

# Customization of cell-free lysates for membrane protein expression and NMR analysis

## Dissertation

zur Erlangung des Doktorgrades  
der Naturwissenschaften

vorgelegt beim Fachbereich  
Biochemie, Chemie und Pharmazie (FB14)  
der Johann Wolfgang Goethe-Universität  
In Frankfurt am Main



von Roman Levin  
aus St. Petersburg

Frankfurt am Main, 2023  
(D30)



Vom Fachbereich Biochemie, Chemie und Pharmazie (FB14)  
der Johann Wolfgang Goethe-Universität  
als Dissertation angenommen

Dekan: Prof. Dr. Clemens Glaubitz

Gutachter: Prof. Dr. Volker Dötsch  
Prof. Dr. Clemens Glaubitz

Datum der Disputation:



## **Eidesstattliche Erklärung**

Ich versichere hiermit, dass ich die vorgelegte Dissertation mit dem Titel „Customization of cell-free lysates for membrane protein expression and NMR analysis“ selbstständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, dass alle Entlehnungen aus anderen Schriften mit Angabe der betreffenden Schrift gekennzeichnet sind.

Ich versichere, die Grundsätze der guten wissenschaftlichen Praxis beachtet, und nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Datum:

Unterschrift:



## Acknowledgements

It was a great time working on my doctoral thesis that I certainly will never forget. I am deeply grateful for the privilege to get to know so many wonderful people who not only contributed to this work, but also gave me advice, shared experiences and always lifted my spirits on gloomy days.

First and foremost, I would like to thank my PIs Frank Bernhard and Volker Dötsch who gave me this opportunity to carry out scientific work in their laboratory. Thank you for your advice and your guidance, but also for providing me with degrees of freedom to pursue different projects and ideas.

Secondly, I would like to give big thanks to my coworkers and friends in the cell-free group Julia Mezhyrova, Zoe Köck and Simon Umbach. It was great working with you, share thoughts, discuss science, but also to let off some steam over a couple of beers. Much love to you and I am sure our paths will cross a lot in the future.

A very special thanks goes out to all the collaboration partners:

- Frank Löhr who measured all samples with me and gave me valuable input on sample quality and pointed out room for improvement.
- Jakob Gebel with whom I worked on the refolding and assignment of CK1δ. Although this “side” project will not be part of this thesis, I highly value his expertise and competence. Jacob also gave me a lot of input and ideas during my work - thanks mate!
- Janosch Martin and René Zangl from AK Morgner who provided me with LILBID-MS data. It was a great pleasure working with you.
- Samuel Seidl from the Glaubitz lab for our work on <sup>19</sup>F labeling of GPCR probes for ssNMR using the CF system. Even though this project is also not part of this thesis, it was a blast working with you. I hope someone will pick up where we left off.
- Theresa Gewering from Arne Möllers group who performed negative stain EM.
- My bachelor students Betül Karakoc, Leah Schüller and Lara Valder who helped me in the lab. It was a great joy working with you, and I wish you all the best in the future!

Of course, a big thanks goes to the whole group too, where I met wonderful people and made a lot of new friends. It was a pleasure working, talking and laughing with you and I will always hold memories dear of our time together. Stay as you are, support each other and invite me to all the legendary AK Dötsch parties. I will try to attend as often as possible!

Furthermore, I'd like to thank Sigrid Oguzer-Fachinger who helped me through the bureaucratic jungle of the Uni, Birgit Schäfer for unfailing support in the lab and the best Midi preps, Natascha Rogova who prepared a lot of SDS gels for me and of course Manfred Strupf – legend has it, that there is nothing he can't fix.

Big thanks to all the former students and colleagues who spent time studying with me, especially my former supervisor Ines Seitzl who always encouraged me to do a PhD. It gave me a lot of confidence that I'm very grateful for.

I would also like to thank my friends and family, who supported me during the years. Big thanks to mom and dad for supporting me. I am super grateful to have such a cool sister - you are pure sunshine!

