expression of the mechanosensitive channel MscL from *E. coli*, a homopentamer of 14 kDa subunits that each consist of two TMSs, resulted in functional protein (14). Conductance measurements on reconstituted MscL protein in giant liposomes were performed by patch clamp assays. Soluble CF produced MscL in presence of TX-100 showed a comparable activity with 8.3 ±1.8 open channels per patch to the in *E. coli* cells overproduced channel with 9.3 ± 5.5 open channels per patch. The addition of an amino-terminal poly(His)$_6$-tag lowered the numbers of channels per patch of MscL protein isolated from both expression systems. The full activity of MscL could be restored by removing the tag upon cleavage with the factor Xa protease. Cross-link experiments indicated the formation of the functional pentameric state of MscL already in detergent micelles. The *in-vitro* expression levels could be increased to a final yield of approx. 3.6 mg MscL protein per ml RM by adding 18 µg/ml *E. coli* lipids.

The bacterial α-helical tellurite transporter TehA (36 kDa, 10 TMSs) and the cysteine transporter Yfik (22 kDa, 6 TMSs) were not produced in detectable amounts in *E. coli* cells. The yields in CECF systems for both protein and of the C-terminal truncated derivative ΔTehA (24 kDa, 7 TMSs) reached levels up to 2.7 mg per ml RM (Fig. 2). No functional analysis of the two proteins has been made but the structural evaluation by NMR spectroscopy revealed strong evidence of folded proteins (11, 26).

Expression in an *E. coli* based system as well as the CF production of the 6.1 kDa α apoprotein of the light harvesting complex (LH1) from *Rhodospirillum rubrum* yielded amounts of 1.2 mg per litre cell-culture and 0.7 mg per ml RM, respectively (15). The CF obtained α-LH1 precipitate could be functionally refolded in a buffer system containing 0.5 to 2.0 % TX-100. Recorded CD-spectra were essentially identical to that of native α–LH1 protein and the *in-vitro* formation of a structural complex with the ß-subunit of the LH1 complex was demonstrated by its specific spectral absorbance pattern.

The CF expression of homologues of eukaryotic glutamate transporters from the four bacterial species *E. coli, Aeropyrum pernix, Pyrococcus furiosus* and *Pyrococcus horikoshii* have been compared (Fig. 3). The approx. 45 kDa proteins have eight proposed TMSs and form putative trimers. All four proteins have been produced as precipitate and could subsequently be solubilized in the detergents zwittergent 3-12 or zwittergent 3-14. Highest expression level was obtained for the *E. coli* homologue with approx. 2 mg per ml RM followed with approx. 30 % of that yield from the homologue of *P. furiosus*. The two other proteins have been produced only at a significantly lower level in the CF system. The data demonstrate that the expression efficiencies of even closely related IMPs can be very different. Optimized codon usage of the target genes, supplementation of rare codon tRNAs or adjusting the amino acid pool according to the composition of the specific target proteins might be considered as further approaches in order to improve the expression yields.

2.2. G-protein coupled receptors

Eukaryotic GPCRs represent the largest single class of receptors and they are predicted to consist of 7 TMSs. Ligand binding induces conformational changes leading to the activation of cytoplasmatic G-proteins. GPCRs and GPCR dependent signalling pathways of higher organisms have become prominent drug targets and there is a rapidly increasing demand for the identification of new GPCR ligand analogues. The structural characterization of these proteins is an essential prerequisite for directed drug screening approaches. However, GPCRs are generally of low natural abundance and efficient overexpression systems are therefore indispensable. *E. coli* strains show in most cases only very low GPCR expression levels with often less than 200 receptors per cell (29, 30). A rare exception with up to 3500 copies per *E. coli* cell (resulting in 0.66 mg per litre culture) of functional muscarinic m_1 receptor is reported (31). However, the overproduced GPCRs generally accumulate in *E. coli* as inclusion bodies. The possibility of post-translational modifications make yeast strains interesting for the overproduction of the eukaryotic GPCRs and improved expression conditions resulted in

the detection of several functionally synthesized GPCRs (32). Reported expression levels were generally approx 1-2 pmol/mg membrane protein with some exceptions of 10 pmol/mg membrane protein in case of the serotonin 5-HT_{5A}, $ß_2$- and α_2-adrenergic and endothelin B receptors (33). The closest alternative to their native environment is the overexpression of GPCRs in mammalian or insect cells especially when post-translational modifications are a prerequisite for functional studies. The expression levels usually do not exceed 5-10 pmol/mg membrane protein in adherent cell-lines and they are often lower in suspension cultures (34).

The human ß2 adrenergic receptor (ß2AR), the human M2 muscarinic acetylcholine receptor (M2) and the rat neurotensin receptor (NTR) could be synthesized in batch CF systems in amounts of 150-200 µg synthesized protein per ml RM (13). The proteins had to be produced as translational fusions C-terminal to thioredoxin and almost no protein production was detected with the native non-fused coding regions. The authors speculated that stabilization of the GPCRs by the N-terminal thioredoxin or improved translation efficiencies might account for that observation. The final size of the recombinant IMPs was in between 53 - 109 kDa and even the larger constructs were synthesized without significant loss of efficiency. In CECF systems, the final yield was improved to approx. 1 mg GPCR per ml RM within 8 hours of reaction. The CF produced thioredoxin fused ß2AR protein was reconstituted into preformed phospholipid vesicles by dialysis and binding studies were performed with the radioactive labelled substrate [^{3}H]dihydroalprenolol (13). The calculated K_d of the CF expressed ß2AR was determined to 5.5 nM and it is comparable to the K_d of 3.9 nM of ß2AR isolated from Sf9 insect cells. Competition studies with non-labelled dihydroalprenolol resulted in EC_{50} values of 2.2 x 10^{-8} M with CF expressed ß2AR and of 1.0 x 10^{-8} M with Sf9-expressed ß2AR in one-site competition analyses.

Amounts of up to 3 mg per ml RM can be obtained for the porcine vasopressin receptor type 2 (V2R) containing only a small amino-terminal T7 tag fusion in an individual CECF set-up (24). CF expression without the 14 amino acid T7 tag fusion resulted only in spurious expression which is in agreement to the above mentioned observations with the CF expression of the GPCRs ß2AR, M2 and NTR. Same observations were made for three further GPCR proteins, the human endothelin B receptor (ETB), the rat corticotropin releasing factor receptor (CRF) and the human V2R (Fig. 2).

2.3. ß-barrel proteins

The outer membrane of gram-negative bacteria forms a protective permeability barrier around the cells and serves as a molecular filter for hydrophilic substances. Outer membrane proteins

(OMPs) often show the general architecture of membrane-spanning ß-barrels which are self-closed ß-sheets. This topology contrasts the typical structure of inner membrane proteins with single or multiple membrane-spanning α-helices. Based on their transport mechanisms, OMPs can be divided into the classes of general porins, substrate specific transporters and active transporters. The substrate-specific transporters contain low-affinity substrate-binding sites that are saturable and allow efficient diffusion of substrates at very shallow concentration gradients. The *E. coli* nucleoside transporter Tsx belongs to this class (35, 36). Tsx is so far the only CF produced OMP and example for the high level production and functional reconstitution of this group of transport proteins (24). Tsx was expressed in standard CF set-ups in levels up to 4 mg per single ml of RM either as precipitate or as soluble proteins in presence of various detergents like Brij-35, Brij-78, TX-100 or others (Fig. 2). Conductance measurements of CF produced Tsx in planar membranes indicated a clear dependence of activity on the mode of expression. Only soluble expressed Tsx in presence of distinct detergents showed pore forming activity. The high-level production of IMPs does therefore not always result in functionally folded protein. Solubilized Tsx precipitate did not show any activity and the efficient folding of soluble produced protein into its active conformation depended also on the presence of specific detergents like TX-100.

3. Structural characterization of cell-free produced membrane proteins

IMPs are grossly underrepresented in the protein structural data base. The rare and often inhomogeneous IMP samples obtained from eukaryotic *in-vivo* expression systems with often incomplete post-translational modifications make structural determinations very difficult. CF expression could now provide access to otherwise difficult to obtain IMPs. The availability of high amounts of overproduced IMPs, the option of IMP expression directly into a detergent micelle representing the desired properties and many other possibilities to stabilize or to modify the synthesized IMPs are valuable properties of the CF technique. The versatility, speediness and the advantages of CF systems in the production of labelled IMPs promise therefore considerable progress for high resolution structural analysis by NMR spectroscopy as well as by X-ray crystallography.

3.1. Crystallization of cell-free produced membrane proteins

Crystallization of membrane proteins is a complicated process due to nonpredictable influences of additives like detergents or lipids and it is mostly limited by the lack of sufficient amounts of protein. More than 25.000 protein structures have been solved by X-ray

crystallography but only some 60 of them represent members of the family of IMPs. Recently the first X-ray structure by taking advantage of a CF produced IMP was presented (16). SeMet-labelled samples of the small multidrug transporter EmrE were produced in a commercially available CF batch system (Invitrogen, Carlsbad, USA) based on *E. coli* lysates. The RM contained 4.5 mM SeMet and was supplemented with 2 mg DMPC per ml and with substrates like TPP^+. EmrE was produced as soluble protein in presence of the preformed DMPC liposomes at an average of 0.2 mg per one ml RM. EmrE crystals used for structure determination and SeMet-EmrE-TPP^+ crystals were grown using similar conditions. The analysis revealed an antiparallel EmrE homodimer with one bound TPP^+ that is consistent with other biochemical studies. The two subunits adopt slightly different tertiary folds which seem to play an important role in substrate transport and H^+ antiport. This first report demonstrates that the quality of CF produced IMP samples is sufficient for crystallization and the technique might therefore attract increasing attention for the sample preparation for X-ray analysis studies.

3.2. Cell-free expression as tool for high resolution NMR spectroscopy

This section describes strategies in sample preparation of IMPs by CF expression in combination with high resolution NMR spectroscopy. Sample preparation has considerable influence on spectrum quality. An indispensable prerequisite for the NMR spectroscopy of larger proteins is the preparation of samples that have been labelled individually or with combinations of the stable isotopes ^2H, ^{13}C and ^{15}N. The full labelling of a protein requires the growth of the expression strain in defined medium supplemented with precursors that are labelled with the desired isotope. For amino acid specific labelling approaches, usually *E. coli* strains have to be employed containing auxotrophic mutations that correspond to the provided labelled amino acid. Conventional expression systems can cause considerable problems upon protein labelling like incomplete label incorporation, label scrambling, metabolic conversion of the labelled precursors or reduced yields if compared with the production of the unlabelled recombinant protein. Especially combinatorial approaches in which several different amino acid types have to be labelled can become extremely difficult and inefficient. CF expression provides generally an interesting alternative tool for the fast and efficient production of labelled protein samples. The complete label incorporation is ensured as no unlabelled amino acids are present in the reaction. Moreover, no auxotrophic mutations are needed as any amino acid can be simply replaced in the reaction mixtures by a labelled derivative. This option is very attractive for combinatorial labelling approaches and protein derivatives with

any label combinations can be generated with the same efficiencies as the unlabelled proteins as the overall reaction conditions have not to be modified. Sample preparation with *in-vivo* expression systems can furthermore take several days to weeks (37, 38) while CF reactions are completed within 12 hours.

A further aspect is the preparation of highly deuterated protein samples in order to attenuate transverse relaxation rates and to improve sensitivity and resolution in NMR spectra. Deuterated protein samples are usually generated by growing the expression strains in almost pure D_2O supplemented with the required nutrients. However, state of the art NMR methodologies for protein backbone assignment rely on the back-exchange of ^2H to ^1H or at least on a high occupancy of protons at the amide positions of deuterated proteins during their isolation in aqueous solutions (39). An incomplete back-exchange may lead to significant loss of signals and poor back-exchange may be aggravated for hydrophobic IMPs and in particular for residues in the hydrophobic TMS regions (40). Therefore, unfolding and refolding of isolated proteins that have been produced in conventional expression systems by employing denaturants such as 6 M urea or Guanidium HCl are often necessary (41). With CF expression, deuterated amino acids can be provided and all labile deuterated positions will already efficiently back-exchange in the aqueous reaction mixtures before the amino acids will become inserted into the growing peptide chain. A complete back-exchange of deuterated IMPs and other proteins at their amide positions is therefore generally ensured and no critical denaturation steps are necessary.

In order to study protein structure and dynamics by NMR spectroscopy, the backbone of a protein must first be assigned. Given the limitations involved in the analysis of IMPs, certain strategies must be implemented for an efficient designation. A significant challenge in assignment inherent in NMR studies remains in the size limitation for proteins. Membrane proteins tend to be large with multiple domains, and the necessity for subsequent solubilization in detergent may further significantly increase the final size of the protein/micelle composite, leading to considerable line broadening due to restricted molecular tumbling. Fortunately, this problem can be alleviated to a certain extent by using suitable detergents that obviously allow relatively free movements of inserted protein. The detergent class of lyso-phosphoglycerols including LMPG and LPPG has been found especially suitable for this purpose (26, 42, 43). Three *E. coli* IMPs, the cysteine exporter Yfik (21.2 kDa), the C-terminally truncated tellurite transporter ΔTehA (24 kDa), and SecE (13.6 kDa), a component of the secretion machinery, have given promising NMR heteronuclear single quantum correlation (HSQC) spectra when dissolved in LMPG (Fig. 4A-C). In all cases, more than 90

% of the expected signals are detectable and backbone assignments as a prerequisite for their structural analysis appear therefore feasible. A limitation with most IMPs is their often high content of α-helical secondary structures giving rise to poor chemical shift dispersion and leading to considerable spectral overlap if compared with other structural elements such as β-sheets. However, spectral overlap especially in the range between 7.5 ppm and 8.5 ppm can significantly be diminished by the analysis of amino acid type specifically labelled samples (Fig. 4D-I). Limited spectral dispersion has further been addressed by the use of certain combinatorial labelling strategies combined with sample analysis in high field spectrometers (26, 44, 45). An important advancement is the development of the Transverse Relaxation Optimized Spectroscopy (TROSY) experiment, which reduces transverse relaxation and allows to obtain high resolution spectra for molecules with higher molecular weights (46). One of the primary steps in assignment strategies is the measurement of 3D or 4D triple resonance ($^{13}C/^{15}N$-labelled) spectra and often protein backbones can be assigned by measuring the TROSY-type experiments such as HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HNCO and HN(CA)CO that allow the sequential identification of backbone residues.

The 24 kDa truncated version of the putative bacterial tellurite transporter TehA with seven TMSs was the first IMP whose backbone assignment was approached by taking advantage of CF expression strategies (26). Standard non-selective triple resonance experiments resulted only in approx. 55 % assignment and high degrees of signal overlap in the α-helical IMP required further strategies for the full backbone assignment. The analysis of 10 further samples having each a different amino acid type specifically labelled resulted only in approx. 10 % additional assignments. The reason for this low rate of success was accredited to redundancy, such that in many cases an identified amino acid type was associated with more than one N or C-terminal connectivity, leading to ambiguous sequence assignments. Of much higher success was the application of a combinatorial selective labelling method which relies on the ^{13}C labelling of certain amino acid types concomitant with the simultaneous ^{15}N labelling of others (44). In this approach, three samples were produced having each two to three different ^{15}N labelled amino acids types together with two ^{13}C labelled types (Fig. 5). The rationale behind this strategy is that by using this combinatorial base together with 2D versions of the HNCO experiment and $^{15}N/^1H$ TROSY spectra, ^{15}N-labelled amino acids types that are preceded by ^{13}C-labelled types can readily be selected (Fig. 5). If the identified combination of two amino acids in the protein has a singular occurrence, the specific site can be unambiguously assigned and may act as an anchor point for further sequential assignments.

It is predicted that 40-50 % of all amino acids within a protein are part of a unique amino acid pair, and algorithm programs which assist with the selection of amino acid types to be used for such combinatorial specific labelling schemes have been made available: http://www.Biophyschem.uni-frankfurt.de/AKDoetsch/ projects/download/combilabel.m).

In combination with the information obtained from the non-selective and from the amino acid specific labelled samples resulted the combinatorial approach in a total of 85 % assignments that is sufficient for structural analysis. However, data re-analysis revealed that the combinatorial approach alone along with the non-selective labelling experiments would have been enough to yield these 85 % assignment of ΔTehA. Expensive and time consuming single selective labelling experiments might therefore not be necessary for future assignments of IMPs. Accordingly, five combinatorial ^{15}N labelled samples of the C-terminal 16 kDa domain of the τ subunit of the *E. coli* DNA polymerase III holoenzyme yielded the same information as derived from 19 individual amino acid type selectively labelled samples (43). Given the considerable degree of signal overlap in IMPs it can be expected that the use of CF expression systems for the generation of combinatorial and non-specifically labelled samples in combination with ^{15}N-HSQC-TROSY spectra and their analysis by the differential HNCO strategy may significantly assist backbone assignments in the next future.

3.3. Applications of cell-free expression for solid state NMR

Solid state NMR (ssNMR) like the more frequent solution state NMR technique offers the possibility to study the atomic structure, exchange processes, spin diffusion and molecular dynamics of proteins. Additional information is provided on torsion angles, atomic orientation and very precise internuclear distances up to 15 Å by using the available chemical shift anisotropy, quadrupole coupling and strong dipolar couplings. The ssNMR technique is very versatile as no inherent size limit on molecules for investigation exists. However, ssNMR is hampered by fast relaxation, insensitivity (> 1µmol protein are required for 2D spectroscopy) and broader linewidth if compared with solution state NMR. Well resolved spectra containing only selected interactions can nevertheless be obtained using magic angle sample spinning (MAS) NMR recoupling techniques. These remove anisotropic interactions by fast sample rotation (in practice 5-20 kHz) about the magic angle (54.7° with respect to the magnetic field) and reintroduce selectively desired interactions using appropriate pulse sequences (47).

The ssNMR enables the study of proteins in a multitude of different states including frozen solution (48), protein aggregates (49), 2D crystals (50), 3D crystals (51) and in proteoliposomes (52). Proteoliposomes are clearly the samples of choice for IMPs as they are

closest to the proteins native environment. Lipid reconstituted samples also offer the option of investigating the dynamics of an IMP upon ligand binding (53) or in response to changes in the lipid environment (54). One of the key problems in the study of uniformly labelled large IMPs is spectral overlap. However, ssNMR is not only focused on structural biology but offers unique possibilities for functional biophysical studies and it has been widely used in order to study membrane bound peptides. Those studies require the efficient residue- and site-selective labelling of e.g. active sites in IMPs. For these labelling schemes, the CF expression system offers the so far largest possible flexibility. Preliminary ssNMR spectra of CF expressed and amino acid selectively labelled SMR proteins are shown in figure 6. CF generated ^{15}N amino acid specifically labelled samples of EmrE and its homologue Hsmr have been prepared as precipitate, frozen DDM micelles and as reconstituted protein in *E. coli* lipid mixture and analysed by ssNMR (Fig. 6I). Best resolution was obtained of the reconstituted Hsmr protein and some signals of individual amino acid residues are already visible. Among SMR proteins the residue Glu14 is highly conserved and suggested to be involved in substrate binding (28). A ^{13}C glutamate selective labelling of EmrE was performed by CF expression in order to monitor the active site (Fig. 6II). The protein was reconstituted in 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) liposomes and signal overlap due to the natural abundance of ^{13}C signals from the lipid and from side chains of the protein could be reduced by applying specific double-quantum filtering techniques. Only signals of the labeled Glu residues remain visible and possible applications of this approach are the observation of the glutamate protonation state and its modulation by substrate binding as well as precise distance measurements between the active site glutamate and bound substrates.

References

1. Martemyanov KA, Shirokov VA, Kurnasov OV, Gudkov AT & Spirin AS (2001) Cell-free production of biologically active polypeptides: Application to the synthesis of antibacterial peptide cecropin. *Protein Expr Purif* **21**, 456-461.
2. Alakov YB, Baranov VI, Ovodov SJ, Ryabova LA, Spirin AS & Morozov IJ. (1995) Method of preparing polypeptides in cell-free translation system in, U.S. Patent 5,478,730.
3. Kim DM & Choi CY (1996) A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. *Biotechnol Prog* **12**, 645-9.
4. Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T & Yokoyama S (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. *FEBS Lett* **442**, 15-9.
5. Kim DM & Swartz JR (1999) Prolonging cell-free protein synthesis with a novel ATP regeneration system. *Biotechnol Bioeng* **66**, 180-8.
6. Kim DM & Swartz JR (2000) Oxalate improves protein synthesis by enhancing ATP supply in a cell-free system derived from Escherichia coli. *biotechnol letters* **22**, 1537-1542.
7. Kim DM & Swartz JR (2000) Prolonging cell-free protein synthesis by selective reagent additions. *Biotechnol Prog* **16**, 385-90.
8. Jewett MC & Swartz JR (2004) Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol Bioeng* **86**, 19-26.
9. Jewett MC & Swartz JR (2004) Substrate replenishment extends protein synthesis with an in vitro translation system designed to mimic the cytoplasm. *Biotechnol Bioeng* **87**, 465-72.
10. Michel-Reydellet N, Calhoun K & Swartz J (2004) Amino acid stabilization for cell-free protein synthesis by modification of the Escherichia coli genome. *Metab Eng* **6**, 197-203.
11. Klammt C, Lohr F, Schafer B, Haase W, Dotsch V, Ruterjans H, Glaubitz C & Bernhard F (2004) High level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem* **271**, 568-80.
12. Torizawa T, Shimizu M, Taoka M, Miyano H & Kainosho M (2004) Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. *J Biomol NMR* **30**, 311-25.

13. Ishihara G, Goto M, Saeki M, Ito K, Hori T, Kigawa T, Shirouzu M & Yokoyama S (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expr Purif* **41**, 27-37.
14. Berrier C, Park KH, Abes S, Bibonne A, Betton JM & Ghazi A (2004) Cell-free synthesis of a functional ion channel in the absence of a membrane and in the presence of detergent. *Biochemistry* **43**, 12585-91.
15. Shimada Y, Wang ZY, Mochizuki Y, Kobayashi M & Nozawa T (2004) Functional expression and characterization of a bacterial light-harvesting membrane protein in Escherichia coli and cell-free synthesis systems. *Biosci Biotechnol Biochem* **68**, 1942-8.
16. Pornillos O, Chen YJ, Chen AP & Chang G (2005) X-ray structure of the EmrE multidrug transporter in complex with a substrate. *Science* **310**, 1950-3.
17. Kigawa T, Yabuki T, Matsuda N, Matsuda T, Nakajima R, Tanaka A & Yokoyama S (2004) Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. *J Struct Funct Genomics* **5**, 63-8.
18. Ozawa K, Jergic S, Crowther JA, Thompson PR, Wijffels G, Otting G & Dixon NA (2005) Cell-free protein synthesis in an autoinduction system for NMR studies of protein-protein interactions. *J Biomol NMR* **32**, 235-41.
19. Zubay G (1973) In vitro synthesis of protein in microbial systems. *Annu Rev Genet* **7**, 267-87.
20. Liu DV, Zawada JF & Swartz JR (2005) Streamlining Escherichia coli S30 extract preparation for economical cell-free protein synthesis. *Biotechnol Prog* **21**, 460-5.
21. Yamane T, Ikeda Y, Nagasaka T & Nakano H (2005) Enhanced cell-free protein synthesis using a S30 extract from Escherichia coli grown rapidly at 42 degrees C in an amino acid enriched medium. *Biotechnol Prog* **21**, 608-13.
22. Endo Y & Sawasaki T (2003) High-throughput, genome-scale protein production method based on the wheat germ cell-free expression system. *Biotechnol Adv* **21**, 695-713.
23. Elbaz Y, Steiner-Mordoch S, Danieli T & Schuldiner S (2004) In vitro synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. *Proc Natl Acad Sci U S A* **101**, 1519-24.
24. Klammt C, Schwarz D, Fendler K, Haase W, Dotsch V & Bernhard F (2005) Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *Febs J* **272**, 6024-38.

25. Garavito RM & Ferguson-Miller S (2001) Detergents as tools in membrane biochemistry. *J Biol Chem* **276**, 32403-32406.
26. Trbovic N, Klammt C, Koglin A, Lohr F, Bernhard F & Dotsch V (2005) Efficient strategy for the rapid backbone assignment of membrane proteins. *J Am Chem Soc* **127**, 13504-5.
27. Winstone TL, Duncalf KA & Turner RJ (2002) Optimization of expression and the purification by organic extraction of the integral membrane protein EmrE. *Protein Expr Purif* **26**, 111-21.
28. Gutman N, Steiner-Mordoch S & Schuldiner S (2003) An amino acid cluster around the essential Glu-14 is part of the substrate- and proton-binding domain of EmrE, a multidrug transporter from Escherichia coli. *J Biol Chem* **278**, 16082-7.
29. Massotte D (2003) G protein-coupled receptor overexpression with the baculovirus-insect cell system: a tool for structural and functional studies. *Biochim Biophys Acta* **1610**, 77-89.
30. Sarramegna V, Talmont F, Demange P & Milon A (2003) Heterologous expression of G-protein-coupled receptors: comparison of expression systems fron the standpoint of large-scale production and purification. *Cell Mol Life Sci* **60**, 1529-46.
31. Hulme EC & Curtis CA (1998) Purification of recombinant M1 muscarinic acetylcholine receptor. *Biochem Soc Trans* **26**, S361.
32. André N, Cherouati N, Prual C, Steffan T, Zeder-Lutz G, Magnin T, Pattus F, Michel H, Wagner R & Reinhart C (2006) *Protein Sci* **15**, 1115-1126.
33. Reilander H & Weiss HM (1998) Production of G-protein-coupled receptors in yeast. *Curr Opin Biotechnol* **9**, 510-7.
34. Stanasila L, Pattus F & Massotte D (1998) Heterologous expression of G-protein-coupled receptors: human opioid receptors under scrutiny. *Biochimie* **80**, 563-71.
35. Bannwarth M & Schulz GE (2003) The expression of outer membrane proteins for crystallization. *Biochim Biophys Acta* **1610**, 37-45.
36. Ye J & van den Berg B (2004) Crystal structure of the bacterial nucleoside transporter Tsx. *Embo J* **23**, 3187-95.
37. Kigawa T, Muto Y & Yokoyama S (1995) Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. *J Biomol NMR* **6**, 129-34.
38. Ozawa K, Dixon NE & Otting G (2005) Cell-free synthesis of 15N-labeled proteins for NMR studies. *IUBMB Life* **57**, 615-22.

39. Tugarinov V, Muhandiram R, Ayed A & Kay LE (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase. *J Am Chem Soc* **124**, 10025-35.
40. Chill JH, Louis JM, Miller C & Bax A (2006) NMR study of the tetrameric KcsA potassium channel in detergent micelles. *Protein Sci* **15**, 684-698.
41. Tugarinov V, Hwang PM & Kay LE (2004) Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu Rev Biochem* **73**, 107-46.
42. Krueger-Koplin RD, Sorgen PL, Krueger-Koplin ST, Rivera-Torres IO, Cahill SM, Hicks DB, Grinius L, Krulwich TA & Girvin ME (2004) An evaluation of detergents for NMR structural studies of membrane proteins. *J Biomol NMR* **28**, 43-57.
43. Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R & Choe S (2005) NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* **307**, 1317-21.
44. Parker MJ, Aulton-Jones M, Hounslow AM & Craven CJ (2004) A combinatorial selective labeling method for the assignment of backbone amide NMR resonances. *J Am Chem Soc* **126**, 5020-1.
45. Wu PS, Ozawa K, Jergic S, Su XC, Dixon NE & Otting G (2006) Amino-acid type identification in 15N-HSQC spectra by combinatorial selective 15N-labelling. *J Biomol NMR* **34**, 13-21.
46. Pervushin K, Riek R, Wider G & Wuthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* **94**, 12366-71.
47. Laws DD, Bitter HM & Jerschow A (2002) Solid-state NMR spectroscopic methods in chemistry. *Angew Chem Int Ed Engl* **41**, 3096-129.
48. Luca S, White JF, Sohal AK, Filippov DV, van Boom JH, Grisshammer R & Baldus M (2003) The conformation of neurotensin bound to its G protein-coupled receptor. *Proc Natl Acad Sci U S A* **100**, 10706-11.
49. Ritter C, Maddelein ML, Siemer AB, Luhrs T, Ernst M, Meier BH, Saupe SJ & Riek R (2005) Correlation of structural elements and infectivity of the HET-s prion. *Nature* **435**, 844-8.
50. Hiller M, Krabben L, Vinothkumar KR, Castellani F, van Rossum BJ, Kuhlbrandt W & Oschkinat H (2005) Solid-state magic-angle spinning NMR of outer-membrane protein G from Escherichia coli. *Chembiochem* **6**, 1679-84.

51. Lorch M, Fahem S, Kaiser C, Weber I, Mason AJ, Bowie JU & Glaubitz C (2005) How to Prepare Membrane Proteins for Solid-State NMR: A Case Study on the alpha-Helical Integral Membrane Protein Diacylglycerol Kinase from E. coli. *Chembiochem* **6**, 1693-1700.
52. Mason AJ, Siarheyeva A, Haase W, Lorch M, van Veen HW & Glaubitz C (2004) Amino acid type selctive isotope labelling of the multidrug ABC transporter LmrA for solid-state NMR studies. *FEBS Letters* **568**, 117-121.
53. Patching SG, Brough AR, Herbert RB, Rajakarier JA, Henderson PJ & Middleton DA (2004) Substrate affinities for membrane transport proteins determined by 13C cross-polarization magic-angle spinning nuclear magnetic resonance spectroscopy. *J Am Chem Soc* **126**, 3072-80.
54. Yamaguchi S, Tuzi S, Bowie JU & Saito H (2004) Secondary structure and backbone dynamics of Escherichia coli diacylglycerol kinase, as revealed by site-directed solid-state 13C NMR. *Biochim Biophys Acta* **1698**, 97-105.

Legends to the figures

Fig. 1. Cell-free expression modes. The CF expression of IMPs can be run in three different modes: IMPs can be expressed as precipitate (Mode A), solubilized in an appropriate detergent and further reconstituted into proteoliposomes. Alternatively the IMPs can be soluble expressed in presence of detergents (Mode B), purified and reconstituted into proteoliposomes, or expressed in presence of preformed liposomes (Mode C) and isolated from the RM by density gradient centrifugation. CD, circular dichroism; EM, electron microscopy; NMR, nuclear magnetic resonance.

Fig. 2. Overproduction of integral membrane proteins by cell-free expression. SDS-polyacrylamide gel analysis of IMPs from different families that have been CF produced either as precipitate or in soluble form in presence of detergents. 1 µl samples of the soluble or of the resuspended insoluble part of the RM were analyzed on 10 % tricine SDS gels (lanes 1-3, 10), 12 % SDS gels (lanes 4, 13-16) or 16.5 % SDS gels (lanes 5-9, 11, 12). The overproduced IMPs are indicated by arrows, proposed secondary structures are illustrated on top and detergents are given in parenthesis if applicable. Lane 1, EmrE; lane 2, SugE; lane 3, Tbsmr; lane 4, Hsmr (DPC) purified by Ni-NTA and additional arrows indicate oligomers; lane 5, RM control; lane 6, YfiK (Brij-58); lane 7, YfiK (Brij-98); lane 8, YfiK, lane 9, TehA, lane 10, ΔTehA; lane 11, SecE; lane 12, Tsx; lane 13, porcine V2R; lane 14, human V2R; lane 15, rat CRF; lane 16, human ETB. P, precipitate, S, soluble expressed, p, purified.

Fig. 3. Cell-free expression of bacterial glutamate transporter. CF expression of different eukaryotic glutamate transporter homologues as precipitate. After CF reaction, the IMP pellets were suspended in a volume equal to the RM and 5 µl samples were analyzed on a 12 % SDS gel. Sizes of marker proteins (kDa) are indicated. The arrows indicate the synthesized glutamate transporter homologues.

Fig. 4. Liquid-state NMR of cell-free produced transporters. ^{1}H-^{15}N TROSY-HSQC spectra of CF expressed ^{15}N labelled IMPs. The proteins were expressed as precipitates and solubilized in LMPG. A, U-^{15}N labelled SecE; B, U-^{15}N labelled YfiK; C, U-^{15}N labelled ΔTehA; D, ^{15}N-threonine labelled ΔTehA; E, ^{15}N-tryptophan labelled ΔTehA; F, ^{15}N-isoleucine labelled ΔTehA; G, ^{15}N-methionine labelled ΔTehA; H, ^{15}N-phenylalanine labelled ΔTehA; I, ^{15}N-arginine labelled ΔTehA. The protein precipitates were dissolved in 25 mM

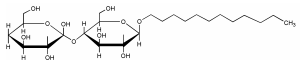
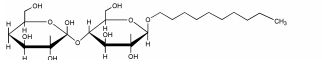
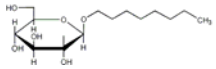
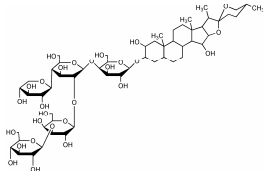
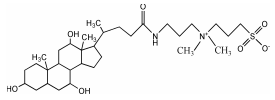
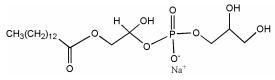
potassium phosphate buffer (pH 6.0) containing 5 % LMPG (A, B, D-I) or dissolved in LMPG, purified by Ni-NTA chromatography and equilibrated in 20 mM MES/Bis-Tris (pH 6.0) with 2 % LMPG (C). The spectra were taken with protein concentrations of 1 mM for SecE, 1 mM for YfiK, 0.1 mM for ΔTehA and 0.9 mM for selectively ΔTehA samples (D-I) at 40°C on a 600 MHz (A), 700 MHz (B), 900 MHz (C) or on a 800 MHz NMR spectrometer (D-I), all equipped with cryogenic ^{1}H [^{13}C/^{15}N] triple-resonance probes.

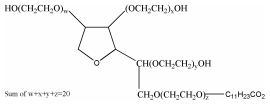
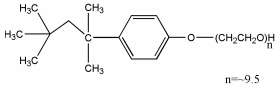
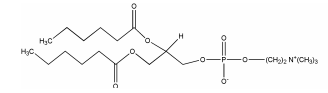
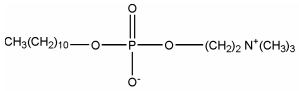
Fig. 5. Combinatorial labelling approach. Example of a combinatorial labelling scheme with three differentially ^{15}N and ^{13}C labelled samples of the tellurite transporter ΔTehA. The labelling scheme is shown in the table on top. In the [^{15}N, ^{1}H]-TROSY spectra, backbone amide protons of the ^{15}N-labelled amino acids are visible. The crossed lines indicate the peak to become identified. The indicated peak at position 117.5/7.78 is identified as phenylalanine as it is present in samples 1 and 2, but not in sample 3. The corresponding HNCO spectra show the amide crosspeaks after carbonyl transfer (big arrow) and indicate that the preceding residue of this phenylalanine must be a glycine, as no crosspeaks are visible in sample 2 and 3 without ^{13}C-glycine. The analysed phenylalanine residue can now be localized in a Gly-Phe pair in the primary sequence which in that example was identified as Phe97 of ΔTehA. All spectra were recorded at a Bruker Avance 600 MHz spectrometer.

Fig. 6. Solid-state NMR spectroscopy of cell-free produced multidrug transporter. **I.** ^{15}N-cross polarisation (CP)-MAS spectra of ^{15}N-Leu-EmrE directly after CF synthesis (A), after solubilization in DDM (B) and of ^{15}N-Phe-Hsmr after CF expression, solubilization in DDM and reconstitution into *E. coli* lipid liposomes. Experiments were performed at a 5 kHz sample spinning rate at 253 K using a Bruker 7mm MAS probe at 40.5 MHz for EmrE and at 8 kHz sample spinning rate at 230 K using a Bruker 4 mm MAS probe at 60.84 Hz for Hsmr. The CP contact time was 1.5 ms, 50.000 scans were accumulated and for Hsmr the CP contact time was 0.75 ms and 30.000 scans were accumulated.

II. ^{13}C-CP-MAS spectra of ^{13}C-Glu-EmrE in DOPC. (A) After reconstitution into DOPC, most resonances are obscured by the overall naturally abundant ^{13}C signals of lipid and protein. (B) Applying double-quantum filtering techniques, natural abundance background can be suppressed allowing selective observation of the labelled sites (Reproduced with permission from M. Lorch, I. Lehner, A. Siarheyeva, D. Basting, N. Pfleger, T. Manolikas and C. Glaubitz, 2005, Biochemical Society Transactions, 33, 873-877).

Table 1. Efficiency of selected detergents for the cell-free production of membrane proteins

Detergent	Solubilization[1] conc. (%)[2]	CF expression[1] conc. (%)[2]	CMC mM (%)[3]	Reference
Alkyl-glucosides				
DDM 	I-III (2.0)	I-III (0.1)	0.12 (0.007)	13, 14, 23, 24
DM 	n.d.	I-III (0.2)	1.8 (0.9)	24
ß-OG 	I (2.0)	0-I (0.8)	19 (0.58)	13, 14, 24
Steroid-derivatives				
Digitonin 	I (2.0)	II-III (0.8)	0.73 (0.09)	13, 24
CHAPS 	0-III (2.0)	0-II (0.75)	5.5 (0.5)	13, 14, 24
Long chain-phosphoglycerols				
LMPG 	III (2.0)	I-II (0.01)	0.05 (0.002)	24
Polyoxyethylene-alkyl-ether				
Brij-35 $C_{12}H_{20}(O\text{-}CH_2CH_2)_nOH$ n= ~23	I (2.0)	II-III (0.8)	0.08 (0.01)	13, 24
Brij-58	n.d.	III (1.5)	0.075 (0.009)	14, 24

$C_{16}H_{33}(O\text{-}CH_2CH_2)_nOH$ n= ~20				
Tween20 	0-II (2.0)	0-III (0.08)	0.059 (0.007)	13, 14, 24
Polyethylene-glycol derivatives				
TX-100 	I-III (2.0)	II-III (0.2)	0.23 (0.015)	13-15, 24
Thesit $C_{12}H_{25}(O\text{-}CH_2CH_2)_nOH$ n= ~4	0-II (2.0)	0-II (0.1)	0.1 (0.006)	24
Bi-/Mono-chain-phosphocholines				
DHPC 	I-II (2.0)	I-III (0.2)	1.4 (0.07)	24
DPC 	I-III (2.0)	0-I (0.1)	1.2 (0.05)	24

[1] the efficiency of solubilization of precipitates or for the soluble expression of IMPs is given according to the following rating: III, high; II, intermediate; I, poor; 0, not suitable; n.d., not determined. Ranges of ratings result from evaluations in combination with different IMPs.

[2] highest applied detergent concentrations according to the cited references (%).

[3] concentration in % that corresponds to 1 x CMC.

Fig. 1

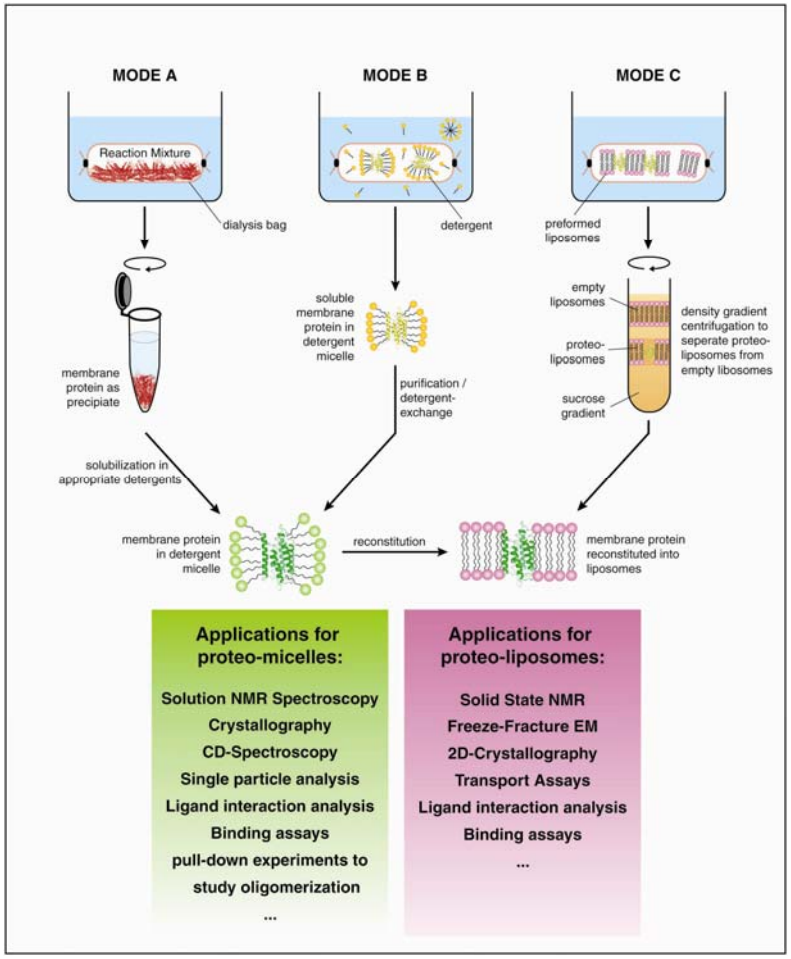


Fig. 2

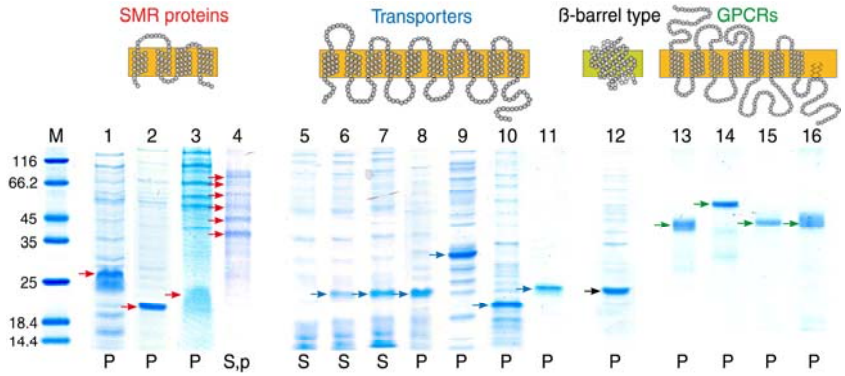


Fig. 3

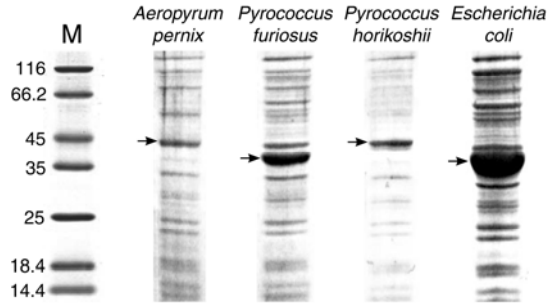


Fig. 4

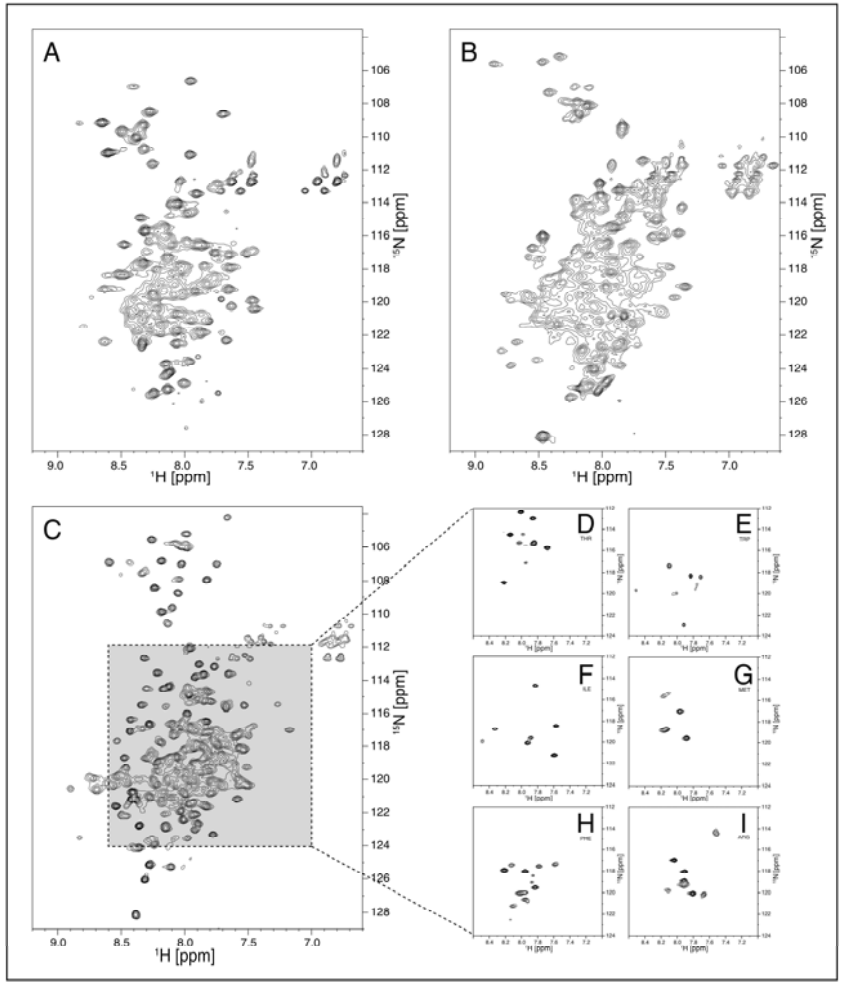
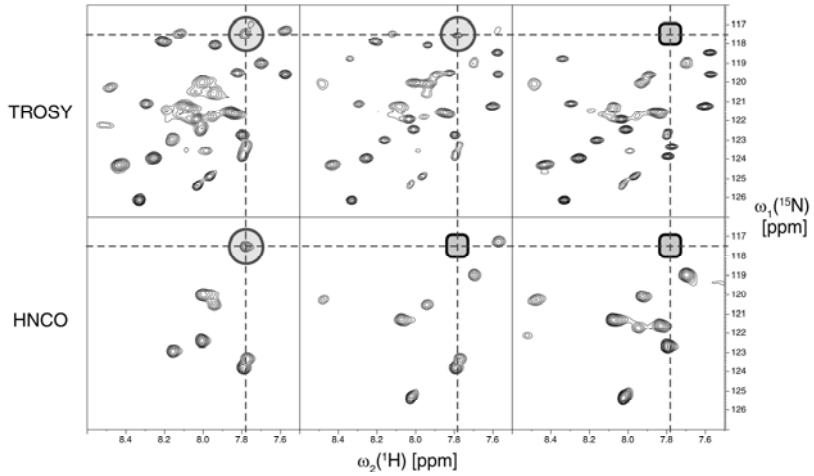


Fig. 5

amino acid	sample 1	sample 2	sample 3
Ala	^{15}N	^{15}N	^{15}N
Phe	^{15}N	^{15}N	
Ile		^{15}N	^{15}N
Ser	^{13}C	^{13}C	
Leu		^{13}C	^{13}C
Gly	^{13}C		
Val			^{13}C



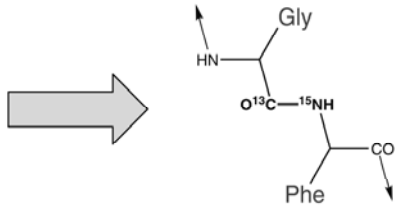
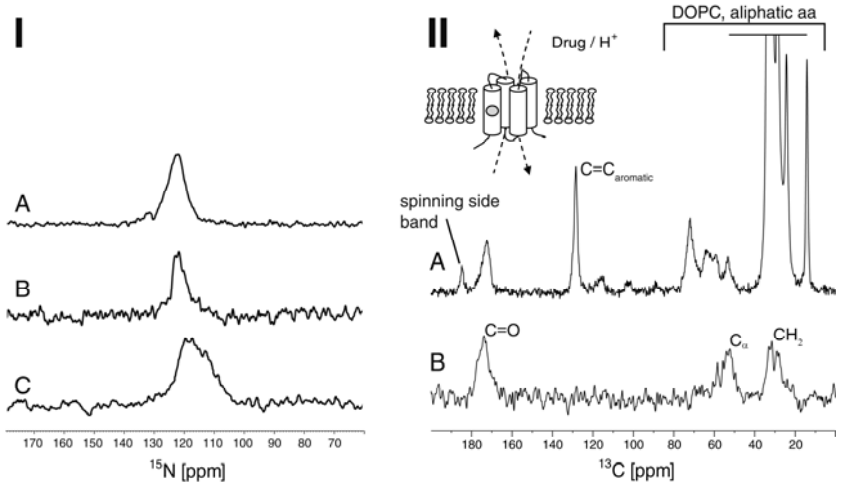


Fig. 6



(P 06)

Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies. D. Schwarz, C. Klammt, A. Koglin, F. Löhr, B. Schneider, V. Dötsch and F. Bernhard. *Methods*, in press. (2006)

Available online at www.sciencedirect.com
ScienceDirect
Methods xxx (2006) xxx–xxx
METHODS
www.elsevier.com/locate/ymeth

Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies

Daniel Schwarz, Christian Klammt, Alexander Koglin, Frank Löhr, Birgit Schneider, Volker Dötsch, Frank Bernhard *

Centre for Biomolecular Magnetic Resonance, University of Frankfurt/Main, Institute for Biophysical Chemistry, Marie-Curie-Str. 9, D-60439 Frankfurt/Main, Germany

Accepted 13 July 2006

Abstract

Cell-free expression techniques have emerged as promising tools for the production of membrane proteins for structural and functional analysis. Elimination of toxic effects and a variety of options to stabilize the synthesized proteins enable the synthesis of otherwise difficult to obtain proteins. Modifications in the reaction design result in preparative scale production rates of cell-free reactions and yield in milligram amounts of membrane proteins per one millilitre of reaction volume. A diverse selection of detergents can be supplied into the reaction system without inhibitory effects to the translation machinery. This offers the unique opportunity to produce a membrane protein directly into micelles of a detergent of choice. We present detailed protocols for the cell-free production of membrane proteins in different modes and we summarize the current knowledge of this technique. A special emphasize will be on the production of soluble and functionally folded membrane proteins in presence of suitable detergents. In addition, we will highlight the advantages of cell-free expression for the structural analysis of membrane proteins especially by liquid state nuclear magnetic resonance spectroscopy and we will discuss new strategies for structural approaches.

Keywords: Integral membrane proteins; Structural analysis; Solubilization; Reconstitution; Detergent; NMR spectroscopy; G-protein coupled receptors; β-barrel proteins; Multidrug resistance; Cell-free expression; Detergent micelles; Transporter; Liposomes; Labelling of proteins; Stable isotopes

1. Introduction

Membrane proteins (MPs) define the link between physiological pathways in the cytoplasm and the extracellular environment. Essential processes like perception and transduction of external signals, import or export of substances through the membrane or the generation of energy is associated with MPs. Many modern drugs are directed against MPs and this class of proteins is therefore an important target for medical and pharmaceutical research. However, high resolution structures of MPs are still the very exception. A major bottleneck for structural analysis is the limited availability of sufficient amounts of protein samples. The bacterial *Escherichia coli* expression system is most frequently used for the production of recombinant proteins. Its simplicity, low costs, the wealth of elaborated protocols, the fast growth rates and often high productivity makes this system highly competitive against most other expression systems. However, *in vivo* expression systems based on prokaryotic or eukaryotic cells do not work for a wide range of MPs, toxins or other problematic targets [1]. MPs often affect the physiology of the cell by insertion into the cellular membranes or by blocking protein trafficking systems. Low expression rates, aggregation or unfolding of the recombi-

* Corresponding author. Fax: +49 69 798 29632.
E-mail address: fbern@bpc.uni-frankfurt.de (F. Bernhard).

1046-2023/$ - see front matter © 2006 Published by Elsevier Inc.
doi:10.1016/j.ymeth.2006.07.001

nant MPs and even toxic effects to the host cells upon overproduction are therefore frequent problems when MPs have to be produced [2].

Cell-free (CF)[1] protein synthesis provides a recently developed and powerful alternative tool for protein production [3–5]. A unique advantage of CF systems is the open access to the reaction at any time of the experiment. This enables the addition of beneficial compounds at defined concentrations and at any stage of the protein synthesis. No membranes have to be penetrated and no selection of substances occurs by specific transport systems. Metabolic conversion or even degradation of added substances is furthermore reduced due to the restricted enzymatic activity of the CF extracts. The only limitation is that the supplemented compounds must not affect the transcription and translation machinery of the expression system. Degradation of proteins or nucleic acids could be prevented by addition of corresponding inhibitor cocktails. Addition of chaperones like the GroEL/ES or DnaK/J-GrpE systems could facilitate the folding efficiency of heterologous proteins. Specific cofactors, substrates or inhibitors might help to stabilize the synthesized recombinant proteins. The addition of detergents or lipids enables the direct translation of MPs into defined hydrophobic environments. Disulfide bridges that are essential for the functional folding of many eukaryotic proteins are likely to become formed due to the easy access of oxygen to the CF reaction [6–8].

The elimination of cytotoxic effects is presumably one of the major reasons for the rapidly increasing number of diverse MPs of prokaryotic and eukaryotic origin that can be produced in CF systems [9–14]. While still limited, the reported examples of CF produced MPs already comprises α-helical and β-barrel type MPs of prokaryotic as well as of eukaryotic origins. One common characteristic is their almost complete membrane integrated topology with only small proposed external loop regions. Especially MPs that will not or only at minor levels be synthesized in living *E. coli* cells might therefore become targets for CF expression. This feature highlights CF expression as a promising future technique for the high level production of otherwise difficult to obtain MPs. Moreover, additional beneficial characteristics like various options to protect and to stabilize recombinant proteins, the possibility to translate MPs into preformed micelles and the considerable advantages upon specific labelling approaches of proteins make CF expression as one of the currently most versatile techniques for the preparative scale production of proteins.

This review summarizes the preparative scale production of MPs in CF systems based on *E. coli* extracts. We will emphasize rather on the set-up of individual CF expression systems than on commercially available systems. To our knowledge, there are currently no reported examples of the preparative scale production of MPs in CF systems based on wheat germ extracts. However, these systems might play an important role in the near future, especially for the synthesis of functionally folded eukaryotic MPs. Therefore, a short overview about the key steps in wheat germ extract preparation and in the reaction design will be provided.

2. Preparative scale cell-free expression systems

2.1. Configuration and productivity of CF-systems

First generation CF expression systems have been batch-formatted reactions containing all compounds in one compartment. The rapid depletion of precursors in combination with the accumulation of inhibitory breakdown products resulted in short reaction times of less than 1 h and consequently in only low product yields of often not more than several micrograms of recombinant protein per one ml of reaction [15]. CF expression systems have therefore been used for a long time only for the analytical scale production of proteins. The splitting of the CF system into a reaction mixture (RM) holding all high molecular weight compounds and into a feeding mixture (FM) containing the low molecular precursors provided the basis for several new reaction designs with considerably improved efficiencies [16]. A common characteristic of preparative scale CF expression systems is the extended supply of fresh precursors combined with the continuous removal of deleterious reaction by-products like pyrophosphate. The reaction times are extended up to approx. 20 h and allow the synthesis of several milligram amounts of protein per 1 ml of RM [17,18]. A frequently used reaction design for the high level production of proteins is the continuous-exchange cell-free (CECF) system [16]. The fixed volume compartments of

[1] *Abbreviations:* AcP, acetyl phosphate; Brij35, polyoxyethylene-(23)-lauryl-ether; Brij56, polyoxyethylene-(10)-cetyl-ether; Brij58, polyoxyethylene-(20)-cetyl-ether; Brij72, polyoxyethylene-(2)-stearyl-ether; Brij78, polyoxyethylene-(20)-stearyl-ether; Brij97, polyoxyethylene-(10)-oleyl-ether; Brij98, polyoxyethylene-(20)-oleyl-ether; CECF, continuous-exchange cell-free; CF, cell-free; CHAPS, 3-[(3-cholamidopropyl)dimethyl ammonio]-1-propansulfonat; CMC, critical micellar concentration; $C_{12}E_8$, polyoxyethylene-(8)-lauryl-ether; DHPC, 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine; diC_6PC, 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine; diC_8PC, 1,2-dioctanoyl-*sn*-glycero-3-phosphocholine; DDM, *n*-dodecyl-β-D-maltoside; DM, *n*-decyl-β-D-maltoside; DMPC, di-myristoyl-phosphatidyl-choline; DPC, dodecyl-phosphocholine; FID, free induction decay; FM, feeding mixture; Genapol C 100, polyoxyethylene-(10)-dodecyl-ether; Genapol X 100, polyoxyethylene-(10)-isotridecyl-ether; GPCR, G-protein coupled receptor; HECAMEG, (6-*O*-(*N*-heptylcarbamoyl)-methyl-α-D-glucopyranoside); HSQC, heteronuclear single quantum correlation; LMPG, 1-myristoyl-2-hydroxy-*sn*-glycerol-3-[phosphor-rac-(1-glycerol)]; LPPG, 1-palmitoyl-2-hydroxy-*sn*-glycerol-3-[phosphor-rac-(1-glycerol)]; MM, master mixture; MP, membrane protein; MWCO, molecular weight cut-off; NG, *n*-nonyl-β-D-glucoside; NMR, nuclear magnetic resonance; NP40, nonylphenyl-polyethylene-glycol; NTP, nucleotide triphosphate; β-OG, *n*-octyl-β-D-glucopyranoside; OMP, outer membrane protein; PCR, polymerase chain reaction; PEP, phosphoenol pyruvate; RM, reaction mixture; SDS, sodium-dodecyl-sulfate; TB, terrific broth; Thesit, polyethylene-glycol 400 dodecylether; TMS, transmembrane segment; Triton X-100, polyethylene-glycol P-1,1,3,3-tetra-methyl-butylphenyl-ether; TROSY, transverse relaxation optimized spectroscopy; Tween20, polyoxyethylene-sorbitan-monolaurate 20; Tween40, polyoxyethylene-sorbitan-monopalmitate 20; Tween60, polyoxyethylene-sorbitan-monostearate 20; Tween80, polyoxyethylene-sorbitan-monoleate 20; UTR, untranslated region.

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

RM and FM are separated by a semipermeable membrane with molecular weight cut-offs (MWCOs) between 10 and 50 kDa that ensures an efficient exchange of compounds (Fig. 1). The individual components of the translation machinery obviously stick together in a macromolecular complex and despite the relatively high MWCO of the membrane, no significant leakage of translation factors is noticed. The reaction is incubated with intensive agitation like stirring, shaking or rotating in order to provide an optimal exchange between the two compartments. Commercial CECF systems (Rapid Translation System (RTS) Roche Diagnostics, Penzberg, Germany) as well as individually prepared systems are highly productive [12,13,17]. However, it should be considered that several parameters of the reaction like ion concentrations, the composition of the energy system and the amino acid pool or even the buffer system can be subject of intensive optimization steps before the high level expression of a new protein target is achieved.

CF lysates are mostly prepared from *E. coli* cells, wheat germs and to a lesser extent from rabbit reticulocytes [4,19,20]. Eukaryotic backgrounds might be preferred for the expression of eukaryotic proteins to provide the optimal environment for their functional folding and to enable posttranslational modifications. Low levels of endogenous mRNAs in wheat germ extracts allow their use in pure translation systems with added mRNA as template for translation [20]. Extended reaction times up to 60 h can yield in 1–4 mg of recombinant protein per 1 ml RM [21]. Most popular is the S30 extract of *E. coli* that contains the soluble fraction of cell lysates after centrifugation at 30,000*g* and including all enzymes necessary for transcription and translation [4]. CF expression systems based on *E. coli* extracts are almost exclusively used as coupled transcription/translation systems by providing DNA as template in combination with the highly specific and efficient T7-RNA polymerase [22].

For specific purposes, the *E. coli* translation machinery could alternatively be reconstituted almost completely *in vitro* by combining the individually purified protein components to isolated ribosomes [23]. All aminoacyl-tRNA-synthetases and translation factors can be overproduced separately in standard *E. coli* expression systems, purified by virtue of terminal poly(His)$_6$-tags and added to *E. coli* S100 extracts containing the relatively pure ribosome fraction. This PURE system (protein synthesis using recombinant elements) enables the CF protein synthesis under defined conditions and allows detailed studies of folding pathways or translation kinetics.

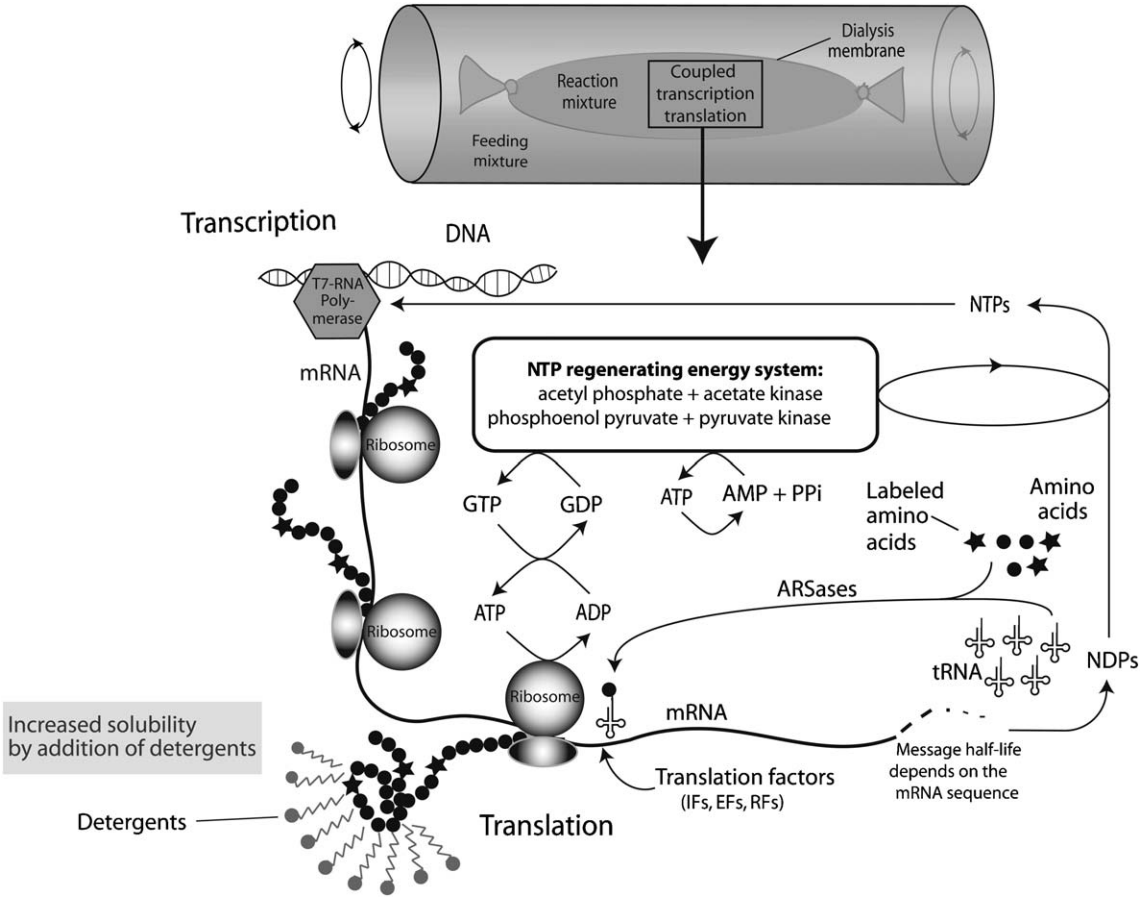


Fig. 1. Schematic configuration of a coupled transcription/translation reaction in a CECF system. The CF reaction can be carried out in simple dialysis tubes placed into suitable plastic vials that hold the FM. The complete set-up is incubated e.g., on a turning device that ensures continuous agitation of the reaction and substance exchange between the two compartments.

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

2.2. Preparation of cell-free lysates

CF lysates provide all the high molecular weight components of the translation machinery. Endogenous low molecular weight substances like amino acids and salts will be removed by extensive dialysis during the preparation procedure. A "run-off" step is furthermore implemented in order to eliminate endogenous cellular mRNA. High salt concentrations cause the dissociation of the ribosomes from endogenous mRNA that will then subsequently become degraded due to the high RNAse content of the extract.

Escherichia coli extracts are relatively easy and fast to prepare and the individual steps of standard protocols include cell fermentation, cell disruption, run-off procedure and buffer exchange by dialysis [4,24]. Common sources for CF extracts are *E. coli* BL21 derivatives or strains devoid of major endogenous RNAses like A19 (Table 1). The cells have to be grown with good aeration until mid-log phase at 37 °C in rich medium like terrific broth (TB), chilled down rapidly and harvested by centrifugation. The time of harvest is somehow crucial and corresponds in TB medium to an OD_{595} of approximately 3.5. Exceeding the optimal time point of harvest can drastically reduce the efficiency of the final CF extract. Rapid chilling of the culture down to below 10 °C upon harvesting stalls further growth and conserves the active state of the cellular physiology. A 10 L fermenter with TB medium should yield 50–70 g wet-weight of bacterial cells. The cell pellet is resuspended and washed three times in ice cold S30-A buffer and it is finally suspended in S30-B buffer pre-cooled at 4 °C (Table 1). The cells should be disrupted by passing through a pre-cooled French-Press and not by sonification, as this treatment could cause the disintegration of ribosomes. Cell-debris is removed by centrifugation at 30,000*g* at 4 °C for 30 min and the upper 2/3rd of the supernatant are transferred into a fresh vial. The centrifugation step and transfer of supernatant is repeated once. For the "run off" step the lysate is adjusted to a final concentration of 400 mM NaCl and incubated at 42 °C for 45 min in a water bath [12]. Besides the elimination of endogenous mRNA, this treatment causes a considerable precipitate. The turbid solution is filled into a dialysis tube (MWCO 14 kDa) and dialyzed at 4 °C against 60 volumes of S30-C buffer with gentle stirring. After one further exchange of the dialysis buffer the *E. coli* S30-extract is harvested by centrifugation at 30,000*g* at 4 °C for 30 min. The clear supernatant is transferred in suitable aliquots into plastic tubes and frozen in liquid nitrogen. The final protein concentration in the extract should be between 30–50 mg/ml and could be adjusted by ultrafiltration. The complete protocol should yield some 50 ml of CF extract out of a 10 L fermentation. We recommend to finish

Table 1
Materials, buffers and substances for the cell-free expression of membrane proteins

Source for S30 *E. coli* lysates:
A19 [*rna19 gdh A2 his95 relA1 spoT1 metB1*] *E. coli* Genetic Stock Center (*E. coli* Genetic Stock Center, New Haven, USA, CGSC No. 5997)
BL21 star [*F-ompT hsdS B* (r_B–m_B-) *gal dcm rne131*] Invitrogen, Karlsruhe, Germany, C6010-03

Materials for S30 extract preparation:
Fermenter
French Press cell disruption device
Dialysis tubes (MWCO 14 kDa)
Devices for protein concentration by ultrafiltration

CF reaction container:
Microdialysers (MWCO 15–25 kDa) (Spectrum Labs, Rancho Dominguez, USA)
Dispodialysers (MWCO 15–25 kDa) (Spectrum Labs, Rancho Dominguez, USA)
Shaking, rolling or stirring device; e.g. Universal Turning Device (Vivascience, Göttingen, Germany; Cat. No. IV-76001061) placed in an incubator with temperature control
Standard glass or plastic vials

Chemicals for S30 extract preparation and set-up for cell-free expression:
Adenosine-5′-triphosphate disodium salt, acetyl-phosphate, amino acids (Sigma–Aldrich, Taufkirchen, Germany), complete mini protease inhibitor (Roche Diagnostics GmbH, Mannheim, Germany), cytidine 5′-triphosphate disodium salt, dithiothreitol, ethylenediamine-tetraaceticacid, folinic acid, guanosine 5′-triphosphate disodium salt, HEPES, KCl, KOAc, $Mg(OAc)_2$, liquid N_2, NaN_3, NaCl, β-mercaptoethanol, phenylmethane-sulfonylfluoride, phosphoenol-pyruvate, polyethyleneglykol 8000, pyruvate kinase, RNAsin (Amersham Biosciences, Freiburg, Germany), total *E. coli* tRNA (Roche Diagnostics GmbH, Mannheim, Germany), T7-RNA polymerase, uridine 5′-triphosphate trisodium salt

Selected detergents:
α-[4-(1,1,3,3-tetramethylbutyl)-phenyl]-ω-hydroxy-poly(oxy-1,2-ethandiyl) (Triton X-100), digitonin, polyethylenglycododecylether (Brij35), polyoxyethylene-(20)-cetyl-ether (Brij58); polyoxyethylene-(20)-stearyl-ether (Brij78); polyoxyethylene-(20)-oleyl-ether (Brij98); (all Sigma–Aldrich, Taufkirchen, Germany); *n*-dodecylphosphocholine (DPC), 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phosphor-rac-(1-glycerol)] (LMPG), (Avanti-Lipids, Alabaster, USA); *n*-dodecyl-β-D-maltoside (DDM); (Glycon Biochemicals, Luckenwalde, Germany)

Buffers/solutions:
S30-A buffer: 10 mM Tris–acetate, pH 8.2, 14 mM $Mg(OAc)_2$, 0.6 mM KCl, 6 mM β-mercaptoethanol
S30-B buffer: 10 mM Tris–acetate, pH 8.2, 14 mM $Mg(OAc)_2$, 0.6 mM KCl, 1 mM DTT, 0.1 mM phenylmethane-sulfonylfluoride
S30-C buffer: 10 mM Tris–acetate, pH 8.2, 14 mM $Mg(OAc)_2$, 0.6 mM KOAc, 0.5 mM DTT. TB-medium (per litre): 24 g yeast extract, 12 g tryptone, 4 ml 100% glycerol, 100 mM potassium phosphate buffer, pH 7.4

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

the complete process consecutively but the interruption after cell harvesting is also possible.

Expression platforms based on eukaryotic cell extracts provide a higher stability of mRNA with reaction times of several days [21]. However, final yields of recombinant protein are still similar if compared to CF systems based on *E. coli* extracts. A major disadvantage is the relatively complicated and time-consuming extract preparation procedure and high variations in the quality of different extract batches. Only a short overview of key steps in wheat germ extract preparation is given below and more extended protocols are available in the literature ([21,25]).

Winter wheat is preferred as source for extract preparation. Fractions containing wheat embryos are isolated from ground seeds by flotation in a mixture of carbon tetrachloride and cyclohexane. Magnifier lenses have to be used to remove damaged embryos and distinct parts of endosperm that will inhibit translation by presence of various inhibitory proteins like e.g., ricin [26]. This step is highly critical and it is the most laborious part. The purified embryos are ground to a fine powder in liquid nitrogen, resolubilized in extraction buffer (40 mM Hepes–KOH, pH 7.6, 100 mM KOAc 5 mM $Mg(OAc)_2$, 2 mM $CaCl_2$, 4 mM DTT, 0.3 mM of each of the 20 amino acids) and pelleted by centrifugation. The supernatant is applied on a PD-10 column pre-equilibrated with extraction buffer to remove inhibitory low molecular weight substances from the extract. The extract is finally concentrated to an A_{280} of at least 200/ml and stored in aliquots at −80 °C [27].

2.3. Design of DNA templates for cell-free expression

The transcription in *E. coli* coupled transcription/translation CF systems is operated by the phage T7-RNA polymerase. The purified enzyme has to be added into the RM at relatively high final concentrations between 4–10 U/μl and detailed protocols for the overproduction and purification of the T7-RNA polymerase have been published [28]. The promotor elements of the target gene have to meet the specific requirements of the T7-RNA polymerase. The ribosomal binding site has to be present in optimal distance to the translational start codon and a transcriptional terminator should be placed 3′ to the reading frame to prevent excessive consumption of NTP precursors. Some suitable commercial vector series are pIVEX (Roche Diagnostics, Penzberg, Germany), pDEST (Invitrogen, Carlsbad, USA) or pET (MerckBioscience, Darmstadt, Germany). These vector systems offer furthermore the option to add a variety of terminal tags to the target protein that might increase the protein expression or that could facilitate purification and detection strategies.

The translation efficiency of mRNA templates in wheat germ systems strongly depends on the 5′- and 3′- untranslated regions (UTRs) [29]. Eukaryotic mRNAs are modified after translation with a 5′ cap (5′7 mGpppG) and a 3′-polyadenylated tail that prevents degradation and results in a better protein expression. The supply of pre-modified mRNAs to CF reactions is useless because of the efficient deadenylation and decapping activity of extract enzymes and higher concentrations of pre-modified mRNAs are even inhibitory for the translation. This problem could be addressed by several strategies like the use of phosphothionate mRNA [30], the modification of the 3′-end of mRNA with adaptor DNA [31] and the immobilization of mRNA on latex beads [32]. Alternatively, different UTRs could be used. Viral UTRs are generally good translation enhancers in the case of the cap-independent initiation. Efficient 5′-UTRs usually contain an increased A/T content including motifs such as (AAC)*n*, (AAAC)*n* or (AAAAC)*n*. Most popular is currently the OMEGA leader derived from the tobacco mosaic virus [33]. The 3′-UTRs stabilize mRNA by the formation of complex secondary structures and also viral 3′-UTRs are frequently used [34].

2.4. Linear DNA as a template for cell-free expression

The possibility to use linear templates generated by PCR in the CF-system eliminates time consuming cloning/subcloning steps and allows the rapid screening of a variety of expression constructs [5]. PCR products can furthermore be directly used for CF expression without prior purification [35]. Multiple-step PCR protocols or the split primer PCR-technique have to be employed to add the required relatively long regulatory elements like T7-promoter and terminator to the coding sequence. Several strategies have been established in order to overcome the degradation of linear DNA templates by endogenous nucleases. The Lambda phage Gam protein is an inhibitor of exonuclease V (*recBCD*) and its addition to the CF reaction stabilized linear DNA [36]. Extract preparation from an engineered *E. coli* strain devoid of the *endA* gene and containing a modified *recBCD* operon resulted in protein yields comparable to that obtained with plasmid templates in batch systems [37]. Stem loop structures at the 3′-end of mRNAs also help to reduce exonuclease mediated degradation [38]. CF expression is predestined for high throughput (HT) applications especially by usage of PCR generated DNA templates [38–40].

2.5. Reaction conditions of E. coli cell-free expression systems

CF expression can be performed in small analytical scale reactions with approximately 50–100 μl RM for optimization reactions and in larger preparative scale reactions of 1–2 ml RM for the production of protein. It should be considered that for almost each new protein target distinct parameters like the concentrations of critical compounds or the generation of an optimal expression construct might need to be optimized in order to find the best conditions for high level expression. The CECF reaction has a very well defined optimum for Mg^{2+} and K^+ ions usually between 13–15 and 280–300 mM, respectively. For each new protein target and also for each new batch of extract at least the

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

concentrations of these two compounds should be optimized for best expression rates. A RM/FM ratio between 1:10 and 1:20 for analytical scale reactions and commercially available microdialysers (Spectrum Laboratories, Rancho Dominguez, USA) with a MWCO between 15 and 25 kDa as a reaction device can be recommended. Expression levels of several milligrams of protein per 1 ml RM can only be obtained if all system components are in optimal conditions. The results can be scaled up into preparative reactions of 1–2 ml RM without significant loss of efficiency. Dispodialyser (Spectrum Laboratories, Rancho Dominguez, USA) or even simple dialysis tubes are recommended as preparative scale reaction containers (Fig. 1).

All stock solutions (Table 2) should be mixed carefully after thawing and the enzymes, tRNA and the S30 extract should be kept on ice. A master-mix (MM) including all shared components of FM and RM should be prepared by first pipetting the higher volume components (e.g. KOAc; $MgO(Ac)_2$; PEG8000; NaN_3). First the FM is completed and pre-incubated in a water-bath at 30 °C. Then the RM is completed, mixed gently and kept on ice. The appropriate volume of FM is filled into standard plastic or glass vials that can be used as FM compartments. The RM is transferred into suitable dispodialyser or washed dialysis tubes (Fig. 1). Air-bubbles that might restrict an efficient exchange between the two compartments should be avoided. The reaction has to be incubated with intensive agitation either on rolling or shaking devices or by magnetic stirring. The incubation temperature is usually between 20 °C and 30 °C and protein synthesis continues up to 20 h (Fig. 2). A RM/FM ratio of 1:17 can already result in the production of several milligrams of protein per 1 ml RM, but the final yields of recombinant protein per reaction can easily be increased by using higher ratios or by refreshing the FM after certain times of incubation.

Due to their low internal RNAse content, CF reactions based on wheat germ extracts can be operated as translation system with supplied mRNA as a template. Transla-

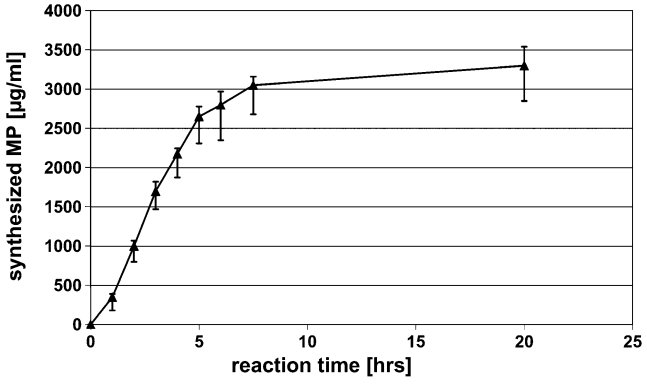


Fig. 2. Kinetics of cell-free production of the nucleoside transporter Tsx. The diagram shows typical kinetics of protein production in a preparative scale CECF system. The nucleoside transporter Tsx was synthesized in a 1 ml CECF standard reaction without detergent over a time period of 21 h at 30 °C.

Table 2
Standard protocol for an individual continuous exchange cell-free reaction

Substance	Stock solution	Final concentration	Notes, references or suggested supplier
RM			
E. coli tRNA	40 mg/ml	0.5 mg/ml	In distilled water; (Roche Diagnostics, Mannheim, Germany; No. 109550)
Pyruvate kinase	10 mg/ml	0.04 mg/ml	(Roche Diagnostics GmbH, Mannheim, Germany; No. 109550)
T7-RNA-polymerase	40 U/µl	6 U/µl	overproduction in *E. coli* [28]
RNAguard porcine	40 U/µl	0.3 U/µl	(Amersham Biosciences, Freiburg, Germany; No. AP27-0816-01)
Plasmid vector	0.15 mg/ml	0.015 mg/ml	T7 promoter regulatory region (e.g. pET vector series; Merck Biosciences, Darmstadt, Germany)
E. coli S30 extract	100%	35%	(modified after: [4,12])
RM + FM			
NTP-Mix	ATP: 360 mM CTP, GTP, UTP: 240 mM each	ATP: 1.2 mM CTP, GTP, UTP: 0.8 mM each	pH 7.0 with NaOH
DTT	500 mM	2 mM	Depends on desired conditions
Folinic acid	10 mg/ml	0.1 mg/ml	Ca^{2+} salt
PEG 8.000	40%	2%	Dissolved at 30 °C in distilled water
Sodium azide	10%	0.05%	In distilled water
20 amino acid mix	4 mM each	1 mM each	Made from individual 100 mM stocks in distilled water, tyrosine as 20 mM stock in distilled water, tryptophan as 100 mM stock in 100 mM Hepes, pH 8.0 (remains turbid).
RCWMDE-Mix	16.7 mM each	1.0 mM each	Arginine, cysteine, tryptophan, methionine, aspartic acid and glutamic acid are limiting due to instability [15,45]
Complete protease inhibitor	50-fold	1-fold	In distilled water (Roche Diagnostics, Mannheim, Germany; No. 1836153)
Acetyl phosphate	1 M	20 mM	K^+ salt in distilled water, pH 7.0, with KOH
PEP	1 M	20 mM	K^+ salt in distilled water, pH 7.0, with KOH
Hepes buffer	2.5 M	100 mM	pH 8.0, with KOH
KOAc	4 M	280–300 mM	Subject of optimization
$Mg(OAc)_2$	1 M	13–15 mM	Subject of optimization

Approx. 140 mM K^+ are provided from other components (e.g. PEP, KOH), approx. 5 mM Mg^{2+} are provided from other components (e.g. extract) and final concentrations might be adjusted according to the amino acid composition of the target protein.

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

tion in a CECF system with wheat germ extract can continue up to several days. In a standard protocol, the final concentration of the components in the RM are: Hepes–KOH (pH 8.0) 40 mM, amino acids 0.2 mM each, $Mg(OAc)_2$ 3.0 mM, glycerine 2%, KOAc 80 mM, ATP 1 mM, GTP/CTP/UTP 0.8 mM each, spermidin 0.15 mM, NaN_3 0.03%, creatinephosphate 16 mM, mRNA 250 pmol/ml, creatinephosphate kinase 0.1 mg/ml, RNAse inhibitor 0.5 U/ml, wheat germ extract at 30% of the RM volume [41]. The composition of the FM is identical with the exception that the high-molecular weight compounds mRNA, creatinephosphate kinase, RNAse inhibitor and wheat germ extract are omitted. The reaction is incubated at 25 °C with intensive agitation. The final concentrations of Mg^{2+} and K^+ ions are subject of optimization and they may vary in the range between 1.5–3.5 and 60–120 mM, respectively. The optimal concentration of mRNA furthermore depends on the specific UTR's and coding sequences [42]. Short expression times in wheat germ systems are mainly attributed to an increased hydrolysis of NTP precursors. Supplying $Cu(OAc)_2$ helps to reduce endogenous ATPase activity and can successfully prolong the expression period [43,44].

2.6. Perspectives for the optimization of cell-free expression systems

The set-up of an efficient CECF reaction is relatively complicated and requires some experience. One focus of further improvements is therefore an increased efficiency of the easier to handle batch format CF systems. The exact adjustment of critical ion concentrations and the increased supply of some rapidly degraded amino acids (e.g. arginine, cysteine, tryptophan, methionine, aspartic acid and glutamic acid) considerably improved the translation efficiency and extended reaction times [45]. More efficient NTP regeneration systems can yield almost milligram amounts of protein per 1 ml reaction in bacterial and wheat germ batch systems [38,46]. Typical energy sources for the regeneration of ATP in CECF systems are phosphoenol pyruvate (PEP) [4], creatine phosphate [47] and acetyl phosphate (AcP) [48] together with the corresponding enzymes pyruvate kinase, creatine kinase and acetate kinase. These energy sources are problematic in batch systems due to the rapid accumulation of inorganic phosphate that inhibits protein synthesis [49]. Few alternative energy systems have therefore been established. The presence of pyruvate oxidase recycles inorganic phosphate by condensation with added pyruvate into acetyl phosphate, which then can be used again as energy source for protein synthesis [49]. Another approach could be the replacement of the energy source PEP by 3-phosphoglycerate. This modification extends batch reactions up to 2 h if compared to 30–45 min with PEP as conventional energy source and results in higher yields [36]. In addition, glucose-6-phosphate has been proposed as further option of a secondary energy source superior to PEP or pyruvate [49]. An economical improvement for CF protein synthesis was recently proposed by using glucose as energy source and nucleotide monophosphates instead of NTPs as precursors [50]. The cost of reagents could thus be lowered by over 75% at similar protein production yields.

The amino acid supply during CF-protein synthesis has a crucial impact on the expression yields. Instability and degradation can produce a rapid bias in the amino acid concentration. Recombinant protein of 500 μg per ml could be obtained in batch reactions by increasing the initial amino acid concentration from 0.5 to 2 mM [49,51]. Repeated addition of amino acids during the reaction also increases protein yields [45]. Genetic engineering of the genome of *E. coli* strains used as extract sources significantly decreased the degradation of arginine, tryptophan and serine in the reaction [52]. Modifications of the reaction conditions provide further potentials for optimization. Omitting PEG and HEPES as well as the addition of the polycations spermidine and putrescine were proposed as improvements of batch systems [53]. Optimized batch systems are already efficient enough to produced sufficient amounts of protein for structural analysis [54].

3. Cell-free preparation of membrane proteins

3.1. Specific characteristics for the CF expression of MPs

The current variety of MPs produced on preparative scales in CF systems comprises several bacterial multidrug transporters [10,12,54,55], a bacterial light harvesting protein [9], the mechanosensitive channel MscL [11], several eukaryotic GPCRs [13,14] and the β-barrel nucleoside transporter Tsx [14]. CF expression allows the production of MPs at levels of several mg/ml RM in two very different modes (Fig. 3). First, MP precipitates are produced in standard CF-systems due to the lack of a hydrophobic environment. Second, hydrophobic media like detergents or probably even lipids could be provided resulting in the production of already solubilized MPs. The CF production of precipitated MPs resembles the formation of inclusion bodies that can frequently be observed upon *in vivo* production of proteins in *E. coli*. However, CF generated MP precipitates might be structurally different from inclusion bodies as they usually do not need strong denaturants like SDS to become solubilized. Mild detergents like alkyl-glycosides, phosphocholines or alkyl-phosphoglycerols could already be sufficient for the quantitative solubilization of CF generated MP precipitates [10–12]. The harvested MP precipitates should be washed for several times in an appropriate buffer (e.g. 15 mM sodium phosphate, pH 6.8, 1 mM DTT) followed by centrifugation (5 min, 5000*g*). Precipitated impurities in the pellet could be removed by washing with a detergent that has poor solubilization properties for the MP (e.g. 3% *n*-octyl-β-glucopyranoside (β-OG)). Finally the MP is solubilized in buffer containing the detergent of choice (e.g. 2% 1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LMPG)) in a volume identical to the volume of the RM. Incubation on a shaker at 30 °C for one hour is usually sufficient for the quantitative solubilization. Residual pellet can be removed by centrifugation.

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

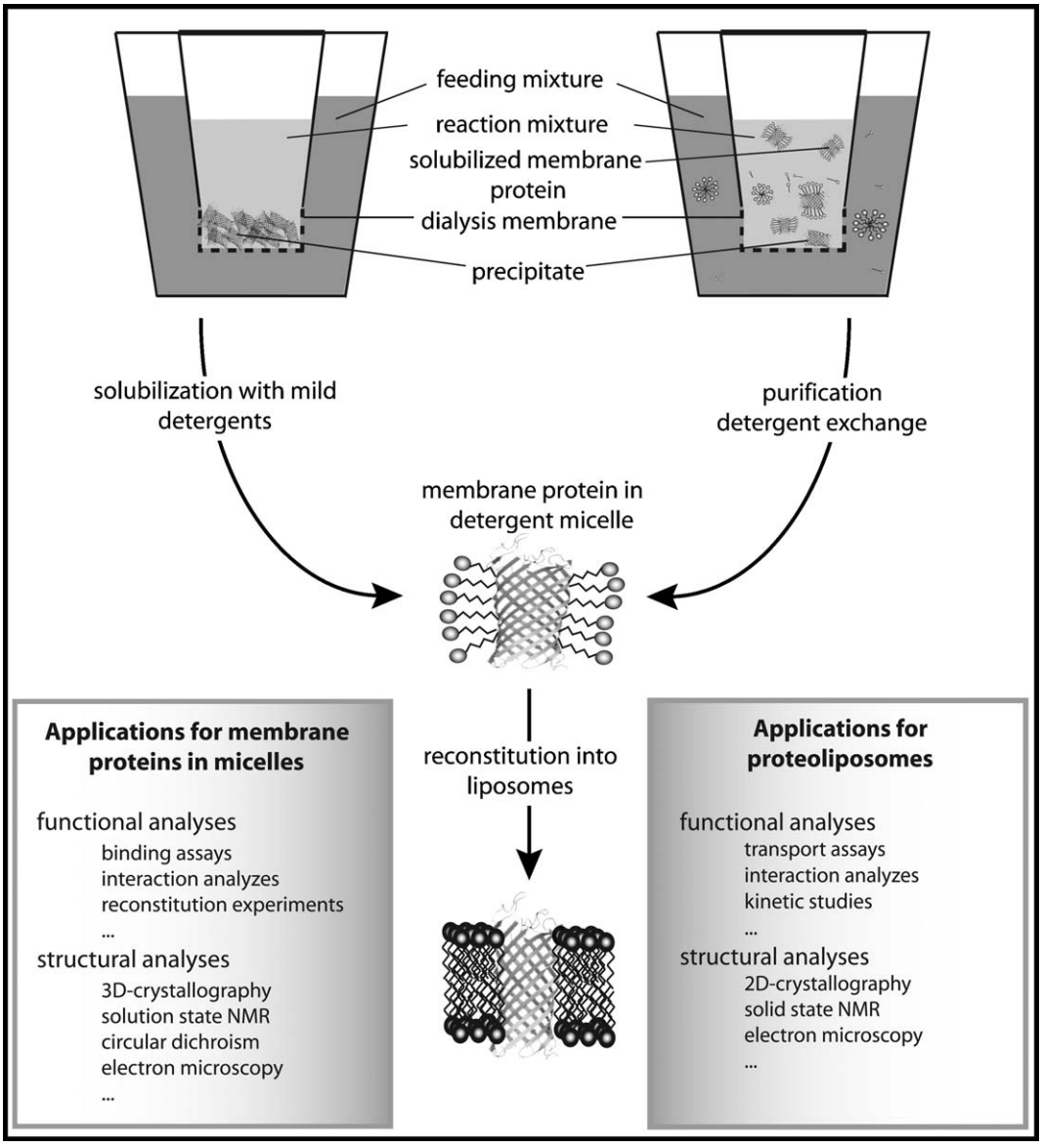


Fig. 3. Two modes of the cell-free production of membrane proteins. Both expression modes result in MPs solubilized in detergent micelles that can be reconstituted into proteoliposomes for further functional and structural studies.

The efficiency of solubilization certainly depends on the specific recombinant MP as well as on the type of detergent. Precipitates from the small α-helical multidrug transporter EmrE can be solubilized in a variety of different detergents while the quantitative solubilization of precipitates from the nucleoside transporter Tsx or from the mechanosensitive channel MscL was restricted to only a small selection of detergents like polyethylene-glycol P-1,1,3,3-tetramethyl-butylphenyl-ether (Triton X-100) or LMPG. Long-chain phosphoglycerols like LMPG and the closely related detergent 1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)] (LPPG) appear to be most versatile for the solubilization of CF produced precipitates of structurally diverse MPs. LMPG was highly efficient in the solubilization of bacterial α-helical and β-barrel type MPs and it proved to be the only detergent suitable for the quantitative solubilization of several eukaryotic G-protein coupled receptors (GPCRs) [14].

3.2. CF expression of MPs in presence of detergents

The open nature of the CF expression set-up enables the addition of defined amounts of detergents directly into the reaction [12–14]. The freshly translated proteins have thus the opportunity to become embedded immediately into preformed detergent micelles. This option to produce MPs in a soluble form associated with a detergent of choice is a unique characteristic for CF expression systems. Proteomicelles could be purified directly out of the RM and critical steps like the destabilization and isolation of MPs from membranes are eliminated. An indispensable prerequisite is that the supplied detergent is tolerated by the CF system even at concentrations exceeding several times the specific critical micellar concentration (CMC) that defines the minimal required effective concentration for the solubilization of proteins. Only few detergent types like dodecyl-phospho-

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

choline (DPC) or β-OG severely inhibit the CF protein production already at low concentrations at or only slightly above their specific CMCs. Fortunately, many other commonly used detergents are tolerated by CF expression systems and the optimal concentration ranges for a relatively large group of detergents have been determined (Table 3)

Table 3
Cell-free expression of membrane proteins in the presence of detergents

Detergent name	Optimal concentration				Protein			References cited
	Short name	[%]	[mM]	[x CMC]				
None								
Alkyl-glucosides								
n-Dodecyl-β-D-maltoside	DDM	0.08	(2.30)	12.1	EmrE	+++++	0	[10]
		0.1	(2.87)	15.1	Tsx	++++	++++	[14]
		<1	(28.69)	151	β2AR	+	+	[13]
		0.066	(1.89)	10	MscL	n.r.	n.r.	[11]
n-decyl-β-D-maltoside	DM	0.2	(4.14)	2.3	EmrE	++++	++	[14]
n-octyl-β-D-glucopyranoside	β-OG	0.75	(25.65)	1.3	EmrE	+	0	[14]
		<1	(34.2)	1.8	M2	0+	0+	[13]
		0.4	(13.68)	0.7	EmrE	n.r.	n.r.	[10]
		0.877	(30)	1.6	MscL	n.r.	n.r.	[11]
(6-O-(*N*-heptylcarbamoyl)-methyl-α-D-glucopyranoside)	HECAMEG	<1	(29.82)	1.8	M2	+	+	[10]
Steroid-derivatives								
Digitonin	Digitonin	0.4	(3.25)	4.5	Tsx	++++	+++	[14]
		<1	(8.13)	11.1	β2AR	++	++	[13]
3-[(3-cholamidopropyl)dimethly-ammonio]-1-propansulfonat	CHAPS	0.75	(12.20)	1.5	Tsx	+	++	[14]
		<1	(16.26)	2	NTR	+	+	[13]
		2.46	(40.01)	5	MscL	n.r.	n.r.	[11]
Long chain-phosphoglycerols								
1-myristoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)]	LMPG	0.01	(0.21)	4.2	Tsx	++	++	[14]
1-palmitoyl-2-hydroxy-*sn*-glycero-3-[phospho-rac-(1-glycerol)]	LPPG	0.025	(0.49)	n.a.	V2R	+	+++++	[14]
Mono-/Bi-chain-phosphocholines:								
1,2-dioctanoyl-*sn*-glycero-3-phosphocholine	diC_8PC	0.1	(1.96)	8.9	EmrE	+++	+++++	[14]
1,2-diheptanoyl-*sn*-glycero-3-phosphocholine	DHPC	0.2	(4.15)	3	Tsx	+++	++	[14]
1,2-dihexanoyl-*sn*-glycero-3-phosphocholine	diC_6PC	0.75	(16.54)	1.2	Tsx	++	++	[14]
Dodecyl-phosphocholine	DPC	0.1	(2.84)	1.5	EmrE	0+	+++	[14]
Polyoxyethylene-alkyl-ether:								
polyoxyethylene-(8)-lauryl-ether, ($C_{12/8}$)	$C_{12}E_8$	<1	(18.52)	260.8	NTR	+	++	[13]
polyoxyethylene-(23)-lauryl-ether, ($C_{12/23}$)	Brij-35	0.1	(0.83)	10.4	EmrE	+++++	+	[14]
		<1	(8.34)	104.2	β2AR	++	++	[13]
polyoxyethylene-(10)-dodecyl-ether, ($C_{12/10}$)	GPC-100	0.1	(1.59)	21.3	EmrE	+	++	[14]
polyoxyethylene-(10)-isotridecyl-ether, ($C_{13/10}$)	GPX-100	0.1	(1.56)	10.4	Tsx	+++	++++	
polyoxyethylene-(10)-cetyl-ether, ($C_{16/10}$)	Brij-56	0.01	(0.15)	4.2	V2R	+	0+	[14]
polyoxyethylene-(20)-cetyl-ether, ($C_{16/20}$)	Brij-58	1.5	(13.36)	178.1	V2R	+++++	+	[14]
		0.84	(0.75)	10	MscL	n.r.	n.r.	[11]
polyoxyethylene-(2)-stearyl-ether, ($C_{18/2}$)	Brij-72	0.2	(5.57)	n.a.	V2R	+	+	[14]
polyoxyethylene-(20)-stearyl-ether, ($C_{18/20}$)	Brij-78	1	(8.68)	188.8	V2R	+++++	+	[14]
polyoxyethylene-(10)-oleyl-ether, ($C_{18\text{-}1/10}$)	Brij-97	0.2	(2.82)	13	EmrE	0	+++++	[14]
polyoxyethylene-(20)-oleyl-ether, ($C_{18\text{-}1/20}$)	Brij-98	0.2	(1.74)	69.6	V2R	+++++	+	[14]
polyoxyethylene-sorbitan-monolaurate 20	Tween 20	<1	(8.14)	138	NTR	+	++	[13]
		0.1	(0.81)	13.8	Tsx	0	+++++	[14]
		0.072	(0.59)	10	MscL	n.r.	n.r.	[11]
polyoxyethylene-sorbitan-monopalmitate 40	Tween 40	<1	(7.79)	288.5	NTR	+	++	[13]
polyoxyethylene-sorbitan-monostearate 60	Tween 60	<1	(7.62)	304.9	M2	+	+	[13]
polyoxyethylene-sorbitan monooleate 80	Tween 80	<1	(7.63)	636.1	NTR	+	++	[13]
Polyethylene-glycol derivatives:								
polyethylene-glycol P-1,1,3,3-tetramethyl-Butylphenyl-ether	TX-100	0.2	(3.09)	13.4	MscL	+++++	n.r.	[11]
		0.1	(1.55)	6.7	Tsx	++++	++++	[14]
		<1	(15.46)	67.2	NTR	+	++	[13]
Polyethylene-glycol 400 dedecyl-ether	Thesit	0.1	(1.72)	17.2	Tsx	++	+++	[14]
Nonylphenyl-polyethylene-glycol	NP40	0.1	(1.66)	9.8	EmrE	+	++++	[14]
		<1	(16.58)	97.6	β2AR	+	++	[13]

S, soluble fraction; P, insoluble precipitated fraction; n.a., not available; n.r., not reported, 0, no detectable expression; 0+, less than 10 μg/ml; +, 10–100 μg/ml; ++, 101–500 μg/ml; +++, 501–1000 μg/ml; ++++, 1001–2000 μg/ml; +++++, more than 2001 μg/ml.

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

[13,14]. The kinetics of MP solubilization versus the detergent concentration shows two phases. Initially, the efficiency of MP solubilization is linear to the amount of added detergent until a certain threshold concentration has been achieved (Fig. 4). Then the yield of solubilized MP remains constant with further increased detergent concentrations and only additional empty micelles will be formed until the detergent becomes toxic to the CF expression system due to the inactivation of essential compounds. The production kinetics has therefore a plateau-like appearance and the maximal yield of soluble MP can be obtained over a distinct concentration range of a specific detergent. It should be considered that the complete solubilization of a MP might not be possible and still some residual MP precipitate could remain even far above the threshold concentration (Fig. 4).

Several elements of CF production of soluble MPs can be subjected to optimization. Basic parameters are concentration, type and chain length of the supplied detergent. For initial screens it might be most straightforward to start with detergent concentrations close to the maximal tolerated levels to receive instantly the highest possible amounts of soluble MP. The detergents should be prepared as highly concentrated stock solutions in water and care should be taken that organic solvents like chloroform have been completely removed e.g. by evaporation, in order to prevent an inhibition of the CF reaction. Soluble protein fractions are separated from precipitates after the reaction by centrifugation at 20,000*g* for 30 min at room temperature. The production of the MP should be quantified in both fractions by SDS–PAGE analysis or by immunoassays. The final detergent concentration can then be decreased in subsequent optimization reactions if desired.

The structure of the supplied detergent type can have a major impact on the efficiency of solubilization as well as on the functional folding of the synthesized MP. Some proteins like the α-helical multidrug transporter EmrE can be expressed in soluble form with a diverse variety of structurally different detergents like alkyl-glucosides, phosphocholines, polyethylene-glycol derivatives or polyoxyethylenes. However, the soluble expression in preparative scales of the majority of the MPs seems to be restricted to a much smaller selection of detergents (Table 3). The supply of many popular detergents that have been used in recent times for the structural analysis of MPs like the alkyl-glucoside *n*-dodecyl-β-D-maltoside (DDM) or the polyethylene-glycol derivative Triton X-100 result in the high level soluble expression of specific MPs. CF expression in presence of DDM yielded milligram amounts of EmrE and of the nucleoside transporter Tsx, but the detergent was rather ineffective for the soluble expression of different GPCRs [13,14]. Tsx could furthermore only partially be solubilized in DDM and approx. 50% of the synthesized protein still remained as precipitate. Clearly outstanding with respect to their ability to efficiently solubilize structurally diverse MPs are the steroid-derivative Digitonin and various detergents from the family of polyoxyethylene-alkyl-ethers like polyoxyethylene-(23)-lauryl-ether (Brij-35), polyoxyethylene-(20)-cetyl-ether (Brij-58), polyoxyethylene-(20)-stearyl-ether (Brij-78) and polyoxyethylene-(20)-oleyl-ether (Brij-98) [13,14]. Out of more then 20 detergents, only Brij derivatives have been successful in the quantitative soluble expression of the human vasopressin type 2 receptor [14]. Brij derivatives having less than 10 polyoxyethylene groups were not effective. In addition, the solubilization of the vasopressin type 2 receptor was modulated by the length of the alkyl chain. As a general guideline, the efficiencies and the optimal concentrations of the most useful detergents for the solubilization of structurally different MPs are compiled in Table 3.

In practice it is recommended to initially perform a set of CF expression reactions in presence of the most promising

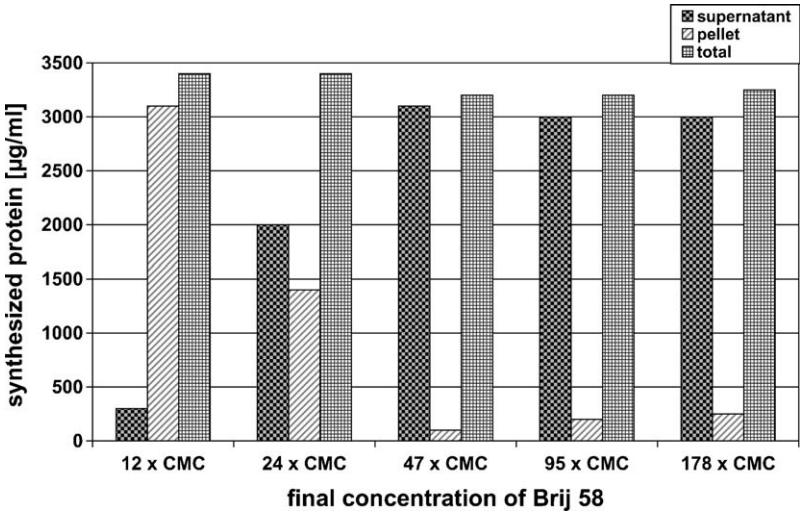


Fig. 4. Soluble expression of the nucleoside transporter Tsx at increased detergent concentrations. The yield of soluble expressed MP increases with the final concentration of Brij58 and reaches a plateau at approx. 47-fold CMC. The amount of soluble Tsx then remains constant upon further increased detergent concentrations up to 178-fold CMC. The total amount of produced Tsx protein (soluble and precipitate) remains constant at each condition.

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

detergents with each new MP target (Fig. 5). In case of the nucleoside transporter Tsx, the yield of soluble protein in presence of the detergents Brij58, Brij78 and Brij97 is comparable to that of the Tsx precipitate isolated out of a standard CF reaction without any detergent (Fig. 5). However, generally some lower amounts of protein might be obtained when choosing the soluble mode of expression if compared to the expression as a precipitate.

After high level CF production strategies have been established, the functional folding of the synthesized MP should be analysed [10–14,56,57]. The modes of expression, the origin of the CF extract, the solubilization procedures of precipitates and the supplied detergent types are important factors that could influence the folding pathway of a protein. Circular dichroism spectra or two-dimensional heteronuclear single-quantum correlation (HSQC) spectra by solution nuclear magnetic resonance (NMR) spectroscopy can give first evidences of the presence of structural elements [12]. However, the development of functional assays would be a very precious tool to prove the three-dimensional folding of a protein into an active conformation. The transport of the substrate ethidium by EmrE could be verified with samples that have been either produced in the soluble mode in presence of DDM [10] or with proteoliposomes reconstituted from in DDM solubilized precipitates [12]. On the other hand, the nucleoside transporter Tsx could only be reconstituted in a highly active form after its CF expression in the soluble mode in presence of Triton X-100 [14]. While the soluble expression in presence of Brij35 still resulted in some residual activity, it was not possible to detect activity from protein that has been produced as a precipitate and solubilized in LMPG. This example demonstrates the importance of initial expression screens that should consider both, the high level production and the functional folding of the MP. The available information of CF expressed MPs that have been analysed by functional assays is still limited to a few examples. However, it is already evident that structurally very different prokaryotic as well as eukaryotic MPs can be produced in CF systems based on lysates of *E. coli* cells in both expression modes as functionally active proteins.

4. New perspectives for the structural analysis of cell-free produced membrane proteins

The CF expression of MPs offers a high potential for their structural analysis by NMR spectroscopy as well as by X-ray crystallography. Protein samples that have been efficiently labelled with stable isotopes or with non-natural amino acids like selenomethionine can now be obtained in less than 24 hours [54,58–60]. The recently presented crystal structure of the small multidrug transporter EmrE at 3.7 Å resolution was solved by taking advantage of CF expression. The proposed active dimer was CF expressed directly into di-myristoyl-phosphatidyl-choline (DMPC) liposomes and crystals which were used for structure determination were obtained in the presence of the detergent *N*-nonyl-β- D-glucoside (NG) [54].

Most of the few known NMR structures from MPs were solved from bacterial β-barrel proteins that have been produced as inclusion bodies in *E. coli* followed by refolding procedures [61,62]. The CF expression technique might open new avenues for the determination of MP structures by NMR spectroscopy as well as by X-ray crystallography already solely because of the possibility to produce high yields of functionally active MPs that are difficult to obtain with other expression systems. In addition, the speediness of the reaction is highly competitive as the protein samples are basically generated over night. Moreover, no special equipment is required and the technique can principally be established in standard biochemical labs within few days. The most important feature however is the easiness and efficiency in the synthesis of labelled protein samples. The CF extract is completely devoid of amino acids due to extensive dialysis steps during the preparation procedure. As the operator has the complete control over the amino acid pool of the CF reaction, any non-labelled amino acid type can just be exchanged by its labelled derivative at the

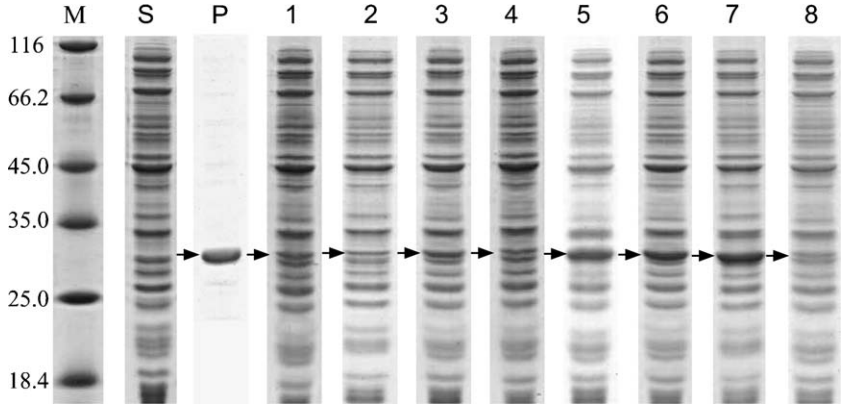


Fig. 5. Detergent screen for the cell-free production of the nucleoside transporter Tsx. Coomassie-stained SDS–Page of CF produced Tsx protein: M, marker; s, supernatant of standard reaction; p, pellet of standard reaction. Lanes 1–8 represent 0.8 μl of supernatant of CF expression in the presence of different type detergents: 1, 0.1% DDM; 2, 0.2% DHPC; 3, 0.4% Digitonin; 4, 0.1% Genapol X100; 5, 0.2% Brij58; 6, 1% Brij78; 7, 0.1% Brij97; 8, 0.5% Brij98. The arrow indicates the overexpressed nucleoside transporter Tsx.