Particularly I would like to thank my grandpa Efim. When I was a child, you have already started to cultivate my curiosity for this world. You understood my child brain and taught me how to use it, always giving me something to think about. During my study years you were always sincerely interested in my work. You supported me emotionally and encouraged me to keep going. You are probably the most admirable person I ever met, and you can't even imagine what inspiration you are to me. I hope that we have many more years ahead together. Thank you!

Finally, I want to thank my precious Verena. Your unfailing love and support helped me to get through my thesis years - and I know I can be a piece of work when I am stressed out. You have listened to me when I came home enthusiastic about my progress, but also helped me to get through less optimistic periods. Life is so much better with you, and I am looking forward to all the years ahead.







---

**Table of contents**

<b>Table of contents</b>	<b>I</b>
<b>List of Figures</b>	<b>III</b>
<b>List of Tables</b>	<b>IV</b>
<b>Abbreviations</b>	<b>V</b>
<b>Summary</b>	<b>IX</b>
<b>Zusammenfassung</b>	<b>XIII</b>
<b>1. INTRODUCTION</b>	<b>1</b>
1.1 THE CELL-FREE SYSTEM	1
1.1.1 Discovery and development	1
1.1.2 <i>E. coli</i> lysates	1
1.2 EXPRESSION AND PURIFICATION OF MEMBRANE PROTEINS	4
1.2.1 MP solubilization in mimetic lipid bilayers	5
1.2.2 Apolipoprotein A-1 derivative based nanoparticles	5
1.2.3 SapA based nanoparticles for cotranslational MP solubilization	6
1.2.4 CF expression of MPs	6
1.3 MEMBRANE PROTEIN NMR	8
1.3.1 State of the art and challenges	8
1.3.2 Advances in NMR studies of CF expressed MPs	10
1.3.3 Metabolic background activities in <i>E. coli</i> lysates	11
1.3.4 Precursor based methyl labeling of proteins for NMR studies	13
1.3.5 Precursor based methyl labeling using the CF system	15
<b>2 MATERIALS</b>	<b>17</b>
2.1 EQUIPMENT	17
2.1.1 General equipment	17
2.1.2 Equipment for protein purification and analysis	17
2.2 STRAINS, CLONING AND DNA PREPARATION	18
2.2.1 <i>E. coli</i> strains	18
2.2.2 Knock-out and screening Primers	18
2.2.3 Equipment for cloning and DNA preparation	20
2.3 CHEMICALS AND REAGENTS	20
2.3.1 General chemicals	20
2.3.2 Isotopes	21
2.3.3 Detergents and lipids	22
2.3.4 Enzymes, protein and antibodies	22
2.4 CULTURE MEDIA AND ANTIBIOTICS	22
2.4.1 Culture media and antibiotic stocks	22
2.5 BUFFERS AND STOCK SOLUTIONS	23
2.5.1 Cell-free expression and lysate processing	23
2.5.2 Protein purification and analysis	23
2.6 SOFTWARE	26
<b>3 METHODS</b>	<b>27</b>
3.1 MOLECULAR BIOLOGY	27
3.1.1 Agarose gel electrophoresis	27
3.1.2 DNA transformation by heat shock	27
3.1.3 DNA transformation by electroporation	27
3.1.4 DNA preparation	27
3.1.5 Restriction and ligation cloning	28
3.1.6 Genome engineering in <i>E. coli</i>	28
3.2 PROTEIN EXPRESSION PURIFICATION FROM <i>E. COLI</i>	30
3.2.1 Expression and purification of T7 RNAP	30
3.2.2 Expression and purification of membrane scaffold proteins	31
3.2.3 Expression and purification of SapA	32
3.2.4 Expression and purification of BCAAT	33