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

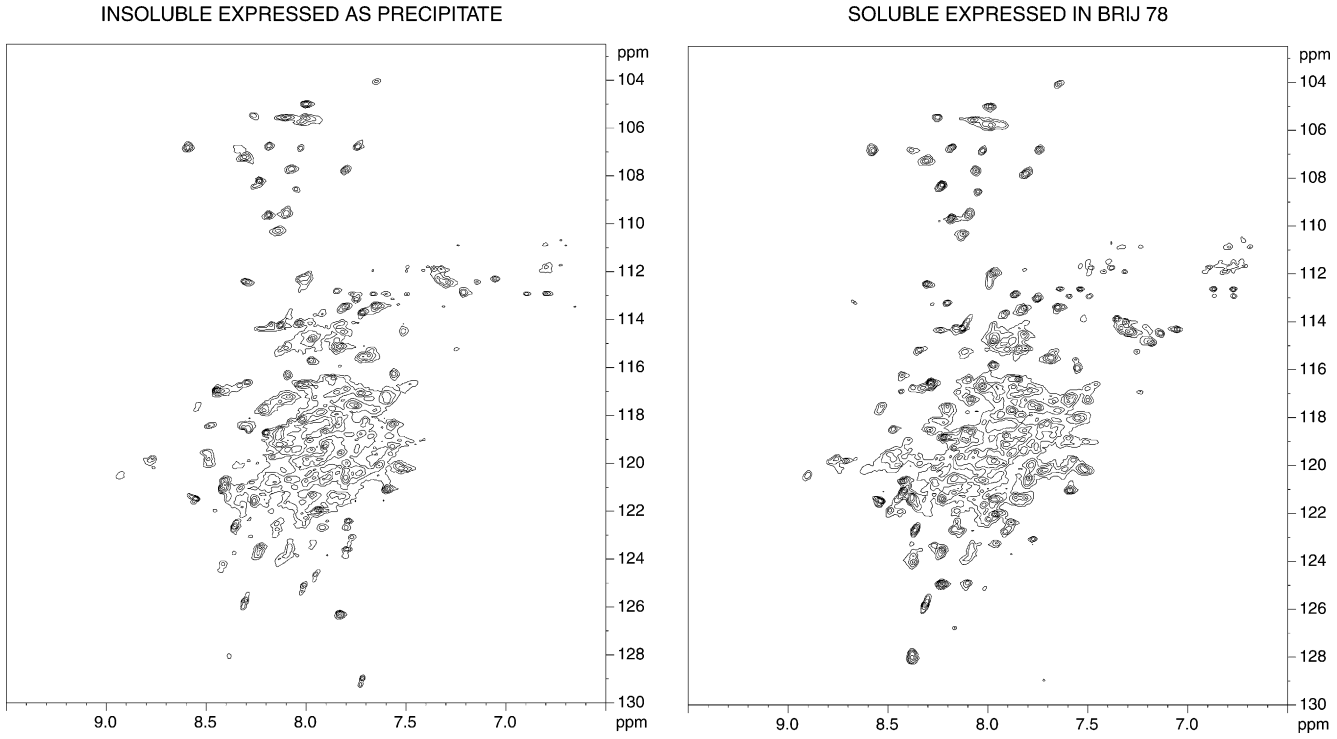


Fig. 6. ^{1}H–^{15}N TROSY-HSQC spectra of cell-free expressed ^{2}H–^{15}N-ΔTehA. The protein was expressed either in the unsoluble mode as precipitate and resolubilized in LMPG or expressed in the soluble mode in the presence of Brij78, purified and the detergent was exchanged to LMPG. The concentration of both samples was 0.5 mM, dissolved in 25 mM sodium phosphate buffer (pH 6.0) containing 3% LMPG. The spectra were measured at 318 K on an Avance 900 MHz NMR spectrometer equipped with a cryogenic probe with eight scans per FID and 512 increments in the indirect dimension for the insoluble expressed ΔTehA and with four scans per FID and 300 increments in the indirect dimension for the soluble expressed transporter.

initial set-up of the experiment. This instantly ensures the 100% label incorporation into the synthesized protein. The specific labelling of any amino acid type and also of any amino acid combinations of an expressed protein is thus as efficient as the production of non-labelled proteins [63]. Auxotrophic strains and minimal media that have to be employed in conventional *in vivo* expression systems and that often considerably reduce the yield of the recombinant protein can be avoided. Scrambling problems are furthermore minimized as the metabolic activity of the CF extract is very low [22]. The target protein is the only synthesized protein in the system as all endogenous mRNA of the extract was eliminated during the extract preparation. The labelled protein could therefore be analysed by NMR spectroscopy directly in the RM without prior purification as no labelled background is present [64].

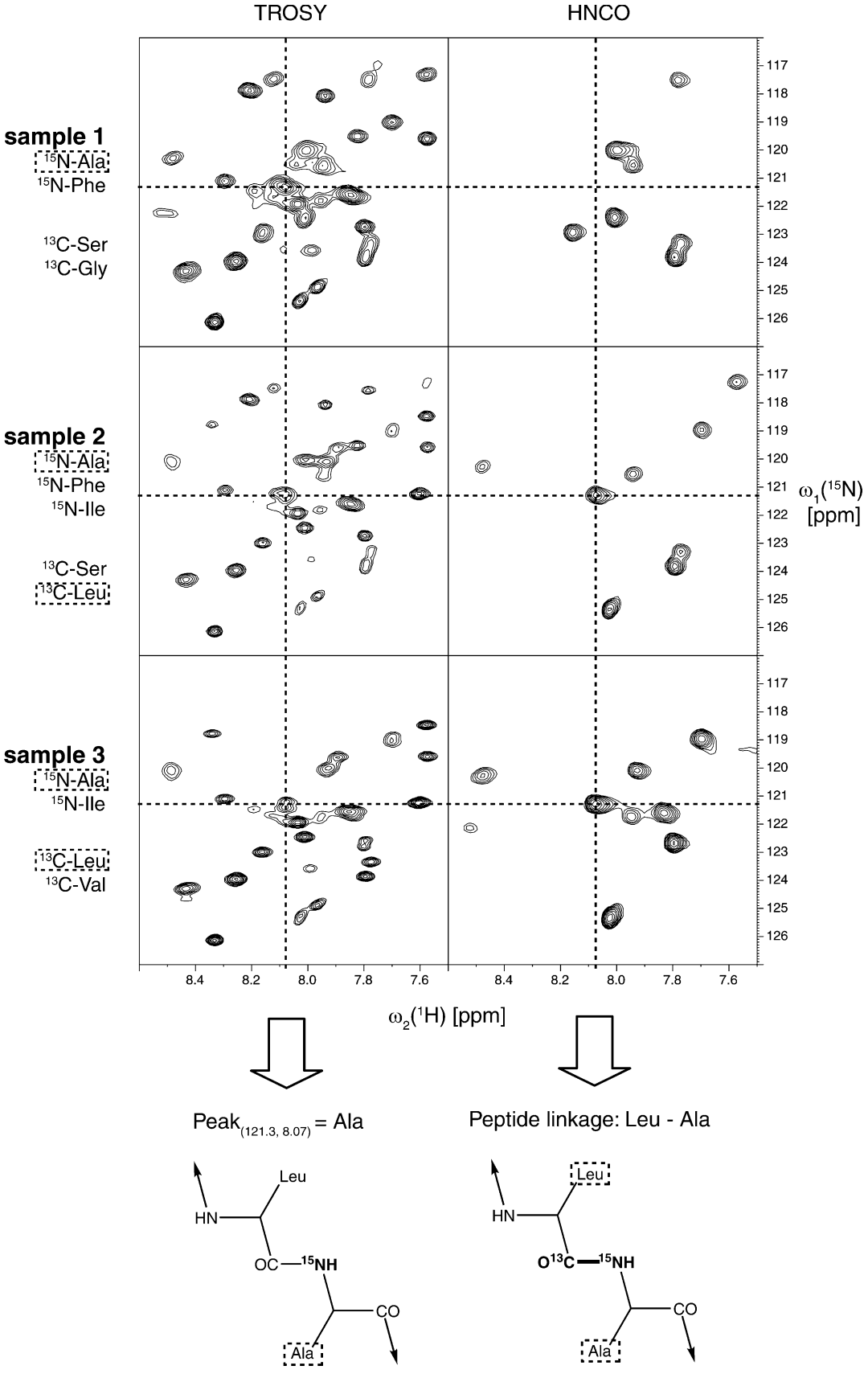
One major restriction for the determination of MP structures is the size limitation of proteins for NMR samples. The rate of rotational tumbling of proteins decelerates with increasing size, resulting in line broadening and lower resolution. Because MPs need to be analysed in detergent micelles, the signal resolution will be even more critical due to the increased size of the protein/detergent complex. Furthermore, MPs are mostly α-helical proteins that show rather narrow chemical shift dispersions with an extensive signal overlap. The hydrophobic transmembrane segments (TMSs) of MPs often contain cluster of similar amino acids with equal chemical shifts that additionally contributes to lower resolutions of corresponding NMR spectra. These problems in combination with the low production rates of many MPs in conventional *in-vivo* expression systems prevented so far in most cases serious attempts for the structural analysis of MPs by NMR.

Several limitations can now be addressed by virtue of CF expression. First, the size limitation of the solubilized protein samples can be approached by using specific detergents. In a systematic screen for liquid-state NMR compatible detergents with regard to long sample lifetimes,

Fig. 7. Combinatorial labelling scheme of membrane proteins by cell-free expression. Example of an amino acid selective combinatorial labelling scheme with three differentially [^{15}N] and [^{13}C]labelled samples of the tellurite transporter TehA. The labelling scheme is shown on the left side. The crossed dotted lines indicate the peak to become identified. In the [^{15}N, ^{1}H]-TROSY spectra, backbone amide protons of the [^{15}N]-labelled amino acids are visible. The indicated peak at position 121.3/8.07 is identified as an alanine as it is present in all three spectra and only the amino acid alanine has been [^{15}N]-labelled in all three samples. The corresponding HNCO spectra show the amide crosspeaks after carbonyl transfer and indicate that the preceding residue of this alanine must be a leucine because a HNCO peak is only observed for samples 2 and 3 containing ^{13}C-leucine, but not for sample 1 without ^{13}C-leucine. The corresponding alanine (peak HN:8.07; N:121.3) can now be localized in a Leu–Ala pair in the primary sequence of the protein which in that example was identified as Ala206 of TehA. All spectra were recorded at an Avance 600 MHz spectrometer.

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

maximal signal detection as well as high spectral resolution, the class of lyso-phosphoglycerols including LMPG and LPPG have been found to provide outstanding benefits [65,66]. LMPG micelles obviously do not restrict the tumbling rates of the inserted solubilized MP and therefore do not result in increased line broadenings. LMPG and its



Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

derivatives were also found to be highly suitable for the solubilization of CF produced MP precipitates [14]. This detergent type might be therefore one of the first choices for the structural analysis of larger MPs by liquid-state NMR spectroscopy. A CF produced truncated 24 kDa fragment of the putative tellurite and multidrug transporter TehA of *E. coli* containing seven TMSs was solubilized in LMPG and analysed by NMR. The amide protons of the protein backbone could be almost completely assigned by using a rationally designed combinatorial labelling approach [63]. The spectral quality of TehA further depends on the mode of CF expression. ^{1}H–^{15}N-HSQC spectra of TehA samples produced in the soluble mode with 1% Brij-78 followed by a buffer exchange against 3% LMPG showed a significantly better resolution if compared with samples that have been prepared as CF precipitates and re-solubilized in LMPG (Fig. 6).

The spectral overlap due to the α-helical structure of many MPs could be approached by a mixture of amino acid specific and combinatorial labelling strategies [63]. A simultaneous labelling of MPs with selected [^{15}N]-labelled amino acids in combination with distinct [^{13}C]-labelled amino acids can easily performed by CF expression (Fig. 7). The application of two dimensional versions of HNCO experiments helps to identify only those [^{15}N]-labeled amino acids that where N-terminally proceeded by a [^{13}C]-labeled amino acid type [63]. This strategy enables the unambiguous identification of consecutive amino acid pairs that can be subsequently used as anchor points for further backbone resonance assignments (Fig. 7). The TehA protein represents one of the largest α-helical proteins currently analysed by NMR and its structural approach demonstrates the powerful synergy of liquid-state NMR and CF expression.

References

[1] P. Renesto, D. Raoult, Ann. N.Y. Acad. Sci. 990 (2003) 642–652.
[2] M. Bannwarth, G.E. Schulz, Biochim. Biophys. Acta 1610 (2003) 37–45.
[3] J.M. Betton, Curr. Protein Pept. Sci. 4 (2003) 73–80.
[4] G. Zubay, Annu. Rev. Genet. 7 (1973) 267–287.
[5] A.S. Spirin, Cell-Free Translation Systems, Springer, New York, 2002.
[6] L.A. Ryabova, D. Desplancq, A.S. Spirin, A. Pluckthun, Nat. Biotechnol. 15 (1997) 79–84.
[7] G. Yin, J.R. Swartz, Biotechnol. Bioeng. 86 (2004) 188–195.
[8] M. Ali, H. Suzuki, T. Fukuba, X. Jiang, H. Nakano, T. Yamane, J. Biosci. Bioeng. 99 (2005) 181–186.
[9] Y. Shimada, Z. Y Wang, Y. Mochizuki, M. Kobayashi, T. Nozawa, Biosci. Biotechnol. Biochem. 68 (2004) 1942–1948.
[10] Y. Elbaz, S. Steiner-Mordoch, T. Danieli, S. Schuldiner, Proc. Natl. Acad. Sci. USA 101 (2004) 1519–1524.
[11] C. Berrier, K. Park, S. Abes, A. Bibonne, J.M. Betton, A. Ghazi, Biochemistry 43 (2004) 12585–12591.
[12] C. Klammt, F. Löhr, B. Schäfer, W. Haase, V. Dötsch, H. Rüterjans, C. Glaubitz, F. Bernhard, Eur. J. Biochem. 271 (2004) 568–580.
[13] G. Ishihara, M. Goto, M. Saeki, K. Ito, T. Hori, T. Kigawa, M. Shirouzu, S. Yokoyama, Prot. Express. Purif. 41 (2005) 27–37.
[14] C. Klammt, D. Schwarz, K. Fendler, W. Haase, V. Dötsch, F. Bernhard, FEBS Journal 272 (2005) 6024–6038.
[15] D.M. Kim, J.R. Swartz, Biotechnol. Bioeng. 66 (1999) 180–188.
[16] A.S. Spirin, V.I. Baranov, L.A. Ryabova, S.Y. Ovodov, Y.B. Alakov, Science 242 (1988) 1162–1164.
[17] Y.B. Alakov, V.I. Baranov, S.J. Ovodov, L.A. Ryabova, A.S. Spirin, I.J. Morozov, Method of preparing polypeptides in cell-free translation system U.S. Patent no. 5, 478, 73 (1995).
[18] D.M. Kim, C.Y. Choi, Biotechnol. Prog. 12 (1996) 1–8.
[19] Y. Endo, T. Sawasaki, Biotechnol. Adv. 21 (2003) 695–713.
[20] B.E. Roberts, B.M. Paterson, Proc. Natl. Acad. Sci. U S A 70 (1973) 2330–2334.
[21] K. Madin, T. Sawasaki, T. Ogasawara, Y. Endo, Proc. Natl. Acad. Sci. USA 97 (2000) 559–564.
[22] K. Ozawa, N.E. Dixon, G. Otting, IUBMB Life 57 (2005) 615–622.
[23] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, Nat. Biotechnol. 19 (2001) 751–755.
[24] T. Yamane, Y. Ikeda, T. Nagasaka, H. Nakano, Biotechnol. Prog. 21 (2005) 608–613.
[25] A.H. Erickson, G. Blobel, Biochemistry 22 (1983) 5201–5205.
[26] T. Ogasawara, T. Sawasaki, R. Morishita, A. Ozawa, K. Madin, Y. Endo, EMBO J. 18 (1999) 6522–6531.
[27] T. Sawasaki, M.D. Gouda, T. Kawasaki, T. Tsuboi, Y. Tozawa, K. Takai, Y. Endo, Methods Mol. Biol. 310 (2005) 131–144.
[28] Y. Li, E. Wang, Y. Wang, Protein Expr. Purif. 16 (1999) 355–358.
[29] F. Mignone, C. Gissi, S. Liuni, G. Pesole, Genome Biol. 3 (2002).
[30] T. Ueda, H. Tohda, N. Chikazumi, F. Eckstein, K. Watanabe, Nucleic Acids Res. 19 (1991) 547–552.
[31] I. Hirao, S. Yoshizawa, K. Miura, Nucleic Acids Symp. Ser. 22 (1990) 75–76.
[32] E. Kobatake, A. Ebisawa, O. Asaka, Y. Yanagida, Y. Ikariyama, M. Aizawa, Apl. Biochem. Biotechnol. 76 (1999) 217–227.
[33] D.R. Gallie, D.E. Sleat, J.W. Watts, P.C. Turner, T.M. Wilson, Nucleic Acids Res. 15 (1987) 3257–3273.
[34] V. Leathers, R. Tanguay, M. Kobayashi, D.R. Gallie, Mol. Cell Biol. 13 (1993) 5331–5347.
[35] T. Sawasaki, T. Ogasawara, R. Morishita, Y. Endo, Proc. Natl. Acad. Sci. USA 99 (2002) 14652–14657.
[36] K. Sitaraman, D. Esposito, G. Klarmann, S.F. Le Grice, J.L. Hartley, D.K. Chatterjee, J. Biotechnol. 110 (2004) 257–263.
[37] J.H. Ahn, H.S. Chu, T.W. Kim, I.S. Oh, C.Y. Choi, G.H. Hahn, C.G. Park, D.M. Kim, Biochem. Biophys. Res. Commun. 338 (2005) 1346–1352.
[38] Y. Endo, T. Sawasaki, J. Struct. Funct. Genomics 5 (2004) 45–57.
[39] T. Sawasaki, Y. Hasegawa, M. Tsuchimochi, N. Kamura, T. Ogasawara, T. Kuroita, Y. Endo, FEBS Lett. 514 (2002) 102–105.
[40] Y. Endo, T. Sawasaki, Methods Mol. Biol. 310 (2005) 145–167.
[41] L.A. Shaloiko, A. Gorokhovatskii, E.E. Maksimov, I.B. Alakhov, Bioorg. Chim. 24 (1998) 539–543.
[42] V.A. Shirokov, P.N. Simonenko, S.V. Biryukov, A.S. Spirin, Cell-Free Translation Systems, in: A.S. Spirin (Ed.), Springer, 2002, 227-234.
[43] Y. Kawarasaki, H. Nakano, T. Yamane, Plant Sci. 119 (1996) 67–77.
[44] Y. Kawarasaki, H. Nakano, T. Yamane, J. Biotech. 61 (1998) 199–208.
[45] D.M. Kim, J.R. Swartz, Biotechnol. Prog. 16 (2000) 385–390.
[46] M.C. Jewett, J.R. Swartz, Biotechnol. Bioeng. 86 (2004) 19–26.
[47] T. Kigawa, T. Yabuki, S. Yokoyama, FEBS Lett. 442 (1999) 15–19.
[48] L. Ryabova, L. Vinocurov, E. Shekhovtsova, Y. Alakhov, A. Spirin, Anal. Biochem. 226 (1995) 184–186.
[49] D.M. Kim, J.R. Swartz, Biotechnol. Bioeng. 74 (2001) 309–316.
[50] K.A. Calhoun, J.R. Swartz, Biotechnol. Prog. 21 (2005) 1146–1153.
[51] M.C. Jewett, A. Voloshin, J. Swartz, M.P. Weiner, Q. Lu (Eds.), Gene Cloning and Expression Technologies, Springer, 2002, 391-411.
[52] N. Michel-Reydellet, K. Calhoun, J. Swartz, Metab. Eng. 6 (2004) 197–203.
[53] M.C. Jewett, J.R. Swartz, Biotechnol. Bioeng. 87 (2004) 465–472.
[54] O. Pornillos, Y.J. Chen, A.P. Chen, G. Chang, Science 310 (2005) 950–1953.
[55] M. Lorch, I. Lehner, A. Siarheyeva, D. Basting, N. Pfleger, T. Manolikas, C. Glaubitz, Biochem. Soc. Trans. 33 (2005) 873–877.
[56] S. Sonar, N. Patel, W. Fischer, K.J. Rothschild, Biochemistry 32 (1993) 13777–13781.

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

[57] L.K. Lyford, R.L. Rosenberg, J. Biol. Chem. 274 (1999) 25675–25681.
[58] T.L. Hendrickson, V. de Crécy-Lagard, P. Schimmel, Annu. Rev. Biochem. 73 (2004) 147–176.
[59] T. Kigawa, E. Yamaguchi-Nunokawa, K. Kodama, T. Matsuda, T. Yabuki, N. Matsuda, R. Ishitani, O. Nureki, S. Yokoyama, J. Struct. Funct. Genomics. 2 (2002) 29–35.
[60] K. Ozawa, M.J. Headlam, D. Mouradov, S.J. Watt, J.L. Beck, K.J. Rodgers, R.T. Dean, T. Hubert, G. Otting, N.E. Dixon, FEBS J. 272 (2005) 3162–3171.
[61] C. Fernandez, C. Hilty, G. Wider, P. Güntert, K. Wüthrich, J. Mol. Biol. 336 (2004) 1211.
[62] C. Fernandez, K. Adeishvili, K. Wüthrich, Proc. Natl. Acad. Sci. USA 98 (2001) 2358–2363.
[63] N. Trbovic, C. Klammt, A. Koglin, F. Lohr, F. Bernhard, V. Dotsch, J. Am. Chem. Soc. 127 (2005) 13504–13505.
[64] L. Guignard, K. Ozawa, S.E. Pursglove, G. Otting, N.E. Dixon, FEBS Lett. 524 (2002) 159–162.
[65] T.P. Roosild, J. Greenwald, M. Vega, S. Castronovo, R. Riek, S. Chos, Science 307 (2005) 1317–1321.
[66] R.D. Krueger-Koplin, P.L. Sorgen, S.T. Krueger-Koplin, I.O. Rivera-Torres, S.M. Cahill, D.B. Hicks, L. Grinius, T.A. Krulwich, M.E. Girvin, J. Biomol. NMR 17 (2004) 43–57.

Please cite this article as: Daniel Schwarz et al., Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies, Methods (2006), doi:10.1016/j.ymeth.2006.07.001.

(P 07)

DNA recombination and protein expression. F. Bernhard, C. Klammt, and H. Rüterjans. In: *Comprehensive Medicinal Chemistry II*, in press. (2006)

A0005 # 3.06 Recombinant DNA and Protein Expression

F Bernhard, C Klammt, and H Rüterjans, University of Frankfurt/Main, Germany

S0005 ## 3.06.1 Design and Generation of Vectors for High-Level Protein Expression in Bacteria

P0005 Vectors are generally defined as the basic vehicles that transport and deliver the target genes to be expressed into a suitable host cell. They can be relatively complicated autonomously replicating elements much as plasmids, bacteriophages, or viruses or they might consist of less complex DNA molecules that integrate into the host chromosome or that are even not stable in the cellular background and provide only a transient expression of the target gene. However, common characteristics are always a few basic requirements like selection marker, copy number, host range, and origin of replication. The copy number is a key feature as it is linearly related to the gene dosage and thus clearly affects the expression yields. In general, replicating vectors are present in multiple copies up to several hundred molecules per single cell in case of common bacterial expression plasmids. The copy number together with the host range of a DNA molecule is determined by the origin of replication, a sequence motif that is recognized by a specific DNA polymerase. In addition, plasmids are classified into separate incompatibility groups based on their origin of replication. Members of the same group are not compatible and cannot be permanently maintained in the same cell. Integration vectors are devoid of any origins of replication and only present in one or few copies, although for specific

cases techniques have been developed to increase the copy number by selective amplification.[1,2] Vectors usually contain selection markers that are indispensable to identify successfully transformed cells. Antibiotic resistance genes encoding, e.g. for the enzyme β-lactamase, are frequently used especially for selection in bacteria. Genes encoding for key enzymes in essential anabolic pathways are also suitable as selection markers if used in combination with specifically engineered auxotrophic host cells. A variety of comprehensive reviews of expression vectors suitable for the most commonly used bacterial host *Escherichia coli* is available.[3–6]

S0010 ### 3.06.1.1 Promoters and Control of Gene Expression

P0010 Expression cassettes containing strong inducible promoters, a convenient multiple cloning site, and an efficient terminator of transcription, and located on a high copy vector are needed in order to obtain high yields of a recombinant protein. Modifications are expression cassettes that provide dual promoters or multisystem expression vectors containing promoters that are active in bacterial as well as in mammalian cells.[7] The possibility of efficiently inducing the expression at a certain time point during the fermentation process is important as it minimizes metabolic burdens and potential toxic effects of the protein product. Fine-tuning the induction conditions can significantly optimize the expression of foreign proteins.[8] Common inducers employ low-cost chemicals or sudden modifications of the growth conditions. In fact, many promoters are suitable for producing proteins at high levels but they are difficult to completely switch off.[9,10] Prototypes are the *E. coli* P_{lac} promoter and its cAMP-independent P_{lacUV5} derivative that are frequently in use for protein expression. Induction is performed by the nonhydrolyzable lactose analog isopropyl-β-D-1-thiogalactopyranoside (IPTG) which releases the specific promoter-bound LacI repressor. Considerably increased expression levels, e.g., of up to 30% of the total cellular protein, can be obtained with the related artificial P_{tac} and P_{trc} promoters consisting of synthetic fusions of the P_{lac} with the P_{trp} promoter.[6,9] The intrinsic leaky expression from *lac*-derived promoters in noninduced cells can become cumbersome, sometimes causing a complete loss of protein expression or a reduced viability of the culture.[7] The tight repression of weaker promoters can be achieved in host strains containing the $lacI^Q$ allele which ensures enhanced synthesis rates of the LacI repressor. Furthermore, the *lacI* or $lacI^Q$ genes can be placed on the expression vector in order to additionally increase the LacI copy numbers. Alternatively, an expression system under control of the promoters P_{T7} from phage T7, P_{BAD} from the *E. coli* arabinose operon, or P_L and P_R from the λ phage might be preferred. The P_L and P_R promoters are among the strongest known in *E. coli*. They are tightly regulated and can be induced by temperature switches from 30 °C to 42 °C by virtue of a temperature-sensitive version of the λcI repressor (cI857) and they allow a free choice of host strains.[11,12]

P0015 The T7-promoter included in the pET system expression plasmids (Merck, Darmstadt, Germany) has become very popular.[13,14] This promoter is specifically recognized only by the T7-RNA polymerase encoded by the T7 gene 1 which can be provided either on a phage or plasmid vector or as integrated chromosomal copy in engineered host strains carrying the prophage λDE3 under control of the IPTG inducible P_{lacUV5} promoter.[7,13,15] A variety of DE3 derivatives in different cellular backgrounds (e.g., BL21, HMS174) is available (Merck, Darmstadt, Germany; Invitrogen, Carlsbad, CA). Virtually no background expression from the P_{T7} promoter can be observed in *E. coli* cells devoid of any T7-polymerase. Residual T7-RNA polymerase produced by λDE3 host strains due to the leakiness of P_{lacUV5} can be neutralized by coexpression of T7 lysozyme that efficiently binds to T7-RNA polymerase from compatible plasmids like pLysS and pLysE (Merck, Darmstadt, Germany). Other mechanisms for the control of leaky expression have also been proposed.[7] T7-RNA polymerase is extremely effective in the initiation of transcription and has a very processive transcription elongation rate. This could result in the accumulation of large amounts of messenger RNA (mRNA) causing a shortage of the cellular ribosomes. The unprotected mRNA is highly susceptible to degradation and can be stabilized by using the RNaseE mutant strain BL21 (DE3) Star (Merck, Darmstadt, Germany) for expression.

P0020 The activity of many strong promoters cannot be modulated after induction. Recombinant proteins can therefore accumulate very fast in the cell and the functional folding of the proteins might then not be able to keep pace. This causes a high risk of unfolded protein precipitating in the host cell as nonsoluble inclusion bodies. In order to solve this problem, a possible option could be the use of the P_{BAD} promoter in combination with its regulator protein AraC. It shows a very fast and tuneable response to changing concentrations of its inducer L-arabinose while heterogeneous cell populations might also account for that effect.[16,17] This allows the modulation of the stability and folding properties of a recombinant protein by controlling the rate of expression.[18] Corresponding vectors are commercially available (Invitrogen, Carlsbad, CA). The P_{BAD} promoter is subject to catabolite repression and a very tight repression can be achieved in the presence of glucose, fructose, and similar metabolites. The rapid kinetics of regulation allows a fast and efficient switch of protein expression from on to off.

S0015 ### 3.06.1.2 Regulatory DNA Sequences Important for Protein Expression

P0025 Regulatory sequences especially in the 5′ untranslated regions (UTR) are very different in *E. coli* if compared with eukaryotic control regions. A ribosome binding site (RBS, Shine–Dalgarno sequence) with the highly conserved consensus sequence 5′-UAAGGAGG-3′ needs to be present at 9 ± 3 bp upstream of the first codon that defines the initiation of translation. The RBS is complementary to the 3′ end of the 16S ribosomal RNA (rRNA) and an interaction is crucial for an efficient initiation of translation. Stable mRNA secondary structures covering the RBS or the initiation codon are usually detrimental to gene expression as they can interfere with ribosome binding. Expression constructs therefore usually represent transcriptional fusions where the regulatory sequences of promoters or terminators are provided by the expression cassette of the vector and only the coding sequence from the target gene has to be inserted.

P0030 Mammalian genes especially almost routinely require extensive modification of additional regulatory regions in their primary sequences prior to a high-level expression in *E. coli*. Strong translation initiation signals are an indispensable prerequisite for high-level expression as the initiation rate can dramatically determine the final expression level.[19] The sequence context immediately surrounding the translational start codon (up to about the first ten codons) is often critical.[20–22] This area should preferably be free from any rigid secondary structure formation since those could completely abolish any expression of the heterologous protein. Raising the numbers of adenosine residues in that area generally reduces the probability of secondary structure formation. A detailed analysis and systematic modification of the translation initiation region can significantly improve the expression. A practical approach for circumventing any problems with a suboptimal initiation of translation would be to produce a heterologous protein as a fusion C-terminal to a small leader peptide that is already codon-optimized. One possibility would be the T7 tag present in the cloning regions of many commonly used expression vectors of the pET series (Merck, Darmstadt, Germany).

S0020 ### 3.06.1.3 Codon Usage

P0035 A problem that frequently affects the yield of an expressed protein is the different codon usage of individual species.[23] A strong codon bias is most evident when prokaryotic and eukaryotic systems are compared. Codons that translate into proline and arginine are particularly affected in expression of human genes in *E. coli*.[24] The transfer RNAs (tRNAs) corresponding to rare codons are only found in minor populations in *E. coli* cells and coding sequences containing several rare codons (especially in iterative arrangements) are likely to become poorly expressed. Besides premature termination the misincorporation of a lysine residue in place of arginine at rare codon sites is a well-recognized problem in the structural analysis of proteins resulting in undesired heterogeneous populations in recombinant protein samples.[24–28] One option for addressing this problem is the genetic manipulation of coding sequences according to the codon preferences of the desired host cell. This approach certainly should be considered in the case of the de novo synthesis of a gene in vitro.[29] However, the subsequent introduction of silent mutations in an already cloned gene can be time-consuming as it probably requires multiple steps of mutagenesis. Alternatively, a eukaryotic gene can be expressed in *E. coli* in the presence of a second compatible plasmid encoding for an additional set of tRNAs complementary to rare codons.[30, 31] Specialized strains that are used for the high-level expression of rare codon containing genes are BL21 CodonPlus (Stratagene, La Jolla, CA) and the Rosetta-2 derivatives of BL21 (Merck, Darmstadt, Germany). The two strains harbor a plasmid containing extra copies of several rare codons. tRNAs that recognize.

S0025 ### 3.06.1.4 Translational Fusion Constructs for Optimized Protein Expression

P0040 In many cases it is preferable to express a target protein as a single polypeptide chain covalently linked with a second already well-defined protein. Fusion strategies provide the advantage that the engineered fusion products combine the properties of the individual proteins and thus permit high-level production of otherwise poorly synthesized recombinant proteins (**Table 1**). Target proteins with almost or completely unknown functions can be linked with the beneficial binding characteristics or optimal expression and stability properties of a fusion partner. The potential benefits of a fusion system include significantly enhanced expression levels, fast purification by means of affinity chromatography based on the binding properties of the fusion partner, increased probability of proper folding of the attached target protein, prevention of inclusion body formation, and protection from proteolytic degradation.[34] In addition, fusion partners can enable the immobilization of a protein to, e.g., biosensors for functional characterizations. The majority of the common fusion systems place the target protein at the C-terminus of a fusion partner that most likely ensures the transfer of better expression and stability characteristics. However, it should be borne in mind that interference of the fusion partner with the activity of the target protein can occur in some cases and its localization could make a difference. If possible, the effects of an attachment to either end of the target protein should therefore be compared.[35]

P0045 Fusion partners can be classified into small peptide tags and into larger proteins that can have additional benefits for the solubility or the stability of the target protein (**Table 1**). Small peptide tags can principally become attached to either end of a target protein while their accessibility can certainly differ depending on the specific tertiary structure of the target protein. Those tags usually cover about ten amino acid residues and they are efficiently used for the fast and convenient purification and/or detection of target proteins.[36]An advantage of using small peptide tags is that they usually do not affect the activities of the recombinant protein and they need not be removed upon a structural and functional characterization of the target protein. Extended processing and purification steps can therefore be avoided. Frequently employed purification tags include poly(His)$_x$ tag consisting mostly of six consecutive histidine residues and various forms of the biotin binding streptavidin peptide. Proteins containing poly(His)$_x$ tag derivatives usually exhibit a high affinity for Cu^{2+}, Ni^{2+}, Co^{2+}, or Zn^{2+} ions and can be purified from crude extract preparations by one-step affinity chromatography using metal-chelate resins even under denaturating conditions.[37] Also the exceptional strong interaction between biotin and streptavidin (K_D 10^{-15} M) ensures the highly selective binding of a fusion protein. Specific antibodies directed against poly(His)$_x$ tags and Strep-tags are available (Qiagen, Hilden, Germany; IBA GmbH, Göttingen, Germany) and the tags can thus be used for the purification as well as for the detection of a protein. An additional popular peptide tag used for the rapid identification of expressed recombinant proteins by an antibody reaction is the T7 tag inserted into many pET vectors (Merck, Darmstadt, Germany). Small peptide tags are furthermore often combined with larger fusion partners in order to optimize the rapid purification of the recombinant protein.[38]

P0050 Fusion proteins like glutathione *S*-transferase (GST), thioredoxin, ubiquitin, and others can significantly increase the expression yield of the attached target proteins (**Table 1**). Thioredoxin is a small monomer encoded by the *trxA* gene which facilitates the soluble expression of a number of mammalian proteins including growth factors and cytokines.[39,40] An affinity resin for the fast purification of thioredoxin fusions is commercially available (Invitrogen, Carlsbad, CA). The ubiquitin of baker's yeast *Saccharomyces cerevisiae* has been successfully used as a fusion tag for small proteins in *E. coli*.[41] A further advantage of that system might be the use of the specific yeast ubiquitin hydrolase in order to remove the tag after isolation of the fusion protein.[42] GST is popular as a purification tag because of its high affinity to glutathione sepharose, and it is additionally effective as an enhancer for the translation efficiency and for the solubility of the target protein.[35,43,44] The *E. coli* maltose binding protein (MBP) has an exceptional and relatively general ability to promote the solubility and stability of the fused recombinant proteins.[38,45,46] The fusion proteins can be easily purified by virtue of the MBP affinity to immobilized amylose resin.[47] The MBP fusion system is available as a kit (New England Biolabs, Frankfurt, Germany) which provides vectors for cytoplasmic or even periplasmic expression by taking advantage of the native MBP leader peptide.

T0005 **Table 1** Characteristics of popular fusion tags for the optimized expression of recombinant proteins.

Fusion tag	*Size*[a]	*Purification*	*Solubilization*	*Stabilization*	*Source*
(His)x-tag	6–12	+	−	−	Various suppliers
Strep-tag	6–12	+	−	−	IBA GmbH[b]
Protein D	110	−	±	+	Reference 32
Protein G(B1)	55	−	±	+	Reference 33
Glutathione *S*-transferase (GST)	230	+	±	+	Various suppliers
Thioredoxin	109	+	±	+	Various suppliers
DsbA	208	−	+	+	Merck[c]
DsbC	235	−	+	+	Merck[c]
MBP	395	+	+	+	New England Biolabs[d]
NusA	495	−	+	+	Merck[c]

[a] Size in amino acid residues; the exact size depends on the individual cloning strategy.
[b] IBA GmbH, Göttingen, Germany.
[c] Merck, Darmstadt, Germany.
[d] New England Biolabs, Frankfurt, Germany.

T0010 **Table 2** Frequently used restriction proteases suitable for the cleavage of fusion proteins.

Restriction protease	*Recognition site*[a]	*Source*
Enterokinase	D-D-D-D-K-↓	New England Biolabs[b]
Factor Xa	I-E/D-G-R-↓	Various providers
Thrombin	L-V-P-R-↓-G-S	Various providers
Tobacco etch virus (TEV) protease	E-N-L-Y-F-Q-↓-G	Invitrogen[c]
PreScission protease	L-E-V-L-F-Q-↓-G-P	Amersham Biosciences[d]

[a] The recognition sequence in the one letter code is given; the arrow indicates the site of proteolytic cleavage.
[b] New England Biolabs, Frankfurt, Germany.
[c] Invitrogen, Carlsbad, CA.
[d] Amersham Biosciences, Freiburg, Germany.

P0055 Large fusion partners are usually not desired if a structural analysis by nuclear magnetic resonance (NMR) spectroscopy of the expressed protein is under consideration, as the increased number of signals would result in spectra that are too crowded. The relatively small monomeric 11.6 kDa bacteriophage λ head protein D might be considered for this purpose. It shows excellent expression properties, has high thermal stability, and mediates optimal initiation of translation while reducing the risk of inclusion-body formation and protein degradation.[32] Also the small protein G (B1 domain) has been presented as solubility-enhancement tag especially for NMR purposes.[33]

P0060 It is often desired to remove the fusion partner after purification in order to obtain a nonmodified target protein. For that purpose, the recognition site for a highly specific "restriction protease" can be introduced in the linker between the two proteins (**Table 2**). An elegant way that eliminates any additional incubation steps after purification is the intracellular cleavage of the fusion protein directly after translation by virtue of the low-level coexpression of a suitable protease.[48,49] It should be considered that some proteases cleave in between their recognition sites and thus still leave residues attached to the N-terminal end of the target protein (**Table 2**). In addition, it is important to realize that the removal of the fusion partners can reintroduce the initial stability problems, and it may happen that the proteolytic processing is accompanied by unfolding and precipitation of the target protein. The production of fusion proteins also must not necessarily result in functional proteins. However, they still can provide a convenient means of generating antibodies that are specific for the protein under study.

S0030 3.06.2 Preparative-Scale Protein Expression Techniques

P0065 The design and set-up of an ideal expression system depends on multiple factors and each step can influence the final success. Most system components can be considered as more or less independent modules that can be selected according to the specific requirements (**Table 3**). High level of production of a protein often requires combinatorial approaches where several individual parameters of an expression system have to be optimized. The expression in cellular backgrounds especially by using *E. coli* strains is probably the first method of choice in most cases: no other system is comparable in speediness, simplicity, and cost-effectiveness. However, a valuable toolbox of microbial and eukaryotic expression systems exists and the individual characteristics should carefully be considered in order to make the optimal choice (**Table 4**). Multiple parallel approaches in order to identify the optimal expression technique for a given protein might be the quickest route to success.[50–54] Specific applications like the production of modified eukaryotic proteins or the expression of difficult protein groups like membrane proteins or disulfide-bonded proteins often require special adaptations and the most powerful and highly promising techniques have to be discussed.

S0035 3.06.2.1 Protein Expression in *Escherichia coli*

Protein expression in *E. coli* is easy to handle and often reveals a cost-effective and high-level production of heterologous proteins.[51,54] The cells can be grown to very high densities in inexpensive complex media within a short time and yields of recombinant protein up to 1 $g l^{-1}$ have been obtained.[17,55] Numerous elaborated protocols for the fast and efficient manipulation of *E. coli* have been established, a variety of options are possible in order to adapt the expression system to the specific requirements (**Table 3**), and a wide collection of expression vectors with well-defined and strong promoters is available. Many different *E. coli* host strains optimized for a high level of expression of recombinant

T0015 **Table 3** Rational strategies for the design of a procaryotic expression system.

Parameters	*General options*
System components	
Host species	*Escherichia coli* (general), *Bacillus subtilis* (e.g., export proteins), *Lactobacillus lactis* (e.g., membrane proteins)
Host strain	Protease deficiency, oxidizing cytoplasm, engineered genotype (e.g., DE3, lacI^q)
Expression vector	Copy number, selection marker, host range, compatibility
Expression cassette	Promoter, induction conditions, available cloning site, peptide tags, or fusion partners
Fermentation conditions	Medium components, addition of cofactors, incubation temperature, oxygen supply
Target gene sequence	
Initation of translation	Optimized environment of the translational start codon
Codon usage	Silent mutagenesis, coexpression of rare codon tRNAs
Targeting	Addition of export signal sequences
Translational fusions	Addition of peptide tags or fusion proteins for better purification and/or stability
Protein characteristics	
Low solubility/stability	Prevention of inclusion bodies (e.g. coexpression of chaperones, addition of solubility tags)
Membrane associated	Construction of fusion proteins, specific host strains
Proteolytic degradation	Construction of fusion proteins, specific host strains
Disulfide bridges	Coexpression of chaperones, engineered host strains, periplasmic expression
Modification	Coexpression of modifying enzymes, addition of cofactors

proteins are available from various suppliers (Invitrogen, Carlsbad, CA; Merck, Darmstadt, Germany; Stratagene, La Jolla, CA). While it is true that some groups of proteins are more likely to become produced in *E. coli* than others, a wealth of published experience exists that this organism is of importance for almost any application. Depending on the desired protein product, a variety of proven optimization strategies are documented that can considerably accelerate the high-level production of a recombinant protein in the desired form.[56]

S0040 ### 3.06.2.1.1 Inclusion Body Formation and Coexpression of Chaperoneas

P0070 Unfortunately a frequent observation in *E. coli*, especially in production of eukaryotic proteins, is the formation of inclusion bodies, which are precipitates composed of unfolded and inactive recombinant protein.[4] Protein folding in prokaryotes and eukaryotes follows different pathways. The biosynthesis of proteins in prokaryotes usually proceeds very fast and can thus cause the precipitation of proteins that require a long time to achieve their functional folding. In addition, specific cofactors, chaperones, or accessory proteins essential in order to complete the folding process of a eukaryotic protein might be absent in the prokaryotic environment. A shortage of ligands or substrates could additionally increase the instability of an overproduced protein. Heterologous proteins expressed in *E. coli* therefore do not fold spontaneously into a native and functional state even though they are produced at significant levels.

P0075 Several protocols exist that result in the in vitro refolding of proteins out of isolated inclusion bodies.[57,58] However it is generally intended to produce recombinant proteins in a folded and active state. Several pathways for the proper folding of proteins in prokaryotes as well as in eukaryotes are well characterized and key enzymes have been identified and cloned. Best studied are the prokaryotic GroEL/GroES and DnaK/DnaJ/GrpE systems.[59,61] GroEL/ES forms large

T0020 **Table 4** Comparative evaluation[a] of expression systems.

	Expression system				
Parameter	*Bacterial*	*Yeast*	*Mammalian*	*Viral*	*Cell-free*
System characteristics					
Cost-effectiveness	+ + +	+ +	+	+	+[b]
Robustness	+ + +	+ + +	+	+	+
Variety of vectors, host strains	+ + +	+ +	+ +	+ +	+
Speediness of protein production	+ + +	+ +	+	+	+ + +
Set-up of reaction	+ + +	+ +	+	+	+ +
High expression levels	+ + +	+ + +	+	+ +	+ +
Inclusion body formation	+ +	+	+	+	+
Stabilization of recombinant proteins[c]	+	+	+	+	+ +
Specific applications					
Posttranslational modifications	−	+ +	+ +	+ +	+
Secretion of proteins	+ +	+ +	+	+ +	−
Labeling of proteins	+ +	+	±	±	+ + +
Membrane proteins	+	+	+	±	+ + +
Disulfide-bridged proteins	+	+ +	+ +	+ +	+ +
Toxins	±	−	−	+ +	+ + +
Proteins requiring cofactors	+	+	+	+	+ +

[a] The rating ranges from '+ + +', best choice to '−', not recommended.
[b] Commercial systems are relatively expensive.
[c] Stabilization by low intrinsic protease levels, by coexpression of additional proteins, or by the addition of stabilizing compounds.

multisubunit cylinders that actively promote the folding of many proteins through an ATP-dependent cycle.[61,62] The DnaJ activated DnaK interacts with hydrophobic regions in nascent peptide chains and prevents intra- and intermolecular aggregations before the freshly synthesized proteins have adopted their native structure.[60] The nucleotide exchange factor GrpE mediates the final complex dissociation and proteins are released for a folding cycle. The coexpression of cloned chaperones or other beneficial "helper" proteins can be advantageous in enhancing the stability or the general expression of recombinant proteins.[63–66] A variety of plasmids have been constructed that could be used in order to express the most prevalent chaperone systems in combination with a target protein.[67,68] Boosting the intracellular level of chaperones in *E. coli* can significantly enhance the expression of soluble recombinant proteins.[69]

P0080 Coexpression can be addressed principally either by bicistronic or by compatible dual-vector approaches.[70] Bicistronic vectors could contain several reading frames arranged as an operon that is controlled by a single promoter. Unfortunately, genes that are located more distantly to the promoter are commonly much less expressed.[71] A single dual-promoter expression vector that provides two cloning sites each having a separate promoter could alternatively be considered.[72] The use of different induction mechanisms would additionally allow the fine-tuning of the individual expression levels. Coexpression by a dual-vector system usually employs two different compatible vectors for each cloned target protein. However, one vector often becomes dominant because of differences in the copy numbers.[73] Such a bias in copy number may be balanced if different selection pressures are maintained.[74] Dual-vector systems can further be used to express several individual subunits of a heteromeric protein complex in a single host cell.[73,74]

S0045 3.06.2.1.2 **Solubility Tags and Stable Expression of Recombinant Proteins**

P0085 The expression of fusion proteins with a highly stable fusion partner at the N-terminal end of the target protein can often provide a successful approach in order to produce soluble and folded heterologous proteins in *E. coli*.[44,75] A fusion partner with a good probability for transmitting a higher solubility to covalently attached target proteins is the *E. coli* MBP.[44] Precursors of mammalian proteinases show enhanced solubility even with C-terminal attached MBP.[76] Hydrophobic areas at the surface of MBP are supposed to confer a rather unspecific chaperone-like assistance in the folding or refolding of aggregation-prone proteins in *cis*. Other fusion partners like thioredoxin, GST, phage λ protein D, or derivatives of DsbA can act in a similar way to prevent formation of inclusion bodies or to enhance the solubility of specific target proteins, but they seem to be not nearly as generally efficient as MBP.[37,38,44,46,75]

P0090 The fermentation temperature may have a pronounced effect on yield, and on the folding and stability of recombinant proteins.[77] Temperatures below 30 °C can significantly improve the solubility of a protein. A reduction of the temperature down to 25 °C about 1 h before induction of expression can generally be recommended for the production of critical proteins. The lack of essential cofactors such as hemes or flavins or simply ions like Mg^{2+} can further prevent the folding process of a recombinant protein upon overexpression. Hence, the accumulation of unstable folding intermediates may finally result in formation of inclusion bodies. The yields of soluble protein should be increased by either enhancing the endogenous cofactor production or by an exogenous supply to the culture.

S0050 3.06.2.1.3 **Targeting of Recombinant Proteins**

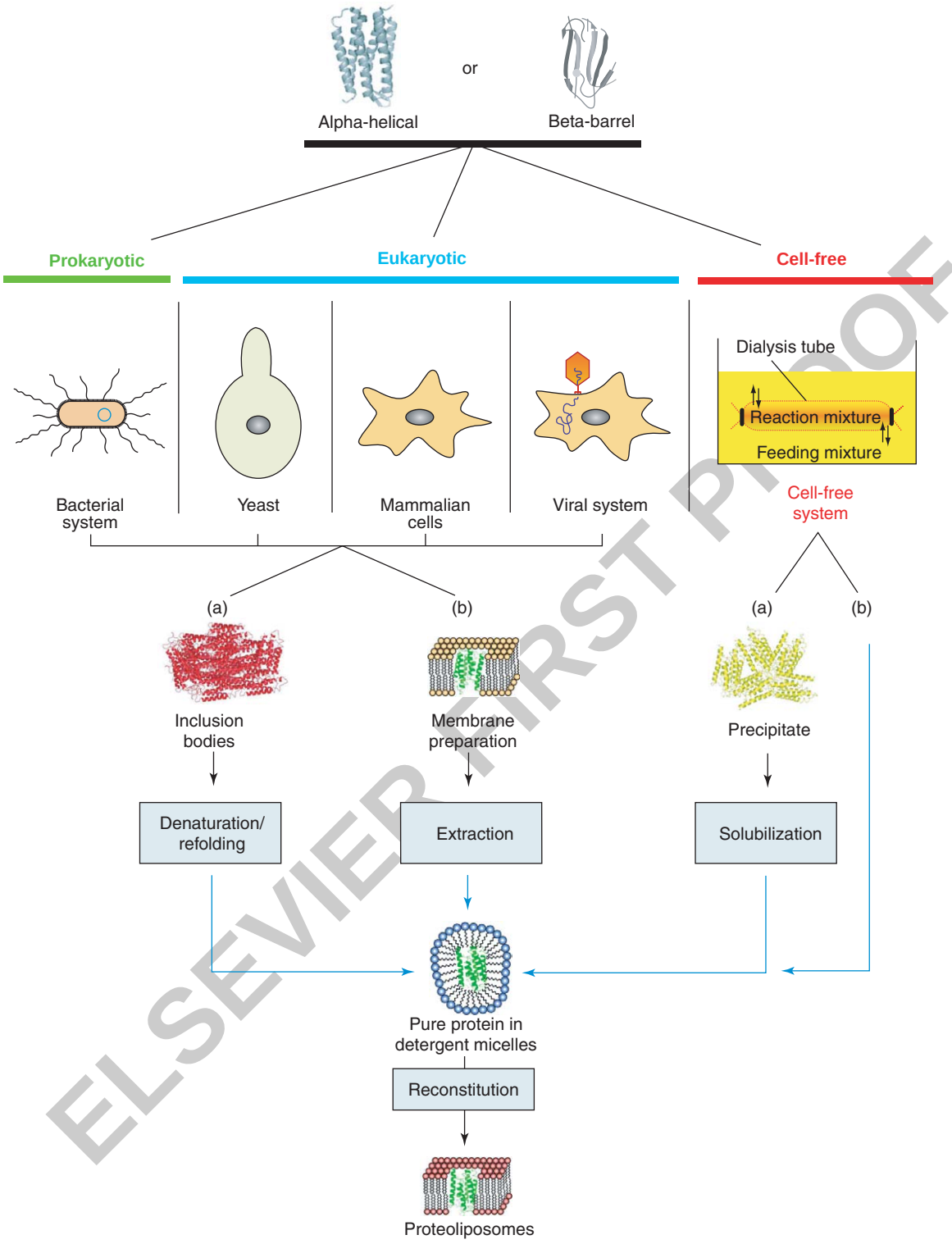
P0095 In general, three compartments can be considered in *E. coli* for the targeting of the expressed protein: the cytoplasm, the periplasm, and the export of the protein into the extracellular medium. Production of proteins in the periplasm has the advantage that correct disulfide bonds can be formed due to its oxidizing environment and due to the presence of disulfide isomerases.[78,79] The microenvironment of the periplasm can be manipulated much more easily with regard to pH, redox state, and composition in order to optimize the folding of a recombinant protein or to prevent degradation.[80,81] The majority of proteins exported to the periplasm or to the outer membrane of *E. coli* use the general secretory pathway consisting of an array of secretory (Sec) proteins located in the cytoplasm and in the inner membrane.[82–84] One major targeting mechanism in *E. coli* is exemplified by the attachment of specific[20–25] amino-acid-long hydrophobic extensions known as signal or leader sequences to the N-terminus of a protein. The leader will be cleaved off upon translocation thus leaving recombinant proteins with their natural N-terminal ends.[34] Direction of a recombinant protein into the periplasm can simply be anticipated by fusing a naturally occurring leader sequence of an efficiently exported protein like OmpA, β-lactamase, or alkaline phosphatase to its N-terminus. However, the translocation process of the recombinant protein often interferes with the translocation machinery due to premature folding, aggregation, or sterical blocking by bulky secondary structures. Successful and efficient export can therefore hardly be predicted. It is generally difficult to facilitate the secretion of large amounts of protein and a decrease in the yield down to at least 10% to that of the cytoplasmatic expression should be calculated. Secretion of proteins is also usually not very efficient in *E. coli*. Approaches to direct heterologous proteins to the extracellular space rely on gene fusion technologies, on the use of dedicated translocators like hemolysin,[85] or on coexpression of proteins that damage the outer membrane like Kil,[86] bacteriocin release proteins,[87,88] and the membrane protein TolAIII.[89]

S0055 3.06.2.1.4 **Expression of Membrane Proteins**

Difficulties in obtaining sufficient amounts of recombinant protein are most pronounced with membrane proteins (**Figure 1**). Mechanisms of membrane targeting and processes leading to the membrane insertion of proteins significantly differ between prokaryotes and eukaryotes. Attachment and targeting by the signal recognition protein, a prevalent mechanism in eukaryotes, does not occur in prokaryotes. Moreover, toxic effects of the overproduced proteins frequently cause a reduced growth or even the death of the host cells. Selection strategies have resulted in the isolation of the specially adapted mutant strains C41 (DE3) and C43 (DE3) from the common host strain BL21 (DE3), which better tolerate the overproduction of membrane proteins.[90] These strains contain proliferated membranes, and recombinant proteins are further directed at elevated levels into inclusion bodies. Culture conditions and genetic manipulations can therefore be crucial for obtaining high yields of a recombinant membrane protein.[91,92] Furthermore, the assembly of membrane proteins in *E. coli* requires complex translocation machineries that may provide promising targets for the optimization of expression.[93]

S0060 3.06.2.1.5 **Proteolytic Degradation of Recombinant Proteins**

P0100 Abundant amounts of heterologous proteins are often prone to proteolytic degradation as they are recognized as abnormal by the proteolytic machinery of the host cell. Early degradation is mainly connected with the five major ATP-



F0005 **Figure 1** Pathways for the production of membrane proteins in different expression systems. a, production as unsoluble protein, followed by solubilization and reconstitution. The solubilization from inclusion bodies often requires the initial denaturation of the protein. b, isolation or production as soluble protein. Blue arrows indicate the individual purification steps that result in pure membrane protein solubilized in micelles.

dependent systems: Lon/La, FtsH/HflB, ClpAP, ClpXP, and ClpYQ/HslUV.[94] Expression in protease-deficient strains could thus be advantageous in order to improve the yields of a recombinant protein, but those mutations often cause growth deficiencies.[95] Protease-sensitive proteins could be targeted to the periplasmic space or into inclusion bodies that generally are protected from proteolysis. A preferred strategy in order to prevent degradation is the production of hybrid proteins by the generation of translational fusions.[96] A heterologous protein can thus be shielded by a covalently attached protein that is normally abundant and very stable in the cell.

S0065 #### 3.06.2.1.6 Expression of Disulfide-Bonded Proteins

P0105 The reducing environment of the *E. coli* cytoplasm and the presence of thioredoxins (TrxA) and glutaredoxins (GrxA-C) that continuously reduce transiently formed disulfide bridges usually prevent the high-level production of oxidized proteins.[79] Consequently, efficient oxidation of recombinant proteins can be achieved by using strains that are defective for some of the major reductases.[97] The BL21 Origami derivatives (Merck, Darmstadt, Germany) have been engineered for a superior expression of disulfide-bonded proteins. These strains carry chromosomal mutations in the thioredoxin reductase (*trxB*) and glutaredoxin reductase (*gor*) genes which are both necessary for the maintenance of a reducing environment in the cytoplasm. Fermentation at reduced temperatures can also favor the oxidation of proteins. An alternative approach is the export of recombinant proteins into the oxidizing environment of the periplasm. The presence of disulfide oxidoreductase and oxidoisomerase (DsbA-C) systems that catalyze the formation and reshuffling of disulfide bonds can considerably facilitate the oxidative folding steps of a recombinant protein in the periplasmic compartment.[64,78,98,99] A further strategy would be the coexpression of oxidizing and disulfide-bridge shuffling chaperones like TrxA, DsbA, DsbC, or the human protein disulfide isomerase (PDI).[100,101]

S0070 ### 3.06.2.2 Protein Expression in Bacterial Hosts other than *E. coli*

P0110 A variety of bacterial expression hosts other than *E. coli* has been described, and high-level production of recombinant proteins can be achieved with some of them. Limitations in the availability of genetic tools and of expression vectors usually restrict more extensive applications. Nevertheless, obvious advantages of choosing non-*E. coli* hosts include: enhanced expression rates due to a different codon usage, more industrial applications of recombinant proteins after their production in a considered-as-safe organism, and increased efficiencies in the secretion of the target protein. Several proteins could be produced in food-grade microorganisms such as lactic acid bacteria or *Corynebacterium glutamicum*.[102–104] *Lactococcus lactis* has recently been proposed as a promising host for the overproduction of functional membrane proteins of prokaryotic and eukaryotic origin and may represent an interesting alternative to other systems (**Figure 1**).[105] Growth rates are similar to that of *E. coli* and essential techniques for genetic manipulations are available. Expression vectors for lactic acid bacteria are often based on broad-range replicons that also replicate in *E. coli* and a selection of inducible strong promoters is available.[104–107] The efficient secretion systems of *Bacillus subtilis* and probably also of some *Streptomyces* species make these organisms suitable as expression hosts especially for the production of exported proteins.[108–110]

S0075 ### 3.06.2.3 Yeast Expression Systems

P0115 Eukaryotic expression systems are generally more difficult to handle. Slow growth rates, higher costs, and lower efficiencies are often attributes of eukaryotic expression systems and they are therefore usually consulted when the expression of a recombinant protein in *E. coli* is unlikely for some reason or already has failed. Important and obvious advantages of eukaryotic expression systems are the possibility of obtaining native-like modification patterns of the recombinant protein like glycosylation, phosphorylation, and lipidation. The folding pathway of eukaryotic proteins may in addition depend on specific chaperone systems that do not occur in prokaryotes. The activation of proteins by an attachment of cofactors is also often not possible in *E. coli*. An interesting alternative might be the use of expression systems based on yeast cells. They resemble bacterial systems in many features but still contain important characteristics of eukaryotic expression systems. Many basic mechanisms for the processing and targeting of proteins are very similar to those in mammalian cells. Yeast cells contain an endoplasmatic reticulum and a membrane topology similar to mammalian cells. Efficient pathways for the posttranslational modification of proteins are present and they are frequently essential in order to maintain the stability or the activity of a protein. However, it should be considered that the type of glycosylation in yeast cells sometimes differs from the pattern obtained in mammalian cells. The yeast system can also be very powerful especially with regard to the production of large quantities of secreted proteins.[111–113]

P0120 Several yeast species can be used for the production of recombinant proteins. Most popular are baker's yeast *Saccharomyces cerevisiae*[114,115] the species *Pichia pastoris*,[116,117] and *Hansenula polymorpha*.[118] Other species like *Klyveromyces*

sp., *Yarrowia lipolytica*,[119] and *Zygosaccharomyces bailii*[120] might gain increasing interest in the next future. An evident advantage is the unicellular organization of yeasts, which enables the direct transfer of many techniques and manipulations commonly applied for bacteria. Comprehensive protocols exist for the transformation of yeast with foreign DNA. Furthermore, yeasts can be grown in relatively simple and inexpensive media and they reach high cell densities in reasonable times.

P0125 An elaborate collection of expression vectors for yeasts is available. Common features are regulatory regions for replication in *E. coli* and a prokaryotic selection marker. This enables the construction and propagation in bacterial cells. The vectors are classified based on their mode of replication into the groups of episomally replicating vectors and chromosomal integration vectors. Stable maintenance of high copy number requires regulatory sequences from the endogenous yeast 2 micron plasmid, while origins of replication from yeast autonomous replicating sequence (ARS) elements are much less stable with loss rates of up to 10% per generation. Integrative vectors are devoid of any origins of replication and they can only be distributed to daughter cells if they integrate into the yeast chromosome by homologous recombination. The efficiency of integration can thus considerably be increased if sequences homologous to chromosomal areas are provided and multiple integrations can be frequent.[117,121] Transformed yeast cells can be screened by a number of selection markers mainly based on auxotrophic deficiencies like HIS3, LEU2, LYS2, TRP1, and URA3.[114] A set of inducible strong promoters is available. Commonly used are the methanol inducible alcohol oxidase 1 (AOX1) and alcohol dehydrogenase promoters (ADH2), the metallothioneine promoter, and galactose inducible promoters (GAL1, GAL7, Gal10).[117,122] Shuttle vector systems suitable for the dual protein expression in *P. pastoris* as well as in *E. coli* are available for the convenient evaluation and rapid comparison of different expression hosts without time-consuming subcloning steps.[123]

P0130 Optimized fermenter cultures of *P. pastoris* are able to produce more than $1\,g\,l^{-1}$ of recombinant protein and the ability to grow on defined minimal media makes this species suitable for the efficient labeling of expressed proteins for a structural analysis.[117,122,124] Interesting applications of yeast expression systems are the high secretion levels of recombinant protein,[125,126] the production of disulfide-bonded proteins,[127] and their potential for the high-level production of membrane proteins (**Figure 1**).[128]

S0080 3.06.2.4 Mammalian Expression Systems

P0135 Expression in mammalian cells provides an environment mostly closely related to the native origin of many eukaryotic proteins. A variety of genetically and phenotypically diverse mammalian cell lines have been used for protein production.[129] Frequently used are baby hamster kidney (BHK) cells,[130] Chinese hamster ovary (CHO) cells,[131] human embryonic kidney (HEK) 293 cells,[132] mouse L-cells,[133] and various myeloma cell lines, which are used to establish relatively stable protein expression systems. Further cell lines for more specific applications are NIH 3T3 cells,[134] murine erythroleukemia (MEL) cells,[135] or lymphoblastoid cell lines like Namalwa[136] and RPMI 1788.[137] Main parameters for the high-level expression of recombinant proteins are the selected cell lines and the specific vector characteristics.[138] Stable productive cell lines are established principally by selecting homogenous cell populations from heterogeneous cell pools. The yield of recombinant protein can often increased by selective amplification with agents like methotrexate (MTX).[139]

P0140 DNA can be transferred into mammalian cells by transfection of coated or encapsulated DNA particles which become incorporated into the cells through endocytosis. Coprecipitation with calcium phosphate[140] or diethylaminoethyl dextran[141] is also relatively simple and highly effective. Lipofection of compact liposome/nucleic acid complexes and electroporation of cells are further powerful techniques.[142,143]

S0085 3.06.2.4.1 Vector Design for Gene Transfer and Expression in Mammalian Cells

P0145 Transcriptional and translational control elements, RNA processing, gene copy numbers, stability of mRNA, the site of chromosomal integration, and the impact of recombinant proteins in the host cell are important parameters for efficient gene expression in mammalian cells. A large number of vectors are available and key elements are strong promoters and enhancers of the promoter activity.[144,145] Conserved and essential sequences of eukaryotic promoters are the TATAA box and the CAAT box that are located approximately 30 and 80 bp upstream of the mRNA initiation site, respectively. Frequently used constitutive promoters are the adenovirus major late promoter, the human cytomegalovirus immediate early promoter or the simian virus 40 (SV40) and Rous sarcoma virus (RSV) promoters. However, similar to bacterial expression systems, an inducible system is desired in most cases as the produced proteins might become toxic to the host cell.[146] The well-known bacterial *lac* promoter–repressor system was therefore adapted for its use in mammalian cells yielding up to a 10 000-20 000-fold induction of gene expression.[147–149] Selection markers for successfully transfected mammalian cells include the genes encoding for dihydrofolate reductase (DHFR), for aminoglycoside

phosphotransferase providing neomycine or G418 resistance,[150] or for enzymes providing resistance against puromycin, hygromycin, or placitidine.[144,151]

P0150 Optimized splicing of introns present in most genes of higher eukaryotes is important. The presence of introns in mRNAs may lead to a 10- to 20-fold increased expression rate.[152] Eukaryotic mRNAs have poly(A) tails attached to their 3′ end which modulate their stability and efficiency of translation.[153,154] Effective poly(A) attachment signals in mammalian expression vectors are derived from the genes of mouse β-globulin,[155] herpes simplex virus thymidine kinase,[156] bovine growth hormone,[157] or the SV40 early transcription unit.[158] The initiation of translation depends on a conserved sequence enclosing the AUG start codon (CC(A/G)CCaugG), named the Kozak sequence.[159] Stable hairpin structures of GC-rich regions in the 5′ UTR can further negatively influence the efficiency of transcription.[160]

S0090 #### 3.06.2.4.2 Modified Mammalian Cell Expression Systems

P0155 The copy number of a heterologous gene may significantly increase if it is physically linked to a chromosomal DNA region that is subject of extensive amplification processes. The associated target gene will then become coamplified, resulting in higher yields of recombinant protein due to the increased gene copy numbers. Spontaneous gene amplification in mammalian cells is unfortunately only a rare event, but a variety of artificial gene amplification techniques have been developed that may ultimately lead to considerable gene amplification frequencies.[161] A broad selection of agents and treatments have been found to be active in inducing the amplification of genomic regions. Reagents like hydroxyurea, aphidicolin, hypoxia, and carcinogens or the treatment by Ultraviolet- or γ-irradiation have been used.[162,163] Frequently used gene amplification systems take advantage of DHFR as a selectable marker present on the expression vector. DHFR catalyzes the reduction of 5,6-dihydrofolate to 5,6,7,8-tetrahydrofolate, a biocatalyst for the essential synthesis of glycine and thymidine monophosphate, and for purine biosynthesis. CHO cells carrying a DHFR deficiency due to chemical mutagenesis are not able to grow in a medium depleted of nucleosides unless they carry a functional DHFR gene through transfection.[139] The folic acid antagonist MTX binds and inhibits DHFR stoichiometrically. Therefore, stepwise-increased MTX concentrations in the medium can be used to effectively select transfectants that amplify the DHFR gene, resulting in cell populations containing several hundreds of gene copies.[164,165] High levels of recombinant heterologous protein can be achieved by cotransfecting and cointegrating plasmids carrying the DHFR marker and the gene of interest at the same site in the genome.[164] Also an expression cassette including a promotorless DHFR gene can be used. The inserted gene of interest is in that case transcribed from a strong promoter upstream of the DHFR open reading frame, producing a bicistronic message and thus protein production dependent on MTX resistance.[166,167]

P0160 A sophisticated technique for antibody production is the fusion of unrestricted and fast-growing myeloma cells with other cells. Fused myeloma and plasma cells create antigen-specific hybridoma cell lines. These fast-growing secretory cell lines are well suited for the expression of recombinant transfected genes and predestinated for the synthesis of monoclonal antibodies.[168,169] A further positive aspect of hybridomas is their natural growth in suspension cultures that saves time-consuming adaptation procedures in the case of large-scale production. The mouse myeloma cell line Sp2/0 Ag14 is a preferred choice for recombinant protein production. It contains deficiencies for synthesizing and secreting endogenous immunoglobulin, is easy to transfect and can proliferate in large-scale serum-free cultures.[170]

P0165 The use of powerful viral transcription/translation machineries is an elegant technique for the protein synthesis in a AU :1 mammalian expression system. COS cell lines are most exclusively used for the transient expression of heterologous genes dependent on expression of SV40 large T-antigen, although several cell lines are suitable.[171–175] Three cell lines (COS-1, COS-3, and COS-7) have been established by transforming the African green monkey cell line CV-1, permissive for the lytic growth of SV40, with an origin-defective SV40 genome and thus creating a cell line harboring the SV40 large T-antigen expression. Transfection with an expression plasmid containing a functional SV40 origin of replication and the gene of interest enables the interaction between the SV40 large T-antigen and the SV40 origin of replication and can finally yield high recombinant protein titers.[176–178] The plasmid replication in COS cells is highest at about 48 h after transfection and the system is not able to produce high amounts of protein over a prolonged period of time.[129] This problem can be solved by using the COS system in extended and modified batch modes that allow multiple harvests and that can result in the cumulative production of preparative amounts of recombinant protein.[179,180]

S0095 #### 3.06.2.5 Viral Expression Systems

P0170 Viral expression systems are elaborate ways to synthesize heterologous proteins. The systems take advantage of the viral nature to introduce foreign DNA with high efficiency into host cells by infection combined with high replication rates due to strong viral promoters. The most popular viral systems include the already mentioned monkey tumor virus

SV40,[175] the baculovirus/insect cells system,[181] and the Semliki Forest virus (SFV) that is used with a wide range of mammalian host cells.[182] Further less prevalent systems use the Epstein–Barr virus,[183] the cytomegalovirus (CMV), RSV, or the SFV-related Sindbis virus.[184] Besides their application in expression systems viral vectors are widely used for vaccines and gene therapies.[185]

S0100 3.06.2.5.1 **The Baculovirus Expression Vector System**

P0175 Protein production in insect cells in combination with the baculovious expression vector system (BEVS) presents several advantages like the ease of culture, a high tolerance to by-product concentrations, and high expression levels.[181] Baculoviruses belong to a family of double-stranded DNA viruses with large circular genomes of 120–180 kbp.[186] They are invertebrate-specific and the host range is restricted mainly to arthropods with about 600 insect species reported to be targets of infection.[129] The characteristic production of large amounts of protein during the late phase of infection has implicated their adaptation as vectors for heterologous gene expression. Probably best characterized is the *Autographa californica* nuclear polyhydrosis virus (AcNPV). The two very late gene products, polyhedron and p10, are expressed from strong promoters at levels up to 50% of the total protein content of the infected cell.[187] These proteins are not essential for the formation of viral particles and are thus predestinated to become replaced by heterologous gene products. Diverse expression vectors have been developed that depend on specific host cell lines for infection and propagation.[188–191] The most frequently used insect cell lines are derived from *Spodoptera frugiperda* (SF_9 and SF_{21}) and from *Trichoplusia ni* (Hi_5 and MG_I), having generation times up to 24 h and requiring complex culture media.[187] Various approaches of large-scale protein synthesis in insect cells have been reviewed.[181] Of great interest is the capability of insect cells to perform posttranslational modifications. Cell lines optimized in order to obtain the most favorable glycosylation patterns have been developed.[192,193] The demonstrated high level of expression of several functional G protein-coupled receptors (GPCRs) makes baculovirus-based expression systems very attractive for the preparative production of targets from this important class of membrane proteins (**Figure 1**).[194,195] Activation of the "late" polyhedron promoter lags significantly behind the infection process and makes the insect/baculovirus system also potentially useful for the expression of cytotoxic recombinant proteins. Recombinant proteins can even effectively become secreted into the medium by adding a signal sequence like the honey bee (*Apis mellifera*) prepromelittin secretory sequence.[196]

S0105 3.06.2.5.2 **Semliki Forest Virus Expression Systems**

P0180 An effective and well-studied expression system for a rapid and high-level gene delivery is based on replication-deficient alphavirus vectors. These systems benefit from an enormous RNA replication capacity in the cytoplasm that consequently results in extreme expression levels. Most common are vectors of SFV, but similar expression systems have been engineered for the Sindbis virus[197] or the Venezuelan equine encephalitis virus.[198] The pathogenic properties of wild-type SFV, an enveloped single-stranded RNA virus, include induction of neuronal apoptosis, demyelination in the central nervous system, and teratogenesis.[199] These viruses are capable of infecting a broad range of mammalian, amphibian, reptilian, insect, avian, and fish cell lines. Mutant strains with drastically reduced virulence have been made in order to use these pathogenic vectors for experimental gene delivery approaches.

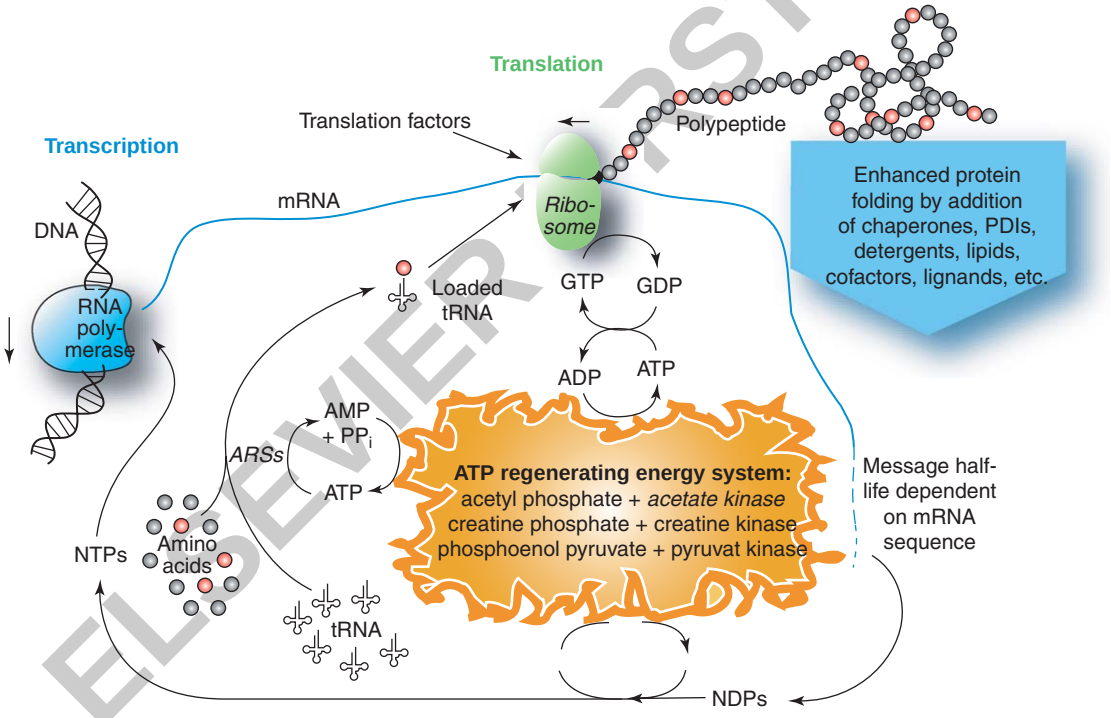
P0185 A typical expression system consists of two plasmids based on complementary (cDNA) copies of the SFV genome.[200] The expression vector, containing the SFV nonstructural genes (nsP1-4) and the strong SFV 26S promoter, has a multilinker cloning region into which foreign genes can be inserted. The second helper vector contains the genes for capsid and envelope proteins. The cotransfection of both vectors into BHK cells generates a high titer (10^9–10^{10} particles ml^{-1}) of recombinant and infectious SFV particles.[201] The use of the two-vector system guarantees the production of replication-deficient particles since the RNA packaging signal is only present in the recombinant RNA of the expression vector and no helper RNA will be packed. The infection of host cells will therefore lead to a rapid and high-level expression of recombinant protein without the generation of virus succession. A further advantage is the broad host range of SFV as parallel expression studies can be performed in different cell lines.

P0190 The SFV system suffers from safety risks and cytotoxicity, and novel less cytotoxic and temperature-sensitive mutant vectors that are inactive at the blood temperature of 36–37 °C have been developed.[182,202] The addition of a translation enhancement signal to the capsid gene resulted in five- to tenfold increased expression levels.[203] High amounts of topologically different proteins have successfully been expressed with SFV vectors. Besides β-galactosidase with a total cellular protein yield of approximately 25%, more than 50 GPCRs and several ion channels have been expressed at high levels (**Figure 1**).[201]

S0110 ### 3.06.2.6 Cell-Free Protein Expression

P0195 The described conventional in vivo technologies for protein production depend on the cellular integrity and are only suitable for the production of proteins that do not affect the physiology of the host cell.[204,205] As already discussed, these methods are limited for the expression of many proteins, e.g., by the formation of inclusion bodies,[206,207] by protein instability due to proteolysis,[205,208] or by too low yields in case of most membrane proteins. High-level cell-free (CF) expression systems are a promising new tool for the preparative production of difficult proteins (**Figure 2**). CF systems are principally independent of the cell physiology and they allow direct and immediate control of the reaction at any time. A wide range of critical reaction parameters such as pH, redox potential, and ionic strength can be chosen and adjusted according to the requirements of the specific target protein. Furthermore, any additives that might help to stabilize the recombinant protein after translation can be supplied directly into the reaction and no transport problems through cellular membranes have to be considered. Metabolic conversions or even the breakdown of added substances is usually very low or even not detectable. Options of possible beneficial compounds can include cofactors, ligands, inhibitors, ions, chaperones, and even detergents. The complete and time-independent control in combination with the high flexibility of the reaction conditions provides a challenging opportunity for the preparative production of formerly highly problematic proteins such as membrane proteins, toxins, or unstable proteins.

P0200 A unique characteristic of CF expression techniques is the possibility of quickly and easily introducing specific labels into a protein (**Figure 2**). Labeling of proteins with stable isotopes is indispensable for the structural and functional analysis by NMR techniques and for drug screening and ligand interaction studies. Labeling of proteins with spectrally enhanced amino acids can further be very helpful for the analysis of protein interaction studies.[209] The composition and concentration of all low molecular weight substances in the CF reaction is fully defined and the operator has therefore complete control over the amino acid pool of the reaction. Any type of amino acid can thus easily be replaced by a labeled derivative and a 100% label incorporation into the recombinant protein is ensured. The kinetics and



F0010 **Figure 2** Schematic illustration of cell-free protein synthesis in a bacterial coupled transcription and translation system. The cell extract contains ribosomes, translation factors acetate kinase, and aminoacyl-tRNA synthetases (ARSs). T7 RNA polymerase and substrates like amino acids, the energy regenerating system components, nucleotide triphosphates (NTPs), tRNAs, and salts are supplied and protein synthesis is initiated by adding template DNA. The incorporation of selected isotopic labeled amino acids (red) can easily be achieved in the cell-free system, leading to a selectively isotopic labeled protein. Regeneration of NTPs is accomplished by an ATP regenerating energy system (yellow) based on the hydrolysis of high-energy substrates in the presence of their cognate kinases. To assist stability and folding of the target proteins, detergents, chaperones, and other supplements may be added to the reaction mixture.

efficiency of the production of labeled proteins equals those of the nonlabeled reactions and no laborious optimizations have to be carried out. In contrast to conventional in vivo labeling systems, no switch to auxotrophic host strains and to fermentations in minimal media is necessary. It should be emphasized that there is virtually no background labeling as the target protein is the only protein that is synthesized in substantial amounts during CF synthesis. High-throughput applications attract increasing attention due to the demands of proteomics associated research. CF synthesis also offers a powerful approach as linear DNA templates simply generated by conventional polymerase chain reaction (PCR) can directly be used for the expression of proteins.[210]

S0115 #### 3.06.2.6.1 Design of Cell-Free Expression Systems

P0205 CF expression systems can be set up as pure translation systems with purified mRNA as a template, or alternatively as coupled transcription–translation systems by adding plasmid or linear DNA as a template.[211,212] The simplest design of CF expression is a batch mode reaction with one compartment holding a fixed volume of reaction mixture (RM) in a test-tube. While batch systems are easy to set up and highly suitable for high-throughput applications, they are limited by the rapid accumulation of deleterious by-products like free phosphates from nucleotide consumption that apparently form complexes with magnesium ions. In addition, substrates like nucleotide triphosphates (NTPs) and high-energy phosphate donors are rapidly consumed even in the absence of protein synthesis.[213] The results are obtained in relatively short reaction times that normally do not exceed 1 h. Recent modifications of the batch system by using a novel NTP regeneration system can help to prevent the accumulation of inorganic phosphate,[213] and the supplementation of additional compounds can significantly extend the reaction time and increase the yield of protein synthesis.[214–219] Optimized CF batch systems have the potential to reach a high productivity and may become a powerful technique in the future.

P0210 Extended reaction times in CF expression can be achieved by using a continuous-flow CF (CFCF) translation device.[220] A key feature is the continuous supply of energy and substrates concomitant with the removal of reaction by-products. The immobilized RM is continuously perfused by a feeding mixture (FM) containing all low molecular weight precursors and substrates and removing deleterious by-products of the protein production. The RM can alternatively be separated from the FM by a membrane with a variable molecular weight cut-off between 10 and 300 kDa.[221] However, reduced exchange rates by blocked membranes can cause significant problems. The protein synthesis in a CFCF system can continue for more than 20 h and yields of up to 1 mg recombinant protein per ml RM are possible. The CFCF system can either be operated in the translating mode by adding mRNA,[222,223] or in the coupled transcription–translation approach by the addition of phage RNA polymerases.[221,222,224] Many successfully synthesized proteins like the bacteriophage MS2 coat protein,[220] the brome mosaic virus coat protein,[220] globin,[225] calcitonin,[222] DHFR,[222,226–228] chloramphenicol acetyltransferase,[222,221,229–231] interleukin-2,[232] and interleukin-6[223] demonstrate the enormous potential of CFCF protein synthesis.

P0215 The relatively complex reaction set-up of the continuous flow mode is simplified in continuous-exchange CF (CECF) expression systems.[230,233] The RM and FM compartments have fixed volumes and are separated by a dialysis membrane. The simplest device for a CECF set-up is a dialysis bag holding the RM that is placed in a suitable container with the FM,[233] but commercially available dialyzers devices, such as the MicroDialyzer® and DispoDialyzer® (Spectrum Laboratories, Breda, The Netherlands), can also be used successfully. An additional advantage of the CECF set-up is the accumulation of the synthesized recombinant protein in the RM. High-level synthesis of up to 6 mg recombinant protein per ml RM has been reported.[230,234] Yields of 1–4 mg per ml RM for several functionally active proteins like the green fluorescent protein, DHFR, luciferase, and RNA replicase of tobacco mosaic virus have been obtained with eukaryotic wheat germ extracts.[235] The CECF system is commercially available as a kit (Roche Diagnostics, Mannheim, Germany) but can also be set up individually.[233,236–238]

S0120 #### 3.06.2.6.2 Components of the Cell-Free Reaction

P0220 All elements involved in gene expression and protein synthesis have to be added to the RM where transcription and translation takes place (**Figure 2**). Components like DNA, a highly processive RNA polymerase like the enzyme encoded by the T7 bacteriophage, NTPs, mRNA, tRNA, aminoacyl tRNA synthetases (ARSs), ribosomes, transcription and translation factors as well as amino acids have to be combined in an optimal pH and salt environment. The required high amounts of free energy for the transcription and translation processes are provided by hydrolysis of the triphosphates ATP and GTP. Crucial for each CF system is therefore an efficient ATP regenerating energy system in order to maintain the NTP concentrations over a long period of time. Conventional energy systems are based on high-energy phosphate donors such as phosphoenol pyruvate in combination with pyruvate kinase,[211,239] creatine phosphate and creatine kinase,[234] or acetyl phosphate with acetate kinase.[240]

S0125 3.06.2.6.3 **Preparation of Cell-Free Extracts**

P0225 The quality of the cell extract is crucial for the success of a CF system. The extract represents usually a crude cell lysate which contains most of the essential high molecular weight components for translation. Only T7 RNA polymerase and certain enzymes for the energy regeneration have to be supplied. While many organisms could potentially serve as an extract source, lysates based on the cell types of *E. coli*, wheat germ, and rabbit reticulocytes have been well established.

P0230 The bacterial source of choice for the preparation of CF extracts are RNAse-deficient *E. coli* strains, e.g., A19[236] or BL21.[237,241] Extracts of *E. coli* S30 represent the soluble fraction after centrifugation of crude lysates at 30 000 *g*. Endogenous mRNA is removed during a preincubation step of the cell extract either with high salt[236] or with added nucleotides and amino acids. This ''runoff'' step releases endogenous mRNA from the ribosomes that will then subsequently become destroyed by endogenous ribonucleases.[242] Alternatively, isolated ribosomes can be added to an S100 extract, centrifuged at 100 000 *g*.[243,244] Extracts of *E. coli* are used in coupled transcription–translation CF systems due to the favorable use of T7 RNA polymerase. Extracts of *E. coli* work well in a wide temperature range between 24 and 38 °C with an optimum at 37 °C.[242,244,245] Due to the relatively simple extract preparation procedure combined with high productivity, the *E. coli* CF system is the most commonly used in vitro protein expression technique and up to 6 mg of protein per ml RM can be synthesized.[234]

P0235 A well-defined system using 31 individually purified enzymes isolated with conventional expression systems together with purified ribosomes is termed the ''protein synthesis using recombinant elements'' (PURE) system.[246] The advantage of the PURE system is the absence of any inhibitory substances such as nucleases, proteases, and enzymes that hydrolyze nucleoside triphosphates.

P0240 The most convenient and promising eukaryotic CF translation system is based on wheat germs isolated from dry wheat seeds.[247] Recent modifications resulted in extracts with a high degree of stability and activity.[235] Important for an enhanced translation efficiency in the wheat germ CF system are the 5′ and 3′ UTRs of eukaryotic mRNAs that play a crucial role in the regulation of gene expression by controlling mRNA translational efficiency, stability, and localization. An optimal 5′ UTR that should therefore be added to mRNAs is the so-called omega sequence (Ω71) of tobacco mosaic virus. For the 3′ UTR the length is more important for an efficient translation than the sequence.[248] Wheat germ extracts possess only low levels of endogenously expressed mRNAs and therefore can be directly used for the expression of templates. [249] The optimal reaction temperature is in the range of 20–27 °C,[250,251] but can be increased to up to 32 °C for higher expression of some templates.[252] The reaction continues up to 60 h and amounts of 1-4 mg of recombinant protein per ml RM can be obtained.[235]

P0245 Lysates of rabbit reticulocytes are obtained from blood cells of anemic rabbits that provide a high number of reticulocytes or proerythrocytes. Endogenous mRNA is removed by treatment with micrococcal Ca^{2+}-dependent RNase.[253] The yields of recombinant protein can also be in the range of mg per ml RM, whereas the expression yields of the wheat germ system are usually higher.[254] This system works in an optimal temperature range of 30–38 °C.[255]

P0250 In principle, the choice of extract source, either prokaryotic or eukaryotic, for CF synthesis should be chosen according to the origin and biochemical nature of the protein of interest. In general *E. coli*-based systems gain in terms of their higher translation rates, better compatibility with combined transcription–translation formats, easier preparation of extracts as well as reaction set-up, and the availability of mutants with reduced degradative activities. On the other hand they suffer from high degradation of genetic messages, shorter lifetimes, and a great tendency of protein aggregation. Eukaryotic extracts are mainly limited by lower translation rates and the complexity of the genetic constructs that are required for an effective expression. Positive aspects of eukaryotic CF systems are their higher stability and longer lifetime in addition to a better compatibility with eukaryotic mRNAs and the synthesis of eukaryotic proteins.

P0255 CECF expression kits based on *E. coli* and wheat germ extracts are commercially available (Roche Diagnostics, Mannheim, Germany). The successful production of more than 40 proteins of different origins, including various enzymes, receptors, hormones, antibodies, and regulatory proteins has been shown by various laboratories.[256,257] Other commercial systems like Expressway™ (Invitrogen, Carlsbad, CA) focus on the batch mode with expression amounts of up to 1 mg ml^{-1}.

S0130 3.06.2.6.4 **Cell-Free Synthesis of Membrane Proteins**

P0260 Membrane proteins today represent less than 1% of the available three-dimensional protein structures, although this is in contrast to their immense medical importance as an estimate of 60-70% of current drug targets are based on membrane proteins. The key bottleneck has been the lack of reliable technologies that ensure the production of a broad variety of recombinant membrane proteins in the required amounts.[258] Toxicity to the host cell, protein

aggregation, and miss-folding of overproduced membrane proteins very often result in low yields. Furthermore, the overexpression of integral membrane proteins often causes cell death by overloading the cytoplasmic membranes or by disrupting the membrane integrity. Two protein groups of outstanding pharmaceutical relevance are multidrug resistance transporters of bacterial pathogens and GPCRs as the basic elements of the eukaryotic signal transduction machinery. Preparative-scale expression of GPCRs has in most cases not been obtained and it is always subject of tedious and long-lasting optimizations.[194,195,259] A promising perspective is that the newly developed CF translation systems offer a powerful alternative for overcoming the tremendous expression barriers for membrane proteins (**Figure 1**). It has recently been shown that functionally active integral membrane proteins, especially small multidrug transporters, GPCR proteins, a light-harvesting membrane protein, and ion channels were expressed in high yields of mg amounts per ml RM in an *E. coli*-based CECF system.[236,238,260–262] The proteins could also be functionally reconstituted into proteoliposomes and isotopically labeled for NMR investigations.[236] Even the addition of mild detergents does not interfere with the translation activity of the CECF systems and results directly in the soluble and functional expression of several integral membrane proteins.[238,260,261] The combination of isotopic labeling and membrane protein expression demonstrates the high potential of the CF method for functional and structural membrane protein research. Finally it should be mentioned that the preparation of membrane protein samples ready for a structural analysis by, e.g., NMR techniques is possible in less than 2 days by using CF expression systems and the structural characterization of even very difficult protein families like the GPCR proteins becomes now feasible.

S0135 P0265

3.06.2.6.5 Cell-Free Synthesis of Disulfide-Bridged Proteins

Disulfide-bonded proteins are rarely expressed in traditional expression systems due to their requirements for oxidizing conditions which in eukaryotes are only found in the lumen of the endoplasmatic reticulum (or in the periplasm of prokaryotes). The open character of CF systems allows the direct addition of purified chaperones like the already discussed eukaryotic protein disulfide isomerase (PDI) or bacterial Dsb derivatives with success of the functional expression of single chain antibodies.[263] CF expression reactions are usually operated in the presence of reducing agents like dithiothreitol that stabilize the protein transcription–translation machinery but provide less favorable conditions for the synthesis of disulfide-bonded proteins. However, the use of a dithiothreitol-deficient wheat germ extract in the presence of the PDI chaperone efficiently synthesizes a single-chain antibody variable fragment with dual disulfide bonds.[264] More recently, the problem of the CF expression of disulfide-bonded proteins was overcome by using a combination of iodoacetamide-treated extract, a suitable glutathione redox buffer, and the addition of disulfide bond shuffling chaperones like Skp and DsbC.[265] A recombinant plasminogen activator protein with nine disulfide bonds could productively be expressed with this approach.

References

1. Peredelchuk, M. Y.; Bennett, G. N. *Gene* **1997**, *187*, 231–238.
2. Olson, P.; Zhang, Y.; Olsen, D.; Owens, A.; Cohen, P.; Nguyen, K.; Ye, J. J.; Bass, S.; Mascarenhas, D. *Protein Expr. Purif.* **1998**, *14*, 160–166.
3. Das, A. *Methods Enzymol.* **1990**, *182*, 93–112.
AU :2 4. Makrides S. C. Microbiol. Rev. **1996**, 00, 512–538.
5. Hannig, G.; Makrides, S. C. *Trends Biotechnol.* **1998**, *16*, 54–60.
6. Baneyx, F. *Curr. Opin. Biotechnol.* **1999**, *10*, 411–421.
7. Mertens, N.; Remaut, E.; Fiers, W. *Gene* **1995**, *164*, 9–15.
8. Donovan, R. S.; Robinson, C. W.; Glick, B. R. *J. Ind. Microbiol.* **1996**, *16*, 145–154.
9. DeBoer, H. A.; Comstock, L. J.; Vasser, M. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 21–25.
10. Sawers, G.; Jarsch, M. *Appl. Microbiol. Biotechnol.* **1996**, *46*, 1–9.
11. O'Connor, C. D.; Timmis, K. N. *J. Bacteriol.* **1987**, *169*, 4457–4462.
12. Elvin, C. M.; Thompson, P. R.; Argall, M. E.; Hendry, P.; Stamford, N. P.; Lilley, P. E.; Dixon, N. E. *Gene* **1990**, *87*, 123–126.
13. Tabor, S.; Richardson, C. C. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 1074–1078.
14. Studier, F. W.; Rosenberg, A. H.; Dunn, J. J.; Dubendorff, J. W. *Methods Enzymol.* **1990**, *185*, 60–89.
15. Studier, F. W.; Moffatt, B. *J. Mol. Biol.* **1986**, *189*, 113–130.
16. Guzman, L. M.; Belin, D.; Carson, M. J.; Beckwith, J. *J. Bacteriol.* **1995**, *177*, 4121–4130.
17. Siegele, D. A.; Hu, J. C. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 8168–8172.
18. Wycuff, D. R.; Matthews, K. S. *Anal. Biochem.* **2000**, *277*, 67–73.
19. Gold, L. *Methods Enzymol.* **1990**, *185*, 11–14.
20. Zabeau, M.; Stanley, K. K. *EMBO J.* **1982**, *1*, 1217–1224.
21. Stanssens, P.; Remaut, E.; Fiers, W. *Gene* **1985**, *36*, 211–223.
22. Grisshammer, R.; Tate, C. G. *Q. Rev. Biophys.* **1995**, *28*, 315–422.
23. Wada, K.; Wada, Y.; Ishibashi, F.; Gojobori, T.; Ikemura, T. *Nucleic Acids Res.* **1992**, *20*, 2111–2118.
24. Kane, J. F. *Curr. Opin. Biotechnol.* **1995**, *6*, 494–500.
25. Calderone, T. L.; Stevens, R. D.; Oas, T. G. *J. Mol. Biol.* **1996**, *262*, 407–412.

26. Day, A. J.; Aplin, R. T.; Willis, A. C. *Protein Expr. Purif.* **1996**, *8*, 1–16.
27. Zhan, K. *J. Bacteriol.* **1996**, *178*, 2926–2933.
28. Forman, M. D.; Stack, R. F.; Masters, P. S.; Hauer, C. R.; Baxter, S. M. *Protein Sci.* **1998**, 7, 500–503.
29. Gustafsson, C.; Govindarajan, S.; Minshull, J. *Trends Biotechnol.* **2004**, *22*, 346–353.
30. Brinkmann, U.; Mattes, R. E.; Buckel, P. *Gene* **1989**, *85*, 109–114.
31. Baca, A. M.; Hol, W. G. *Int. J. Parasitol.* **2000**, *30*, 113–118.
32. Forrer, P.; Jaussi, R. *Gene* **1998**, *224*, 45–52.
33. Zhou, P.; Lugovskoy, A. A.; Wagner, G. *J. Biomol. NMR* **2001**, *20*, 11–14.
34. LaVallie, E. R.; McCoy, J. M. *Curr. Opin. Biotechnol.* **1995**, *6*, 501–506.
35. Sharrocks, A. D. *Gene* **1994**, *138*, 105–108.
36. Nilsson, J.; Stahl, S.; Lundeberg, J.; Uhlen, M.; Nygren, P. A. *Protein Expr. Purif.* **1997**, *11*, 1–16.
37. Hochuli, E.; Döbeli, H.; Schacher, A. *J. Chromatogr.* **1987**, *411*, 177–184.
38. Pryor, K. D.; Leiting, B. *Protein Expr. Purif.* **1997**, *10*, 309–319.
39. Frorath, B.; Abney, C. C.; Berthold, M.; Scanarini, M.; Northemann, W. *Biotechniques* **1992**, *12*, 558–563.
40. LaVallie, E. R.; DiBlasio, E. A.; Kovacic, S.; Grant, K. L.; Schendel, P. F.; McCoy, J. M. *Biotechnology* **1993**, *11*, 187–193.
41. Kohne, T.; Kusunoki, H.; Sato, K.; Wakamutsu, K. *J. Biomol. NMR* **1998**, *12*, 109–121.
42. Koken, M. H.; Odijk, H. H.; VanDuin, M.; Fornerod, M.; Hoeijmakers, J. H. *Biochem. Biophys. Res. Commun.* **1993**, *195*, 643–653.
43. Smith, D. B.; Johnson, K. S. *Gene* **1988**, *67*, 31–40.
44. Nygren, P. A.; Stahl, S.; Uhlen, M. *Trends Biotechnol.* **1994**, *12*, 184–188.
45. Di Guan, C.; Li, P.; Riggs, P. D.; Inouye, H. *Gene* **1988**, *67*, 21–30.
46. Kapust, R. B.; Waugh, D. S. *Protein Sci.* **1999**, *8*, 1668–1674.
47. Riggs, P. *Mol. Biotechnol.* **2000**, *15*, 51–63.
48. Kapust, R. B.; Waugh, D. S. *Protein Expr. Purif.* **2000**, *19*, 312–318.
49. Shih, Y. P.; Wu, H. C.; Hu, S. M.; Wang, T. F.; Wang, A. H. *Protein Sci.* **2005**, *14*, 936–941.
50. Friedberg, T.; Wolf, C. R. *Adv. Drug Rev.* **1996**, *22*, 187–213.
51. Swartz, J. R. *Curr. Opin. Biotechnol.* **2001**, *12*, 195–201.
52. Andersen, D. C.; Krummen, L. *Curr. Opin. Biotechnol.* **2002**, *13*, 117–123.
53. Hunt, I. *Protein Expr. Purif.* **2005**, *40*, 1–22.
54. Sorensen, H. P.; Mortensen, K. K. *J. Biotechnol.* **2005**, *115*, 113–128.
55. Lesley, S. A. *Protein Expr. Purif.* **2001**, *22*, 159–164.
56. Weickert, M. J.; Doherty, D. H.; Best, E. A.; Olins, P. O. *Curr. Opin. Biotechnol.* **1996**, 7, 494–499.
57. Misawa, S.; Kumagai, I. *Biopolymers* **1999**, *51*, 297–307.
58. Carrió, M. M.; Villaverde, A. *J. Biotechnol.* **2002**, *96*, 3–12.
59. Bukau, B.; Horwich, A. L. *Cell* **1998**, *92*, 351–366.
60. Feldman, D. E.; Frydman, J. *Curr. Opin. Struct. Biol.* **2000**, *10*, 26–33.
61. Sigler, P. B.; Xu, Z.; Rye, H. S.; Burston, S. G.; Fenton, W. A.; Horwich, A. L. *Annu. Rev. Biochem.* **1998**, *67*, 581–608.
62. Wang, J. D.; Weissman, J. S. *Nat. Struct. Biol.* **1999**, *6*, 597–600.
63. Georgiou, G.; Valax, P. *Curr. Opin. Biotechnol.* **1996**, 7, 190–197.
64. Thomas, J. G.; Ayling, A.; Baneyx, F. *Appl. Biochem. Biotechnol.* **1997**, *66*, 197–238.
65. Schlieker, C.; Bukau, B.; Mogk, A. *J. Biotechnol.* **2002**, *96*, 13–21.
66. DeMarco, A.; DeMarco, V. *J. Biotechnol.* **2004**, *109*, 45–52.
67. Nishihara, K.; Kanemori, M.; Kitagawa, M.; Yanaga, H.; Yura, T. *Appl. Environ. Microbiol.* **1998**, *64*, 1694–1699.
68. Nishihara, K.; Kanemori, M.; Yanagi, H.; Yura, T. *Appl. Environ. Microbiol.* **2000**, *66*, 884–889.
69. Cole, P. A. *Structure* **1996**, *4*, 239–242.
70. Castanié, M. P.; Bergès, H.; Oreglia, J.; Prère, M. F.; Fayet, O. *Anal. Biochem.* **1997**, *254*, 150–152.
71. Rucker, P.; Torti, F. M.; Torti, S. V. *Protein Eng.* **1997**, *10*, 967–973.
72. Kim, K. J.; Kim, H. E.; Lee, K. H.; Han, W.; Yi, M. J.; Jeong, J.; Oh, B. H. *Protein Sci.* **2004**, *13*, 1698–1703.
73. Johnston, K.; Clements, A.; Venkatamanrani, R. N.; Trievel, R. C.; Marmorstein, R. *Protein Expr. Purif.* **2000**, *20*, 435–443.
74. Yang, W.; Zhang, L.; Lu, Z.; Tao, W.; Zhai, Z. *Protein Expr. Purif.* **2001**, *22*, 472–478.
75. Zhang, Y.; Olsen, D. R.; Nguyen, K. B.; Olson, P. S.; Rhodes, E. T.; Mascarenhas, D. *Protein Expr. Purif.* **1998**, *12*, 159–165.
76. Sachdev, D.; Chirgwin, J. M. *Protein Expr. Purif.* **1998**, *12*, 122–132.
77. Baneyx, F. In vivo folding of recombinant proteins in *Escherichia coli*. In *Manual of Industrial Microbiology and Biotechnology*, 2nd ed.; Davies, J. E., Demain, A. L., Cohen, G., Hershberger, C. L., Forney, L. J., Holland, I. B., Hu, W. S., Wu, J.-H. D., Sherman, D. H., Wilson, R. C., Eds.; Publisher: Washington, DC, 1999; pp 551–565.
78. Missiakas, D.; Raina, S. *J. Bacteriol.* **1997**, *179*, 2465–2471.
79. Aslund, F.; Beckwith, J. *J. Bacteriol.* **1999**, *181*, 1375–1379.
80. Bowden, G. A.; Georgiou, G. *Biotechnol. Prog.* **1988**, *4*, 97–101.
81. Baneyx, F.; Ayling, A.; Palumbo, T.; Thomas, D.; Georgiou, G. *Appl. Microbiol. Biotechnol.* **1991**, *36*, 14–20.
82. Pugsley, A. P. *Microbiol. Rev.* **1993**, *57*, 50–108.
83. Sandkvist, M.; Bagdasarian, M. *Curr. Opin. Biotechnol.* **1996**, 7, 505–511.
84. Driessen, A. J.; Fekkes, P.; vanderWolk, J. P. *Curr. Opin. Microbiol.* **1998**, *1*, 216–222.
85. Blight, M. A.; Holland, I. B. *Trends Biotechnol.* **1994**, *12*, 450–455.
86. Robbens, J.; Raeymaekers, A.; Steidler, L.; Fiers, W.; Remaut, E. *Protein Expr. Purif.* **1995**, *6*, 481–486.
87. VanderWal, F. J.; Luirink, J.; Oudega, B. *FEMS Microbiol. Rev.* **1995**, *17*, 381–399.
88. VanderWal, F. J.; Koningstein, G.; TenHagen, C. M.; Oudega, B.; Luirink, J. *Appl. Environ. Microbiol.* **1998**, *64*, 392–398.
89. Wan, E. W.; Baneyx, F. *Protein Expr. Purif.* **1998**, *14*, 13–22.
90. Miroux, B.; Walker, J. E. *J. Mol. Biol.* **1996**, *260*, 289–298.
91. Wang, D. N.; Safferling, M.; Lemieux, M. J.; Griffith, H.; Chen, Y.; Li, X. D. *Biochim. Biophys. Acta* **2003**, *1610*, 23–36.
92. Montigny, C.; Penin, F.; Lethias, C.; Falson, P. *Biochim. Biophys. Acta* **2004**, *1660*, 53–65.
93. Drew, D.; Fröderberg, L.; Baares, L.; deGier, J. W. L. *Biochim. Biophys. Acta* **2003**, *1610*, 3–10.

AU :3

94. Gottesman, S. *Annu. Rev. Genet.* **1996**, *30*, 465–506.
95. Roman, L. J.; Sheta, E. A.; Martasek, P.; Gross, S. S.; Liu, Q.; Siler-Masters, B. S. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8428–8432.
96. Murby, M.; Uhlén, M.; Stahl, S. *Protein Expr. Purif.* **1996**, 7, 129–136.
97. Bessette, P. H.; Aslund, F.; Beckwith, J.; Georgiou, G. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13703–13708.
98. Bardwell, J. C. A. *Mol. Microbiol.* **1994**, *14*, 199–205.
99. Wülfing, C.; Plückthun, A. *Mol. Microbiol.* **1994**, *12*, 685–692.
100. Humphreys, D. L.; Weir, N.; Mountain, A.; Luind, P. A . *J. Biol. Chem.* **1995**, *270*, 28210–28215.
101. Yasukawa, T.; Kanei-Ishii, C.; Maekawa, T.; Fujimoto, J.; Yamamoto, T.; Ishii, S. *J. Biol. Chem.* **1995**, *270*, 25328–25331.
102. Billman-Jacobe, H. *Curr. Opin. Biotechnol.* **1996**, 7, 500–504.
103. Kuipers, O. P.; deRuyter, P. G. G. A.; Kleerebezem, M.; deVos, W. M. *Trends Biotechnol.* **1997**, *15*, 135–140.
104. DeVos, W. M. *Curr. Opin. Microbiol.* **1999**, *2*, 289–295.
105. Kunji, E. R. S.; Slotboom, D. J.; Poolman, B. *Biochim. Biophys. Acta* **2003**, *1610*, 97–108.
106. Sorvig, E.; Grönqvist, S.; Naterstad, K.; Mathiesen, G.; Eijsink, V. G. H.; Axelsson, L. *FEMS Microbiol. Lett.* **2003**, *229*, 119–126.
107. Miyoshi, A.; Jamet, E.; Commissaire, J.; Renault, P.; Langella, P.; Azevedo, V. *FEMS Microbiol. Lett.* **2004**, *239*, 205–212.
108. Wong, S. L. *Curr. Opin. Biotechnol.* **1995**, *6*, 517–522.
109. Binnie, C.; Cossar, J. D.; Stewart, D. I. H. *Trends Biotechnol.* **1997**, *15*, 315–320.
110. Lam, K. H.; Chow, K. C.; Wong, W. K. *J. Biotechnol.* **1998**, *63*, 167–177.
111. Van Heeke, G.; Ott, T. L.; Strauss, A.; Ammaturo, D.; Bazer, F. W. *J. Interferon Cytokine Res.* **1996**, *16*, 119–126.
112. Austin, A. J.; Jones, C. E.; van Heeke, G. *Protein Expr. Purif.* **1998**, *13*, 136–142.
113. Boetner, M.; Prinz, B.; Holz, C.; Stahl, U.; Lang, C. *J. Biotechnol.* **2002**, *99*, 51–62.
114. Romanos, M. A.; Scorer, C. A.; Clare, J. J. *Yeast* **1992**, *8*, 423–488.
115. Schuster, M.; Einhauer, A.; Wasserbauer, E.; Süßenbacher, F.; Ortner, C.; Paumann, M.; Werner, G.; Jungbauer, A. *J. Biotechnol.* **2000**, *84*, 237–248.
116. Hollenberg, C. P.; Gellissen, G. *Curr. Opin. Biotechnol.* **1997**, *8*, 554–560.
117. Cereghino, J. L.; Cregg, J. M. *FEMS Microbiol. Rev.* **2000**, *24*, 45–66.
118. VanDijk, R.; Faber, K. N.; Kiel, J. A. K. W.; Veenhuis, M.; vanderKlei, I. *Enzyme Microb. Technol.* **2000**, *26*, 793–800.
119. Madzak, C.; Gaillardin, C.; Beckerich, J. M. *J. Biotechnol.* **2004**, *109*, 63–81.
120. Branduardi, P.; Valli, M.; Brambilla, L.; Sauer, M.; Alberghino, L.; Porro, D. *FEMS Yeast Res.* **2004**, *4*, 493–504.
121. Sreekrishna, K.; Brankamp, R. G.; Kropp, K. E.; Blankenship, D. T.; Tsay, J. T.; Smith, P. L.; Wierschke, J. D.; Subramaniam, A.; Birkenberger, L. A. *Gene* **1997**, *190*, 55–62.
122. Hensing, M. C. M.; Rouwenhorst, R. J.; Heijnen, J. J.; vanDijken, J. P.; Pronk, J. T. *Antonie van Leeuwenhoek* **1995**, *67*, 261–279.
123. Lueking, A.; Holz, C.; Gotthold, C.; Lehrach, H.; Cahill, D. *Protein Expr. Purif.* **2000**, *20*, 372–378.
124. Vad, R.; Nafstad, E.; Dahl, L. A.; Gabrielsen, O. S. *J. Biotechnol.* **2005**, *116*, 251–260.
125. Parekh, R. N.; Wittrup, K. D. *Biotechnol. Prog.* **1997**, *13*, 117–122.
126. Cereghino, G. P. L.; Cereghino, J. L.; Ilgen, C.; Cregg, J. M. *Curr. Opin. Biotechnol.* **2002**, *13*, 329–332.
127. White, C. E.; Kempi, N. M.; Komives, E. A. *Structure* **1994**, *2*, 1003–1005.
128. Reiländer, H.; Wei, H. M. *Curr. Opin. Biotechnol.* **1998**, *9*, 510–517.
129. Geisse, S.; Gram, H.; Kleuser, B.; Kocher, H. P. *Protein Expr. Purif.* **1996**, *8*, 271–282.
130. Wirth, M.; Bode, J.; Zettlmeissl, G.; Hauser, H. *Gene* **1988**, *73*, 419–426.
131. Page, M. J.; Sydenham, M. A. *Biotechnology* **1991**, *9*, 64–68.
132. Berg, D. T.; McClure, D. B.; Grinnell, B. W. *Biotechniques* **1993**, *14*, 972–978.
133. Gopal, T. V.; Polte, T.; Arthur, P.; Seidman, M. *In vitro Cell Dev. Biol.* **1989**, *25*, 1147–1154.
134. Kane, S. E.; Reinhard, D. H.; Fordis, C. M.; Pastan, I.; Gottesman, M. M. *Gene* **1989**, *84*, 439–446.
135. Needham, M.; Gooding, C.; Hudson, K.; Antoniou, M.; Grosveld, F.; Hollis, M. *Nucleic Acids Res.* **1992**, *20*, 997–1003.
136. Okamoto, M.; Nakayama, C.; Nakai, M.; Yanagi, H. *Biotechnology* **1990**, *8*, 550–553.
137. Lopez, C.; Chesnay, A. D.; Tournamille, C.; Ghanem, A. B.; Prigent, S.; Drouet, X.; Lambin, P.; Cartron, J. P. *Gene* **1994**, *148*, 285–291.
138. Wenger, R. H.; Moreau, H.; Nielsen, P. J. *Anal. Biochem.* **1994**, *221*, 416–418.
139. Urlaub, G.; Chasin, L. A. *Proc. Natl. Acad. Sci. USA* **1980**, 77, 4216–4220.
140. Graham, F. L.; van der Eb, A. J *Virology* **1973**, *52*, 456–467.
141. Vaheri, A.; Pagano, J. S. *Virology* **1965**, *27*, 434–436.
142. Wong, T. K.; Neumann, E. *Biochem. Biophys. Res. Commun.* **1982**, *107*, 584–587.
143. Felgner, P. L.; Gadek, T. R.; Holm, M.; Roman, R.; Chan, H. W.; Wenz, M.; Northrop, J. P.; Ringold, G. M.; Danielsen, M. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 7413–7417.
144. Makrides, S. C. *Protein Expr. Purif.* **1999**, *17*, 183–202.
145. Novina, C. D.; Roy, A. L. *Trends Genet.* **1996**, *12*, 351–355.
146. Angrand, P. O.; Woodfoofe, C. P.; Buchholz, F.; Stewart, A. F. *Nucleic Acids Res.* **1998**, *26*, 3263–3269.
147. Hu, M. C.; Davidson, N. *Cell* **1987**, *48*, 555–566.
148. Figge, J.; Wright, C.; Collins, C. J.; Roberts, T. M.; Livingston, D. M. *Cell* **1988**, *52*, 713–722.
149. Ward, G. A.; Stover, C. K.; Moss, B.; Fuerst, T. R. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 6773–6777.
150. Colbere-Grapin, F.; Horodniceanu, F.; Kourilsky, P.; Garapin, A. C. *J. Mol. Biol.* **1981**, *150*, 1–14.
151. Karreman, C. *Gene* **1998**, *218*, 57–61.
152. Buchman, A. R.; Berg, P. *Mol. Cell Biol.* **1988**, *8*, 4395–4405.
153. Jackson, R. J.; Standart, N. *Cell* **1990**, *62*, 15–24.
154. Gray, N. K.; Wickens, M. *Annu. Rev. Cell Dev. Biol.* **1998**, *14*, 399–458.
155. Pandey, N. B.; Chodchoy, N.; Liu, T. J.; Marzluff, W. F. *Nucleic Acids Res.* **1990**, *18*, 3161–3170.
156. Schmidt, E. V.; Christoph, G.; Zeller, R.; Leder, P. *Mol. Cell. Biol.* **1990**, *10*, 4406–4411.
157. Goodwin, E. C.; Rottman, F. M. *J. Biol. Chem.* **1992**, *267*, 16330–16334.
158. Van den Hoff, M. J.; Labruyère, W. T.; Moorman, A. F.; Lamers, W. H. *Nucleic Acids Res.* **1993**, *21*, 4987–4988.
159. Kozak, M. J. *Mol. Biol.* **1987**, *196*, 947–950.
160. Grens, A.; Scheffler, I. E. *J. Biol. Chem.* **1990**, *265*, 11810–11816.