## Table of contents

---

3.3 CF EXPRESSION OF PROTEINS.....	33
3.3.1 S30 extract preparation .....	33
3.3.2 Preassembly of MSP NDs .....	34
3.3.3 Preassembly of SapNPs.....	34
3.3.4 Preconversion of methyl labeled precursors .....	35
3.3.5 CF expression setup.....	35
3.3.6 Strep purification of CF expressed MPs .....	35
3.3.7 Expression and purification of cyclophilin D .....	36
3.3.8 Expression and purification of PR .....	36
3.3.9 FFAR <sub>2</sub> expression and purification for NMR analysis.....	36
3.4 PROTEIN ANALYSIS .....	37
3.4.1 GFP fluorescence assay .....	37
3.4.2 Quantification of PR.....	37
3.4.3 SDS-PAGE .....	37
3.4.4 Coomassie staining .....	37
3.4.5 Western blot .....	37
3.4.6 Radioligand filter binding assays .....	38
3.4.7 Size exclusion chromatography.....	38
3.4.8 LILBID native-MS .....	38
3.4.9 Negative stain electron microscopy .....	39
3.4.10 CypD NMR experiments.....	39
3.4.11 PR NMR experiments.....	39
3.4.12 FFAR <sub>2</sub> NMR experiments .....	39
<b>4. RESULTS.....</b>	<b>41</b>
4.1 SAPNPs FOR COTRANSLATIONAL STABILIZATION OF MPs .....	41
4.1.1 Spontaneously formed SapNPs during purification .....	41
4.1.2 Preassembly of SapNPs for MP expression.....	43
4.1.3 Coassembly of SapA with lipids and MPs .....	46
4.1.4 Coexpression of SapA with PR .....	51
4.1.5 Applicability of SapNPs in the CF system .....	52
4.2 LYSATE ENGINEERING FOR REDUCED <sup>15</sup> N LABEL SCRAMBLING .....	57
4.2.1 Identification of mutation targets.....	57
4.2.2 Engineering of <i>E. coli</i> A19 and characterization of strains .....	59
4.2.3 NMR analysis of CypD expressed with engineered lysates .....	63
4.2.4 Analysis of AA scrambling in A19 Stablelabel by GFP expression .....	65
4.2.5 Precursor preconversion for methyl labeling of BCAA .....	66
4.2.6 First NMR attempts of CF synthesized GPCR in NDs.....	69
<b>5. DISCUSSION .....</b>	<b>71</b>
5.1 Introducing SapNPs in the bacterial CF system .....	71
5.2 The future of membrane mimetics in CF systems .....	72
5.3 Engineered lysates for stable isotope labeling and beyond .....	74
5.4 Outlook on NMR investigations with cotranslationally solubilized GPCR.....	76
<b>6. REFERENCES.....</b>	<b>79</b>
<b>7. APPENDIX.....</b>	<b>93</b>
7.1 DECLARATION OF CONTRIBUTIONS.....	93
7.2 CURRICULUM VITAE .....	97

## List of Figures

Nr.	Title	Chapter	Page
1	Cell free protein synthesis in a two-compartment system	1.1.2	3
2	Examples of devices for CECF expression	1.1.2	4
3	Different CF expression modes for MP synthesis.	1.2.4	8
4	Deuterium-hydrogen back exchange at the methyl group of L-Ala	1.3.3	12
5	Metabolic pathways in <i>E. coli</i> for AA synthesis from precursors	1.3.4	15
6	IlvE catalyzed conversion of $\alpha$ -ketoacid precursors	1.3.5	16
7	Purification of SapNPs (EC) and application for PR solubilization	4.1.1	42
8	Analysis of preassembled SapNPs (DOPG)	4.1.2	43
9	PR solubilization with preassembled SapNP (DOPG) particles	4.1.2	45
10	Negative stain EM images of NDs, PR/NDs, SapNPs and PR/SapNPs	4.1.2	46
11	Coassembly kinetics of SapA and lipids in the CF mixture with or without PR expression	4.1.3	47
12	Coassembly of PR/SapNPs with nascent PR, purified SapA and DOPG liposomes	4.1.3	48
13	Sample quality of coassembled PR/SapNPs at different template and lipid ratios	4.1.3	49
14	PR/SapNP sample quality analysis after synthesis by coassembly	4.1.3	50
15	Analysis of SapNP/PR particles obtained by coexpression of PR and SapA in the CF reaction	4.1.4	51
16	Cotranslational solubilization of GFP chimeras of various MPs	4.1.5	52
17	Expression and purification of bacterial multidrug efflux proteins SugE-GFP-his and EmrE-strep	4.1.5	53
18	Quality control of Tts $\beta_1$ AR-GFP-his purified via IMAC	4.1.5	54
19	Expression and purification of thermostabilized Hts $\beta_1$ AR variants	4.1.5	55
20	Investigation of AA conversions in the CF reaction	4.2.1	57
21	Scheme of enzymes involved in most prominent $^{15}\text{N}$ scrambling in CF reactions	4.2.1	58
22	Screening for gene deletion mutants of <i>E. coli</i> A19 and recovery of kanR marker	4.2.2	60
23	Auxotrophic mutations determining growth requirements of created mutants	4.2.3	62
24	Growth kinetics and CF protein production performance of selected engineered <i>E. coli</i> A19 lysates	4.2.2	63
25	Selective labeling of the previously fully assigned CypD (43-207) variant	4.2.3	64
26	Expression of GFP with A19 (white bars) and M1-3-4-6-8-9 "Stablelabel" (grey bars) lysates in absence of L-Asp, L-Asn, L-Glu, L-Gln or L-Ala, respectively	4.2.4	65
27	Preconversion of methyl labeled precursor KIV to L-Val and MOV to L-Leu	4.2.5	67
28	Conversion of BCAA precursors prior to labeling	4.2.5	68
29	Preliminary NMR data of FFAR $_2$ and PR in MSP1D1 $\Delta$ H5 NDs (POPG)	4.2.6	69