161. Wright, J. A.; Smith, H. S.; Watt, F. M.; Hancock, M. C.; Hudson, D. L.; Stark, G. R. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 1791–1795.
162. Schimke, R. T. *J. Biol. Chem.* **1988**, *263*, 5989–5992.
163. Stark, G. R.; Debatisse, M.; Giulotto, E.; Wahl, G. M. *Cell* **1989**, *57*, 901–908.
164. Kaufman, R. J.; Wasley, L. C.; Spiliotes, A. J.; Gossels, S. D.; Latt, S. A.; Larsen, G. R.; Kay, R. M. *Mol. Cell. Biol.* **1985**, *5*, 1750–1759.
165. Kaufman, R . J. *Methods Enzymol.* **1990**, *185*, 537–566.
166. Kaufman, R. J.; Sharp, P. A. *J. Mol. Biol.* **1982**, *159*, 601–621.
167. Kaufman, R. J.; Murtha, P.; Davies, M. V. *EMBO J.* **1987**, *6*, 187–193.
168. Traunecker, A.; Oliveri, F.; Karjalainen, K. *Trends Biotechnol.* **1991**, *9*, 109–113.
169. Greene, G. L.; Fitch, F. W.; Jensen, E. V. *Proc. Natl. Acad. Sci. USA* **1980**, 77, 157–161.
170. Shulman, M.; Wilde, C. D.; Kohler, G. *Nature* **1978**, *276*, 269–270.
171. Boast, S.; LaMantia, G.; Lania, L.; Blasi, F. *EMBO J.* **1983**, *2*, 2327–2331.
172. Gerard, R. D.; Gluzman, Y. *Mol. Cell. Biol.* **1985**, *5*, 3231–3340.
173. Asselbergs, F. A. *J. Biotechnol.* **1992**, *26*, 301–313.
174. de Chasseval, R.; de Villartay, J. P. *Nucleic Acids Res.* **1992**, *20*, 245–250.
175. Gluzman, Y. *Cell* **1981**, *23*, 175–182.
176. Mellon, P.; Parker, V.; Gluzman, Y.; Maniatis, T. *Cell* **1981**, *27*, 279–288.
177. Edwards, C. P.; Aruffo, A. *Curr. Opin. Biotechnol.* **1993**, *4*, 558–563.
178. Trill, J. J.; Shatzman, A. R.; Ganguly, S. *Curr. Opin. Biotechnol.* **1995**, *6*, 553–560.
179. Ridder, R.; Geisse, S.; Kleuser, B.; Kawalleck, P.; Gram, H. *Gene* **1995**, *166*, 273–276.
180. Wurm, F.; Bernard, A. *Curr. Opin. Biotechnol.* **1999**, *10*, 156–159.
181. Ikonomou, L.; Schneider, Y. J.; Agathos, S. N. *Appl. Microbiol. Biotechnol.* **2003**, *62*, 1–20.
182. Lundstrom, K.; Rotmann, D.; Hermann, D.; Schneider, E. M.; Ehrengruber, M. U. *Histochem. Cell Biol.* **2001**, *115*, 83–91.
183. Brickell, P. M.; Patel, M. S. *Mol. Biotechnol.* **1995**, *3*, 199–205.
184. Yamanaka, R.; Xanthopoulos, K. G. *DNA Cell Biol.* **2004**, *23*, 75–80.
185. Polo, J. M.; Gardner, J. P.; Ji, Y.; Belli, B. A.; Driver, D. A.; Sherrill, S.; Perri, S.; Liu, M. A.; Dubensky, T. W. *Dev. Biol.* **2000**, *104*, 181–185.
186. Ayres, M. D.; Howard, S. C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R. D. *Virology* **1994**, *202*, 586–605.
187. Sarramegna, V.; Talmont, F.; Demange, P.; Milon, A. *Cell Mol. Life Sci.* **2003**, *60*, 1529–1546.
188. Kitts, P. A.; Possee, R. D. *Biotechniques* **1993**, *14*, 810–817.
189. Luckow, V. A. *Curr. Opin. Biotechnol.* **1993**, *4*, 564–572.
190. Davies, A. H. *Biotechnology* **1994**, *12*, 47–50.
191. Jones, I.; Morikawa, Y. *Curr. Opin. Biotechnol.* **1996**, 7, 512–516.
192. Ogonah, O. W.; Freedman, R. B.; Jenkins, N.; Patel, K.; Rooney, B. C. *Biotechnology* **1996**, *14*, 197–202.
193. Palomares, L. A.; Joosten, C. E.; Hughes, P. R.; Granados, R. R.; Shuler, M. L. *Biotechnol. Prog.* **2003**, *19*, 185–192.
194. Massotte, D. *Biochim. Biophys. Acta* **2003**, *1610*, 77–89.
195. Tate, C. G.; Haase, J.; Baker, C.; Boorsma, M.; Magnani, F.; Vallis, Y.; Williams, D. C. *Biochim. Biophys. Acta* **2003**, *1610*, 141–153.
196. Chai, H.; Vasudevan, S. G.; Porter, A. G.; Chua, K. L.; Oh, S.; Yap, M. *Biotechnol. Appl. Biochem.* **1993**, *18*, 259–273.
197. Xiong, C.; Levis, R.; Shen, P.; Schlesinger, S.; Rice, C. M.; Huang, H. V. *Science* **1989**, *243*, 1188–1191.
198. Davis, N. L.; Willis, L. V.; Smith, J. F.; Johnston, R. E. *Virology* **1989**, *171*, 189–204.
199. Atkins, G. J.; Sheahan, B. J.; Liljestrom, P. *J. Gen. Virol.* **1999**, *80*, 2287–2297.
200. Liljestrom, P.; Garoff, H. *Biotechnology* **1991**, *9*, 1356–1361.
201. Lundstrom, K. *Biochim. Biophys. Acta* **2003**, *1610*, 90–96.
202. Lundstrom, K.; Schweitzer, C.; Rotmann, D.; Hermann, D.; Schneider, E. M.; Ehrengruber, M. U. *FEBS Lett.* **2001**, *504*, 99–103.
203. Sjoberg, E. M.; Suomalainen, M.; Garoff, H. *Biotechnology* **1994**, *12*, 1127–1131.
204. Henrich, B.; Lubitz, W.; Plapp, R. *Mol. Gen. Genet.* **1982**, *185*, 493–497.
205. Goff, S.; Goldberg, A. L. *J. Biol. Chem.* **1987**, *262*, 4508–4515.
206. Charles, I. G.; Chubb, A.; Grill, R.; Clare, J.; Lowe, P. N.; Holmes, L. S.; Page, M.; Keeling, J. G.; Moncada, S.; Riveros-Moreno, V. *Biochem. Biophys. Res. Commun.* **1993**, *191*, 1481–1489.
207. Chrunky, B. A.; Evans, J.; Lillquist, J.; Young, E.; Wetzel, R. *J. Biol. Chem.* **1993**, *268*, 18053–18061.
208. Maurizi, M. R. *J. Biol. Chem.* **1987**, *262*, 2696–2703.
209. Sengupta, K.; Klammt, C.; Bernhard, F.; Ruterjans, H. Incorporation of fluorescence labels into cell-free produced proteins. In *Cell-Free Protein Expression*; Swartz, J. R., Ed.; Springer Verlag: Berlin, Germany, 2003, pp 81–88.
210. Ohuchi, S.; Nakano, H.; Yamane, T. *Nucleic Acids Res.* **1998**, *26*, 4339–4346.
211. De Vries, J. K.; Zubay, G. *Proc. Natl. Acad. Sci. USA* **1967**, *57*, 1010–1012.
212. Lederman, M.; Zubay, G. *Biochim. Biophys. Acta.* **1967**, *149*, 253–258.
213. Kim, D. M.; Swartz, J. R. *Biotechnol. Bioeng.* **1999**, *66*, 180–188.
214. Kim, D. M.; Swartz, J. R. *Biotechnol. Prog.* **2000**, *16*, 385–390.
215. Kim, D. M.; Swartz, J. R. *Biotechnol. Lett.* **2000**, *22*, 1537–1542.
216. Kim, D. M.; Swartz, J. R. *Biotechnol. Bioeng.* **2001**, *74*, 309–316.
217. Jewett, M. C.; Swartz, J. R. *Biotechnol. Prog.* **2004**, *20*, 102–109.
218. Jewett, M. C.; Swartz, J. R. *Biotechnol. Bioeng.* **2004**, *86*, 19–26.
219. Jewett, M. C.; Swartz, J. R. *Biotechnol. Bioeng.* **2004**, *87*, 465–472.
220. Spirin, A. S.; Baranov, V. I.; Ryabova, L. A.; Ovodov, S. Y.; Alakhov, Y. B. *Science* **1988**, *242*, 1162–1164.
221. Kigawa, T.; Yokoyama, S. *J. Biochem.* **1991**, *110*, 166–168.
222. Spirin, A. S. Cell-free protein synthesis bioreactor. In *Frontiers in Bioprocessing II*; Todd, P., Sikdar, S. K., Beer, M., Eds.; American Chemical Society: Washington, DC, 1991, pp 31–43.
223. Volyanik, E. V.; Dalley, A.; McKay, I. A.; Leigh, I.; Williams, N. S.; Bustin, S. A. *Anal. Biochem.* **1993**, *214*, 289–294.
224. Baranov, V. I.; Spirin, A. S. *Methods Enzymol.* **1993**, *217*, 123–142.
AU :4 225. Ryabova, L. A.; Ortleb, S. A.; Baranov, V. I. *Nucleic Acids Res.* **1989**, 17, 4412–0000.
226. Baranov, V. I.; Morozov, I. Y.; Ortlepp, S. A.; Spirin, A., S. *Gene* **1989**, *84*, 463–466.
227. Endo, Y.; Otsuzuki, S.; Ito, K.; Miura, K. *J. Biotechnol.* **1992**, *25*, 221–230.

228. Kudlicki, W.; Kramer, G.; Hardesty, B. *Anal. Biochem.* **1992**, *206*, 389–393.
229. Ryabova, L. A.; Volyanik, E.; Kurnasov, P.; Spirin, A. S.; Wu, Y.; Kramer, F. R. *J. Biol. Chem.* **1994**, *269*, 1501–1505.
230. Kim, D. M.; Choi, C. Y. *Biotechnol. Prog.* **1996**, *12*, 645–649.
231. Kitaoka, Y.; Nishimura, N.; Niwano, M. *J. Biotechnol.* **1996**, *48*, 1–8.
232. Kolosov, M. I.; Kolosova, I. M.; Alokhov, V. Y.; Ovodov, S. Y.; Alokhov, Y. B. *Biotechnol. Appl. Biochem.* **1992**, *16*, 125–133.
233. Alakov, Y. B.; Baranov, V. I.; Ovodov, S. J.; Ryabova, L. A.; Spirin, A. S.; Morozov, I. J. Method of Preparing Polypeptides in Cell-Free Translation System. U.S. Patent 5,478,73x, 1994.
234. Kigawa, T.; Yabuki, T.; Yoshida, Y.; Tsutsui, M.; Ito, Y.; Shibata, T.; Yokoyama, S. *FEBS Lett.* **1999**, *442*, 15–19.
235. Madin, K.; Sawasaki, T.; Ogasawara, T.; Endo, Y. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 559–564.
236. Klammt, C.; Lohr, F.; Schafer, B.; Haase, W.; Dotsch, V.; Ruterjans, H.; Glaubitz, C.; Bernhard, F. *Eur. J. Biochem.* **2004**, *271*, 568–580.
237. Torizawa, T.; Shimizu, M.; Taoka, M.; Miyano, H.; Kainosho, M. *J. Biomol. NMR* **2004**, *30*, 311–325.
238. Ishihara, G.; Goto, M.; Saeki, M.; Ito, K.; Hori, T.; Kigawa, T.; Shirouzu, M.; Yokoyama, S. *Protein Expr. Purif.* **2005**, *41*, 27–37.
239. Nirenberg, N. W.; Matthaei, J. H. *Proc. Natl. Acad. Sci. USA* **1961**, *47*, 1588–1602.
240. Ryabova, L. A.; Vinokurov, L. M.; Shekhovtsova, E. A.; Alakhov, Y. B.; Spirin, A. S. *Anal. Biochem.* **1995**, *226*, 184–186.
241. Kigawa, T.; Yabuki, T.; Matsuda, N.; Matsuda, T.; Nakajima, R.; Tanaka, A.; Yokoyama, S. *J. Struct. Funct. Genomics* **2004**, *5*, 63–68.
242. Zubay, G. *Annu. Rev. Genet.* **1973**, 7, 267–287.
243. Gold, L. M.; Schweiger, M. *Proc. Natl. Acad. Sci. USA* **1969**, *62*, 892–898.
244. Gold, L. M.; Schweiger, M. *Methods Enzymol.* **1971**, *20*, 537–542.
245. Kim, D. M.; Kigawa, T.; Choi, C. Y.; Yokoyama, S. *Eur. J. Biochem.* **1996**, *239*, 881–886.
246. Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T. *Nat. Biotechnol.* **2001**, *19*, 751–755.
247. Johnston, F. B.; Stern, H. *Nature* **1957**, *179*, 160–161.
248. Sawasaki, T.; Ogasawara, T.; Morishita, R.; Endo, Y. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14652–14657.
249. Roberts, B. E.; Paterson, B. M. *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 2330–2334.
250. Anderson, C. W.; Straus, J. W.; Dudock, B. S. *Methods Enzymol.* **1983**, *101*, 635–644.
251. Kawarasaki, Y.; Nakano, H.; Yamane, T. *Biosci. Biotechnol. Biochem.* **1994**, *58*, 1911–1913.
252. Tulin, E. E.; Ken-Ichi, T.; Shin-Ichiro, E. *Biotechnol. Bioeng.* **1995**, *45*, 511–516.
253. Pelham, H. R. B.; Jackson, R. J. *Eur. J. Biochem.* **1976**, *67*, 247–256.
254. Katzen, F.; Chang, G.; Kudlicki, W. *Trends Biotechnol.* **2005**, *23*, 150–156.
255. Woodward, W. R.; Ivey, J. L.; Herbert, E. *Methods Enzymol.* **1974**, *30*, 724–731.
256. Betton, J. M. *Curr. Protein Pept. Sci.* **2003**, *4*, 73–80.
257. Spirin, A. S. *Trends Biotechnol.* **2004**, *22*, 538–545.
258. Bannwarth, M.; Schulz, G. E. *Biochim. Biophys. Acta* **2003**, *1610*, 37–45.
259. Tate, C. G. *FEBS Lett.* **2001**, *504*, 94–98.
260. Berrier, C.; Park, K. H.; Abes, S.; Bibonne, A.; Betton, J. M.; Ghazi, A. *Biochemistry* **2004**, *43*, 12585–12591.
261. Elbaz, Y.; Steiner-Mordoch, S.; Danieli, T.; Schuldiner, S. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 1519–1524.
262. Shimada, Y.; Wang, Z. Y.; Mochizuki, Y.; Kobayashi, M.; Nozawa, T. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 1942–1948.
263. Ryabova, L. A.; Desplancq, D.; Spirin, A. S.; Pluckthun, A. *Nat. Biotechnol.* **1997**, *15*, 79–84.
264. Kawasaki, T.; Gouda, M. D.; Sawasaki, T.; Takai, K.; Endo, Y. *Eur. J. Biochem.* **2003**, *270*, 4780–4786.
265. Yin, G.; Swartz, J. R. *Biotechnol. Bioeng.* **2004**, *86*, 188–195.

AU :5

Biographies

From left to right: Heinz Rüterjans, Christian Klammt, Frank Bernhard **Heinz Rüterjans** studied chemistry at the University of Muenster, Germany. Since 1979 he has been full professor of biophysical chemistry at the J. W. Goethe University of Frankfurt, Germany. From 1981 to 1982 and from 1992 to 1993 he was Dean of the Faculty of Biochemistry at the University of Frankfurt. In 1995, he became head of the European Large Scale Facility for Biomolecular NMR at the University of Frankfurt. In the same year, he was awarded the Humboldt Prize of the French Government. In 2002,

he established the Centre for Biomolecular Magnetic Resonance at the University of Frankfurt. He further acted as president of the European Biophysics Societies Association from 2001 to 2003. Heinz Rüterjans was leading scientist of a variety of RTD projects of the European Union. His research interests cover the structural and functional analysis of biological macromolecules by high-resolution NMR spectroscopy, the dynamics of protein structures, protein/nucleic acid and protein/lipid interactions, and various issues of the cellular metabolism.

AU :6

Christian Klammt studied biochemistry at the University of Frankfurt, Germany. Since 2002 he has worked on the structural analysis of integral membrane proteins at the Institute of Biophysical Chemistry at the University of Frankfurt. Christian Klammt is associated with the Centre for Membrane Proteomics in Frankfurt and an expert for the preparative scale cell-free expression of proteins. This new techniques, based on cell-free extracts, for the high-level production of difficult membrane proteins like transporters or G-protein coupled receptors have led to his recently award from the Federation of European Biochemical Societies for the best scientific contribution by a young scientist in the year 2004.

Frank Bernhard studied biology and chemistry at the University of Heidelberg, Germany. During his Ph.D. studies at the Max-Planck Institute for Medical Research in Heidelberg and at Columbus State University in Ohio, USA he analyzed signal transduction mechanisms of bacterial pathogens. After finishing his Ph.D. in 1991 he worked as a group leader at the Technical University in Berlin, Germany on the characterization and genetic engineering of peptide antibiotic synthetases. In 1995 he became laboratory leader at the Institute of Crystallography of the Free University in Berlin where he focused on the structural analysis of proteins. Since 2000 he has been leader of the protein expression unit at the Institute of Biophysical Chemistry at the University of Frankfurt, Germany. Frank Bernhard has more than ten years of experience in the large-scale expression of diverse varieties of proteins for structural research. In recent years, he has concentrated on strategies for efficient labeling of proteins for structural analysis by NMR techniques and for ligand binding studies.

Comprehensive Medicinal Chemistry II
ISBN (set): 0-08-044513-6
ISBN (Volume 3) 0-08-044516-0; pp. 1–22

Non-Print Items

Author's Contact Information

AU :7

Summary

The main purpose of preparative production of recombinant proteins is as the first prerequisite for drug target screening purposes, but in addition it has many applications in structural and functional research. Despite numerous attempts to streamline and to rationalize heterologous gene expression the vast majority of samples still have to be treated more or less empirically and a number of systems and hosts have to be tested. Only about 20% of the nonmembrane proteins in an organism may be instantly expressed in generalized standard systems in the amounts and qualities sufficient for a structural analysis: a high-level expression of membrane proteins is still the exception. The optimal expression of proteins is therefore often simply a result of trial and error. The ideal choice of an expression system with regard to control elements, vector characteristics, host environment, fermentation condition, and kinetics of expression is crucial, as it determines the yield and the quality of the recombinant protein. Expression systems are composed of relatively independent modules that can be optimized separately, and a good compromise between workload and costs have to be found in a reasonable time. A wealth of options and strategies exists and usually combined approaches have to be employed to become successful. This chapter is intended to serve as a guideline for making decisions in order to obtain the highest yields of recombinant proteins and it provides an overview about the currently most promising expression systems in bacterial and eukaryotic cells as well as in cell-free environments. We concentrate on critical details and features that have to be considered upon the in vitro recombination of expression vectors and during the process of protein production. Key elements of expression systems will be discussed individually; and modified techniques that have been approved for the production of distinct and difficult groups of proteins, e.g., membrane proteins or disulfide bonded proteins, will be especially highlighted. A special focus is also the preparative scale cell-free production of recombinant proteins as this relatively new technique provides several unique characteristics with a high potential for the future. However, this chapter does not cover downstream procedures in the preparation of recombinant proteins like purification strategies, nor does it include specific high-throughput approaches.

Keywords: baculovirus expression systems, cell-free expression, cell-lines, coexpression of proteins, disulfide bridge formation, expression systems, expression vectors, fusion proteins, mammalian cell expression, membrane proteins, promoter, protein overproduction, protein production in *Escherichia coli*, recombinant protein, semliki forest virus expression systems, solubility tags, vector design, viral expression systems, yeast expression systems

Non-Print Items

Nomenclature

ARSs	aminoacyl-tRNA synthetases
BHK	baby hamster kidney
CECF	continuous-exchange cell-free
CF	cell-free
CFCF	continuous-flow cell-free
CHO	Chinese hamster ovary
DHFR	dihydrofolate reductase
FM	feeding mixture
GPCR	G protein-coupled receptors
GST	glutathione *S*-transferase
IPTG	isopropyl-β-D-1-thiogalactopyranoside
mRNA	messenger RNA
MTX	methotrexate
NMR	nuclear magnetic resonance
NTP	nucleotide triphosphates
PDI	protein disulfide isomerase
RBS	ribosome binding site
RM	reaction mixture
rRNA	ribosomal RNA
SFV	Semliki Forest virus
tRNA	transfer RNA
UTR	untranslated region

(P 08)

^{13}C- and ^{15}N-Isotopic Labeling of Proteins. C. Klammt, F. Bernhard and H. Rüterjans. In: *Molecular Biology in Medical Chemistry*, T. Dingermann, G. Folkers, H. Steinhilber (eds.), pp. 269-292, Wiley-VCH, Weinheim. (2004)

10
^{13}C- and ^{15}N-Isotopic Labeling of Proteins

Christian Klammt, Frank Bernhard and Heinz Rüterjans

10.1
Introduction

Nuclear magnetic resonance (NMR) spectroscopy has experienced a tremendous growth over the past decade, and has been developed as a widely used technique with enormous potential for the study of the structure and dynamics of biological macromolecules. In particular, structural analysis of proteins has greatly benefited from recent advancements of high field solution NMR. Improvements in the hardware and the development of NMR spectrometers with field strengths up to 900 MHz considerably increased the sensitivity and reduced the requirements for high amounts of samples. Furthermore, major milestones in the study of proteins have been the use of isotope labeling and the development of multidimensional techniques. The structure of small proteins below a limit of approximately 10 kDa could be analyzed without the need for special isotope labeling after the development of homonuclear ^{1}H two-dimensional (2-D) experiments in the late 1970s and early 1980s. However, beyond that limit, the increased complexity due to line broadening and ^{1}H chemical shift overlap prevents detailed structural analysis of larger proteins by these techniques. The structural analysis of large proteins therefore requires multinuclear labeling and special strategies in order to replace the prevalent NMR inactive isotopes ^{14}N and ^{12}C in proteins by the NMR active isotopes ^{13}C and ^{15}N [1]. Protocols have been developed to generate either uniformly ^{13}C/^{15}N-labeled proteins or to introduce the label only at specific sites. In addition, heteronuclear multidimensional NMR experiments have been designed that utilize the relatively large one- and two-bond scalar couplings introduced by the label [2, 3]. Using the ^{13}C and/or ^{15}N chemical shifts, the overlapping ^{1}H resonances can now be routinely separated into three or four dimensions. The combination of these techniques addresses the limitations imposed on the study of large proteins and permits structure determination of proteins up to a molecular weight of approximately 22 kDa [4, 5]. This limit could be extended up to 40 kDa by the random incorporation of ^{2}H into nonexchangeable sites of proteins [6–9].

In this chapter, recently developed strategies for the incorporation of ^{15}N and ^{13}C labels into proteins resulting in highly sensitive spectra and facilitating struc-

Molecular Biology in Medicinal Chemistry. Edited by Th. Dingermann, D. Steinhilber, G. Folkers

ISBN: 3-527-30431-2

tural analysis of high-molecular-weight systems are discussed. Conventional labeling strategies require the overproduction of proteins in bacterial or eukaryotic expression systems and we give an overview of the most advanced systems as they are indispensable prerequisites for the efficient incorporation of isotopes into proteins. In addition to structural studies, NMR plays an increasingly important role in the description of protein dynamics, and relationships between dynamics and function. We therefore also address techniques of selective isotope labeling for the study of protein side-chain dynamics and protein–ligand interactions. The second part of the chapter focuses on newly developed *in vitro* techniques, utilizing cell-free extracts for the generation of protein samples. While cell-free production of proteins in an analytical scale has been possible for several years, preparative protein production using extracts from *Escherichia coli* or other sources has become one of the most powerful tools for the production of labeled samples suitable for NMR analysis. We therefore will especially address the recent advancements in cell-free expression and labeling strategies.

10.2
Expression Systems for the *In Vivo* Incorporation of ^{13}C and ^{15}N Labels into Proteins

Apart from technical barriers, the quality and stability of protein samples often limit their analysis by NMR technologies. The relatively low sensitivity of most NMR spectrometers requires sample concentrations starting from 0.2 mM and many proteins aggregate long before they reach that limit. In addition, samples must remain stable for at least several days until a complete set of measurements is finished. In order to obtain optimal resolution, NMR analysis usually requires relatively high temperatures of 15–25°C, which could also affect protein stability. Due to the high costs involved in the preparation of multinuclear-labeled proteins, one sample should be used for several experiments and aggregation, denaturation or even degradation of the proteins due to spurious amounts of proteases or due to autoproteolysis needs to be reduced to a minimum. Reasonably high yields of sample preparations are therefore required, which in conventional *in vivo* expression systems should reach a minimum of several milligrams of protein per liter of culture medium. In order to obtain the best protein production rates, the choice of the optimal expression system is therefore of primary importance. The most commonly used organism for labeled recombinant protein production is still the enterobacterium *E. coli*. An immense variety of expression systems and vectors have been designed for protein production in *E. coli*, and this large selection is certainly one of the major advantages for choosing an *E. coli* expression system. The predominant eukaryotic systems for protein production are based on the highly productive species *Pichia pastoris*. Labeling of proteins in mammalian cells is only rarely considered because of the high costs due to low yields concomitant with the need for expensive labeling precursors.

Labeled protein samples were initially generated by growing the cells in defined minimal media supplemented with specifically labeled compounds. Common label

precursors include $^{15}NH_4Cl$, $(^{15}NH_4)_2SO_4$, $K^{15}NO_3$ and $Na^{15}NO_3$ for nitrogen labeling, and $NaH^{13}CO_3$, [^{13}C]glycerol, [^{13}C]glucose and [^{13}C]acetate for carbon labeling. A variety of supplements such as trace element mixtures and vitamin cocktails have been tested to enhance the cellular growth in minimal media and to improve the productivity. Furthermore, single- or double-labeled mixtures of amino acids or complex algal or microbial hydrolyzates have been added to culture media [10–13] in order to produce uniformly labeled proteins. In order to reduce the isotope costs, cells may be grown in unlabeled rich media to high cell densities, harvested and then suspended in low volumes of labeled defined medium for protein production. High levels of isotope incorporation concomitant with a 4- to 8-fold increase in yield were obtained in *E. coli* with this technique [14, 15].

10.2.1
Protein Production and ^{13}C and ^{15}N-labeling in *E. coli* Expression Systems

Techniques for isotope enrichment in *E. coli* have been available for several years [16–18] and the vast majority of labeled proteins analyzed by NMR spectroscopy are still being produced by heterologous expression in various *E. coli* hosts yielding up to 50% of the total cell protein. Despite considerable improvements in eukaryotic and cell-free expression systems, bacterial expression remains the fastest and most economical system for the production of proteins. Standard methods of generating heteronuclear-labeled samples in *E. coli* use M9 minimal medium [19] or modified derivatives of it – supplemented with [^{13}C]glucose and [^{13}C]glycerol for carbon labeling, and $(^{15}NH_4)_2SO_4$ and $^{15}NH_4Cl$ for nitrogen labeling. An important advantage when using *E. coli* expression is the elaborated selection of vectors and host strains available. Upon overproduction of proteins, it is often essential that the transcription of the target gene can be repressed as tightly as possible until the cells reach the most suitable growth phase for induction. Leaky expression prior to induction could either favor the accumulation of mutations in the target gene in order to suppress the production of unwanted proteins, or even damage or kill the cells due to toxic effects.

The most popular inducible promoters in *E. coli* expression systems are listed in Tab. 10.1. Repression of the strong *E. coli lac* promoter is performed by interaction

Tab. 10.1 Commonly used promoters for protein overproduction in *E. coli*.

Promoter	*Origin*	*Repression*	*Induction*
P_{lac}	*E. coli*	LacI	isopropyl β-thiogalactoside (IPTG)
P_{ara}	*E. coli*	AraR	L-arabinose
P_{tet}	*E. coli*, Tn10	TetR	tetracycline
$\phi1$–$\phi10$	phage T7 of T7 RNA polymerase	repression/inactivation	requires T7 RNA polymerase, IPTG
λ_{PL}	phage λ	λ *cI* repressor	heat inducible, temperature shift of 42°C

of its repressor LacI with specific operator sequences in P_{lac}, thus preventing the *E. coli* RNA polymerase from binding. The LacI protein is released from its operator by the addition of isopropyl *β*-thiogalactoside (IPTG), a nonmetabolizable analog of the natural inducer. A disadvantage of the *lac* promoter is its relatively high background expression. This could be partly suppressed by an increasing copy number of LacI in expression hosts either by introducing additional copies of the *lacI* gene on a plasmid or by increasing the expression of the chromosomal *lacI* gene by an up mutation of its native promoter, called $lacI^q$. Several derivatives, like the *tac* or *trc* promoter, have been constructed by fusion of P_{lac} with other promoters, e.g. from the tryptophan operon.

In expression systems based on the highly specific T7 promoters, the protein production is induced by the supply of T7 RNA polymerase [20]. In principal, the T7 RNA polymerase can be provided in several ways. A common method is the expression in specially designed strains carrying a chromosomal copy of T7 gene 1, encoding T7 RNA polymerase. This construct, termed ''DE3'', represents a transcriptional fusion of gene 1 with the IPTG-inducible *lac* promoter. To ensure tight repression, gene 1 is accompanied by an extra copy of the $lacI^q$ gene. In order to eliminate any background of T7 RNA polymerase due to leaky expression, the plasmid-borne gene encoding for T7 lysozyme can be provided. T7 lysozyme efficiently binds and inactivates T7 RNA polymerase, and this expression system even allows the production and labeling of toxic proteins in *E. coli*.

While the structural analysis of the native protein is predominantly the major task of a labeling approach, most samples were initially produced as translational fusions to other proteins or to artificial peptide-tags (Tab. 10.2). Those strategies can provide valuable advantages for the fast and efficient purification of the labeled protein or for its stabilization [21]. Furthermore, the addition of a suitable fusion partner to the N-terminus of the target protein can significantly enhance its production. Fusion proteins usually include a recognition site for a highly specific protease like enterokinase or the TEV protease in the linker between the two proteins. If necessary, the fusion partners could then be cleaved off from the target proteins after purification, leaving the nonmodified labeled protein for analysis.

Tab. 10.2 Popular tags and carrier proteins for protein overproduction in fusion systems.

Fusion	*Effect on solubility*[a]	*Purification*
Poly$(His)_6$-tag	no effect	immobilized metal chelate chromatography
Strep-tag	no effect	affinity chromatography
Maltose-binding protein	++	affinity chromatography
NusA	++	
Glutathione S-transferase	+	affinity chromatography
Thioredoxin	+	
Protein G (B1 domain)	+	
λ head protein D	+	
Protein A (Z domain)	+	affinity chromatography

[a] Effect on the solubility of the target protein.

The most frequent peptide-tag used for an easy purification is a poly$(His)_6$-tag added to the N- or C-terminus of the protein. Such small tags usually do not disturb the conformation or activity of proteins and their removal after purification might not even be necessary. Proteins containing a poly$(His)_6$-tag can easily be purified from a crude cell extract by forming a chelate complex with metal ions like Ni^{2+} or Cu^{2+} immobilized on a chromatography column. The imidazole moieties of adjacent histidine residues are crucial for firm chelate formation and the bound proteins can later be eluted from the column with a competing imidazole gradient. The production of proteins with a poly$(His)_6$-tag has become a routine technique for the heterologous expression of recombinant genes in any host.

Only approx. 25% of overproduced proteins are initially suitable or stable enough for structural analysis [22] and most experiments must be optimized in order to obtain sufficient yields of correctly folded proteins. General parameters of optimization are the selection of suitable expression hosts and vectors, variation of growth conditions and inducer concentrations, construction of fusion proteins, and coexpression of helper proteins. A frequently encountered problem upon heterologous expression in *E. coli* is the formation of inclusion bodies – large aggregates of unfolded inactive protein. Protein targeting as insoluble inclusion bodies could be a strategy to circumvent bactericidal activities of the overproduced protein [23]. Several protocols have been described to solubilize inclusion bodies after purification in strong denaturants and to refold denatured proteins into their correct 3-D conformation [24, 25]. However, for each protein the refolding strategy has to be optimized and might represent an intensive laborious trial-and-error procedure giving less than satisfactory results. Thus, it is often intended to maximize the production of proteins in completely soluble form. Fortunately, in many cases, the formation of inclusion bodies during expression can be prevented by C-terminal linkage of the target protein to a highly soluble fusion partner. Several carrier proteins are able to enhance the solubility of an otherwise insoluble proteins (Tab. 10.2); in systematic screens for proteins suitable to confer solubility on insoluble target proteins fused to their C-terminal ends, the *E. coli* maltose-binding protein MalE [26] and NusA [27] proved to be the most efficient. As large carrier proteins have to be removed through specific proteases prior to NMR analysis, the cleavage of the fusion proteins can reintroduce problems with solubility and the released target proteins sometimes tend to precipitate soon after they are no longer covalently attached to a carrier. In addition, the use of large carrier proteins would waste a considerable amount of the precious label precursors. Smaller carrier proteins like the phage λ head protein D [28] or domains of *Staphylococcus* Protein A [29] or *Streptococcus* Protein G [30] might therefore be used as noncleavable tags to enhance the stability of smaller proteins or peptides. Fusions with the B1 domain of Protein G even improved NMR spectra of labeled proteins [30]. As well as enhancing solubility, carrier proteins can be used to facilitate the production of small peptides in *E. coli* and to stabilize them against degradation [23].

Misfolding can be a particular problem with the expression of eukaryotic proteins in *E. coli*, especially if they have several disulfide bonds, consist of multiple subunits or contain prosthetic groups. A popular strategy to address such problems

Tab. 10.3 Modifications of *E. coli* expression systems.

Problems/requirements	*Modifications*
Disulfide bridge formation	expression in *trx*$^-$ and *gor*$^-$ strains, providing an oxidizing environment in the cytoplasm
	coexpression of disulfide isomerases like DsbA, DsbC or PDI
	periplasmic expression using signal peptides
Chaperone-dependent folding	coexpression of chaperones like GroELS and/or DnaKJ/GrpE
Rare codon usage	coexpression of corresponding tRNAs

is the coexpression of accessory helper proteins (Tab. 10.3). If folding of the overproduced protein depends on chaperones, the increase of chaperone concentrations by coexpression of GroEL/ES, DnaKJ/GrpE or other chaperone systems could be highly advantageous. Low production rates due to translational pausing or misincorporation of amino acids as a result of rare codon usage of the overproduced protein could be optimized by the coexpression of corresponding tRNAs. A frequent problem, which in many cases still cannot be completely solved when using *E. coli* expression systems, is the correct formation of disulfide bonds in proteins. Strategies for the optimization of disulfide bond formation *in vivo* include the export of overproduced proteins into the oxidizing periplasm by addition of appropriate export signal sequences, expression in specially designed strains lacking thioredoxin and glutathione reductase and thus providing an oxidizing cellular environment [31], and coexpression of disulfide isomerases like the DsbA and DsbC proteins of *E. coli* or the eukaryotic chaperone protein disulfide isomerase (PDI).

10.2.2
^{13}C- and ^{15}N-labeling of Proteins in *P. pastoris*

Yeast cells combine many advantages of prokaryotic and eukaryotic expression systems. They are small, robust and easy to handle. The growth rate of yeast cells is much faster compared with other eukaryotic cells and the fermentation of yeast cells is usually finished within a couple of days. Furthermore, yeasts do not require complex supplements in the growth media and can be propagated in simple inexpensive media or on agar plates like *E. coli* cells. Their excellent growth on defined minimal media make them especially attractive for the production of labeled proteins for NMR analysis. In addition, due to the secretion capabilities of *P. pastoris*, recombinant proteins might be easily purified in a single step [32]. Many proteins of eukaryotic origin need to undergo specific posttranslational modifications like glycosylation or lipidation, disulfide bridge formation, or posttranslational processing. The modifications are often essential for the stability or enzymatic activity of proteins [33], or for the adoption of their native conformation. The failure to form correct disulfide bridges can retard or even block the folding process of proteins, resulting in the formation of inclusion bodies or in their degradation. The production of modified proteins cannot be addressed in a satisfactory manner by using

prokaryotic hosts like *E. coli*. Expression in yeast cells like the methylotrophic *P. pastoris* may therefore be the method of choice when searching for a fast-growing host with the potential for the production of modified and glycosylated proteins [34, 35]. However, the fact that the protein modification pattern obtained after expression in yeast might differ from that usually present in the native protein has to be considered, e.g. yeasts are only capable of attaching mannose-rich glycans, representing the core polysaccharide of glycoproteins.

Labeling approaches should take into account that heterologous protein production in *P. pastoris* is correlated with a high cell density. Up to 100-fold increased yields can be obtained by fermentation of the cells in bioreactors as compared to shake flask cultures [36–38]. The induction of heterologous protein production should be started after reaching high cell densities. Yields of more than 100 mg protein/l culture medium have been reported [34, 38]. In shake flasks, the reported yields of uniformly labeled heterologous proteins vary from 2 mg/l in case of the Vaccinia virus complement control protein [39] to up to 27 mg/l for tick anticoagulant protein [36]. The average protein yield in yeast using shake flasks therefore only reaches the lower limit of comparable protein production in *E. coli*. It is furthermore important to realize that if compared with protein production in *E. coli*, much higher amounts of labeled isotopes are required to obtain a reasonable growth of the yeast cells, and at least 5- to 10-fold higher concentrations of $(^{15}NH_4)_2SO_4$ and [^{13}C]glucose are routinely used when growing the cells in shake flasks. To produce 90 mg of a ^{13}C-labeled fragment of thrombomodulin in *P. pastoris*, 143 g [^{13}C]glycerol or 162 g [^{13}C]glucose had to be used by growing the cells in a 1.25-l fermentor [34]. Therefore, based on economic aspects, labeling approaches with yeasts are only competitive in cases when the proteins cannot be produced in *E. coli*.

In defined minimal medium, the sole nitrogen source for *P. pastoris* usually is NH_4OH, which in addition serves as a base to neutralize considerable amounts of acid produced by the yeast during growth [34]. However, due to economic reasons, NH_4OH has to be replaced by $(^{15}NH_4)_2SO_4$ for ^{15}N-labeling of proteins in *P. pastoris*. The required high cell density demands the addition of high amounts of $(^{15}NH_4)_2SO_4$ and more than 40-folds concentration as compared with fermentations of *E. coli* had to be used. To avoid any negative effects on cell growth arising from the increased ionic strength of the medium through the formation of K_2SO_4, the $(^{15}NH_4)_2SO_4$ had to be added in several portions at different times [38], and optionally also in combination with an exchange of medium before induction of protein expression [34].

For ^{13}C-labeling of proteins using the strong *AOX1* promoter, two usually carbon sources are required during fermentation – glycerol or glucose to obtain a high cell density and methanol for the induction of the *AOX1* promoter. Although *P. pastoris* is able to grow on methanol as the sole carbon source, the slow doubling time would prevent high cell densities and requires the use of alternate carbon sources like glycerol or glucose. While glucose could certainly help to increase the cell density, it represses the strong *AOX1* promoter, leading to a decrease in heterologous protein production. However, [^{13}C]glucose could be used in labeling experi-

ments as an initial carbon source until the desired high cell density is reached. The cells may then be further fermented with [^{13}C]glycerol as the carbon source and induced with [^{13}C]methanol [34]. Up to 30% of the carbon from the labeled protein originates from methanol and therefore it is crucial to induce the protein production with [^{13}C]-methanol in order to obtain fully labeled proteins, especially when using a yeast strain with a *mut*$^{+}$ genotype [34].

10.2.3 ^{13}C- and ^{15}N-labeling of Proteins by Expression in Chinese Hamster Ovary (CHO) Cells

Compared to yeast cells, CHO cells represent the alternative of a more advanced system enabling the generation of labeled samples containing all of the possible modifications specific to eukaryotic proteins. Like with yeast cells, recombinant proteins can be efficiently secreted into the medium and protein production rates of more than 100 mg/l from CHO cells have been reported [40]. However, mammalian cells usually require amino acids, vitamins, cofactors and in many cases a complex serum as essential additives to the medium. To reduce costs, CHO cells could also be grown by replacing the highly expensive isotopically labeled amino acids with labeled bacterial hydrolyzates or algal extracts and sufficient yields of recombinant protein highly enriched with isotopes can be obtained [10]. To avoid dilution of the isotope label with unlabeled amino acids from the serum, a dialysis step should be included to remove low-molecular-weight compounds from the serum [10, 41]. A further important and probably cost-intensive factor is the requirement of relatively high concentrations of labeled glutamine, which is essential for the metabolic pathways of several amino acids and nucleic acid components. Although CHO cells represent an expensive host for the generation of labeled proteins, they might be the best choice for the overexpression of complex glycoproteins or of proteins that cannot be produced in *E. coli* [41–43].

10.2.4 ^{13}C- and ^{15}N-labeling of Proteins in Other Organisms

A further option to produce labeled glycosylated proteins can be the overproduction in the slime mould *Dictyostelium discoideum*. This host is able to feed on bacterial cells like *E. coli* that could then be pre-enriched with isotopic labels using standard protocols. Suitable expression vectors for *D. discoideum* are available and a glycosylated 16-kDa protein has already been labeled with ^{13}C/^{15}N to a high extent [44]. As the growth rate of *D. discoideum* is rather low, the complete experiment took more than 10 days. However, if compared to labeling approaches using yeast cells as a host, lower amounts of labeled precursors are needed to obtain sufficient protein for NMR measurements, and *D. discoideum* might therefore be considered as an economic alternative to produce labeled glycoproteins.

While the previously discussed expression systems require relatively expensive mixtures of labeled precursors, the photoautotrophic cyanobacterium *Anabaena* sp.

can be grown in defined media containing $Na^{15}NO_3$ and $NaH^{13}CO_3$ as the sole nitrogen and carbon sources. A 24-kDa domain of the *E. coli* gyrase B subunit was successfully overproduced in *Anabaena* sp. by using a specially designed shuttle vector system and more than 90% $^{13}C/^{15}N$ label incorporation was obtained [45]. The yield of the protein (approximately 3–6 mg/l) was found to be comparable to that obtained in *E. coli*. It is noteworthy that the expressed gene was under control of the *E. coli tac* promoter, which is also functional in *Anabaena*. While ^{15}N-labeling in *Anabaena* can be carried out with standard growth protocols, ^{13}C-labeling is problematic under aerobic conditions because of the ability of endogenous CO_2 fixation, resulting in incomplete ^{13}C-labeling of less than 30% [45]. However, ^{13}C-enrichment of more than 90% could be obtained after growing the *Anabaena* cells with nitrogen gas aeration under controlled anaerobic conditions. The fermentation of *Anabaena* requires illumination and the duration is about 5 times longer compared to *E. coli* because of the slower doubling time. However, as less expensive labeled precursors can be used, the cost could be reduced to only about 10%. Heterologous protein production in *Anabaena* might therefore be an option for proteins with a low production rate for which higher volumes of medium have to be used. In addition, it has to be considered that the non-protonated nature of the final carbon supply might also be advantageous and cost-effective for the generation of perdeuterated protein samples necessary for the structural analysis of large proteins or dynamic studies.

Although heterologous protein production is the most common approach, small peptide products have also been labeled in their natural host. A crucial prerequisite is that the organisms can be adapted to grow on defined media supplemented with suitable labeled precursors. Examples of successful approaches are the labeling of cyclosporin A in *Tolypocladium inflatum* [46], alamethicin in *Trichoderma viride* [47], bellenamine in *Streptomyces nashvillensis* [48] and cyanophycin in *Synechocystis* sp. [49].

10.2.5 Strategies for the Production of Selectively ^{13}C- and ^{15}N-labeled Proteins

NMR spectra of uniformly labeled proteins become increasingly complex with the increasing size of the proteins. Several methods have been established to selectively label only specific domains, amino acids or even single nuclei of the protein in order to simplify the spectral analysis. Using *in vivo* labeling techniques, this can be done by feeding the host cells with specifically designed label precursors, which are obtained in most cases by chemical synthesis. The resulting spectral simplification facilitates the unambiguous assignment of resonances in large proteins. Choosing the correct labeling strategy can therefore be crucial to accelerate the assignment of resonances in proteins.

10.2.5.1 Selective Labeling of Amino Acids

In principal, any amino acid residue in an overproduced protein can be specifically labeled by providing excessive amounts of the amino acid containing the desired

label in the growth medium. Resonances of single amino acids can be easily identified by comparison with spectra of a protein with other labeled amino acids and such approaches could considerably accelerate the resonance assignments of large proteins. However, the high costs for purified labeled amino acids prevent the application of this technique for routine use of structure determinations. The selective labeling strategy is suited for the dynamic and functional analysis of selected amino acid residues e.g. during ligand interactions. Specific labeling is also useful in the analysis of protein denaturation. Unfolded proteins generally have a low dispersion of resonances, but the tracking of a limited number of labeled residues, ideally well distributed throughout the protein, could enable the analysis of the unfolding process. In a previous study, selective labeling with [^{15}N]isoleucine was used to detect folding intermediates of a complex protein [50].

As an alternatively to specific labeling of amino acids, in reverse isotope labeling an excess of one or several nonlabeled amino acids is added to a growth medium in combination with commonly used general labeling compounds like $^{15}NH_4Cl$ or [^{13}C]glucose. An interesting application is the reverse labeling of aromatic residues like phenylalanine or tyrosine as their side-chains are frequently involved in ligand-binding interfaces or positioned in hydrophobic cores of proteins, making distance restraints for the environment of these residues highly valuable [51, 52]. The reverse isotope-labeling approach clearly depends on the complexity of the analyzed protein and is generally useful for smaller proteins containing only few of the analyzed amino acid residues. The application of amino acid-specific or selective isotope-labeling strategies is limited by scrambling effects of the isotope label to other types of residues. Specific auxotrophic mutants of *E. coli* can be used for the overproduction and specific labeling of recombinant proteins in order to minimize this problem, [53].

10.2.5.2 Specific Isotope Labeling with ^{13}C

Amino acids labeled at selected carbon positions can be added to defined growth media for their incorporation into the recombinant protein. These partially labeled compounds were usually generated by chemical synthesis and introduced to optimize the spectral resolution of large proteins in specific NMR experiments. Aromatic side-chain protons may be difficult to assign if the number of aromatic residues increases. Phenylalanine residues ^{13}C-labeled only at the ε position could thus help to rapidly assign the aromatic ring protons [54]. Furthermore, fully ^{13}C-labeled proteins have undesirable $^{13}C^{\alpha}-^{13}C^{\beta}$ scalar couplings. Incorporation of amino acids labeled in the ^{13}C backbone overcome this problem [55]. Fractional ^{13}C-labeling of amino acids can be achieved by growing the protein-producing cells with a defined mixture of ^{13}C- and ^{12}C-labeled carbon sources [56]. The analysis of internal side-chain dynamics within proteins in uniformly ^{13}C-labeled proteins is also complicated due to interference effects between different contributing relaxation interactions as well as by contributions from $^{13}C-^{13}C$ scalar and dipolar couplings. To solve this problem, proteins can be synthesized with an alternating $^{12}C-^{13}C-^{12}C$-labeling pattern by an elaborate isotope labeling procedure using either [2-^{13}C]glycerol or [1,3-$^{13}C_2$]glycerol as the sole carbon source [57]. Further

carbon-selective labeling strategies are highly valuable in combination with selective protonation to analyze the structure of large proteins [52].

10.2.5.3 **Segmental Isotope Labeling**

Despite advances in heteronuclear multidimensional NMR techniques, the increased complexity of spectra due to a lack of resolution and increased overlap of signals having similar chemical shifts still limits the structural analysis of large proteins. However, a promising approach to further extend the actual size limit of the structural evaluation of proteins by NMR spectroscopy is the segmental- or block-labeling technique. In principal, partially labeled full-length proteins can be produced by the posttranslational *in vitro* ligation of nonlabeled domains together with a labeled domain obtained in different expression experiments. Resonances of single domains of larger proteins may be sequentially assigned and the complete structure of the protein could be obtained by a combinatorial approach, dividing a large target protein into parts of manageable size. In contrast to a structural elucidation of isolated protein domains, one should bear in mind that the structure of labeled domains with the segmental-labeling technique is analyzed in the context of the complete protein, thus providing information of the structural and functional interaction between domains while preserving the overall structural character of the protein. The ligation process is catalyzed by self-splicing enzymes – the inteins [58]. Inteins are insertion sequences cleaved off after translation through self-excision, leaving the flanking protein regions, the exteins, joined together through native peptide bond formation.

Depending on the nature of the intein and on the solubility of the target domains, several modified protocols of the segmental labeling strategy have been established.

- The intein is cut in the middle of the sequence within a flexible loop region and the two fragments are expressed separately as fusions to the target domains. The isolated denatured proteins are mixed and the two intein fragments efficiently associate during refolding, resulting in the ligation of the two target domains [59, 60]. This strategy is especially applicable if the intein fusions are expressed as insoluble inclusion bodies and if an efficient protocol for refolding is available. Using different inteins at each splice junction, even proteins containing central labeled domains can be generated [61].
- The chemical ligation of folded proteins under native conditions uses an ethyl α-thioester at the C-terminal end of a recombinant protein, generated by cleavage of an intein fusion, for the ligation with a second protein or peptide containing a cysteine residue at its N-terminal end [62]. Both ligation partners can be combined in a correctly folded conformation at a physiological pH and this strategy principally could be extended to link more than two domains together.
- The mini intein from the *dnaE* gene of the cyanobacterium *Synechocystis* sp. comprises less than 200 amino acid residues, and is divided into N- and C-terminal fragments that are expressed separately. If combined after purification, the two fragments efficiently interact under native conditions to catalyze the

splicing and ligation reaction of covalently attached proteins [63, 64]. Target domains to be ligated have to be expressed as fusion N- and C-terminal to the intein fragments.

Although *in vitro* ligation yields as high as 90% have been obtained, several potential problems should be considered when starting a segmental labeling approach. While most enzymatically important residues are located within the inteins, an efficient excision and ligation mechanism also requires some residues from the extein, and the reaction is therefore not completely independent of the sequence of newly formed protein junctions [64]. The C-terminal extein requires a cysteine, threonine or serine residue at its N-terminal end. This sequence requirement must not affect the structure or activity of the analyzed protein. Refolding and ligation conditions must be optimized for each individual protein, and the splicing/ligation junction should be located in a loop region ensuring high flexibility. In principal, block labeling might apply best to proteins composed of independently folded domains having distinct biochemical properties.

10.3 Cell-free Isotope Labeling

Cell-free protein synthesis is an attractive and promising alternative to the conventional technologies for protein production using bacterial or eukaryotic cell cultures. In contrast to *in vivo* gene expression methods where protein synthesis is carried out in the cellular context surrounded by cell walls and membranes, cell-free protein synthesis provides a completely open system. This allows direct control and access to the reaction at any time. Compounds to improve protein production or to stabilize recombinant proteins, e.g. chaperones, chemicals, detergents or protease inhibitors, can easily be added without considering side-effects of the cellular metabolism or transport problems through the cell membrane. Most cellular functions with the exception of transcription and translation need not be maintained during cell-free protein expression. In principal, applications can therefore be extended to proteins or conditions that would not be tolerated by living organisms. While *in vivo* protein production is often limited by the formation of insoluble inclusion bodies or protein instability caused by intracellular proteases, the cell-free system offers new possibilities for the synthesis of complex proteins (Tab. 10.4). Protein folding and stability can be promoted by direct addition of chaperones, PDIs or necessary cofactors. In contrast, overexpression of chaperones *in vivo* can lead to cell filamentation or other undesirable phenotypes that can be detrimental for the viability of *E. coli* cells and protein expression [65]. Reaction parameters such as pH, redox potential and ionic strength can be determined without concern for harmful effects on the growth and viability of cells, and with the certainty that these parameters will directly influence the relevant reactions. This new opportunity of *in vitro* gene expression allows full control and high flexibility of conditions, and offers new potentials for difficulties associated with cyto-

Tab. 10.4 Potential advantages and disadvantages of cell-free protein synthesis.

Potential advantages	*Disadvantages*
Completely open system (easy access) → customization of reaction conditions to synthesized protein	Often low productivity
Incorporation of labeled, glycoslylated, modified or unnatural amino acids	Complex system compared to *in vivo* protein synthesis
Direct translation of PCR products → high-throughput screening	Expensive reaction compared to *in vivo* expression
Production of toxic proteins	Small number of commercial systems and kits
Production of membrane bound proteins	Difficult standardization because of multitude of reaction components
Expression of proteins requiring cofactors	High costs
Easy addition of chaperones or PDI	
Miniaturization (e.g. 50 µl reactions)	
Working without living organisms → no growth restrictions	
Allowing a direct isolation of products to shorten time required for preparing purified proteins	

toxicity, proteolytic degradation or improper folding and aggregation of synthesized proteins. The production of cytotoxic proteins [66], membrane proteins [67, 68] as well as the production of functional antibodies using PDI and chaperones [69] or functional viruses [70] had been reported. Most of all, cell-free protein synthesis offers completely new possibilities for the incorporation of labeled [71–74], glycosylated [75, 76], modified or unnatural amino acids [77]. Another promising potential for cell-free synthesis can be found in its suitability for high-throughput expression of proteins by direct translation of linear polymerase chain reaction (PCR) products [78]. Current limitations of cell-free systems are connected with their high complexity, high costs and often low productivity.

10.3.1 Components of Cell-free Expression Systems

In cell-free protein production, all components involved in gene expression and protein synthesis have to be added to the reaction mixture (Fig. 10.1). Components like DNA, nucleotide triphosphates (NTPs), messenger RNA (mRNA), transfer RNA (tRNA), aminoacyl-tRNA synthetases (ARSases), polymerases, ribosomes, transcription and translation factors like initiation factors (IFs), elongation factors (EFs) or release factors (RFs) and amino acids have to be optimized with regard to their concentration, and optimal salt and pH environment. The process of transcription/translation requires large amounts of free energy supplied by the hydrolysis of the triphosphates ATP and GTP. Therefore, *in vitro* protein synthesis requires an ATP-regenerating energy system to maintain the triphosphate concentration. For this purpose, high-energy phosphate donors such as acetyl phosphate

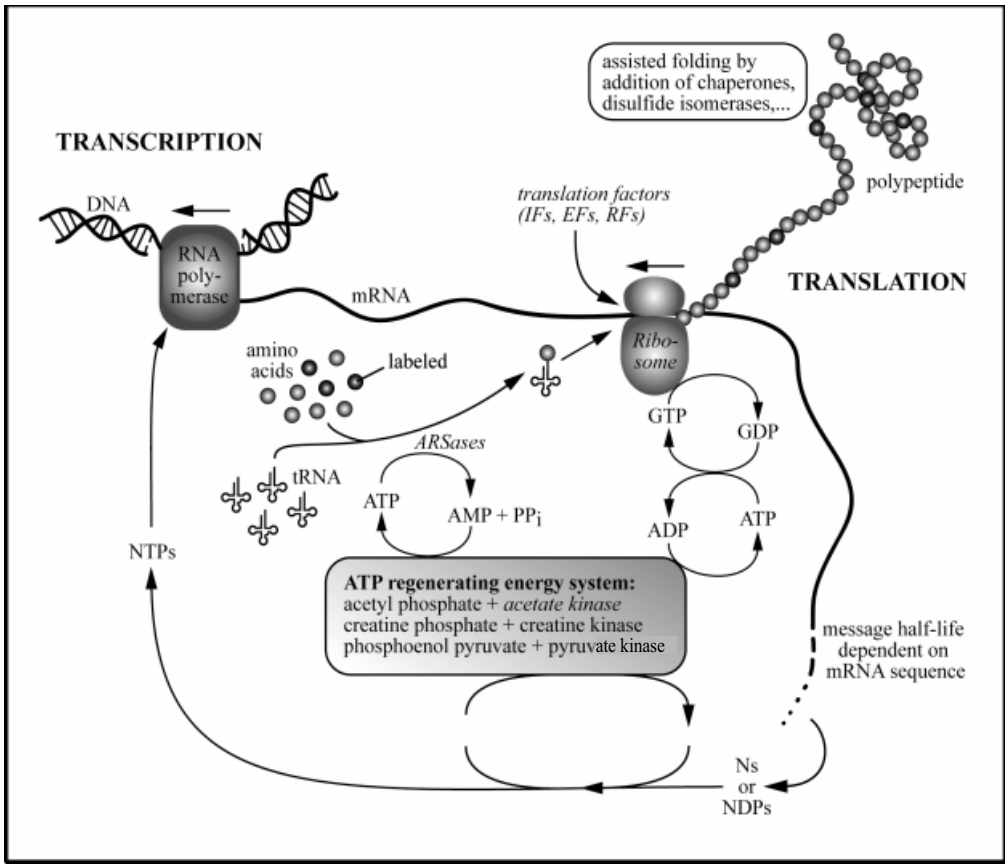


Fig. 10.1 Schematic to illustrate cell-free protein synthesis, showing the coupled process of transcription and translation in a bacterial system. Initially, cells are grown, lysed and the cell extract is prepared containing ribosomes, translation factors like initiation factors (IFs), elongation factors (EFs) and release factors (RFs), acetate kinase and aminoacyl-tRNA synthetases (ARSases). Substrates like amino acids, the energy-regenerating system components or nucleoside triphosphates (NTPs) and salts are then added to the extract, and protein synthesis is initiated by adding the template DNA. The DNA is transcribed by an added RNA polymerase. Added tRNA is loaded with amino acids by ARSases and they are used in the translation of mRNA. The incorporation of some stable isotope-labeled amino acids (dark) can easily be done in the cell-free system, leading to a selective isotope-labeled protein. Regeneration of ATP and GTP and even the NTPs in the cell-free system is achieved by an ATP-regenerating energy system based on the hydrolysis of high-energy substrates in the presence of their kinases. Chaperones can easily be added to the reaction mixture to assist the folding of the target protein.

(AcP), creatine phosphate (CrP) or phosphoenol pyruvate (PEP) in the presence of their kinases (acetate kinase, creatine kinase and pyruvate kinase) have been used.

A crucial component is a cell extract based on crude cell lysate which contains the necessary reaction components such as IFs, EFs, RFs, ARSases, tRNA and

Tab. 10.5 Bacterial and eukaryotic cell-free expression systems.

Bacterial systems		***Eukaryotic systems***	
Escherichia coli		***Wheat germ***	***Rabbit reticulocyte***
S30 extract preparation	***S100 extract preparation***	***Wheat germ extract***	***Rabbit reticulocyte lysates***
Supernatant fraction at 30,000 *g* centrifugation of *E. coli* extract preincubated to detract endogenous DNA and mRNA	Supernatant fraction at 100,000 *g* centrifugation deprived of all nucleic acids, e.g. by DEAE cellulose treatment	Directly used for expression of endogenous or exogenous templates	Treated with micrococcal Ca^{2+} dependent RNase
Contains: RNA polymerase, ribosomes, tRNAs, ARSases, translation factors	Contains: RNA polymerases, ARSases and translation factors Ribosomes, tRNAs added	Contains: Ribosomes, tRNAs, ARSases, translation factors	Contains: Ribosomes, tRNAs, ARSases, translation factors
24–38°C (optimum at 37°C)[a] [79, 82, 83, 132, 133][b]	24–38°C (optimum at 37°C)[a] [80, 81][b]	20–27°C up to 32°C[a] [85–87][b]	30–38°C[a] [82, 89, 98, 113][b]

[a] Reaction temperature.
[b] References.
To be added:
DNA (plasmid with appropriate promotor for SP6, T7 RNA-polymerase) + polymerase (bacteriophage SP6 or T7 RNA polymerase) or mRNA.
Amino acids.
Energy components: ATP and GTP.
NTP-regeneration system: PEP + PK or CP + CrP or AcP.
Formyltetrahydrofolate or its congener, e.g. methenyltetrahydrofolate or folinic acid.
SH-compound: mercaptoethanol or dithiothreitol.
Mg^{2+} and K^{+} in optimal concentrations.
In some cases:
Ca^{2+} and NH_4^{+} in optimal concentrations.
cAMP.
Polyamines, e.g. spermidine.

enzymes for energy regeneration like acetate kinase. Cell-free expression systems are classified according to the origin of their extract. In principle, functional *in vitro* systems can be prepared from any cell type, but many factors contribute to the protein production efficiency. The most common *in vitro* reactions are based on extracts made from *E. coli*, wheat germ or rabbit reticulocyte lysates (Tab. 10.5). *E. coli* extracts consist of the so-called S30 supernatant fraction, named after the soluble fraction when centrifuged at 30,000 *g*, containing endogenous ribosomes, enzymes like acetate kinase and factors necessary for transcription and translation, ARSases, tRNA and mRNA. Endogenous mRNA is removed from the ribosomes during preincubation of the crude cell extract in a "run-off" step and destroyed by endogenous ribonucleases [79]. Another way to deploy bacterial extracts is described in the Gold–Schweiger system [80, 81]. Ribosomes are added to the super-

natant fraction of a S100 extract especially purified from endogenous amino acids and nucleic acids by ion-exchange chromatography. This system provides a very low background due to endogenous synthesis and better-controlled conditions at the expense of more complicated preparation.

The *E. coli* system functions well in a temperature range of 24–38°C with an optimum at 37°C [79, 82, 83]. Wheat germ extract possesses a low level of endogenously expressed messengers and therefore can be directly used for expression of endogenous [84] or exogenous templates. In the wheat germ system, the optimum temperature is in the range 20–27°C [85, 86], but can be increased to up to 32°C for higher expression of some templates [87]. Reticulocyte extract is prepared by directly lysing blood cells of anemic rabbits; this increases the number of proerythrocytes or reticulocytes that are subsequently treated with micrococcal Ca^{2+}-dependent RNase to remove endogenous mRNA [88]. This system works in a temperature range of 30–38°C [89]. With regard to the reaction temperature, it should be noted that apart from any effect on the enzymatic process of transcription/translation and mRNA degradation, the temperature affects the folding of the synthesized protein. To assist the folding of proteins, chaperones like the GroEL/ES system [90] or DnaK and DnaJ [69] can be added to the reaction mixture. Usually, lower temperatures will lead to higher yields of recombinant protein in the absence of chaperones, but not necessarily in their presence [91]. Furthermore, to assist disulfide bond formation, PDI can be added to the transcription/translation mixture [69].

Less common translation systems based on yeast extracts [92], mammalian cells like human HeLa and mouse L-cells [93] or on tobacco chloroplasts, which strongly depend on exogenously added mRNA [94], have high levels of degradation and relatively low protein yields. With regard to the cell-free labeling of proteins, only *E. coli* and wheat germ extracts have been used. Using a very elaborate approach, Shimizu et al. developed a cell-free translation system reconstructed from purified poly(His)-tagged translation factors [95]. Their system, termed the ''protein synthesis using recombinant elements'' (PURE) system, contains 32 individually purified components with high specific activity, allowing efficient protein production. An advantage of the PURE system, apart from the absence of inhibitory substances such as nucleases, proteases and enzymes that hydrolyze nucleoside triphosphates, is the simple purification of the synthesized protein by removing tagged protein factors by affinity chromatography.

One limitation of the cell-free system is the degradation of exogenous added mRNA. Various RNase activities present in cell extracts usually restrict the lifetime of mRNA, and subsequently the efficiency of protein synthesis is inhibited. This problem could be solved by the periodical reintroduction of messengers into the reaction mixture in a simple translation system [96]. Coupled transcription–translation systems, where mRNA is continuously synthesized from DNA templates added to the reaction mixture, can be advantageous to translation systems containing presynthesized messengers. Direct transcription in the reaction mixture may be executed from appropriate promoters by endogenous *E. coli* RNA polymerase, or by exogenous phage RNA polymerase in bacterial systems [97] or in eu-

karyotic systems [82, 98]. To avoid rapid messenger degradation, especially in *E. coli* cell-free systems, partially ribonuclease-depleted extracts or RNase inhibitors are used. A good choice of template DNA in the prokaryotic system is circular plasmid DNA. In the wheat germ system with lower nuclease activities, both plasmid DNAs [99] and linear PCR fragments function well [100–102]. The translational efficiency of mRNA depends on its structural features. Most cDNA sequences can be sufficiently well expressed without addition of translational enhancers. The sequence of interest only needs to be provided with a favorable Kozak sequence in eukaryotic cell-free systems and the Shine–Dalgarno sequence in prokaryotic cell-free systems, which are the respective sequences upstream of the ATG codon responsible for translation initiation.

A further problem in cell-free protein synthesis is the high consumption of biochemical energy provided by ATP and GTP. Creatine phosphate concomitant with creatine kinase is usually used for ATP and GTP regeneration in eukaryotic cell-free systems, whereas the combination of PEP and pyruvate kinase, acetyl phosphate and acetate kinase or a combination of both has been applied for bacterial *in vitro* protein synthesis. The acetyl phosphate energy system may have the advantage that the ATP level is maintained twice as long as in the presence of PEP. Since acetate kinase is present at sufficient levels in bacterial extracts, it does not need to be added exogenously in the *E. coli* system [103]. Studies of the biochemical energy levels in different cell-free systems observed a high rate of triphosphate hydrolysis to mono- and diphosphates during protein synthesis in wheat germ extracts [104] as well as in an *E. coli* S30 extract [105]. It is reported that more than 80% of ATP and GTP hydrolysis in the wheat germ system initially occurs independently of protein synthesis and it was suggested that acid phosphatases are responsible for the nonspecific hydrolysis of the nucleotide triphosphates [86, 104].

Recent extensive studies to increase the protein yield of cell-free reactions have focused on the composition of the reaction mixture, especially the amino acids composition [106, 107], and methods to prepare the cell extract. For example, endogenous phosphatases have been removed by using S30 extract prepared from the spheroplasts of *E. coli* [108] or by immunodepletion of the phosphatase [109]. Approaches to concentrate the extract components by ultrafiltration [110] or by dialysis [74] have also been reported.

10.3.2 Cell-free Expression Techniques

In cell-free reactions carried out in a ''batch'' mode, the reaction conditions change as a result of substrate consumption and the accumulation of products. Translation stops as soon as any essential substrate is exhausted, or any product or by-product reaches an inhibiting concentration. Actually, the bacterial cell-free systems are active only for 10–30 min at 37°C. Systems based on rabbit reticulocyte lysates or wheat germ extract are typically capable of working for up to 1 h. However, *in vitro* protein-synthesizing systems in batch mode work well for most analytical purposes, but short lifetimes and low productivities limit their application for the

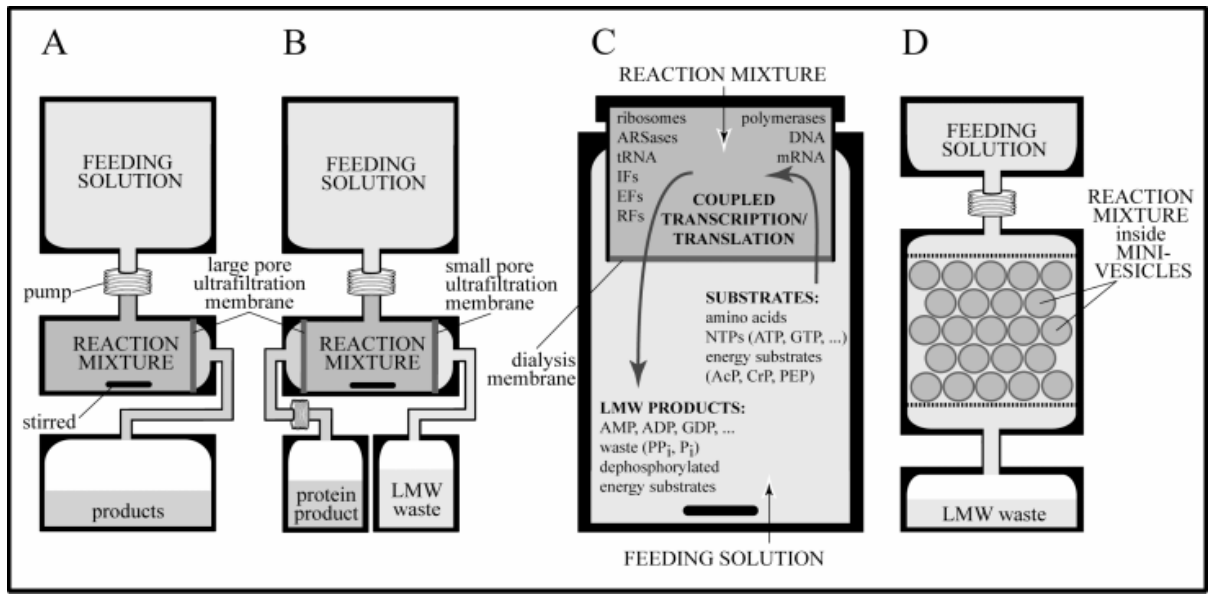


Fig. 10.2 Illustration of cell-free reactors for continuous flow (A and B) and continuous-exchange (C and D) cell-free expression systems. (A) Schematic drawing of a direct-flow CFCF reactor where substrates are supplied and products are removed by the flow of a feeding solution, forced by a pump. (B) Y-flow CFCF reactor, containing two ultrafiltration membranes of different pore size, separating protein product and low-molecular-weight waste outflows [123]. (C) CECF reactor design with explanation of feeding solution (light grey) and reaction mixture components (dark grey). Compartments are separated by a semipermeable dialysis membrane, providing the diffusion exchange of substrates and low-molecular-weight products, and the retention of reaction mixture components. (D) Flow-exchange column reactor consisting of semipermeable mini-vesicles packed in a column, containing the reaction mixture surrounded by feeding solution, changed by flow through the column. Low-molecular-weight products and substrates are changed by diffusion through the mini-vesicle membrane [113]. Modified after Shirokov *et al.* [134].

synthesis of preparative amounts of protein. The reasons for the low yield are degradation of mRNA, depletion of nucleotide triphosphates and accumulation of their hydrolyzates. Prolonged reaction times in cell-free expression were first achieved by Spirin et al. [111–113] by using a continuous-flow cell-free (CFCF) translation device (Fig. 10.2A). The basic idea is to continuously supply amino acids, energy-regenerating components (AcP, CrP or PEP) and NTPs in a feeding solution, and continuously remove small molecule byproducts (mainly products of triphosphate hydrolysis like inorganic phosphates and nucleoside monophosphates) by active (forced) flow of the feeding solution across an ultrafiltration membrane (molecular weight cut-off in the range of 10–300 kDa). In this case, all products, including the protein synthesized, are continuously removed from the reaction compartment if the pore size is large enough (Fig. 10.2A). Permanent stirring of the reaction mixture and in some set-ups upright flow of the feeding solution are applied to minimize membrane clogging [114]. The CFCF system can function for more than 20 hours and results in preparative protein expression of about 0.1–1 mg protein/ml reaction volume or higher. The template for this system can either be mRNA [113, 115], DNA transcribed by endogenous bacterial RNA polymerase or added phage RNA polymerase [82, 113, 114], or self-replicating RNA [116]. Using the CFCF system, proteins like bacteriophage MS2 coat protein [111], brome mosaic virus coat protein [111], calcitonin polypeptide [113], globin [117], functionally active dihydrofolate reductase (DHFR) [112, 113, 118, 119], chloramphenicol acetyltransferase (CAT) [105, 113, 114, 116, 120, 121] interleukin (IL)-2 [122] and IL-6 [115] have successfully been synthesized.

The important advantage of the CFCF system is the continuous removal of synthesized proteins from the reaction mixture, which can result in 80–85% purity of the protein product as previously demonstrated in the outflow of a bacterial CFCF system synthesizing DHFR [119] and in a wheat germ CFCF system synthesizing IL-6 [115]. The synthesized protein diffuses out of the CFCF reactor in a large volume of the effluent and is quite diluted. To reduce the dilution effect, the so-called Y-flow reactor with a split outflow has been proposed [123]. The Y-flow reactor (Fig. 2B) has two membranes with different pore sizes. Initially the low-molecular-weight products are removed through a small-pore membrane at a high rate and the synthesized protein is subsequently collected through a large-pore membrane at a low rate. Here, the flow of the protein product is controlled by a separate pump. Nevertheless, a number of laboratories attempting CFCF expression have had difficulties in establishing this complex system. The main problems are RNA degradation when using bacterial extracts, even in the coupled transcription/translation mode, and low efficiency of initiation complex formation, which might cause leakage and therefore loss of some translation components by ultrafiltration or the blockage of the ultrafiltration membrane [91].

However, there is an alternative way of carrying out prolonged protein synthesis, namely by using diffusion instead of pumping to supply substrates and remove low-molecular-weight products (Fig. 2C). This set-up, using a reaction mixture separated from a feeding solution by applying a dialysis membrane, is called a continuous-exchange cell-free system (CECF) [124] or a semicontinuous-flow cell-

free system (SFCF) [121]. The simplest device for the continuous supply of substrates and the removal of low-molecular-weight products by passive exchange with a feeding mixture is a dialysis bag [124]. In practice, simple homemade dialysis bags or standard commercial dialyzers, such as the MicroDialyzer® and DispoDialyzer® from Spectrum, can be used successfully. The pore size of the reactor membranes is usually in the range of 10–50 kDa. However, better performance of reactors with a larger pore membrane has been reported [74]. Stirring of either the feeding solution or both, the feeding and reaction mixture, is necessary for efficient supply of substrates.

Using this approach, Kim and Choi [121] reported the synthesis of 1.2 mg/ml of CAT over 14 h in *E. coli* S30 extract. The product yield, quantified by ELISA, exceeded the yield of the analogous batch reaction by 10–12 times. More recently, Kigawa et al. [74] succeeded in synthesizing CAT and Ras proteins in amounts up to 6 mg/ml over 18 h in their version of the bacterial CECF system, and Madin *et al.* [96] reached yields of 1–4 mg/ml for several functionally active proteins like DHFR, green fluorescent protein (GFP), luciferase and RNA replicase of tobacco mosaic virus (TMV). Figure 10.3 describes the cell-free synthesis of GFP in a CECF method using a MicroDialyzer®.

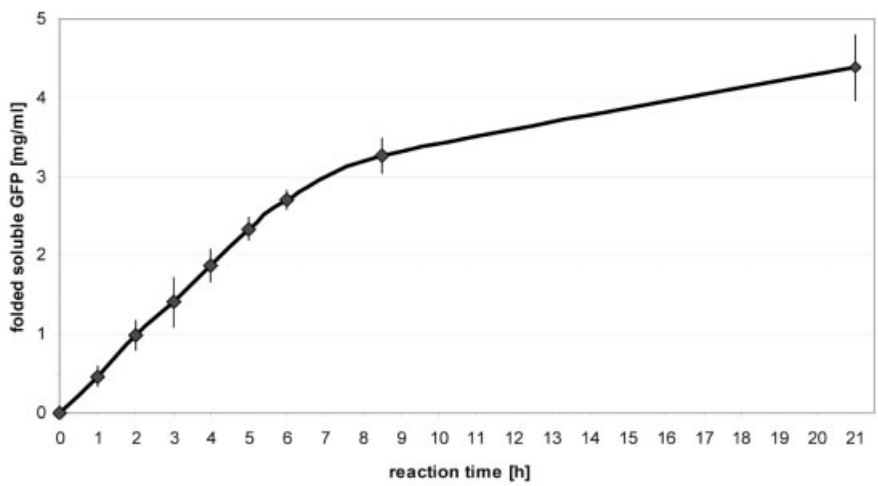


Fig. 10.3 Cell-free synthesis of GFP in a bacterial CECF system using a 100-µl MicroDialyzer®. The kinetic points display the mean of three reactions and error bars indicating the deviation of the mean. The reaction was carried out at 30°C using a 17-fold amount of feeding mixture compared to the reaction mixture volume and under the following reaction conditions: 0.05% NaN_3, 2% phosphoenolglycol, 196 µM folinic acid, 2 mM 1.4-dithiothreitol, 1.2 mM ATP, 0.8 mM each of CTP, GTP and UTP, 20 mM PEP, 20 mM acetyl phosphate, 1 mM of each amino acid, except the amino acids arginine, aspartic acid, cysteine, glutamic acid, methionine and tryptophan, for which 2 mM was used, 100 mM HEPES–KOH (pH 8.0), 2.8 mM EDTA, 1× complete protease inhibitor (Roche), 280 mM K^+, 13 mM Mg^{2+}, 40 µg/ml pyruvate kinase, 500 µg/ml tRNA (*E. coli*), 3 U/µl T7-RNA polymerase, 0.3 U/µl RNasin, 35% S30 extract (*E. coli*) and 15 µg/ml plasmid.

The advantage of the CECF system is the accumulation of the synthesized product in the reaction mixture. The synthesis of proteins and polypeptides fused with GFP provide a direct and demonstrative way to visualize product accumulation by fluorescence of the GFP moiety. The synthesis of a HIV protein, the so-called Nef antigen, fused with GFP [125] and an antibacterial polypeptide Cecropin P1 fused with GFP [126], both in the bacterial CECF T7 transcription/translation system, have been reported.

Reactors combining both exchange and flow have also been developed. In one version, the reaction mixture is encapsulated into polysaccharide minivesicles that can be packed into a column, where feeding solution is passed through the column (Fig. 10.2D). In this case, product–substrate exchange across the vesicle walls takes place during the flow [113]. More sophisticated versions of the CECF reactor are being developed in order to meet the demands of scientists and biotechnologists. The first commercial CECF reactor has recently been launched on the market by Roche Diagnostics. The Roche CECF system promises a high protein yield of more than 2 mg/ml. Other commercial systems focus on the batch mode, where recent improvements have been made. For example, a novel NTP regeneration system, avoiding accumulation of inorganic phosphate by adding pyruvate oxidase, which generates AcP from pyruvate and inorganic phosphate directly in the reaction mixture, has been proposed by Kim and Swartz [127]. Later they showed that addition of oxalate, a potent inhibitor of PEP synthetase, substantially increases the yield of CAT synthesis through the enhanced supply of ATP by about 47% [128]. Furthermore, Kim and Swartz developed a ''fed-batch'' mode where coordinated addition of PEP, magnesium, and the amino acids arginine, cysteine and tryptophan resulted in a final concentration of cell-free synthesized CAT that was more than 4-fold compared to a batch reaction [107]. As a result of these improvements it became possible to synthesize about 350 μg/ml CAT [107] or 450 μg/ml recombinant DNA human protein thrombopoietin [106] in a batch reaction.

10.3.3
Specific Applications of the Cell-free Labeling Technique

Conventional *in vivo* labeling techniques are often accompanied by low protein yields due to retarded growth in a minimal medium and, in the case of selective isotope labeling, by scrambling effects that drastically reduce the efficiency and selectivity of labeling. A major advantage of cell-free labeling techniques therefore is the absence of any scrambling effects or metabolic conversion of labeled amino acids [71, 73, 129]. The selective incorporation of isotope-labeled amino acids in cell-free synthesized proteins for NMR research has already been demonstrated by several groups [71, 73, 74, 129–131]. Similarly, in the reverse isotope-labeling technique, most amino acids are single or dual (^{15}N, ^{13}C) labeled, except for a few residues [74]. Figure 10.4(A) illustrates the 2-D ^{1}H–^{15}N heteronuclear single-quantum correlation spectroscopy (HSQC) NMR spectra of the cell-free reverse isotope-labeled C-terminal DNA-binding domain of the transcriptional response regulator RcsB (cRcsB) which has been uniformly ^{15}N-labeled except for aspar-

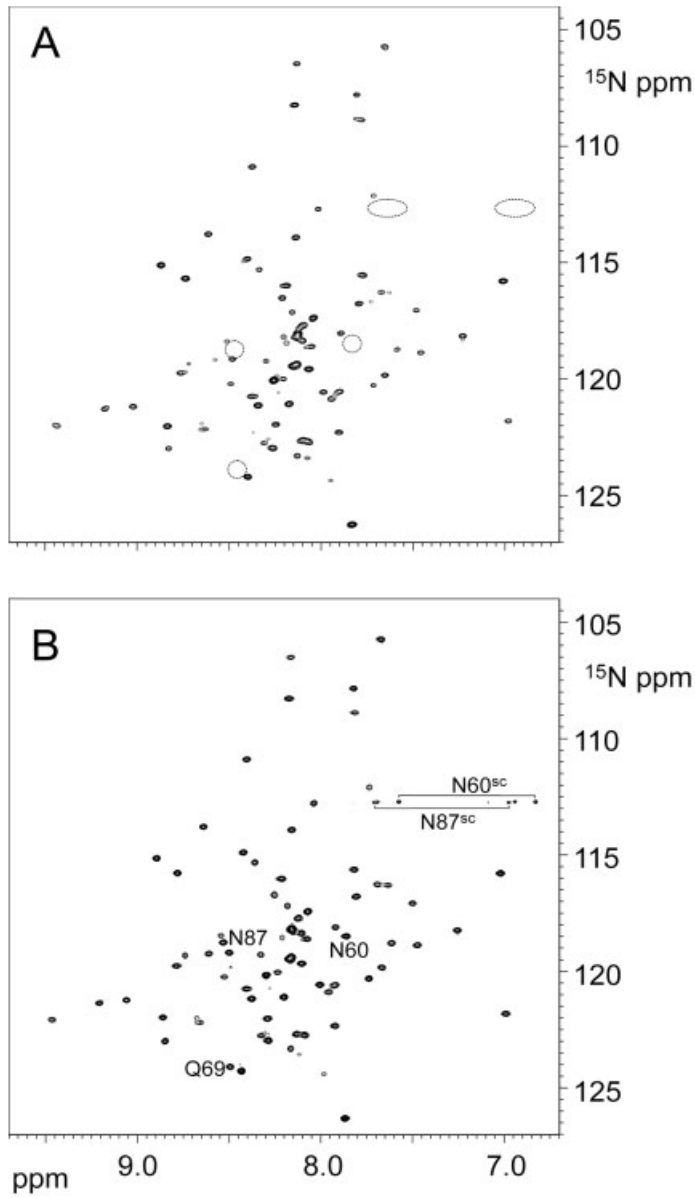


Fig. 10.4 Comparison of the 2-D ^{1}H–^{15}N HSQC spectra of the *in vitro* and *in vivo* generated purified C-terminal domain of the bacterial transcriptional regulator RcsB (cRcsB) using a Bruker DRX-800 MHz NMR spectrometer with a cryoprobe. (A) Cell-free ^{15}N-reverse isotopic-labeled cRcsB synthesized in an optimized CECF system using a DispoDialyzer®. All amino acids except asparagine (N) and glutamine (Q) were ^{15}N-labeled. As expected, the N and Q residues (circles) are absent in the 2-D ^{1}H–^{15}N HSQC spectrum. (B) Uniformly ^{15}N isotopic-labeled cRcsB expressed *in vivo*. This spectrum is in good agreement with that of the reverse cell-free isotopically labeled cRcsB, and N and Q residues can easily be identified. (NMR spectra kindly provided by Frank Löhr.)

agine (N) and glutamine (Q) residues. The uniformly ^{15}N-labeled cRcsB expressed *in vivo* is shown in Fig. 10.4(B). The 2-D ^{1}H–^{15}N HSQC spectrum of the labeled cRcsB protein synthesized *in vitro* is consistent with that of the uniformly labeled protein synthesized *in vivo* (only the asparagine and glutamine peaks are missing).

Cell-free expression is highly suited for the generation of protein samples labeled only at distinct residues. Site-specific labeling is extremely useful to simplify the observation and resonance assignment procedures for specified amino acid residues of particular interest, and can also be used for analyzing local structures of large proteins and protein–protein interactions. Ellman et al. demonstrated for the first time the incorporation of a particular ^{13}C-labeled residue at a suppressible termination codon by translating the modified sequence in an *in vitro* system supplemented with the charged suppressor tRNA. Subsequently it became possible to track the labeled residues upon denaturation and refolding of the protein by NMR spectroscopy [130]. Milligram quantities of site-specific isotope labeled protein can be obtained in a cell-free system involving the amber suppression strategy [73]. The *E. coli* amber suppressor tRNA can be prepared by *in vitro* transcription with T7 RNA polymerase and later aminoacylated with the appropriate purified *E. coli* amino ARSase and the cognate labeled amino acid. The codon for the selected amino acid residue in the protein was previously changed into an amber codon by standard techniques.

Segmental labeling *in vivo* is limited by specific requirements like organization of the target protein into domains, presence of specific residues and folding problems. Pavlov et al. suggested a cell-free technique based on *in vitro* translation of matrix-coupled mRNAs, which principally is devoid of any sequence and conformational requirements [72]. The size of the labeled region is controlled by codon usage and no intrinsic upper limit to the size of proteins that can be isotope-labeled in selected regions exists. This method is based on the usage of translation mixtures depleted of either one amino acid and/or its tRNA and/or its amino ARSase and consists of three steps. Using column-coupled template RNA the reaction mixture can easily be exchanged. Initially, the unlabeled N-terminal region, using an unlabeled reaction mixture, is synthesized up to the first codon without a matching amino acyl-tRNA in the extract. Here the ribosomes pause and the translation mixture can be exchanged against a mix containing isotope-labeled amino acids, now deficient for a different amino acid, tRNA or amino ARSase. Translation resumes, thereby labeling the region until the ribosomes encounter the next codon without the corresponding tRNA. In the last step, the C-terminal part is synthesized without any label and the protein is released from the ribosome. However, the technique results in low protein yields, as it can, at the very best, produce protein stoichiometric to the immobilized mRNA.

A very promising advantage of cell-free isotope labeling is the possibility of *in situ* NMR measurement as described by Guignard et al. They showed NMR analysis of *in vitro*-synthesized proteins without any chromatographic purification and with minimal sample handling using an optimized CECF reaction combined with the sensitivity of a cryoprobe [131]. As expected, they observed no cross peak for any excess ^{15}N amino acid and only the target protein was enriched with the

isotope-labeled residues. They suggest a new possibility for inexpensive high-throughput protein analysis applicable in large-scale proteomics, where selectively labeled proteins can be expressed in 0.5 ml of reaction medium using small quantities of labeled amino acids and analyzed by NMR. All steps from the expression to the completed NMR spectra were done in less than 24 h.

Due to the exceptional advantages of cell-free protein synthesis, isotope labeling can be achieved for proteins that are normally difficult to express. Using cell-free labeling techniques, it might become possible to synthesize and isotope label proteins that are toxic, require cofactors or chaperones for adopting an active conformation, or even membrane proteins. Subunit labeling of oligomers, further research on disulfide bridge formation or protein oxidation should be possible in the near future.

References

1 Dötsch V, Wagner G. 1998. New approaches to structure determination by NMR spectroscopy. *Curr Opin Struct Biol* **8**, 619–23.

2 Pervushin K, Riek R, Wider G, Wüthrich K. 1997. Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci USA* **94**, 12366–71.

3 Tjandra N, Bax A. 1997. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **278**, 1111–4.

4 Clore GM, Gronenborn AM. 1991. Structures of larger proteins in solution: three- and four-dimensional heteronuclear NMR spectroscopy. *Science* **252**, 1390–9.

5 Bax A. 1994. Multidimensional nuclear-magnetic-resonance methods for protein studies. *Curr Opin Struct Biol* **4**, 738–44.

6 Clore GM, Gronenborn AM. 1997. NMR structures of proteins and protein complexes beyond 20,000 M_r. *Nat Struct Biol* **4**, 849–53.

7 Kay LE, Gardner KH. 1997. Solution NMR spectroscopy beyond 25 kDa. *Curr Opin Struct Biol* **7**, 722–31.

8 Gardner KH, Kay LE. 1998. The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct* **27**, 357–406.

9 Wider G, Wüthrich K. 1999. NMR spectroscopy of large molecules and multimolecular assemblies in solution. *Curr Opin Struct Biol* **9**, 594–601.

10 Hansen AP, Petros AM, Mazar AP, Pederson TM, Rueter A, Fesik SW. 1992. A practical method for uniform isotopic labeling of recombinant proteins in mammalian cells. *Biochemistry* **31**, 12713–8.

11 Sailer M, Helms GL, Henkel T, Niemczura WP, Stiles ME, Vederas JC. 1993. ^{15}N- and ^{13}C-labeled media from *Anabaena* sp. for universal isotopic labeling of bacteriocins: NMR resonance assignments of leucocin A from *Leuconostoc gelidum* and Nisin A from *Lactococcus lactis*. *Biochemistry* **32**, 310–8.

12 Reilly D, Fairbrother WJ. 1994. A novel isotope labeling protocol for bacterially expressed proteins. *J Biomol NMR* **4**, 459–62.

13 Brandner G, Phalip V, Kellenberger E, Zhang CC. 1999. Optimization of a cyanobacterial culture for production of isotope-labeled recombinant proteins. *Biotechnol Tech* **13**, 177–80.

14 Cai M, Huang Y, Sakaguchi K, Cloe GM, Gronenborn AM, Craigie R. 1998. An efficient and cost-effective isotope labeling protocol for proteins

expressed in *Escherichia coli*. *J Biomol NMR* **11**, 97–102.

15 Marley J, Lu M, Bracken C. 2001. A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* **20**, 71–5.

16 Muchmore DC, McIntosh LP, Russel CB, Anderson DE, Dahlquist FW. 1989. Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods Enzymol* **177**, 44–73.

17 Ikura M, Kay LE, Bax A. 1990. A novel approach for sequential assignment of ^{1}H, ^{13}C, and ^{15}N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry* **29**, 4659–67.

18 Venters RA, Calderone TL, Spicer LD, Fierke CA. 1991. Uniform ^{13}C isotope labeling of proteins with sodium acetate for NMR studies: application to human carbonic anhydrase II. *Biochemistry* **30**, 4491–4.

19 Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd edn, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

20 Studier FW, Rosenberg AH, Dunn JJ, Dubendorf JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**, 60–89.

21 Nilsson J, Stahl S, Lundeberg J, Uhlen M, Nygren PA. 1997. Affinity fusion strategies for detection, purification and immobilization of recombinant proteins. *Protein Express Purif* **11**, 1–16.

22 Christendat D, Yee A, Dharamsi A, Kluger Y, Savchenko A, Cort JR, Booth V, Mackereth CD, Saridakis V, Ekiel I, Kozlov G, Maxwell KL, Wu N, McIntosh LP, Gehring K, Kennedy MA, Davidson AR, Pai EF, Gerstein M, Edwards AM, Arrowsmith CH. 2000. Structural proteomics of an archaeon. *Nat Struct Biol* **7**, 903–9.

23 Majerle A, Kidric J, Jerala R. 2000. Production of stable isotope enriched antimicrobial peptides in *Escherichia coli*: an application to the production of a ^{15}N-enriched fragment of lactoferrin. *J Biomol NMR* **18**, 145–51.

24 De Bernardez CE. 1998. Refolding of recombinant proteins. *Curr Opin Biotechnol* **9**, 157–63.

25 Lilie H, Scharz E, Rudolph R. 1998. Advances in refolding of proteins produced in *E coli*. *Curr Opin Biotechnol* **9**, 497–501.

26 Kapust RB, Waugh DS. 1999. *Escherichia coli* maltose binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8**, 1668–74.

27 Davis, DD, Elisee C, Newham DM, Harrison RG. 1999. New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnol Bioeng* **65**, 382–8.

28 Forrer P, Jaussi R. 1998. High-level expression of soluble heterologous proteins in the cytoplasm of *Escherichia coli* by fusion to the bacteriophage Lambda head protein D. *Gene* **224**, 45–52.

29 Jansson M, Li YC, Jendeberg L, Anderson S, Montelione BT, Nilsson B. 1996. High-level production of uniformly ^{15}N- and ^{13}C-enriched fusion proteins in *Escherichia coli*. *J Biomol NMR* **7**, 131–41.

30 Zhou P, Lugovskoy AA, Wagner G. 2001. A solubility-enhancement tag (SET) for NMR studies of poorly behaving proteins. *J Biomol NMR* **20**, 11–4.

31 Bessette PH, Aslund F, Beckwith J, Georgiou G. 1999. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc Natl Acad Sci USA* **96**, 13703–8.

32 De Lamotte F, Boze H, Blanchard C, Klein C, Moulin G, Gautier MF, Delsuc MA. 2001. NMR onitoring of accumulation and folding of ^{15}N-labeled protein overexpressed in *Pichia pastoris*. *Protein Express Purif* **22**, 318–24.

33 Wyss DF, Wagner G. 1996. The structural role of sugars in glycoproteins. *Curr Opin Biotechnol* **7**, 409–16.

34 Wood MJ, Komives EA. 1999. Production of large quantities of isotopically labeled protein in *Pichia pastoris* by fermentation. *J Biomol NMR* **13**, 149–59.

35 Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev* **24**, 45–66.

36 Laroche Y, Storme V, De Meutter J, Messens J, Lauwereys M. 1994. High-level secretion and very efficient isotopic labeling of tick anticoagulant peptide (TAP) expressed in the methylotrophic yeast *Pichia pastoris*. *Biotechnology* **12**, 1119–24.

37 White CE, Hunter MJ, Meininger DP, White LR, Komives EA. 1995. Large-scale expression, purification and characterization of small fragments of thrombomodulin: the roles of the sixth domain and of methionine 388. *Protein Eng* **8**, 1177–87.

38 Van den Burg HA, de Wit PJGM, Vervoort J. 2001. Efficient $^{13}C/^{15}N$ double labeling of the avirulence protein AVR4 in a methanol-utilizing strain (Mut^{+}. of *Pichia pastoris*. *J Biomol NMR* **20**, 251–61.

39 Wiles AP, Shaw G, Bright J, Perczel A, Campell ID, Barlow PN. 1997. NMR studies of a viral protein that mimics the regulation of complement activation. *J Mol Biol* **272**, 253–65.

40 Hensley P, McDecitt PJ, Brooks I, Trill JJ, Feild JA, McNulty DE, Connor JR, Griswold DE, Kumar NV, Kopple KD. 1994. The soluble form of E-selectin is an asymmetric monomer. Expression, purification, and characterization of the recombinant protein. *J Biol Chem* **269**, 23949–58.

41 Lustbader JW, Birken S, Pollak S, Pound A, Chait BT, Mirza UA, Ramnarain S, Canfield RE, Brown JM. 1996. Expression of human chorionic gonadotropin uniformly labeled with NMR isotopes in Chinese hamster ovary cells: an advance toward rapid determination of glycoprotein structures. *J Biomol NMR* **7**, 295–304.

42 Archer SJ, Bax A, Roberts AB, Sporn MB, Ogawa Y, Piez KA, Weatherbee JA, Tsang ML, Lucas R, Zheng B, Wenker J, Torchia DA. 1993. Transforming growth factor β1, NMR signal assignments of the recombinant protein expressed and isotopically enriched using Chinese hamster ovary cells. *Biochemistry* **32**, 1152–63.

43 Wyss DF, Dayie KT, Wagner G. 1997. The counterreceptor binding site of human CD2 exhibits an extended surface patch with multiple conformations fluctuating with millisecond to microsecond motions. *Protein Sci* **6**, 534–42.

44 Cubbedu L, Moss CX, Swarbrick JD, Gooley AA, Williams KL, Curmi PMG, Slade MB, Mabbutt BC. 2000. *Dictyostelium discoideum* as expression host: isotopic labeling of a recombinant glycoprotein for NMR studies. *Protein Express Purif* **19**, 335–42.

45 Desplancq D, Kieffer B, Schmidt K, Posten C, Forster A, Oudet P, Strub JM, Van Dorsselaer A, Weiss E. 2001. Cost-effective and uniform ^{13}C- and ^{15}N-labeling of the 24-kDa N-terminal domain of the *Escherichia coli* gyrase B by overexpression in the photoautotrophic cyanobacterium *Anabaena* sp. PCC 7120. *Protein Express Purif* **23**, 207–17.

46 Senn H, Weber C, Kobel H, Traber R. 1991. Selective ^{13}C-labelling of cyclosporin A. *Eur J Biochem* **199**, 653–8.

47 Yee AA, O'Neil JD. 1992. Uniform ^{15}N labeling of a fungal peptide: the structure and dynamics of an alamethicin by ^{15}N and ^{1}H NMR spectroscopy. *Biochemistry* **31**, 3135–43.

48 Ikeda Y, Naganawa H, Kondo S, Takeuchi T. 1992. Biosynthesis of bellenamine by *Streptomyces nashvillensis* using stable isotope labeled compounds. *J Antibiot (Tokyo)* **45**, 1919–24.

49 Suarez C, Kohler SJ, Allen MM, Kolodny NH. 1999. NMR study of the metabolic ^{15}N isotopic enrichment of cyanophycin synthesized by the cyano-

bacterium *Synechocystis* sp. strain PCC 6308. *Biochim Biophys Acta* **1426**, 429–38.

50 Biekofsky RR, Martin SR, McCormick JE, Masino L, Fefeu S, Bayley PM, Feeney J. 2002. Thermal stability of calmodulin and mutants studied by $^{1}H-^{15}N$ HSQC NMR measurements of selectively labeled [^{15}N]Ile proteins. *Biochemistry* **41**, 6850–9.

51 Vuister GW, Kim SJ, Wu C, Bax A. 1994. 2D and 3D NMR-study of phenylalanine residues in proteins by reverse isotope labeling. *J Am Chem Soc* **116**, 9206–10.

52 Goto NK, Kay LE. 2000. New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Curr Opin Struct Biol* **10**, 585–92.

53 Waugh DS. 1996. Genetic tools for selective labeling of proteins with alpha-^{15}N-amino acids. *J Biomol NMR* **8**, 184–92.

54 Wang H, Janowick DA, Schkeryantz JM, Liu X, Fesik SW. 1999. A method for assigning phenylalanines in proteins. *J Am Chem Soc* **121**, 1611–2.

55 Coughlin PE, Anderson FE, Oliver EJ, Brown JM, Homans SW, Pollak S, Lustbader JW. 1999. Improved resolution and sensitivity of triple-resonance NMR methods for the structural analysis of proteins by use of a backbone-labeling strategy. *J Am Chem Soc* **121**, 11871–4.

56 Atreya HS, Chary KVR. 2001. Selective ''unlabeling'' of amino acids in fractionally ^{13}C labeled proteins: an approach for stereospecific NMR assignments of CH_3 groups in Val and Leu residues. *J Biomol NMR* **19**, 267–72.

57 LeMaster DM, Kushlan DM. 1996. Dynamical mapping of *E. coli* thioredoxin via ^{13}C NMR relaxation analysis. *J Am Chem Soc* **118**, 9255–64.

58 Perler FB. 1998. Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. *Cell* **92**, 1–4.

59 Yamazaki T, Otomo T, Oda N, Kyogoku Y, Uegaki K, Ito N, Ishino Y, Nakamura H. 1998. Segmental isotope labeling for protein NMR using peptide splicing. *J Am Chem Soc* **120**, 5591–2.

60 Otomo T, Teruya K, Uegaki K, Yamazaki T, Kyogoku Y. 1999. Improved segmental isotope labeling of proteins and application to a larger protein. *J Biomol NMR* **14**, 105–14.

61 Otomo T, Ito N, Kyogoku Y, Yamazaki T. 1999. NMR observation of selected segments in a larger protein: central-segment isotope labeling through intein-mediated ligation. *Biochemistry* **38**, 16040–4.

62 Xu R, Ayers B, Cowburn D, Muir TW. 1999. Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. *Proc Natl Acad Sci USA* **96**, 388–93.

63 Wu H, Hu Z, Liu XQ. 1998. Protein *trans*-splicing by a split intein encoded in a split DnaE gene of *Synechocystis* sp. PCC6803. *Proc Natl Acad Sci USA* **95**, 9226–31.

64 Evans TC, Martin D, Kolly R, Panne D, Sun L, Ghosh I, Chen L, Benner J, Liu XQ, Xu MQ. 2000. Protein *trans*-splicing and cyclization by a naturally split intein from the *dnaE* gene of *Synechocystis* species PCC6803. *J Biol Chem* **275**, 9091–4.

65 Blum P, Ory J, Bauernfeind J, Krska J. 1994. Physiological consequences of DnaK and DnaJ overproduction in *Escherichia coli*. *J Bacteriol* **174**, 7436–44.

66 Stiege W, Erdmann VA. 1995. The potentials of the *in vitro* biosynthesis system. *J Biotechnol* **41**, 81–90.

67 Falk MM, Buehler LK, Kumar NM, Gilula NB. 1997. Cell-free synthesis and assembly of connexins into functional gap junction membrane channels. *EMBO J* **16**, 2703–16.

68 Huppa JB, Ploegh HL. 1997. *In vitro* translation and assembly of a complete T cell receptor-CD3 complex. *J Exp Med* **186**, 393–403.

69 Ryabova LA, Desplancq D, Spirin AS, Pluckthun A. 1997. Functional antibody production using cell-free translation: effects of protein disulfide

isomerase and chaperones. *Nat Biotechnol* **15**, 79–84.

70 Molla A, Paul AV, Wimmer E. 1991. Cell-free, *de novo* synthesis of poliovirus. *Science* **254**, 1647–51.

71 Kigawa T, Muto Y, Yokoyama S. 1995. Cell-free synthesis and amino acid-selective stable isotope labelling of proteins for NMR analysis. *Biomol NMR* **2**, 129–34.

72 Pavlov MY, Freistroffer DV, Ehrenberg M. 1997. Synthesis of region-labelled proteins for NMR studies by *in vitro* translation of column-coupled mRNAs. *Biochimie* **79**, 415–22.

73 Yabuki T, Kigawa T, Dohmae N, Takio K, Terada T, Ito Y, Laue ED, Cooper JA, Kainosho M, Yokoyama S. 1998. Dual amino acid-selective and site-directed stable-isotope labelling of the human c-Ha-Ras protein by cell-free synthesis. *J Biomol NMR* **11**, 295–306.

74 Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T, Yokoyama S. 1999. Cell-free production and stable-isotope labelling of milligram quantities of proteins. *FEBS Lett* **442**, 15–9.

75 Joseph SK, Boehning D, Pierson S, Nicchitta CV. 1997. Membrane insertion, glycosylation, and oligomerization of inositol triphosphate receptors in a cell-free translation system. *J Biol Chem* **272**, 1579–588.

76 Duszenko M, Kang X, Bohne U, Homke R, Lehner M. 1999. *In vitro* translation in a cell-free system from *Trypanosoma brucei* yields glycosylated and glycosylphosphatidyl-inositol-anchored proteins. *Eur J Biochem* **266**, 789–97.

77 Park Y, Luo J, Schultz PG, Kirsch JF. 1997. Noncoded amino acid replacement probes of the aspartate aminotransferase mechanism. *Biochemistry* **36**, 10517–25.

78 Ohuchi S, Nakano H, Yamane T. 1998. *In vitro* method for the generation of protein libraries using PCR amplification of a single DNA molecule and coupled transcription/translation. *Nucleic Acids Res* **26**, 4339–346.

79 Zubay G. 1973. *In vitro* synthesis of protein in microbial systems. *Annu Rev Genet* **7**, 267–87.

80 Gold LM, Schweiger M. 1969. Synthesis of phage-specific x- and β-glucosyl transferases directed by T-even DNA *in vitro*. *Proc Natl Acad Sci USA* **62**, 892–8.

81 Gold LM, Schweiger M. 1971. Synthesis of bacteriophage-specific enzymes directed by DNA *in vitro*. *Methods Enzymol* **20**, 537–42.

82 Baranov VI, Spirin AS. 1993. Gene expression in cell-free systems on preparative scale. *Methods Enzymol* **217**, 123–42.

83 Kim DM, Kigawa T, Choi CY, Yokoyama S. 1996. A highly efficient cell-free protein synthesis system from *Escherichia coli*. *Eur J Biochem* **239**, 881–6.

84 Roberts BE, Pateron BM. 1973. Efficient translation of tobacco mosaic virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ. *Proc Natl Acad Sci USA* **70**, 2330–4.

85 Anderson CW, Straus JW, Dudock BS. 1983. Preparation of a cell-free protein-synthesizing system from wheat germ. *Methods Enzymol* **101**, 635–44.

86 Kawarasaki Y, Nakano H, Yamane T. 1994. Prolonged cell-free protein synthesis in a batch system using wheat germ extract. *Biosci Biotechnol Biochem* **58**, 1911–3.

87 Tulin EE, Ken-Ichi T, Shin-Ichiro E. 1995. Continuously coupled transcription-translation system for the production of rice cytoplasmic aldolase. *Biotechnol Bioeng* **45**, 511–6.

88 Pelham HRB, Jackson RJ. 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur J Biochem* **67**, 247–56.

89 Woodward WR, Ivey JL, Herbert E. 1974. Protein synthesis with rabbit reticulocyte preparations. *Methods Enzymol* **30**, 724–31.

90 Tsalkova T, Zardeneta G, Kudlicki W, Kramer G, Korwitz PM, Hardesty B. 1993. GroEL and GroES increase the specific enzymatic activity

of newly-synthesized rhodanese if present during *in vitro* transcription/translation. *Biochemistry* **32**, 3377–80.

91 Jermutus L, Ryabova LA, Pluckthun A. 1998. Recent advantages in producing and selecting functional proteins by using cell-free translation. *Curr Opin Biotechnol* **9**, 534–48.

92 Iizuka N, Sugiura M. 1997. Translation-component extract from Saccharomyces cerevisieae: effects of L-A RNA, 5′ cap, 3′ poly(A) tail on translational efficiency of mRNAs. *Methods* **11**, 353–60.

93 Tolunay HE, Yang L, Kemper WM, Safer B, Anderson WF. 1984. Homologous globin cell-free transcription system with comparison of heterologous factors. *Mol Cell Biol* **4**, 17–22.

94 Hirose T, Sugiura M. 1996. *Cis*-acting element and *trans*-acting factors for accurate translation of chloroplast *psb*A mRNAs: development of an *in vitro* translation system from tobacco chloroplasts. *EMBO J* **15**, 1687–695.

95 Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T. 2001. Cell-free translation reconstituted with purified components. *Nat Biotechnol* **19**, 751–5.

96 Madin K, Sawasaki T, Ogasawara T, Endo Y. 2000. A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. *Proc Natl Acad Sci USA* **97**, 559–64.

97 Chen HZ, Zubay G. 1983. Prokaryotic coupled transcription–translation. *Methods Enzymol* **101**, 674–90.

98 Craig D, Howell MT, Gibbs CL, Hunt T, Jackson RJ. 1992. Plasmid cDNA-directed synthesis in a coupled eukaryotic *in vitro* transcription–translation system. *Nucleic Acids Res* **20**, 4987–995.

99 Nevin DE, Pratt JM. 1991. A coupled *in vitro* transcription–translation system for the exclusive synthesis of polypeptides from the T7 promotor. *FEBS Lett* **291**, 259–63.

100 Lesley SA, Brow MA, Burgess PR. 1991. Use of *in vitro* protein synthesis from polymerase chain reaction-generated templates to study interaction of *Escherichia coli* transcription factors with core RNA polymerase and for epitope mapping of monoclonal antibodies. *J Biol Chem* **266**, 2632–8.

101 Burks EA, Chen G, Georgiou G, Iverson BL. 1997. *In vitro* scanning saturation mutagenesis of an antibody binding pocket. *Proc Natl Acad Sci USA* **94**, 412–7.

102 Martemyanov KA, Spirin AS, Gudkov AT. 1997. Direct expression of PCR products in a cell-free translation/transcription system: synthesis of antibacterial peptide cecropin. *FEBS Lett* **414**, 268–70.

103 Ryabova LA, Vinokurov LM, Shekhovtsova EA, Alakhov YB, Spirin AS. 1995. Acetyl phosphate as an energy source for bacterial cell-free translation systems. *Anal Biochem* **226**, 184–6.

104 Yao SL, Shen XC, Suzuki E. 1997. Biochemical energy consumption by wheat germ extract during cell-free protein synthesis. *J Ferment Bioeng* **84**, 7–13.

105 Kitaoka Y, Nishimura N, Niwano M. 1996. Cooperativity of stabilized mRNA and enhanced translational activity in the cell-free system. *J Biotechnol* **48**, 1–8.

106 Patnaik R, Swartz JR. 1998. *E. coli*-based *in vitro* transcription/translation: *in vivo*-specific synthesis rates and high yields in a batch system. *BioTechniques* **24**, 862–8.

107 Kim DM, Swartz JR. 2000. Prolonging cell-free protein synthesis by selective reagent additions. *Biotech Prog* **16**, 385–90.

108 Kang SH, Oh TJ, Kim RG, Kang TJ, Hwang SH, Lee EY, Choi CY. 2000. An efficient cell-free protein synthesis system using periplasmic phosphatase-removed S30 extract. *J Microbiol Methods* **43**, 91–6.

109 Kawarasaki Y, Nakano H, Yamane Y. 1998. Phosphatase-immunodepleted cell-free protein synthesis system. *J Biotechnol* **61**, 199–208.

110 Nakano H, Tanaka T, Kawarasaki Y, Yamane T. 1994. An increased rate of

cell-free protein synthesis by condensing wheat germ extract with ultrafiltration membranes. *Biosci Biotechnol Biochem* **58**, 631–4.

111 Spirin AS, Baranov VI, Ryabova LA, Ovodov SY, Alakhov YB. 1988. A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* **242**, 1162–4.

112 Baranov VI, Morozov IY, Ortlepp SA, Spirin AS. 1989. Gene expression in a cell-free system on the preparative scale. *Gene* **84**, 463–6.

113 Spirin AS. 1991. Cell-free protein synthesis bioreactor. In: Todd P, Sikdar SK, Beer M, eds, *Frontiers in Bioprocessing II*, 31–43. American Chemical Society, Washington, DC.

114 Kigawa T, Yokoyama S. 1991. A continuous cell-free protein synthesis system for coupled transcription–translation. *J Biochem* **110**, 166–8.

115 Volyanik EV, Dalley A, Mc Kay IA, Leigh I, Williams NS, Bustin SA. 1993. Synthesis of preparative amounts of biologically active interleukin-6 using a continuous-flow cell-free translation system. *Anal Biochem* **214**, 289–94.

116 Ryabova L, Volyanik E, Kurnasov P, Spirin AS, Wu Y, Kramer FR. 1994. Coupled replication–translation of amplifiable messenger RNA: A cell-free protein synthesis system that mimics viral infection. *J Biol Chem* **269**, 1501–5.

117 Ryabova LA, Ortleb SA, Baranov VI. 1989. Preparative synthesis of globin in a continuous cell-free translation system from rabbit reticulocytes. *Nucleic Acids Res* **17**, 4412.

118 Endo Y, Otsuzuki S, Ito K, Miura K. 1992. Production of an enzymatic active protein using a continuous flow cell-free system. *J Biotech* **25**, 221–30.

119 Kudlicki W, Kramer G, Hardesty B. 1992. High efficiency cell-free synthesis of proteins: refinement of the coupled transcription/translation system. *Anal Biochem* **206**, 389–93.

120 Ryabova LA, Morozov IY, Spirin AS. 1998. Continuous-flow cell-free translation, transcription–translation, and replication–translation system. I. *Methods Mol Biol* **77**, 179–93.

121 Kim DM, Choi CY. 1996. A semi-continuous prokaryotic coupled transcription/translation system using a dialysis membrane. *Biotechnol Prog* **12**, 645–9.

122 Kolosov MI, Kolosova IM, Alokhov VY, Ovodov SY, Alokhov YB. 1992. Preparative *in vitro* synthesis of bioactive human interleukin-2 in a continuous flow translation system. *Biotech Appl Biochem* **16**, 125–33.

123 Biryukov SV, Simonenko PN, Shirokov VA, Majorov SG, Spirin AS. 1999. Method of preparing polypeptides in cell-free system and devise for its realization. *PCT WO 00 50436.*

124 Alakov YB, Baranov VI, Ovodov SJ, Ryabova LA, Spirin AS, Morozov IJ. 1995. Method of preparing polypeptides in cell-free translation system. *US Patent 5478730.*

125 Chekulayeva MN, Kurnasov OV, Shirokov VA, Spirin AS. 2001. Continuous-exchange cell-free protein-synthesizing system: synthesis of HIV-1 antigen Nef. *Biochem Biophys Res Commun* **280**, 914–7.

126 Martemyanov KA, Shirokov VA, Kurnasov OV, Gudkov AT, Spirin AS. 2001. Cell-free production of biologically active polypeptides: application to the synthesis of antibacterial polypeptide cecropin. *Protein Expr Purif* **21**, 456–61.

127 Kim DM, Swartz JR. 1999. Prolonging cell-free protein synthesis with a novel ATP regeneration system. *Biotech Bioeng* **66**, 180–8.

128 Kim DM, Swartz JR. 2000. Oxalate improves protein synthesis by enhancing ATP supply in a cell-free system derived from *Escherichia coli*. *Biotechnol Lett* **22**, 1537–42.

129 Betton JM, Palibroda N, Namane A, Metzler T, Bârzu O. 2002. Selective labelling of proteins in the RTS 500 System. In Spirin AS, ed., *Cell-free Translation Systems*, 219–5, Springer, Berlin.

130 Ellman JA, Volkman BF, Mendel B, Schultz PG, Wemmer DE. 1992. Site-specific isotopic labeling of proteins for NMR studies. *J Am Chem Soc* **141**, 7959–961.

131 Guignard L, Ozawa K, Pursglove SE, Otting G, Dixon NE. 2002. NMR analysis of *in vitro*-synthesized proteins without purification: a high-throughput approach. *FEBS Lett* **524**, 159–62.

132 Lederman M, Zubay G. 1967. DNA-directed peptide synthesis *in vitro*. I. A comparison of T2 and *E. coli* DNA-directed peptide synthesis. *Biochem Biophys Acta* **149**, 253–8.

133 De Vries JK, Zubay G. 1967. DNA-directed peptide synthesis. II. The synthesis of the α-fragment of the enzyme β-galactosidase. *Proc Natl Acad Sci USA* **57**, 1010–2.

134 Shirokov V, Simonenko PN, Biryukov SV, Spirin AS. 2002. Continuous-flow and continuous-exchange cell-free translation systems and reactors. In Spirin AS, ed., *Cell-free Translation Systems*, 91–108, Springer, Berlin.

(P 09)

Efficient Strategy for the Rapid Backbone Assignment of Membrane Proteins. N. Trbovic, C. Klammt, A. Koglin, F. Löhr, F. Bernhard, and V. Dötsch. *J. Am. Chem. Soc.*, **127**, 13504-13505. (2005)

J|A|C|S
COMMUNICATIONS

Published on Web 09/08/2005

Efficient Strategy for the Rapid Backbone Assignment of Membrane Proteins

Nikola Trbovic, Christian Klammt, Alexander Koglin, Frank Löhr, Frank Bernhard, and Volker Dötsch*

Institute for Biophysical Chemistry, University of Frankfurt and Center for Biomolecular Magnetic Resonance, 60439 Frankfurt, Germany

Received June 17, 2005; E-mail: vdoetsch@em.uni-frankfurt.de

Membrane proteins represent one of the biggest challenges in the area of structural biology based on the fact that they are hard to express, to purify, and difficult to analyze by high-resolution structural methods. Recently, the development of efficient cell-free transcription/translation protocols for the expression of milligram amounts of membrane proteins that cannot be expressed in sufficient quantities in vivo has opened a new avenue toward high-resolution structural investigations by X-ray crystallography and NMR spectroscopy.[1,2] To understand the function of membrane proteins, detailed investigations of their structure and interaction with their binding partners are necessary. NMR spectroscopy is an ideal tool for such investigations since it can provide both the structure and information about binding sites through chemical shift mapping. The basis for such detailed investigations is, however, the assignment of a protein's resonances. Unfortunately, α-helical proteins tend to display narrower chemical shift dispersion[3−5] as compared to that of proteins containing *β*-sheets. Since the majority of integral membrane proteins consists exclusively of α-helices, their NMR spectra tend to show a significant degree of peak overlap. The overlap problem is further aggravated by the often considerable size of the proteins and further enhanced by the fact that the proteins have to be solubilized in detergent micelles, which significantly increases the molecular weight of the protein/micelle particles, resulting in broader line width. Combined, these disadvantages of membrane proteins pose a considerable challenge for the chemical shift assignment, suggesting that new assignment strategies might be necessary in order to make backbone assignment of membrane proteins as routine a task as backbone assignment of soluble proteins. In this communication, we describe our efforts to assign the completely α-helical integral membrane protein TehA based on a combination of standard heteronuclear triple resonance experiments and a combinatorial labeling scheme.

The bacterial protein TehA is a 36 kDa membrane protein that shows limited homology to the family of small multidrug resistance proteins (SMR).[6] Its overexpression in bacteria confers resistance to tellurite compounds as well as to lipophilic cationic dyes. In vivo experiments have demonstrated that a 24 kDa fragment of TehA, which contains seven out of the 10 predicted transmembrane helices, shows the same biological effects as the full length protein. Therefore, we have focused on this 24 kDa fragment. For the expression and labeling with NMR active isotopes, we have employed an in vitro transcription/translation system based on *E. coli* S30 extracts, which yields 3 mg of TehA protein per milliliter of reaction volume.[1] Figure 1 shows the [^{15}N,^{1}H] TROSY spectrum of a ^{2}H/^{13}C/^{15}N triple-labeled sample of the 24 kDa TehA fragment, demonstrating the relatively narrow chemical shift dispersion and peak overlap in the middle of the spectrum.

To assign the backbone of the protein, we have measured HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HNCO, and HN(CA)CO spectra. In addition, we used a ^{2}H/^{15}N-labeled sample to measure a ^{15}N-edited NOESY−TROSY and a [^{15}N,^{1}H] HMQC−NOESY−

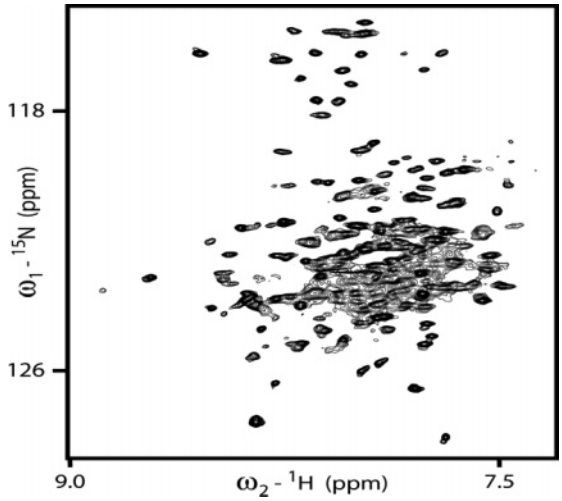


Figure 1. [^{15}N,^{1}H] TROSY spectrum of the 24 kDa fragment of TehA. The concentration of the protein was 0.6 mM, dissolved in 25 mM potassium phosphate buffer (pH 6) containing 5% LMPG (1-myristoyl-2-hydroxy-*sn*-glycero-3-[phosphor-*rac*-(1-glycerol)]). The spectrum was measured at 40 °C on a 800 MHz NMR instrument with 4 scans per FID and 300 increments in the indirect dimension.

TROSY experiment. On the basis of the combination of these experiments, we were able to assign 55% of the protein's backbone unambiguously. However, severe overlap, as well as the absence of some peaks, prevented us from obtaining more assignments. To close these gaps, we have expressed several amino acid type selective-labeled samples. However, despite the labeling of 10 different amino acid types (W, A, V, T, S, R, M, L, I, F), only an additional 10% of unambiguous assignments could be obtained. The main reason for the failure of the selective labeling procedure to result in higher assignment yields was that, in many cases, the N- and C-terminal connectivities to an identified amino acid type were not unambiguous, resulting in more than one possible sequence-specific assignment. To solve this problem, we decided to use a specific labeling procedure based on the simultaneous labeling of certain amino acid types with ^{15}N and other amino acid types with ^{13}C on the backbone carbonyls, which has been used in previous applications for site-specific labeling.[7−9] By measuring two-dimensional versions of an HNCO experiment, it is possible to select only those ^{15}N-labeled amino acids that are N-terminally preceded by a ^{13}C-labeled amino acid type. If that combination occurs only once in the entire protein, that amino acid is site specifically assigned and can be used as an anchor point for further sequence-specific assignments. To optimize this procedure and to minimize the number of samples that have to be produced, we employed a combinatorial approach. As summarized in Table 1, we produced three different samples, each one labeled with two to three different ^{15}N-labeled amino acid types and in addition with two different ^{13}C-carbonyl-labeled amino acid types. By measuring

10.1021/ja0540270 CCC: $30.25

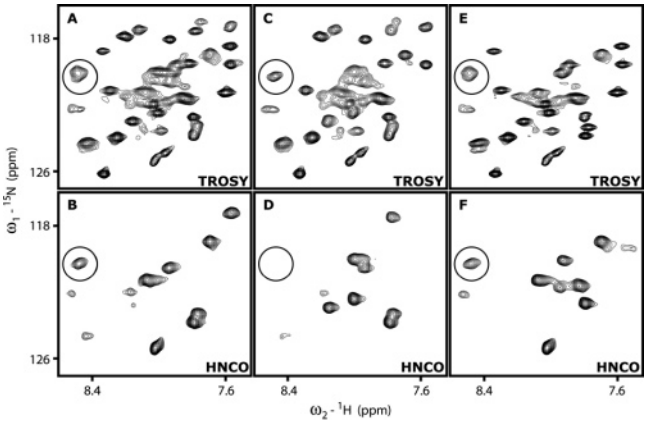


Figure 2. Results of the combinatorial labeling scheme. A, C, and E show TROSY spectra of samples 1−3, and B, D, and F the corresponding HNCO spectra. The circle indicates the resonance position of alanine 48, which in TehA is preceded by leucine 47.

Table 1. Labeling Schemes Used for the Combinatorial Labeling

amino acid type	sample 1	sample 2	sample 3
alanine	^{15}N	^{15}N	^{15}N
phenylalanine	^{15}N	^{15}N	
isoleucine	^{15}N		^{15}N
serine	^{13}C	^{13}C	
leucine	^{13}C		^{13}C
glycine		^{13}C	
valine			^{13}C

a [^{15}N,^{1}H] TROSY spectrum as well as a two-dimensional HNCO spectrum for each of the three samples (Figure 2), the sequence-specific assignment for eight new amino acids could be obtained, which served as specific starting points for more residues, bringing the total backbone assignment to 85% (in addition the assignment for 14 amino acids previously assigned could be confirmed). The amino acid types used for this combinatorial specific labeling approach were selected by an algorithm programmed in Matlab (http://www.biophyschem.uni-frankfurt.de/AK_Doetsch/projects/download/combilabel.m). The input for this algorithm are the amino acid sequence and the unassigned sequence stretches. On the basis of this information, the algorithm calculates the optimal combination of ^{15}N- and ^{13}C-labeled amino acids that will provide the most new specific assignments. Of the remaining 15% of backbone resonances that could not be assigned, 10% are not visible even in a two-dimensional TROSY spectrum while 5% are visible, but do not show sequential connectivities. These 5% can, in principle, be assigned with the specific labeling method. This, however, basically requires one sample per assignment.

Reinvestigation of our data showed that the combination of the nonselective triple resonance experiments with the combinatorial specific labeling strategy would have produced the same level of overall backbone assignment (85%) as the combination of all three assignment strategies (nonselective triple resonance experiments, amino acid type selective labeling, and combinatorial specific labeling). We, therefore, propose as the most straightforward strategy for the backbone assignment of membrane proteins the combination of nonselective triple resonance experiments and a combinatorial specific labeling protocol based on the production of proteins with an in vitro transcription/translation system. Recently, specific labeling in combination with triple resonance experiments has been used to accelerate the assignment process,[8] and a partial assignment procedure based entirely on the use of a combinatorial specific labeling scheme has been proposed for the selective assignment of certain amino acids in soluble proteins.[9] While a pure combinatorial approach is very efficient and useful for applications that only require the assignments of the amide proton and nitrogen frequencies (such as binding assays), full structure determinations are increasingly based on the use of ^{13}C backbone chemical shifts as structural parameters. In particular, for membrane proteins, the use of ^{13}Cα and ^{13}Cβ chemical shifts as indicators of the secondary structure is very important.[3,5] These chemical shifts are automatically provided by the nonselective triple resonance experiments. Furthermore, typically only 40−50% of all amino acids of a protein are part of a unique amino acid pair within the sequence and can, therefore, be unambiguously assigned solely based on combinatorial specific labeling pattern.[9] In the case of TehA, 95 pairs were unique corresponding to 43.4%; 35 pairs occurred twice, 9 pairs three times, and 6 pairs more than three times. By first assigning as many resonances as possible with the nonspecific triple resonance experiments and using an optimization procedure to pick from the remaining sequence stretches those amino acid combinations with the highest number of unique pairs, this problem of multiple possible assignments can be almost completely avoided.

This combinatorial labeling scheme relies on the use of a cell-free transcription/translation system for the production of the protein samples. In principle, amino acid type selective labeling is also possible in auxotrophic bacteria;[10] however, the available strains are only auxotrophic for certain types of amino acids, thus limiting the potential labeling combinations. In contrast, cross labeling in in vitro transcription/translation reactions is negligible.[11] Furthermore, the small amounts of labeled amino acids that are used make selective labeling in this system far less expensive than in cellular systems. Finally, producing NMR samples with an in vitro transcription/translation system is very fast. A typical NMR sample can be produced in less than 24 h since no complicated cell disruption and purification schemes are involved. In fact, since the produced protein is the only labeled macromolecule in the reaction mixture, NMR spectra can, in principle, be measured without any chromatographic purification.[11] We, therefore, believe that the combination of cell-free transcription/translation with standard NMR triple resonance experiments and combinatorial labeling schemes will provide a very efficient avenue toward the backbone assignment of membrane proteins.

Acknowledgment. Financial support was provided by the DFG (SFB 628).

References

(1) Klammt, C.; Löhr, F.; Schäfer, B.; Haase, W.; Dötsch, V.; Rüterjans, H.; Glaubitz, C.; Bernhard, F. *Eur. J. Biochem.* **2004**, *271*, 568−580.
(2) Ishihara, G.; Goto, M.; Saeki, M.; Ito, K.; Hori, T.; Kigawa, T.; Shirouzu, M.; Yokoyama, S. *Protein Expr. Purif.* **2005**, *41*, 27−37.
(3) Oxenoid, K.; Kim, H. J.; Jacob, J.; Sönnichsen, F. D.; Sanders, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 5048−5049.
(4) Tian, C.; Breyer, R. M.; Kim, H. J.; Karra, M. D.; Friedman, D. B.; Karpay, A.; Sanders, C. R. *J. Am. Chem. Soc.* **2005**, *127*, 8010−8011.
(5) Roosild, T. P.; Greenwald, J.; Vega, M.; Castronov, S.; Riek, R.; Choe, S. *Science* **2005**, *307*, 1317−1321.
(6) Turner, R. J.; Taylor, D. E.; Weiner, J. H. *Antimicrob. Agents Chemother.* **1997**, *41*, 440−444.
(7) Weigelt, J.; van Dongen, M.; Uppenberg, J.; Schultz, J.; Wikstrom, M. *J. Am. Chem. Soc.* **2002**, *124*, 2446−2447.
(8) Shi, J.; Pelton, J. G.; Cho, H. S.; Wemmer, D. E. *J. Biomol. NMR* **2004**, *28*, 235−247.
(9) Parker, M. J.; Aulton-Jones, M.; Hounslow, A. M.; Craven, C. J. *J. Am. Chem. Soc.* **2004**, *126*, 5020−5021.
(10) Ou, H. D.; Lai, H. C.; Serber, Z.; Dötsch, V. *J. Biomol. NMR* **2001**, *21*, 269−273.
(11) Ozawa, K.; Headlam, M. J.; Schaeffer, P. M.; Henderson, B. R.; Dixon, N. E.; Otting, G. *Eur. J. Biochem.* **2004**, *271*, 4084−4093.

JA0540270

(P 10)

Combination of cell-free expression and NMR spectroscopy as a new approach for structural investigation of membrane proteins. A. Koglin, C. Klammt, N. Trbovic, D. Schwarz, B. Schneider, B. Schäfer, F. Löhr, F. Bernhard and V. Dötsch. *Magn. Res. Chem.*, **44**, 17-23. (2006)

MAGNETIC RESONANCE IN CHEMISTRY
Magn. Reson. Chem. 2006; **44**: S17–S23
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/mrc.1833

Combination of cell-free expression and NMR spectroscopy as a new approach for structural investigation of membrane proteins

Alexander Koglin, Christian Klammt, Nikola Trbovic, Daniel Schwarz, Birgit Schneider, Birgit Schäfer, Frank Löhr, Frank Bernhard and Volker Dötsch*

Institute for Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Johann Wolfgang Goethe University, Marie-Curie Str. 9, 60439 Frankfurt, Germany

Received 30 November 2005; Revised 2 March 2006; Accepted 10 March 2006

Despite major technical advance in methods used for structural investigations of proteins structure determination of membrane proteins still poses a significant challenge. Recently, the application of cell-free expression systems to membrane proteins has demonstrated that this technique can be used to produce quantities sufficient for structural investigations for many different membrane proteins. In particular for NMR spectroscopy, cell-free expression provides major advantages since it allows for amino acid type selective and even amino acid position specific labeling. In this mini-review we discuss the combination of cell-free membrane protein expression and liquid state NMR spectroscopy. Copyright © 2006 John Wiley & Sons, Ltd.

KEYWORDS: cell-free expression; membrane proteins; amino acid type selective labeling

INTRODUCTION

Structure determination of membrane proteins still poses a very significant technical challenge despite the rapid technical advances in methods such as X-ray crystallography, electron microscopy and NMR spectroscopy. Problems in working with membrane proteins arise at all levels, starting from expressing quantities sufficient for structural analysis, solubilization and purification as well as obtaining two- or three-dimensional crystals or solutions amenable to liquid state NMR investigations.[1] Recent advances in production systems have at least provided new possibilities for the expression of large amounts of membrane proteins.[2–5] In particular, the use of cell-free expression systems holds promise to solve the protein production problem for many membrane proteins.[6–13] In combination with NMR spectroscopy, cell-free expression systems also provide a very efficient technique for selective labeling of certain amino acid types or even for specific amino acid positions, thus enabling efficient backbone assignment.[13,14] In this mini-review we will summarize the new advances in these cell-free expression systems and discuss the consequences of their application for structural investigations of membrane proteins in combination with NMR spectroscopy.

*Correspondence to: Volker Dötsch, Institute for Biophysical Chemistry and Center for Biomolecular Magnetic Resonance, Johann Wolfgang Goethe University, Marie-Curie Str. 9, 60439 Frankfurt, Germany. E-mail: vdoetsch@em.uni-frankfurt.de

CELL-FREE EXPRESSION SYSTEMS

Expression of large quantities of integral membrane proteins for structural investigations has been achieved with several different organisms. These include the most-often-used organisms, bacteria, in particular, *Escherichia coli*, as well as yeast and higher eukaryotic cells such as SF9 insect cells. In many cases, however, no or only small quantities of integral membrane proteins – only detectable by western blot analysis – are expressed in these cellular systems. Reasons for the low expression levels are, for example, toxic effects of the insertion of membrane proteins into the cellular membrane, inefficient transport of the overexpressed membrane proteins within the cell to the cellular membrane, poor growth of overexpressing strains and the generally unfavorable impact on the cellular metabolism. Obviously, some of these problems such as cellular transport and toxicity are eliminated in cell-free expression systems, making them a favorable alternative to the established cellular expression systems. Cell-free expression systems basically consist of a cellular extract containing the macromolecular transcription and translation machinery as well as additional high molecular weight components (tRNAs, pyruvate kinase and T7 RNA polymerase) and small molecules (amino acids, nucleotides, ions).[15] Currently, extracts for large-scale protein expression are mainly prepared from two different cellular systems. The first one is based on wheat germs[16,17] and the second one on *E. coli*[18–21] which we have used for the experiments described in this review.

From a technical standpoint, *in vitro* cell-free expression systems can be divided into two classes: batch systems and

continuous-exchange systems.[6,16,20,21] In the batch system all the macromolecular components of the extracts are mixed with the low molecular compounds such as amino acids and nucleotides in one reaction chamber. In contrast, in the continuous-exchange system the macromolecular components (reaction mixture) and the low molecular compounds (feeding mixture) are separated into two different compartments that are connected by a dialysis membrane of a certain cutoff size. This arrangement ensures the continuous supply of fresh low molecular weight precursors into the reaction mixture concomitant with the removal of inhibitory by-products like pyrophosphate. In our setup we use a reaction mixture to feeding mixture ratio of 1 : 17, which in our hands provides the optimal efficiency yielding up to 6 mg of membrane protein per single milliliter of reaction mixture. This yield can, in principle, be further increased if the entire feeding mixture gets replaced with a fresh one after approximately 4 to 6 h. However, the yield of protein after exchange of the feeding mixture will be only some additional 20%. In case the amount of reaction mixture (in particular, the extract) is the limiting factor, exchanging the feeding mixture, however, is an option for maximizing the yield.

While the use of cell-free expression systems in the field of structural biology has steadily increased over the last years, it has only recently been applied for the expression of membrane proteins. NMR-based structural investigations have so far mainly focused on bacterial β-barrel proteins which can be expressed in sufficient quantities in *E. coli* as inclusion bodies and can be efficiently refolded and solubilized.[21–25] Only a few α-helical membrane proteins with up to three transmembrane helices have been studied by NMR spectroscopy.[26–28] However, α-elical proteins represent by far the largest group of integral membrane proteins and include such important classes as G-protein-coupled receptors (GPCR) and different categories of transporters. We have, therefore, investigated the possibility of expressing larger α-helical integral membrane proteins in a continuous-exchange, cell-free system. Figure 1 shows the results for the expression of different classes of membrane proteins, including EmrE, SugE (bacterial small multidrug transporter), YfiK (bacterial amino acid transporter), TehA (bacterial transporter involved in heavy-metal resistance) and a eukaryotic GPCR (vasopressin receptor type 2). All these proteins are produced in the cell-free system in quantities ranging from 1 to 5 mg.[8] Comparison of these yields with expression tests of the same constructs in *E. coli* showed that with the exception of EmrE none of the other proteins could be expressed in appreciable amounts. Other reports have described the successful cell-free expression of functional GPCRs (β-adrenergic receptor, muscarinic acetylcholine receptor, neurotensin receptor).[11] Using a commercial cell-free expression system large quantities of small bacterial multidrug transporters, including EmrE,[10] of the mechanosensitive channel MscL[7] and of a bacterial light-harvesting protein[12] have also been obtained.

While use of a cell-free expression system can – at least for a significant number of membrane proteins – solve the problem of having to produce sufficient quantities for structural investigations, it does not automatically solve the problem of solubilization. Membrane proteins

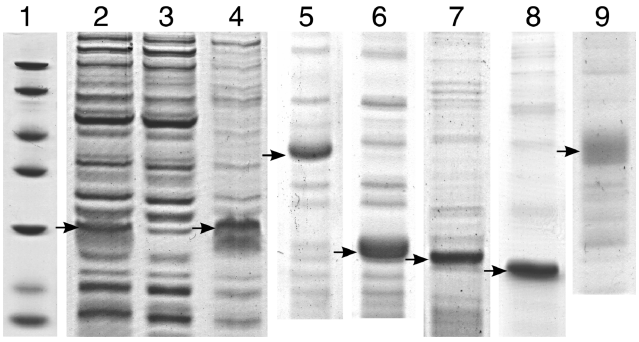


Figure 1. Examples of proteins expressed with the continuous-exchange, cell-free expression system. Lane 1: molecular weight marker; lane 2: crude lysate after expression of YfiK; lane 3: soluble fraction of the YfiK expression reaction; lane 4: pellet of the YfiK expression reaction; lane 5: TehA; lane 6: truncation mutant of TehA, lacking the *C*-terminal three transmembrane helices; lane 7: EmrE; lane 8: SugE; lane 9: porcine vasopressin receptor type 2.

expressed with a cell-free expression system are produced as precipitates.[8] Fortunately, however, these precipitates behave very differently from inclusion bodies obtained in *E. coli*. While proteins produced in inclusion bodies have to be solubilized with denaturation agents like urea or guanidinium hydrochloride and then refolded by buffer exchange, the pellets obtained from the cell-free expression system do not need any denaturation–refolding procedure in order to become solubilized. Instead the precipitates readily dissolve after gentle mixing with mild detergents.[6–13] However, not every detergent can dissolve every membrane protein pellet. Washing the pellet with a detergent that does not dissolve the membrane protein but other proteins that are contained in the pellet can thus be used as an additional purification step. Figure 2 shows such a case. Incubation of the bacterial amino acid transporter YfiK with the detergent *n*-nonyl-β-maltoside (NM) does not dissolve the pellet but removes impurities. LMPG on the other hand, readily dissolves the YfiK precipitate. Similarly, Fig. 2 demonstrates that other integral membrane proteins produced with the cell-free expression system can be resolubilized with different detergents. Further purification of the solubilized membrane proteins can then be achieved by standard protein purification methods such as Ni chelate and size exclusion chromatography. In principle, purification is not even required in the case of isotopically labeled samples for heteronuclear NMR measurements because all other macromolecular components are unlabeled. Only the amino acids that have been added in an isotopically labeled form to the feeding mixture and their potential metabolic products can produce background signals in NMR experiments with the crude reaction mixture of cell-free expression systems.[29] These small molecules can, however, be efficiently removed by dialysis.

One of the most significant advantages of the cell-free system is the fact that it constitutes an 'open system' that allows for the addition of additional factors to the mixture, such as protease inhibitors or RNase inhibitors. In the case of expressing integral membrane proteins the

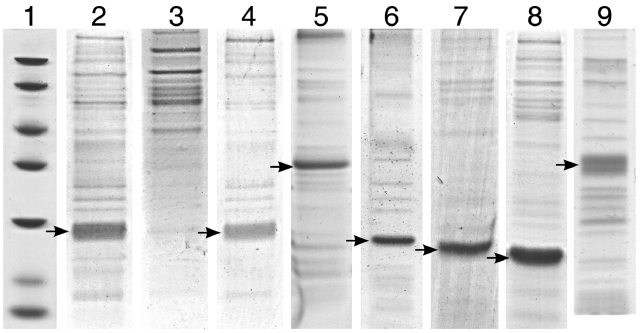


Figure 2. Examples of solubilization of different membrane proteins with mild detergents after expression with a cell-free expression system. Lane 1: molecular weight marker; lane 2: YfiK solubilized in 1% LMPG; lane 3: supernatant after washing the YfiK pellet with 1% NM; lane 4: YfiK solubilized in 1% LMPG after washing with 1% NM showing a reduced amount of impurities; lane 5: TehA in 1% (*n*-dodecyl β-D-maltose) DDM; lane 6: truncation mutant of TehA lacking the *C*-terminal three transmembrane helices solubilized in 5% LMPG; lane 7: EmrE in 2% DDM; lane 8: SugE in 2% DDM; lane 9: porcine vasopressin receptor in 5% LMPG.

addition of detergents to the reaction directly could, in principle, prevent the formation of a precipitate and produce a micelle-solubilized membrane protein sample directly. We have tested the possibility of producing micelle bound membrane proteins by adding different detergents directly to the mixture with three different integral membrane proteins: the α-helical small multidrug transporter EmrE from *E. coli*, the β-barrel outer membrane protein Tsx from *E. coli* and the porcine GPCR vasopressin receptor type 2.[9] While several different types of detergents from different chemical classes are capable of producing high yields of solubilized protein in all three cases, several other detergents inhibit the expression independent of the protein type. Members of this latter class of detergents often have a high critical micellar concentration (CMC). Examples are β-OG and CHAPS, and also phosphocholine derivatives like DPC which have a relatively low CMC. Detergents that can be used for producing large amounts of solubly expressed membrane proteins are, in particular, different members of the long chain polyoxyethylene-alkyl-ether group like Brij35, Brij58, Brij78 and Brij98 as well as the steroid derivative digitonin. Similar results were obtained by Ishihara and coworkers with three different GPCRs.[11] While a wide variety of different detergents are capable of solubilizing membrane proteins, functional tests have to be carried out in each case to ensure that the solubilized protein is active. Investigation of the functional state of the β-barrel protein TSX solubilized with different detergents and reconstituted in black lipid membranes, for example, has shown that it is active when solubilized in TX100 but not in LMPG micelles.[9] For other proteins other detergents will be optimal (Brij type detergents, for example, for some GPCRs[11]). Unfortunately, however, activity assays are not available for all membrane proteins. While binding studies (e.g. with GPCRs) are relatively easy to perform both in micelles as well as reconstituted in membranes,[10,11] transport assays for membrane proteins involved in transport are difficult to establish.[8]

A recent investigation of the suitability of different detergents for NMR-based structural investigation of membrane proteins has shown that LMPG and LPPG are the best suited.[30] Unfortunately, none of these is suitable for soluble expression of membrane proteins in cell-free expression systems. This problem can, however, be overcome by exchanging the detergent during purification.

NMR INVESTIGATIONS OF α-HELICAL MEMBRANE PROTEINS

While high expression levels and efficient procedures for solubilizing membrane proteins are necessary prerequisites for obtaining a high-resolution structure, they are by no means sufficient. In contrast to soluble proteins for which recent technical advances make structure determination either by X-ray crystallography or by NMR spectroscopy often straightforward, applying the same techniques to integral membrane proteins poses additional technical challenges. For NMR spectroscopy the often large size of integral membrane proteins resulting in slow rotational tumbling and concomitant broad linewidth constitutes a significant problem. This problem is aggravated by the fact that membrane proteins have to be solubilized in micelles, which contribute considerably to the overall molecular weight of the protein/micelle particle. Furthermore, α-helical proteins tend to display narrower chemical shift dispersions, resulting in peak overlap, and the transmembrane sections of these proteins consist predominantly of hydrophobic amino acids often leading to clustering of identical amino acids with very similar chemical shifts. An example of a spectrum of an all α-helical integral membrane protein, TehA, is shown in Fig. 3. TehA is a 36 kDa protein that is involved in detoxifying tellurite compounds in bacteria, although the exact mechanism is currently not known.[31,32] In addition, it can transport quaternary organic cations such as crystal violet and ethidium bromide. The wild-type protein consists of ten transmembrane helices. *In vivo* experiments, however, have demonstrated that expression of the seven *N*-terminal helices is sufficient to confer tellurite resistance to the bacterial cell. For our NMR investigations of the TehA protein we have, therefore, used a truncation mutant that contains only the first seven transmembrane helices (219 amino acids).[13] The NMR sample used in the spectrum shown in Fig. 3 was produced in our cell-free expression system by adding a mixture of all 20 amino acids, ^{15}N-labeled and deuterated, to the feeding mixture. Adding 1% of Brij-78 to the reaction mixture resulted in a soluble protein, which was purified by Ni chelate chromatography. After exchanging the detergent with 3% LMPG, the final concentration of the NMR sample was 0.6 mM. The spectrum of the solubilized protein shows many resolved peaks but also regions with very significant peak overlap as expected for an all α-helical protein of 219 amino acids packed into a micelle.

In order to obtain the backbone assignment we initially measured a set of standard triple resonance experiments, including HNCA, HNCOCA, HNCACB, HNCOCACB,

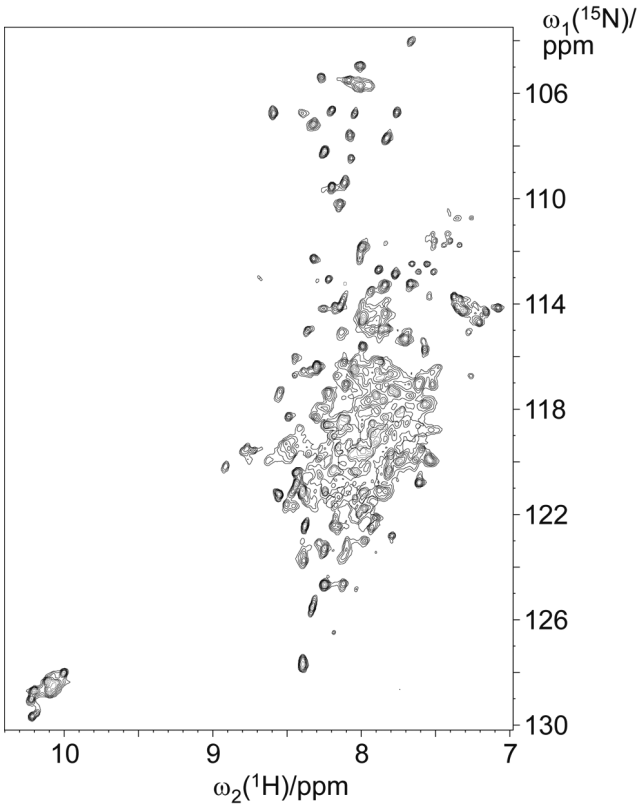


Figure 3. [^{15}N,^{1}H]-TROSY spectrum of the solubly expressed ^{2}H/^{15}N-labeled 24 kDa fragment of TehA. The concentration of the protein was 0.5 mM, dissolved in 25 mM potassium phosphate buffer (pH 6) containing 3% LMPG. The spectrum was measured at 40 °C on a 900 MHz NMR instrument with four scans per free induction decay (FID) and 300 increments in the indirect dimension.

HNCO and HNCACO, as well as a ^{15}N-edited [^{1}H,^{1}H]-NOESY-TROSY experiment. Figure 4 shows [^{13}C,^{1}H] planes taken from an HNCACB spectrum, demonstrating the high quality of the spectra, which allowed us to assign 55% of the amino acid backbone resonances. The main reason why we did not achieve a higher percentage is the extensive overlap in the spectra. One of the advantages of the cell-free system is that it allows for efficient amino acid type selective labeling with almost all amino acid types by minimizing metabolic scrambling. In contrast, amino acid type selective labeling in bacteria is restricted to certain amino acid types, or requires auxotrophic strains.[33–35] Using the cell-free expression system we prepared ten different samples each labeled with a different amino acid type (W, A, V, T, S, R, M, L, I, F). Despite these efforts these samples provided only 10% additional backbone assignments. The main reason for the failure of the selective labeling procedure to result in higher assignment yields was that in many cases the *N*- and *C*-terminal connectivities to an identified amino acid type were not unambiguous. A possible solution of this problem is a site-specific labeling procedure. Such site-specific labeling is possible with the methods developed in the laboratory of Peter Schultz, which are based on the use of special tRNAs that recognize a stop codon to introduce a labeled amino acid at the site of this stop codon.[36] Unfortunately, the preparation of such tRNAs is a relatively complicated technique. NMR spectroscopy provides an alternative to site-specific identification. This method is based on simultaneous labeling of a protein with one type of ^{15}N-labeled amino acid and another type of a ^{13}C-labeled amino acid. By measuring two-dimensional versions of an HNCO experiment it is possible to select only those ^{15}N-labeled amino acids that are *N*-terminally preceded by a ^{13}C-labeled amino acid type.[13,37–40] This method has been used in the past to accelerate the backbone assignment of soluble proteins[39] or to obtain as many assignments as possible without the use of triple resonance experiments.[38] A disadvantage of this specific-site labeling scheme is that it, in principle, requires one sample per assignment which would be, even with the use of a cell-free expression system, quite labor intensive. In order to make site-specific labeling more efficient we have used a combinatorial approach based on simultaneously

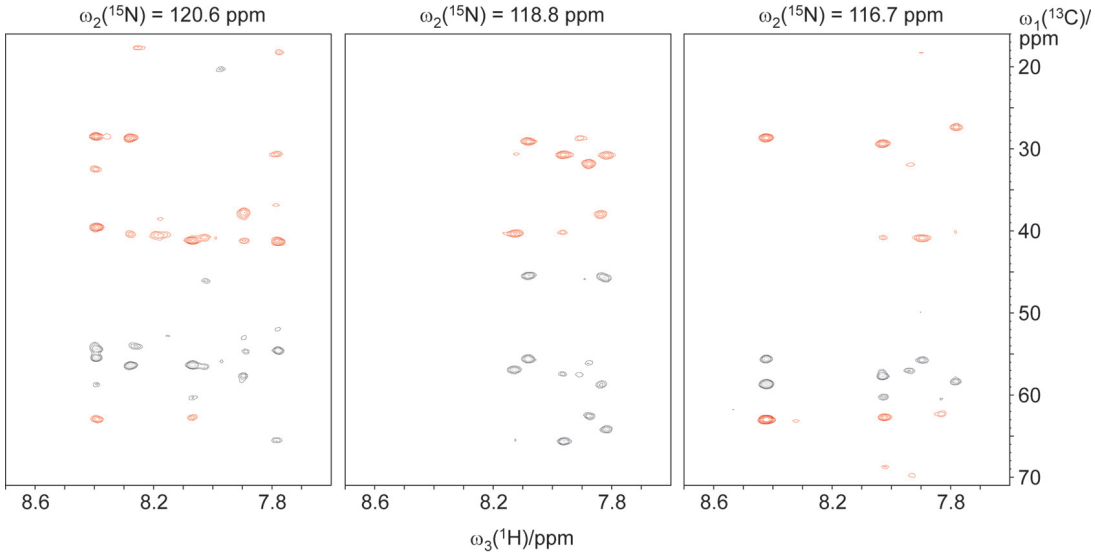


Figure 4. [^{13}C,^{1}H] planes taken from an HNCACB experiment measured with a ^{2}H/^{13}C/^{15}N-labeled sample of the 24 kDa fragment of TehA on a 900 MHz instrument.

labeling several different amino acids with ^{15}N or with ^{13}C. By preparing three different samples, each one labeled in a different combination of ^{15}N or ^{13}C labeled amino acids, the assignment of unique combinations can be achieved by analyzing the three different [^{15}N,^{1}H]-TROSY and the three different two-dimensional HNCO spectra.[13] Altogether, this procedure yielded 22 site-specific assignments, 8 of which were new, while the remaining 14 had already been assigned previously with the help of triple resonance experiments. For these 14 amino acids the site-specific labeling procedure confirmed the previous assignment. Using the eight new assignments as anchor points, the overall percentage of backbone assignment could be extended to 85%. Of the remaining 15% of backbone resonances that are not assigned, 10% are not visible even in 2D TROSY experiments, while 5% could in principle be assigned, which, however, required the preparation of one sample per assignment.

One problem that limits the maximum number of assignments that can be obtained from a combinatorial labeling approach is that many amino acid pairs occur more than once in the protein sequence, thus preventing an unambiguous assignment.[13,38] In the case of TehA 95, amino acid pairs were unique corresponding to 43.4% while 35 pairs occurred twice, 9 pairs three times and 6 pairs more than three times. This statistic is typical for an average protein and has been seen also with soluble proteins.[38] As a consequence, the complete assignment of a protein's backbone cannot be obtained from such a combinatorial labeling approach alone, but needs additional information as input. Therefore, we suggest using a combination of nonselective standard triple resonance experiments and site-specific labeling for the backbone assignment of membrane proteins. In order to minimize the number of samples needed for the combinatorial assignment procedure we further suggest the assignment of as many backbone resonances as possible through nonselective triple resonance experiments in the first stage of the assignment process and the use of the site-specific labeling method to fill in the gaps.

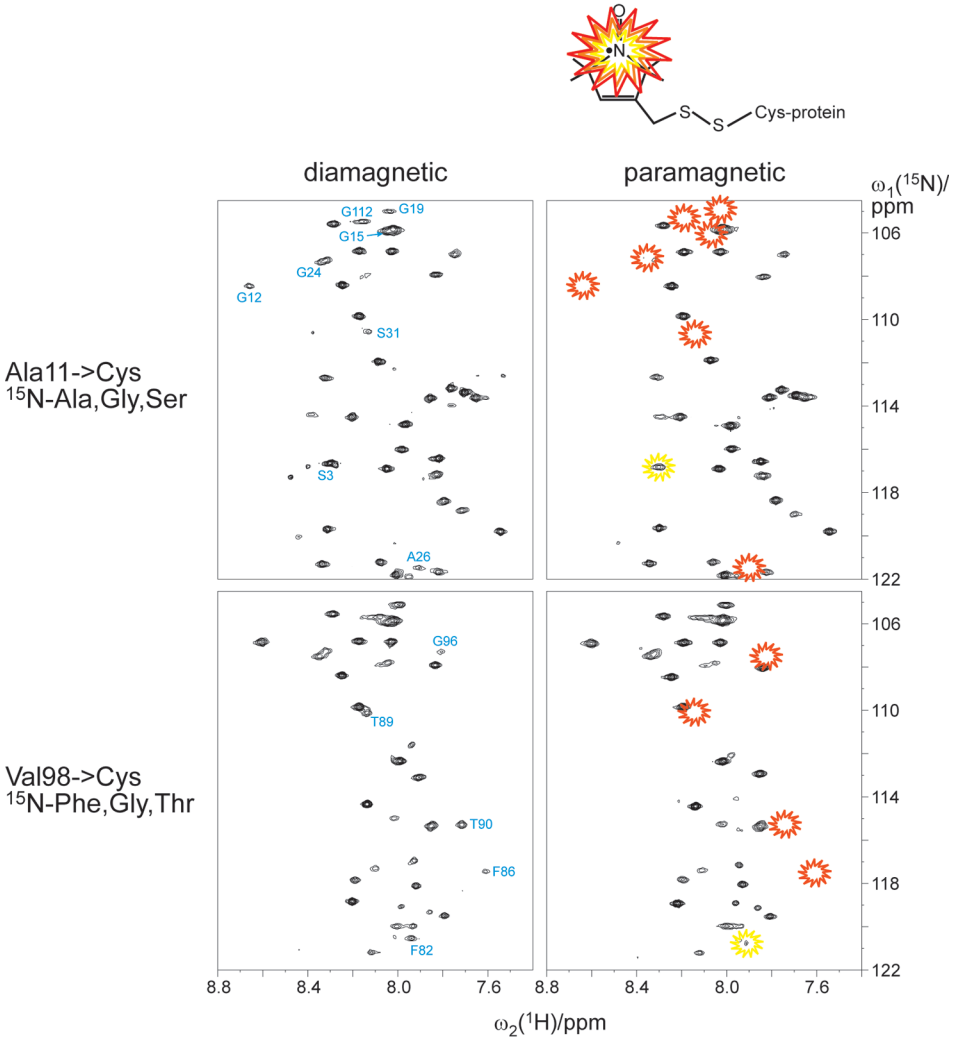


Figure 5. Comparison of TROSY spectra of the 24 kDa fragment of TehA in its reduced, diamagnetic form (left) and in its paramagnetic form with a spin label (MTSL) attached (right). The significant overlap of the fully labeled spectrum makes amino acid type selective labeling necessary. Two different experiments are shown. The site of the spin label attachment as well as the type of labeling pattern is indicated on the left. Peaks that are completely absent in the paramagnetic spectrum are labeled with an orange symbol and those that show significant line broadening with a yellow symbol. The spectra were measured on an 800 MHz instrument. The assignment of these affected peaks is shown in the diamagnetic spectrum.

Following the sequential assignment of 85% of the backbone resonances, we have started to analyze both the secondary structure as well as to explore techniques to obtain information about the three-dimensional structure of the entire protein. Analysis of the secondary structure is mainly based on the $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts as well as the pattern of sequential NOEs in the ^{15}N-edited [^{1}H,^{1}H]-NOESY spectrum, following the same procedures that have been developed for soluble proteins. Investigation of the three-dimensional structure, on the other hand, has to rely on different strategies. Because of the limited chemical shift dispersion observed in the spectra of α-helical integral membrane proteins, assignment of the side chain resonances will be possible for very few residues only. In addition, the high molecular weight of the micelle/protein particles requires deuteration, which eliminates most of the protons that can be used in classical $^{1}H-^{1}H$ NOE-based structure determination procedures. While measuring residual dipolar couplings[26,41,42] as well as incorporating selectively protonated amino acids (e.g. Ile, Leu and Val with protonated methyl groups) into deuterated proteins[43–45] are options for obtaining structural information, distance constraints derived from paramagnetic relaxation enhancement experiments could become important and might replace to a large extent NOE-based distance constraints in structural studies of membrane proteins.[46–48] This type of distance constraints has been developed with soluble proteins that are too big for classical structure determination procedures.[46] In these experiments the line broadening effect of a paramagnetic tag (e.g. 1-oxy(-2,2,5,5-tetramethyl-Δ^3-pyrroline-3-methyl) methanethiosulfonate (MTSL)) that is attached to a protein through a specific cysteine side chain on the surrounding amide protons is investigated. Since this line broadening is proportional to $1/r^6$, a quantitative measurement of the line broadening can be used for obtaining distance constraints. This method has been successfully used in the structure determination of the membrane-associated protein Mystic[48] and in the refinement of the outer membrane β-barrel protein OmpA.[49] We have also started to use paramagnetic relaxation enhancement to obtain distance constraints for the structure determination of TehA. For this purpose we have mutated all three naturally occurring cysteines to alanines and have reintroduced a cysteine residue at specific locations, mostly at the beginning or the end of helices. Because of the significant peak overlap in the [^{15}N,^{1}H]-TROSY spectrum we cannot use a fully ^{15}N-labeled sample for these investigations but have to rely on our cell-free expression system to prepare several samples each labeled selectively with a different amino acid. Figure 5 shows as an example a comparison of the [^{15}N,^{1}H]-TROSY spectra of TehA, labeled with a paramagnetic tag at amino acid position 11 or 98 and the same sample after reducing the spin label to a diamagnetic compound with ascorbic acid.

In this review we have focused on discussing the combination of cell-free expression systems with liquid state NMR spectroscopy for structural investigations of membrane proteins. In particular, for NMR spectroscopy the possibility to obtain amino acid type selective labeled protein samples for almost all amino acid types without metabolic scrambling is a very significant advantage over cellular based expression systems. Additional advantages are the very efficient production of proteins in these cell-free systems, which does not require cell disruption or time-consuming unfolding and refolding procedures.

REFERENCES

1. Arora A, Tamm LK. *Curr. Opin. Struct. Biol.* 2001; **11**: 540.
2. Tate CG. *FEBS Lett.* 2001; **504**: 94.
3. Dobrovetsky E, Lu ML, Andorn-Broza R, Khutoreskaya G, Bray JE, Savchenko A, Arrowsmith CH, Edwards AM, Koth CM. *J. Struct. Funct. Genomics* 2005; **6**: 33.
4. Rapp M, Drew D, Daley DO, Nilsson J, Carvalho T, Melen K, De Gier JW, von Heijne G. *Protein Sci.* 2004; **13**: 937.
5. Drew D, Slotboom DJ, Friso G, Reda T, Genevaux P, Rapp M, Meindl-Beinker NM, Lambert W, Lerch M, Daley DO, Van Wijk KJ, Hirst J, Kunji E, de Gier JW. *Protein Sci.* 2005; **14**: 2011.
6. Spirin AS, Baranov VI, Ryabova LA, Ovodov SY, Alakhov YB. *Science* 1988; **242**: 1162.
7. Berrier C, Park KH, Abes S, Bibonne A, Betton JM, Ghazi A. *Biochemistry* 2004; **43**: 12585.
8. Klammt C, Löhr F, Schäfer B, Haase W, Dötsch V, Rüterjans H, Glaubitz C, Bernhard F. *Eur. J. Biochem.* 2004; **271**: 568.
9. Klammt C, Schwarz D, Fendler K, Haase W, Dötsch V, Bernhard F. *FEBS J.* 2005; **272**: 6024.
10. Elbaz Y, Steiner-Mordoch S, Danieli T, Schuldiner S. *Proc. Natl. Acad. Sci. U.S.A.* 2004; **101**: 1519.
11. Ishihara G, Goto M, Saeki M, Ito K, Hori T, Kigawa T, Shirouzu M, Yokoyama S. *Protein Expr. Purif.* 2005; **41**: 27.
12. Shimada Y, Wang ZY, Mochizuki Y, Kobayashi M, Nozawa T. *Biosci. Biotechnol. Biochem.* 2004; **68**: 1942.
13. Trbovic N, Klammt C, Koglin A, Löhr F, Bernhard F, Dötsch V. *J. Am. Chem. Soc.* 2005; **127**: 13504.
14. Ozawa K, Headlam MJ, Schaeffer PM, Henderson BR, Dixon NE, Otting G. *Eur. J. Biochem.* 2004; **271**: 4084.
15. Baranov VI, Spirin AS. *Methods Enzymol.* 1993; **217**: 123.
16. Vinarov DA, Lytle BL, Peterson FC, Tyler EM, Volkman BF, Markley JL. *Nat. Methods.* 2004; **1**: 149.
17. Tyler RC, Aceti DJ, Bingman CA, Cornilescu CC, Fox BG, Frederick RO, Jeon WB, Lee MS, Newman CS, Peterson FC, Phillips GN Jr, Shahan MN, Singh S, Song J, Sreenath HK, Tyler EM, Ulrich EL, Vinarov DA, Vojtik FC, Volkman BF, Wrobel RL, Zhao Q, Markley JL. *Proteins* 2005; **59**: 633.
18. Spirin AS. *Trends Biotechnol.* 2004; **22**: 538.
19. Martemyanov KA, Shirokov VA, Kurnasov OV, Gudkov AT, Spirin AS. *Protein Expr. Purif.* 2001; **21**: 456.
20. Kim DM, Choi CY. *Biotechnol. Prog.* 1996; **12**: 645.
21. Knapp KG, Swartz JR. *FEBS Lett.* 2004; **559**: 66.
22. Arora A, Abildgaard F, Bushweller JH, Tamm LK. *Nat. Struct. Biol.* 2001; **8**: 334.
23. Fernandez C, Hilty C, Wider G, Güntert P, Wüthrich K. *J. Mol. Biol.* 2004; **336**: 1211.
24. Fernandez C, Adeishvili K, Wüthrich K. *Proc. Natl. Acad. Sci. U.S.A.* 2001; **98**: 2358.
25. Hwang PM, Choy WY, Lo EI, Chen L, Forman-Kay JD, Raetz CR, Prive GG, Bishop RE, Kay LE. *Proc. Natl. Acad. Sci. U.S.A.* 2002; **99**: 13560.
26. Oxenoid K, Kim HJ, Jacob J, Sonnichsen FD, Sanders CR. *J. Am. Chem. Soc.* 2004; **126**: 5048.
27. MacKenzie KR, Prestegard JH, Engelman DM. *Science* 1997; **276**: 131.
28. Girvin ME, Rastogi VK, Abildgaard F, Markley JL, Fillingame RH. *Biochemistry* 1998; **37**: 8817.
29. Guignard L, Ozawa K, Pursglove SE, Otting G, Dixon NE. *FEBS Lett.* 2002; **524**: 159.
30. Krueger-Koplin RD, Sorgen PL, Krueger-Koplin ST, Rivera-Torres IO, Cahill SM, Hicks DB, Grinius L, Krulwich TA, Girvin ME. *J. Biomol. NMR* 2004; **28**: 43.

31. Turner RJ, Taylor DE, Weiner JH. *Antimicrob. Agents Chemother.* 1997; **41**: 440.
32. Taylor DE. *Trends Microbiol.* 1999; **7**: 111.
33. Waugh DS. *J. Biomol. NMR* 1996; **8**: 184.
34. McIntosh LP, Dahlquist FW. *Q. Rev. Biophys.* 1990; **23**: 1.
35. Ou HD, Lai HC, Serber Z, Dötsch V. *J. Biomol. NMR* 2001; **21**: 269.
36. Noren CJ, Anthony-Cahill SJ, Griffith MC, Schultz PG. *Science* 1989; **244**: 182.
37. Weigelt J, van Dongen M, Uppenberg J, Schultz J, Wikstrom M. *J. Am. Chem. Soc.* 2002; **124**: 2446.
38. Parker MJ, Aulton-Jones M, Hounslow AM, Craven CJ. *J. Am. Chem. Soc.* 2004; **126**: 5020.
39. Shi J, Pelton JG, Cho HS, Wemmer DE. *J. Biomol. NMR* 2004; **28**: 235.
40. Kainosho M, Tsuji T. *Biochemistry* 1982; **21**: 6273.
41. Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S. *J. Biomol. NMR* 2000; **18**: 303.
42. Chou JJ, Gaemers S, Howder B, Louis JM, Bax A. *J. Biomol. NMR* 2001; **21**: 377.
43. Rosen MK, Gardner KH, Willis RC, Parris WE, Pawson T, Kay LE. *J. Mol. Biol.* 1996; **263**: 627.
44. Gardner KH, Rosen MK, Kay LE. *Biochemistry* 1997; **36**: 1389.
45. Hilty C, Fernandez C, Wider G, Wüthrich K. *J. Biomol. NMR* 2002; **23**: 289.
46. Battiste JL, Wagner G. *Biochemistry* 2000; **39**: 5355.
47. Gaponenko V, Howarth JW, Columbus L, Gasmi-Seabrook G, Yuan J, Hubbell WL, Rosevear PR. *Protein Sci.* 2000; **9**: 302.
48. Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R, Choe S. *Science* 2005; **307**: 1317.
49. Liang B, Bushweller JH, Tamm LK. *J. Am. Chem. Soc.* 2006; **128**: 4389.

(P 11)

Incorporation of fluorescence labels into cell-free produced proteins. K. Sengupta, C. Klammt, F. Bernhard and H. Rüterjans. In: *Cell-Free Protein Expression*, J.R. Swartz (ed.) pp. 81-88, Springer Verlag, Berlin, Heidelberg, New York. (2003)

Incorporation of Fluorescence Labels into Cell-Free Produced Proteins

10

K. SENGUPTA, C. KLAMMT, F. BERNHARD, H. RÜTERJANS*

Introduction

Natural Trp residues in proteins have an intrinsic fluorescence in the range between λ_{em} 330 and 350 nm. Trp fluorescence is sensitive to general solvent effects and its emission spectrum and quantum yield strongly depends on the protein structure and on the local microenvironment. Conformational changes in proteins, e.g., due to denaturation or ligand and substrate binding, could therefore be principally analyzed by fluorescence spectroscopy. However, the study of protein/protein interactions is limited, as the intrinsic Trp fluorescence of different proteins overlaps, thus making it impossible to assign and interpret fluorescence changes that result from intermolecular associations. An approach to solve this problem is to specifically alter the spectral properties of one protein by the incorporation of spectrally enhanced Trp analogues like 5-fluorotryptophan (5-FW), 5-hydroxytryptophan (5-HW) or 7-azatryptophan (7-AW). While all three analogues differ in their photophysical properties (Ross et al. 1997), the two latter are the most widely used tryptophan analogues which have already been incorporated into a number of proteins (Hogue et al. 1992; Ross et al. 1992; Laue et al. 1993; Soumillion et al. 1995). 5-HW and 7-AW display a red-shifted λ_{em}(max) and the unique spectral features of these analogues allows then to be observed specifically also in presence of other natural Trp residues.

Replacement of natural Trp residues in proteins by appropriate analogues is usually achieved by the tightly controlled overproduction of the protein in a Trp auxotrophic *Escherichia coli* strain growing in minimal medium containing the desired Trp analogue (Ross et al. 1997; Mohammadi et al. 2001). The recently developed techniques for high-level cell-free expression of proteins offer an excellent alternative possibility for efficient label incorporation into recombinant proteins. The compartmentation of the continuous exchange cell-free system (Spirin et al. 1988) in a defined feeding solution containing precursors including all amino acids, and in a reaction solution containing all high molecular weight substances, allows the convenient and uniform labeling of any amino acid type. In addition, common problems of standard *in vivo* labeling

* Dr. Heinz Rüterjans, Institut für Biophysikalische Chemie, Universität Frankfurt, Marie-Curie-Str. 9, 60435 Frankfurt am Main, e-mail: hruet@bpc.uni-frankfurt.de

protocols associated with toxic effects of the label precursors, reduced protein yields or low label incorporation into protein samples can be eliminated.

The RcsB protein of *E. coli* consists of 216 amino acid residues and contains one Trp residue in the N-terminal receiver domain at amino acid position 28. The RcsB protein is a key regulator in enteric and plant pathogenic bacteria. It is the central transcriptional effector of a modified two-component system essential for (1) the induction of exopolysaccharide (EPS) biosynthesis (Stout and Gottesman 1990), (2) regulation of cell division (Gervais et al. 1992; Carballes et al. 1999) and (3) expression of the osmoregulated gene *osmC* (Davalos-Garcia et al. 2001). Two well-conserved sequence motifs, an N-terminal phosphorylation motif and a C-terminal helix-turn-helix DNA-binding motif can be found in RcsB, dividing the 24-kDa protein into an N-terminal "receiver" and probably protein-interacting domain, and a C-terminal DNA interacting "effector" domain. For the activation of EPS biosynthesis, RcsB interacts with the coactivator RcsA and binds as a heterodimer at a 14-bp consensus sequence, the RcsAB box, present in promoters responsible for bacterial capsule production (Kelm et al. 1997; Wehland and Bernhard 2000). Further evidence exists, that RcsB is able to interact with other coactivators as well in order to control distinct biosynthetic operons (Gervais and Drapeau 1992; Virlogeux et al. 1996). On the other hand, RcsB might also form homooligomers during regulation of cell division or by induction of the *osmC* gene. These different modes of regulation by the RcsB protein imply specific recognition mechanisms among RcsB proteins, and also between RcsB and other coactivators. We report the replacement of the natural Trp residue in RcsB by various Trp analogues and the high-level production of modified RcsB derivatives using cell-free expression. The isolated RcsB alloproteins proved to be suitable for protein interaction studies by fluorescence spectroscopy, and we present first evidence of an oligomerization of RcsB due to molecular association of the C-terminal effector domains.

Methods

DNA and Protein Techniques

Standard protocols for PCR amplification, DNA recombination, analysis and purification were used as described (Sambrook et al. 1989). Proteins were analyzed by SDS-PAGE and quantified using the Bradford assay. Plasmid pET21-rcs-B_{EC} was constructed by amplifying the coding region for the *E. coli* RcsB protein by standard PCR techniques using Vent polymerase (NEB) and chromosomal DNA of *E. coli* strain XL1 as a template. The PCR product was cloned into the *NdeI/HindIII* sites of plasmid pET21a(+) (Novagen). DNA fragments encoding for the N-terminal (amino acids 1–144) and C-terminal (amino acids 129–216) domain of RcsB were PCR-amplified and cloned into the *BamHI/HindIII* sites of the vector pQE30 (Qiagen). The full-length RcsB protein and the two domains were purified by Ni^{2+}-chelate chromatography, and the RcsA protein was overproduced and purified, as described (Kelm et al. 1997).

Cell-Free Production of Labeled RcsB Proteins

Proteins were produced with a commercial cell-free rapid translation system (RTS, Roche Applied Science) using the RTS 500 or RTS 9000 *E. coli* HY Kits and the RTS maxi or RTS Proteomaster device for incubation. Plasmid pET21-rcsB$_{EC}$ was used as a template for transcription in RTS. Production of labeled protein samples was performed with a total of 15 µg plasmid DNA in the RTS 500 and of 150 µg in the RTS 9000 device. All amino acids were used in L configuration and the concentrations were adjusted according to the composition of RcsB with 1 mM for Gln, His, Pro, Thr, Tyr, Phe, 2 mM for Leu and 1.5 mM for all others. Trp and Cys were completely omitted from the reaction mixtures. The final concentrations of labeled amino acid analogues were 2 mM for 5-HW and 5-p-fluorophenylalanine (5-pFF), and 4 mM for 7-AW and 5-FW. The latter two were added as a mixture of D and L configuration. All other solutions were prepared according to the manufacturer's recommendations. The reactions were carried out at 30 °C at 900 rpm (RTS 9000 *E. coli* HY) and 150 rpm (RTS 500 *E. coli* HY), respectively.

Electrophoretic Mobility Shift Assay (EMSA)

A 183-bp fragment containing the RcsAB box from the *Erwinia amylovora amsG* promoter (Kelm et al. 1997) was labeled with ^{32}P by the fill-in technique with Klenow polymerase. The labeled DNA was incubated with 30 pmol of the various RcsB derivatives and 300 pmol RcsA protein. The EMSA was done as described (Kelm et al. 1997) and the dried gel was exposed to an X-ray film.

Fluorescence Spectroscopy

Measurements were carried out in 50 mM phosphate buffer, pH 6.4, with 100 mM NaCl. Steady-state emission spectra were recorded with 2 nmol of protein in a quartz cuvette of 1 cm path length in a Jasco FP 6300 spectrofluorometer at 20 °C. The bandpasses for emission and excitation were 4 nm. All spectra were recorded with a step size of 0.5 nm. Excitation maxima of all proteins were determined at 292 nm. Fluorescence emission spectra of native and analogue-containing RcsB were scanned from 320 to 400 nm. For interaction studies, equimolar amounts of samples were mixed thoroughly and equilibrated for 5 min at 20 °C before starting the measurement.

Results and Discussion

In Vitro Expression and Labeling of RcsB

The single Trp residue of RcsB was replaced by the modified analogues 7-AW (7-AW-RcsB), 5-FW (5-FW-RcsB) and 5-HW (5-HW-RcsB). Furthermore, in

one sample modified with 5-HW, all six Phe residues were additionally replaced by 5-pFF (5-HW/5-pFF-RcsB). The 25-kDa RcsB proteins were produced with a N-terminal poly(His)$_6$ tag and could be purified in one step by metal-chelate chromatography. The production rate of RcsB could be considerably increased from approximately 0.5 mg/ml reaction solution up to 3 mg/ml after optimizing the amino acid concentrations in the RTS solutions according to the amino acid composition of the RcsB protein (data not shown). All *in vitro*-produced modified RcsB derivatives were fully soluble and comparable production rates for all proteins have been achieved (Fig. 10.1). Similar yields were obtained in RTS 500 and RTS 9000 devices (data not shown).

The replacement by the various Trp analogues resulted in active RcsB derivatives as determined by their DNA-binding properties in an EMSA (Fig. 10.2). All fluorescence labeled RcsB proteins were able to form a heterodimer with the coactivator RcsA and showed complex formation with a DNA target from the *E. amylovora amsG* promoter containing the RcsAB box (Wehland and Bernhard 2000). This result demonstrates the functional integrity of all modified proteins, indicating that they still retain their native conformation.

Spectral Properties of Modified RcsB Proteins

The fluorescence emission spectra of native RcsB and the analogue-containing derivatives are shown in Fig. 10.3A. A pH value of 6.4 was chosen for all experiments, as RcsB is most stable under these conditions. Three-dimensional synchronous scans from 270 to 330 nm revealed an excitation maximum of all sam-

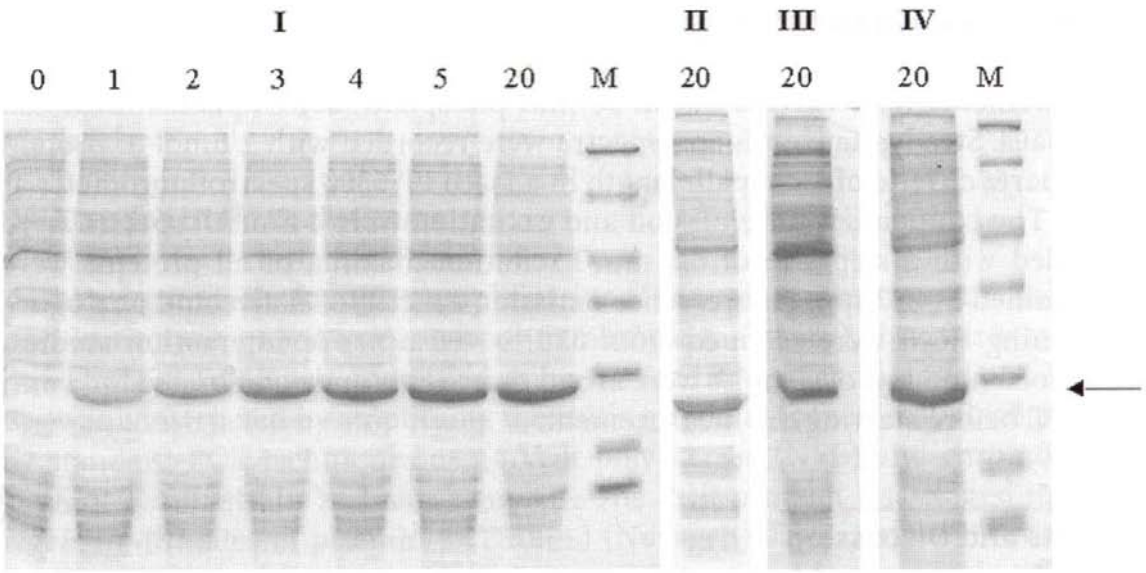


Fig. 10.1. Production of labeled RcsB proteins by cell-free expression in an RTS 500 device. Aliquots were taken after the indicated hours of expression and protein production was analyzed by SDS-PAGE on a 17.5% gel. The *arrow* indicates the synthesized RcsB derivatives: *I* 7-AW-RcsB; *II* 5-HW-RcsB; III 5-FW/5-pFF-RcsB; *IV* 5-FW-RcsB. *M* molecular weight marker, *from bottom to top*: 14.4, 18.4, 25, 35, 45, 66.2, 115 kDa

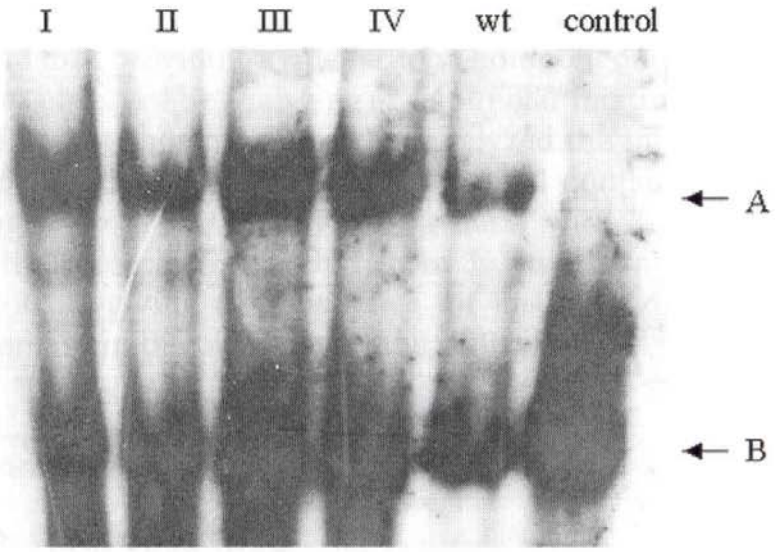
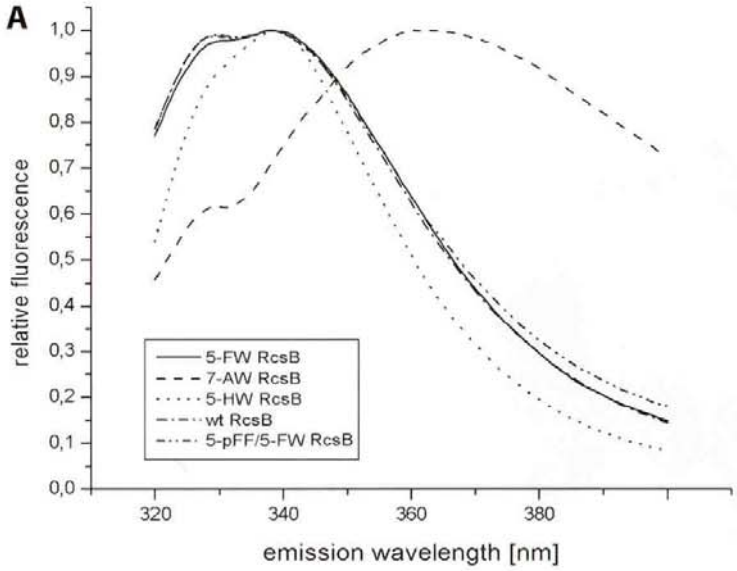


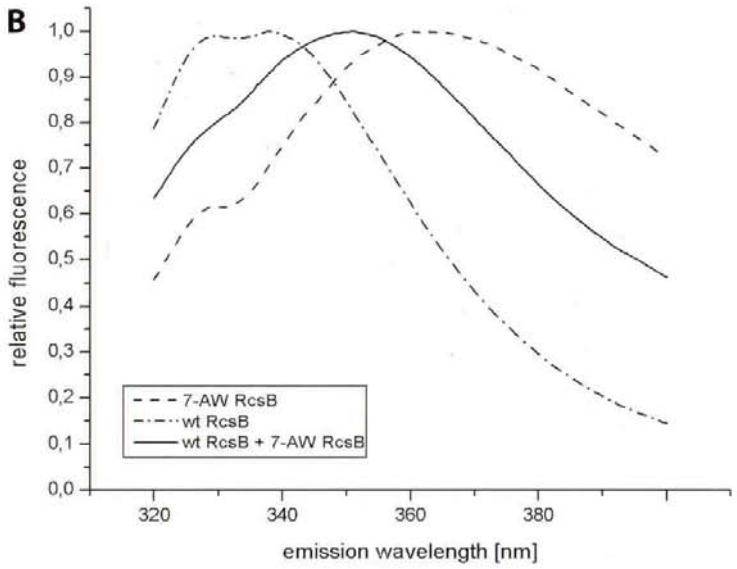
Fig. 10.2. DNA-binding activity of modified RcsB proteins. The interaction with the RcsAB box in a 183-bp DNA fragment was analyzed by the EMSA technique using 30 pmol RcsB and 300 pmol RcsA. The protein/DNA complexes were analyzed by autoradiography after separation on an 8% native polyacrylamide gel. *I* 5-FW/5-pFF-RcsB; *II* 5-FW-RcsB; *III* 7-AW-RcsB; *IV* 5-HW-RcsB; *wt* native RcsB; *control* DNA without protein; *arrows*: *A* RcsAB/DNA complex; *B* free DNA

ples at approximately λ_{ex}(max) of 292 nm. The fluorescence emission λ_{em}(max) for the native RcsB and also for the modified analogues 5-FW-RcsB, 5-HW-RcsB and 5-FW/5-pFF-RcsB upon excitation of λ_{ex} 292 nm is within the range of 335 to 340 nm. An exception is the analogue 7-AW-RcsB with a considerable red-shifted λ_{em}(max) of 362 nm. This agrees with protein spectra obtained after fluorescence labeling of other single Trp-containing proteins (Wong and Eftink 1997). The apparent quantum yield for all samples differs considerably and was lowest for 7-AW-RcsB, followed by native RcsB, and highest quantum yields were detected from 5-FW-RcsB and 5-FW/5-pFF-RcsB. The additional labeling of the Phe residues in 5-FW-RcsB did not result in a significant shift of λ_{ex}(max) or λ_{em}(max).

Analysing Protein Interactions by Steady State Fluorescence Spectroscopy

It is speculated that RcsB is able to activate distinct promoters in *E. coli* through binding of homooligomers, and also the formation of homooligomers in solution can be assumed. We now analyzed RcsB homooligomerization by taking advantage of a shifted λ_{em}(max) after substitution of the internal Trp residue with spectrally enhanced analogues. Most suited for interaction studies was the analogue 7-AW-RcsB, exhibiting a red shift of λ_{em}(max) of approximately 22 nm after excitation at 292 nm. An equimolar mixture of 7-AW-RcsB and native RcsB was analyzed under conditions identical to those for the samples of the individual proteins (Fig. 10.3B). The λ_{em}(max) of the mixture was blue-shifted at 350 nm and just in between the emission maxima of the two pure samples, sug-





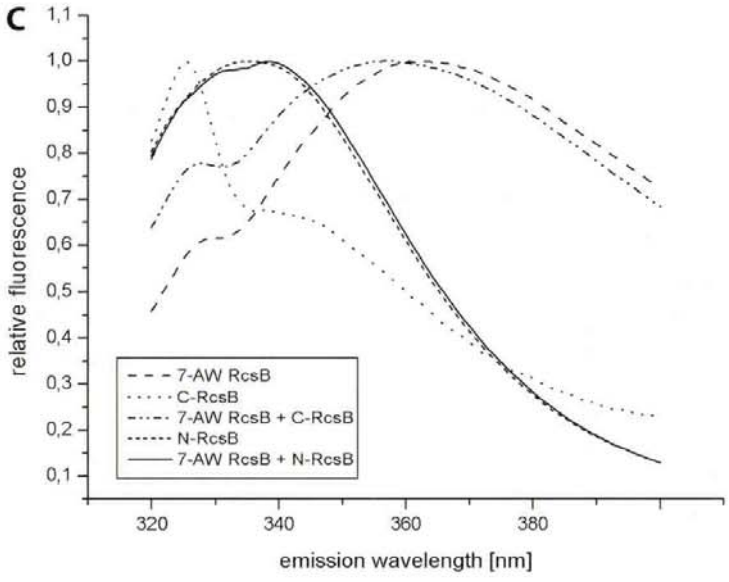


Fig. 10.3 A–C. Interaction of modified RcsB proteins analyzed by fluorescence spectroscopy. Peak normalized fluorescence emission spectra of native (wt) and modified RcsB proteins. Spectra were recorded at pH 6.4 and 20 °C with an excitation at 292 nm. The spectra were normalized to 1.0 at the corresponding λ_{em}(max). **A** Emission spectra of native and modified RcsB proteins. **B** Interaction of 7-AW-RcsB and native RcsB. **C** Interaction of 7-AW-RcsB with the N- and C-terminal RcsB domains

gesting conformational changes of the RcsB proteins upon complex formation, and thus providing evidence for a homooligomerization of RcsB. A similar effect was observed by analyzing an equimolar mixture of 7-AW-RcsB and 5-HW-RcsB (data not shown). We further analyzed the interaction of 7-AW-RcsB with the purified individual N- and C-terminal domains of RcsB (Fig. 10.3C). Interestingly, no spectral shift of λ_{em}(max) with 7-AW-RcsB/N-RcsB was detected, whereas a clear blue shift from 362 to 355 nm was obtained for 7-AW-RcsB/C-RcsB, giving first evidence for an involvement of the C-terminal effector domain in the oligomerization of RcsB.

Conclusions

We have demonstrated that cell-free expression could be a versatile tool for the high-level production of uniformly labeled and spectrally enhanced protein samples. No negative effects due to the utilized amino acid analogues on the kinetics or efficiency of the cell-free protein production have been observed. In addition, the constructed RcsB alloproteins showed changes neither in their interaction with the coinducer RcsA nor in their DNA-binding activities. From the analyzed amino acid derivatives, the incorporation of the analogue 7-AW proved to be best for the spectral enhancement of RcsB and the λ_{em}(max) was red-shifted for more than 20 nm. The spectroscopic characterization of interaction studies with modified RcsB analogues, native RcsB and with the purified N- and C-terminal RcsB domains revealed an oligomerization mechanism which seems to be based on interactions of the C-terminal DNA-binding domains. This would be in contrast to other homologous transcriptional regulators, where the interface of protein dimers is mostly formed by the N-terminal domains. The presented approach could therefore be useful for the fast and easy generation of spectrally enhanced protein samples suitable for protein interaction studies.

Acknowledgements

We are grateful to Clemens Glaubitz and Jakob Lopez for valuable discussions. We further thank Birgit Schäfer and Soma Sengupta for helpful assistance.

References

Carballes F, Bertrand C, Bouche JP, Cam K (1999) Regulation of *Escherichia coli* cell division genes *ftsA* and *ftsZ* by the two-component system *rcsC-rcsB*. Mol Microbiol 25: 442–450

Davalos-Garcia M, Conter A, Toesca I, Gutierrez C, Cam K (2001) Regulation of *osmC* gene expression by the two-component system *rcsB-rcsC* in *Escherichia coli*. J Bacteriol 183: 5870–5876

Gervais FG, Drapeau GR (1992) Identification, cloning, and characterization of *rcsF*, a new regulator gene for exopolysaccharide synthesis that suppresses the division mutation *ftsZ84* in *Escherichia coli* K-12. J Bacteriol 174: 8016–8022

Gervais FG, Phoenix P, Drapeau GR (1992) The *rcsB* gene, a positive regulator of colanic acid biosynthesis in *Escherichia coli*, is also an activator of *ftsZ* expression. J Bacteriol 174: 3964–3971

Hogue CWV, Rasquinha I, Szabo AG, MacManus JP (1992) A new intrinsic fluorescent probe for proteins. FEBS Lett 310: 269–272

Kelm O, Kiecker C, Geider K, Bernhard F (1997) Interaction of the regulator proteins RcsA and RcsB with the promoter of the operon for amylovoran biosynthesis in *Erwinia amylovora*. Mol Gen Genet 256: 72–83

Laue TM, Sklenar DF, Eaton S, Ross JBA (1993) 5-hydroxytryptophan as a new intrinsic probe for investigating protein-DNA interactions by analytical ultracentrifugation. Study of the effect of DNA on self-assembly of the bacteriophage λ cI repressor. Biochemistry 32: 2469–2472

Mohammadi F, Prentice GA, Merrill AR (2001) Protein-protein interaction using tryptophan analogues: novel spectroscopic probes for toxin-elongation factor-2 interactions.Biochem 40: 10273–10283

Ross JBA, Senear DF, Waxman E, Kombo BB, Rusinova E, Huang YT, Laws WR, Hasselbacher CA (1992) Spectral enhancement of proteins: biological incorporation and fluorescence characterization of 5-hydroxytryptophan in bacteriophage λ cI repressor. Proc. Natl Acad Sci USA 89: 12013–12027

Ross JBA, Szabo AG, Hogue CWV (1997) Enhancement of protein spectra with tryptophan analogs: fluorescence spectroscopy of protein–protein and protein–nucleic acid interactions. Methods Enzymol 278: 151–190

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd ed. Ford N, Nolan C, Ferguson M (eds). Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Soumillion P, Jespers L, Vervoort J, Fastrez J (1995) Biosynthetic incorporation of 7-azatryptophan into the phage lamda lysozyme: estimation of tryptophan accessibility, effect on enzyme activity and protein stability. Protein Eng 8: 451–456

Spirin AS, Baranov VI, Ryabova LA, Ovodov SY, Alakhov YB (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. Science 242: 1162–1164

Stout V, Gottesman S (1990) RcsB and RcsC: a two-component regulator of capsule synthesis in *Escherichia coli*. J Bacteriol 172: 659–669

Virlogeux I, Waxin H, Ecobichon C, Lee JO, Popoff MY (1996) Characterization of the *rcsA* and *rcsB* genes from *Salmonella typhi*: *rcsB* through *tviA* is involved in regulation of Vi antigen synthesis. J Bacteriol 178: 1691–1698

Wehland M, Bernhard F (2000) The RcsAB box. Characterization of a new operator essential for the regulation of exopolysaccharide biosynthesis in enteric bacteria. J Biol Chem 275: 7013–7020

Wong CY, Eftink MR (1997) Biosynthetic incorporation of tryptophan analogues into staphylococcal nuclease: effect of 5-hydroxytryptophan and 7-azatryptophan on structure and stability. Protein Sci 6: 689–697

(P 12)

Mit neuer Methode lassen sich Membranproteine “knacken”. Erstmals können ausreichende Mengen begehrter Proteine produziert werden – Auszeichnung für Christian Klammt. *Forschung Frankfurt*, **3**, 9-10. (2005)

Aussagen zur entwicklungsbiologischen Rolle der jeweiligen Gene, weil sämtliche Mutationen in jeder einzelnen Zelle des Gesamtorganismus' vorhanden sind und nicht – wie bei einer Vielzahl menschlicher Erkrankungen – nur in bestimmten Geweben auftreten und sich erst im Laufe des Lebens entwickeln. Im Rahmen des Forschungsprojekts EUCOMM werden nun Verfahren eingesetzt, die zeitlich und gewebespezifisch kontrollierbare (konditionale) Mutationen auslösen.

An dem auf Initiative des Deutschen Genfallenkonsortiums ins Leben gerufenen Projekt sind die wichtigsten europäischen Mausmutagenesezentren beteiligt. In dem EUCOMM arbeiten zehn Forschungsgruppen aus vier europäischen Ländern unter der Leitung des GSF-Forschungszentrums für Umwelt und Gesundheit in Neuherberg und des Sanger Instituts des Wellcome Trusts in Hinxton mit der Biotechnologiefirma GeneBridges zusammen. Darüber hinaus sind Wissenschaftler des Universitätsklinikums Frankfurt, des Max-Planck-Instituts für Molekulare Genetik, Berlin, der Technischen Universität Dresden, des Institute Clinique de la Souris, Straßburg, des Europäischen Laboratoriums für Molekularbiologie (EMBL), Monterontondo, der Mammalian Genetics Unit des Medical Research Councils (MRC), Harwell und des Deutschen Ressourcenzentrums für Genomforschung(RZPD), Heidelberg, in die Zusammenarbeit eingebunden. ◆

Mit neuer Methode lassen sich Membranproteine »knacken«

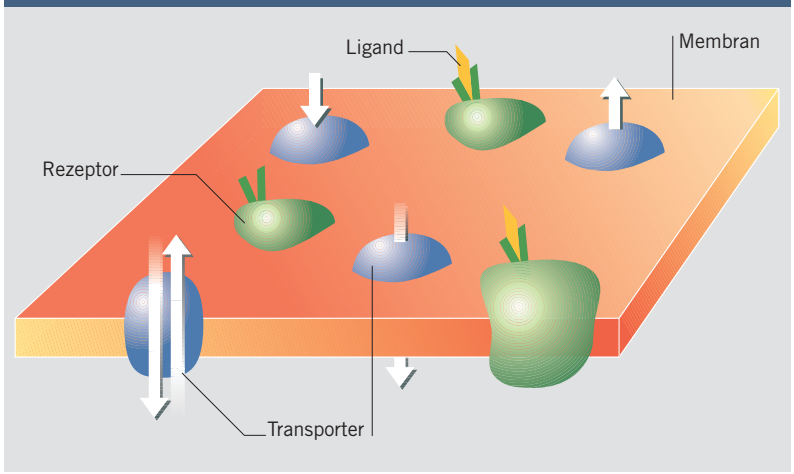
Erstmals können ausreichende Mengen begehrter Proteine produziert werden – Auszeichnung für Christian Klammt

Membranproteine lassen sich in konventionellen zellulären »Proteinfabriken« nur sehr schwer produzieren. Daher gingen Wissenschaftler am Institut für Biophysikalische Chemie der Universität Frankfurt jetzt einen völlig neuen unkonventionellen Weg, um die insbesondere für die Entwicklung zielgenauer Medikamente entscheidenden Schlüsselproteine zu erzeugen. Christian Klammt hat in seiner Doktorarbeit in der Arbeitsgruppe von Dr. Frank Bernhard eine neue Technik zur Produktion größerer Mengen funktioneller Membranproteine etabliert. Für seinen Beitrag zu diesem Thema wurde Klammt mit dem auf 10000 Euro dotierten FEBS Journal Preis für den besten Artikel im Journal der »Federation of European Biochemical Societies« im Jahr 2004 ausgezeichnet. Christian Klammt ist der erste Träger dieser 2004 erstmals ausgelobten Anerkennung.

Membranproteine sind die zentralen Schlüsselelemente in der Kommunikation jeder Zelle mit ihrer Außenwelt. Eingebettet in das Wasser abweisende Milieu von Lipid-Membranen sind sie an der äußersten Zellgrenze lokalisiert und besitzen eine fundamentale Rolle in vielen lebenswichtigen Prozessen: Jegliche Aufnahme externer Reize, sei es hören, fühlen, schmecken oder riechen, wird letztlich durch Membranproteine vermittelt und durch sie in das Zellinnere weitergeleitet. Ebenso werden essenzielle zelluläre Transportmechanismen, zum Beispiel von Nährstoffen, Hormonen, aber auch von Medikamenten, durch Membranproteine kontrolliert. Es erscheint daher nicht verwunderlich, dass gegenwärtig schätzungsweise mehr als 60 Prozent aller modernen Pharmazeutika in ihrer Wirkungsweise unmittelbar an Membranproteinen ansetzen.

Im krassen Gegensatz zu ihrer Bedeutung steht jedoch die wissenschaftliche Erforschung von Membranproteinen. Dabei ist insbesondere die Kenntnis der Morphologie, der dreidimensionalen Struktur der Proteine, eine unerlässliche Voraussetzung für ein Verständnis ihrer Funktion und für die Entwicklung von Arzneistoffen und Therapieansätzen. Während jeder Mensch etwa 10000 verschiedene Membranproteine besitzt, sind jedoch zurzeit weltweit trotz intensivster Bemühungen lediglich die Strukturen von nicht mehr als 40 Membranproteinen aufgeklärt. Anders als normale Proteine besitzen Membranproteine naturgemäß eine stark Wasser abweisende Oberfläche, was ihre konventionelle Synthese schwierig macht. Da die Wirtszellen oft sehr schnell absterben, können Proteinmengen, die für eine Strukturanalyse erforderlich wären, nicht einmal annähernd erreicht werden. »Die Präparation ausreichender Substanzmengen ist eine der größten Barrieren für die Erforschung von Membranprotei-

Modell Membranproteine



Ihre Wasser abweisende Natur und ihr begrenztes Vorkommen in den Lipidschichten von Zellen machen Membranproteine äußerst schwer zugänglich.

Preisverleihung: Dr. Richard Perham, Chairman of the Editorial Board and Editor-in-Chief of the FEBS Journal, zeichnet den Frankfurter Wissenschaftler Christian Klammt (rechts) für den besten Artikel im Journal der »Federation of European Biochemical Societies« im Jahr 2004 aus.

nen«, kommentiert Prof. Dr. Robert Tampé, Sprecher des Sonderforschungsbereichs »Functional Membrane Proteomics« an der Universität Frankfurt.

Der Clou der von Klammt entwickelten Methode besteht in der Verwendung von selbst präparierten zellfreien Extrakten, in denen die gesamte Maschinerie zur Synthese von Proteinen enthalten ist. »Dadurch werden toxische Effekte vollständig ausgeschaltet und wir können nun Membranproteine unterschiedlichster Art quasi über Nacht in großen Mengen herstellen«, erläutert der Preisträger. Die Technik ist zudem relativ preiswert und kann ohne weiteres in jedem durchschnittlichen biochemischen Labor in wenigen Tagen etabliert werden. Entsprechend groß ist auch die internationale Resonanz auf diese Arbeit. »Wir haben nahezu täglich Anfragen von anderen Labors, die Protokolle anfordern oder uns Mitarbeiter zum Erlernen der Technik schicken wollen«, resümiert Dr. Frank Bernhard.

Insbesondere für die Strukturanalyse mit Hilfe der Kernmagnetischen Resonanz Spektroskopie (Nuclear Magnetic Resonance, NMR) – die Universität Frankfurt ist hier ein europäisches Zentrum – bieten sich mit der entwickelten Technologie vielfältige neue Ansätze. »Es ist faszinierend: Proben, die bis vor kurzem nicht machbar waren oder deren Herstellung zumindest Monate dauerte, sind jetzt in wenigen Stunden möglich«, bestätigt Prof. Dr. Volker Dötsch, Professor am Institut für Biophysikalische Chemie und Mitglied des Zentrums für Biomolekulare Magnetische Reonanz (BMRZ). Und sein Kollege Prof. Dr. Clemens Glaubitz fügt hinzu: »Das Potenzial der zellfreien Produktion von Membranproteinen für die NMR ist noch nicht abschätzbar«.

Grundlegende Resultate dieses von der Deutschen Forschungsgemeinschaft im Rahmen des Sonderforschungsbereichs 628 geförderten Projekts wurden im Wissenschaftsjournal der »Federation of European Biochemical Societies« (FEBS Journal, früher European Journal of Biochemistry) publiziert: »High level cell-free expression and specific labelling of integral membrane proteins«, Christian Klammt, Frank Löhr, Birgit Schäfer, Winfried Haase, Volker Dötsch, Heinz Rüterjans, Clemens Glaubitz and Frank Bernhard. European Journal of Biochemistry (2004) 271, Seite 568–580. Für diese Arbeit wurde nun Christian Klammt ausgezeichnet. Das Forschungsvorhaben profitierte wesentlich von einer von Prof. Dr. Heinz Rüterjans initiierte Kooperation des Instituts für Biophysikalische Chemie mit dem »Institute for Protein Research« in Pushchino/Moskau. Die Arbeit ist ein Gemeinschaftsprojekt zwischen den Abteilungen Dötsch (Lösungs-NMR) und Glaubitz (Festkörper-NMR) des Instituts für Biophysikalische Chemie und des Max-Planck-Instituts für Biophysik und dokumentiert das Synergiepotenzial innerhalb des neu gegründeten Sonderforschungsbereichs 628 und des »Center for Membrane Proteomics« am Standort der Universität Frankfurt. ◆

In der ersten Liga dabei

Institut für Religionsphilosophische Forschung gewinnt weltweiten Wettbewerb um Templeton Research Lectures

Beherrscht die Materie den Geist? Biofakt oder Artefakt – sind wir auf dem Weg zu einem neuen Begriff des Lebens? Gibt es eine biologische Basis für den Glauben? Die spannenden Fragen, die Geisteswissenschaften und Naturwissenschaften gleichermaßen beschäftigen, sind formuliert, und damit hat das Frankfurter Organisationskomitee der Templeton Lectures auch das Programm für die kommenden drei Jahre umrissen. Insgesamt hat das Institut für Religionsphilosophische Forschung (IRF) der Johann Wolfgang Goethe-Universität zunächst fast 400 000 Dollar zur Verfügung, um im Dialog mit den Naturwissenschaften Antworten auf die schwierigen Fragen zum menschlichen Bewusstsein und zu den materiellen Bedingtheiten unseres Denksystems zu finden. Diese Summe kann sich durch Verlängerung der Förderung um ein viertes Jahr bis auf 500 000 Dollar erhöhen.

Frankfurt hat, gemeinsam mit der Vanderbilt University, Knoxville (USA), den weltweiten Wettbewerb mit namhaften Universitäten für sich entschieden und wurde mit den Templeton Research Lectures in das Programm des Metanexus Institutes, Philadelphia (USA), aufgenommen. Finanziell unterstützt wird dieses Programm von der Templeton Foundation; sie fördert globale Initiativen, die sich mit Grenzfragen zwischen Theologie und Naturwissenschaften auseinandersetzen. »Dass wir uns als Philosophen und Religionswissenschaftler mit den Nahtstellen zwischen Geistes- und Naturwissenschaften beschäftigen, gehört zur Frankfurter Tradition der Kritischen Theorie. So werden wir auch die Dominanz der Naturwissenschaften, wenn es um die Erklärung der Welt und der individuellen Existenz geht, im kritischen Diskurs beleuchten«, erläutert der Direktor des Instituts für Religionsphilosophische Forschung, Prof. Dr. Thomas M. Schmidt, der gemeinsam mit dem Theologen Dr. Michael Parker den erfolgreichen Antrag für die Templeton Research Lectures gestellt hat. Das Programm trägt den Titel »Ko-Schöpfer oder Produkt der Natur? Die menschliche Person im Licht von Neurophilosophie, Biofaktizität und Evolutionsbiologie«.

Christian Klammt, Westerbachstrasse 268d, 65936 Frankfurt am Main
Telefon: 069-344118, Email: c.klammt@bpc.uni-frankfurt.de

Lebenslauf

Persönliche Informationen

Geburtsdatum:	25. Oktober 1976
Geburtsort:	Frankfurt am Main
Familienstand:	ledig
Nationalität:	Deutsch

Naturwissenschaftliche Ausbildung

Promotion

seit 10/2002 — Promotion am Institut für Biophysikalische Chemie der J.W. Goethe-Universität Frankfurt am Main

Projekt: Präparative zellfreie Expression von funktionellen integralen Membranproteinen und Untersuchung des humanen Endothelin B Rezeptors.

Betreuer: Dr. Frank Bernhard und Prof. Volker Dötsch

Studium

04/2002 — Abschluss: Diplom-Biochemiker an der J.W. Goethe-Universität Frankfurt am Main

10/2001 – 04/2002 — Diplomarbeit am Institut für Biophysikalische Chemie der J.W. Goethe-Universität Frankfurt am Main

Projekt: Etablierung eines zellfreien Systems zur präparativen Proteinsynthese.

Betreuer: Dr. Frank Bernhard und Prof. Heinz Rüterjans

10/1997 – 09/2001 — Studium der Biochemie an der J.W. Goethe-Universität Frankfurt am Main

Akademische Lehrer:
E. Bamberg, J. Bereiter-Hahn, V. Dötsch, E. Egert, J. Engels, K.-D. Entian, H. Fasold, J. Feierabend, K. Fendler, C. Griesinger, M. Göbel, U. Günther, K. Hainer, B. Kohlbesen, A. Kröger, R. Lancaster, B. Ludwig, H. Michel, W. Müller, R. Prinzinger, R. Rüterjans, H. Schwalbe, J. Soppa, H. Steiger, R. Tampé, C. Winter, H. Zimmermann.

Christian Klammt, Westerbachstrasse 268d, 65936 Frankfurt am Main
Telefon: 069-344118, Email: c.klammt@bpc.uni-frankfurt.de

Schule und Zivildienst

08/1996 – 09/1997	Zivildienst beim mobilen sozialen Hilfsdienst des evangelischen Regionalverbandes Frankfurt am Main
06/1996	Abitur
08/1993 – 06/1996	Gymnasiale Obersturfe (Friedrich-Dessauer-Gymnasium in Frankfurt am Main)
08/1987 – 07/1993	Gymnasium (Helene-Lange-Schule in Frankfurt am Main)
08/1983 – 07/1987	Grundschule (Albrecht-Dürer-Schule in Frankfurt am Main)

Auszeichnungen

07/2005 **FEBS Journal Preis für junge Wissenschaftler 2004**

Für den Erstautor in der besten Veröffentlichung im FEBS Journal im Kalenderjahr 2004 für die Veröffentlichung:

C. Klammt, F. Löhr, B. Schäfer, W. Haase, V. Dötsch, H. Rüterjans, C. Glaubitz and F. Bernhard: ***High level cell-free expression and specific labeling of integral membrane proteins.*** *Eur. J. Biochem., 271, 568-580. (2004)*

05/2005 **Posterpreis** für das beste Poster auf dem 5. Transport Kolloquium, Rauischholzhausen/ Deutschland für das Poster:

Functional cell-free expression of G-protein coupled receptors in a preparative scale.

05/2003 **Posterpreis** für das beste Poster auf dem internationalen Kongress für Protein Expression und Protein Funktion, Berlin/ Deutschland, für das Poster:

Cell-free expression of membrane proteins.

Christian Klammt, Westerbachstrasse 268d, 65936 Frankfurt am Main
Telefon: 069-344118, Email: c.klammt@bpc.uni-frankfurt.de

Vorträge

06/2006 „High level and functional expression of integral membrane proteins by individual cell-free expression systems" im Rahmen des Seminars über Membran-Biologie im Max-Planck-Institut für Biophysik in Frankfurt am Main/ Deutschland.

06/2005 "High level functional expression of diverse types of integral membrane proteins by an individual cell-free expression system" auf dem 30. FEBS Kongress und der 9. IUBMB Konferenz in Budapest/ Ungarn.

Konferenzen und Workshops

20. – 25.08.2006 **22nd International Conference on Magnetic Resonance in Biological Systems (ICMRBS)**, Göttingen/ Deutschland.

24. – 29.06.2006 **31st FEBS Congress on Molecules in Health & Disease**, Istanbul/ Türkei.

12. – 13.11.2005 **ESBF Workshop "Crystal and Solution Structures of Proteins: Possibilities and Limitations of Methodology"**, Frankfurt am Main, Deutschland.

26. – 29.09.2005 **GDCH Magnetic Resonance Division (High-Field Bio-NMR Symposium & MR meets Theory)**, Mainz, Deutschland.

02. – 07.07. 2005 **30th FEBS Congress and 9th IUBMB Conference**, Budapest, Ungarn.

20. – 21.05.2005 **5th Transport Colloquium**, Rauischholzhausen, Germany.

16. – 21.01.2005 **21st International Conference on Magnetic Resonance in Biological Systems (ICMRBS)**, Hyderabad, Indien.

12. – 16.07.2004 **IUBMB/ASBMB 2004**, Boston, USA.

02. – 04.03.2004 **CMP workshop on BN-PAGE, BN/SDS-, and doubled SDS PAGE by Prof. Hermann Schägger**, Frankfurt am Main, Deutschland.

14.01.2004 **GPCRs as Target for the Treatment of CNS Diseases**, Frankfurt am Main, Deutschland.

21. – 22.11.2003 **7th European Large Scale Facility User Meeting 2003**, Oosterbeek, Niederlande.

05. – 09.07.2003 **4th European Biophysics Congress**, Alicante, Spanien.

25. – 28.05.2003 **International Congress on Protein Expression & Protein Function**, Berlin, Deutschland.

29. – 31.01.2003 **EMBO/ILL Workshop on Deuterium Labeling for NMR and Neutron Diffraction 2003**, Grenoble, Frankreich.

17. – 20.10.2002 **6th Large Scale Facility (LSF) user meeting 2002**, Montecatini, Italien.

Publikationen

JOURNAL ARTIKEL

C. Klammt, D. Schwarz, N. Eifler, A. Engels, J. Piehler, V. Dötsch and F. Bernhard: **High-level cell-free expression of diverse GPCRs for structural studies.** *J. Struct. Biol.*, in preparation. (2006)

C. Klammt, A. Srivastava, N. Eifler, D. Schwarz, H. Michel, V. Dötsch and F. Bernhard: **High-level cell-free expression of functionally folded human endothelin B receptor: Identification of the first transmembrane segment as a primary determinant for ligand binding and homodimerization.** *EMBO Journal***,** in preparation. (2006)

C. Klammt, D. Schwarz, K. Fendler, W. Haase, V. Dötsch and F. Bernhard: **Evaluation of detergents for the soluble expression of α-helical and β-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system.** *FEBS Journal*, **272**, 6024-6038. (2005)

N. Trbovic, **C. Klammt**, A. Koglin, F. Löhr, F. Bernhard, and V. Dötsch: **Efficient Strategy for the Rapid Backbone Assignment of Membrane Proteins.** *J. Am. Chem. Soc.*, **127**, 13504-13505. (2005)

Christian Klammt, Westerbachstrasse 268d, 65936 Frankfurt am Main
Telefon: 069-344118, Email: c.klammt@bpc.uni-frankfurt.de

C. Klammt, F. Löhr, B. Schäfer, W. Haase, V. Dötsch, H. Rüterjans, C. Glaubitz and F. Bernhard: **High level cell-free expression and specific labeling of integral membrane proteins.** *Eur. J. Biochem.*, **271**, 568-580. (2004)

REVIEW ARTIKEL

C. Klammt, D. Schwarz, V. Dötsch and F. Bernhard: **Cell-free production of integral membrane proteins on a preparative scale.** *Meth. Mol. Biol.*, accepted. (2006)

D. Schwarz, **C. Klammt**, A. Koglin, F. Löhr, B. Schneider, V. Dötsch and F. Bernhard. **Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies.** *Methods*, in press. (2006)

C. Klammt, D. Schwarz, F. Löhr, B. Schneider, V. Dötsch and F. Bernhard: **Cell-free expression as an emerging technique for the large scale production of integral membrane protein.** *FEBS Journal*, **273**, 4141-4153. (2006)

A. Koglin, **C. Klammt**, N. Trbovic, D. Schwarz, B. Schneider, B. Schäfer, F. Löhr, F. Bernhard and V. Dötsch: **Combination of cell-free expression and NMR spectroscopy as a new approach for structural investigation of membrane proteins.** *Magn. Reson. Chem.*, **44**, 17-23. (2006)

BUCH KAPITEL

C. Klammt, D. Schwarz, B. Schneider, F. Löhr, I. Lehner, C. Glaubitz, V. Dötsch and F. Bernhard**: Cell-free expression of integral membrane proteins for structural studies.** In: *Cell-free expression techniques*, A. Spirin (ed.), Wiley-VCH, Weinheim, submitted. (2006)

F. Bernhard, **C. Klammt**, and H. Rüterjans: **DNA recombination and protein expression.** In: *Comprehensive Medicinal Chemistry II*, in press. (2006)

C. Klammt, F. Bernhard and H. Rüterjans: **^{13}C- and ^{15}N-Isotopic Labeling of Proteins**, In: *Molecular Biology in Medicinal Chemistry*, T. Dingermann, G. Folkers, H. Steinhilber (eds.), pp. 269-292, Wiley-VCH, Weinheim. (2003)

K. Sengupta, **C. Klammt**, F. Bernhard and H. Rüterjans: **Incorporation of fluorescence labels into cell-free produced proteins.** In: *Cell-Free Protein Expression*, J.R. Swartz (ed.) pp. 81-88, Springer Verlag, Berlin, Heidelberg, New York. (2003)

Eidesstattliche Versicherung

Ich erkläre hiermit an Eides statt, dass ich die vorgelegte Dissertation über „Functional and structural analysis of cell-free produced transporters and G-protein coupled receptors: Development of new techniques for the fast and efficient production of integral membrane proteins (Funktionelle und strukturelle Analyse von zellfrei produzierten Transportern und G-Protein gekoppelten Rezeptoren: Entwicklung neuer Techniken zur schnellen und effizienten Produktion von integralen Membranproteinen)“ selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe. Ich erkläre weiterhin, dass Entlehnungen aus Schriften, soweit sie in der Dissertation nicht ausdrücklich als solche bezeichnet sind, nicht stattgefunden haben. Ich habe bisher an keiner anderen Universität ein Gesuch um Zulassung zur Promotion eingereicht oder die vorliegende oder eine andere Arbeit als Dissertation vorgelegt.

Frankfurt am Main, den 21. September 2006

_____

Christian Klammt