**List of tables**

<b>Nr.</b>	<b>Title</b>	<b>Chapter</b>	<b>Page</b>
1	CF expressed MPs for NMR analysis.	1.3.2	10
2	Molar masses and extinction coefficients of different MSP derivatives	3.3.2	34
3	SEC columns and operational parameters	3.4.7	38
4	Enzymes with potential AA scrambling activity in standard A19 S30 lysates	4.2.1	58
5	Constructed chromosomal mutants in strain A19	4.2.2	61

## Abbreviations

AA	Amino acid
AcP	Acetyl phosphate
AMP	Adenosin monophosphate
ADP	Adenosin diphosphate
AOA	Aminooacetate
ATP	Adenosin triphosphate
APS	Ammonium persulphate
BCAA	Branched chain amino acid
BCAT	Branched chain amino acid aminotransferase
CECF	Continuous exchange cell-free
CF	Cell-free
CFPS	Cell-free protein synthesis
cm	Centimeter
cNDs	Circularized nanodiscs
conv	Converted
CTP	Cytidine triphosphate
CV	Column volume
D-CF	Detergent cell-free expression mode
DF	Disc formation
DH7PC	1,2-diheptanoyl-sn-glycero-3-phosphocholine
DIBMA	Diisobutylene maleic acid
DM	D-malate
DMPC	1,2-dimyristol-sn-glycero-3-phosphocholine
DNA	Desoxyribonucleic acid
DON	6-diazo-5-oxo nucleucine
DONV	5 diazo-4-oxo norvaline
DOPG	1,2-dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
DPC	<i>n</i> -dodecylphosphocholine
DTT	dithiothreit
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
EL	Elution
EM	Electron microscopy
FFAR <sub>2</sub>	Free fatty acid receptor 2
FM	Feeding mixture
fw	Forward
FRT	Flippase recognition target
FRUIT	Flexible recombineering using integration of <i>thyA</i>
FT	Flow-through
G protein	Guanine nucleotide-binding proteins
GDN	Glyco diosgenin
GFAT	L-glutamine-D-fructose-6-phosphate aminotransferase
GFP	Green fluorescent protein
GOI	Gene of interest
GPCR	G-protein coupled receptor
GSH	Glutathione reduced
GSSG	Glutathione oxidized
GTP	Guanosine triphosphate
h	Hour
His-tag	Polyhistidine tag
HCl	Hydrochloric acid
HEPES	2-(4-(2-hydroxyethyl)-1-piperazinyl)-ethanosulfonic acid
HMQC	Heteronuclear multiple quantum correlation
HSQC	Heteronuclear single quantum coherence
Htsβ <sub>1</sub> AR	Thermostabilized β <sub>1</sub> adrenergic receptor from human
IMAC	Ion metal affinity chromatography
IPTG	Isopropyl-D-thiogalactopyranosid
IVTT	<i>In vitro</i> transcription and translation
K	Kelvin

## Abbreviations

---

KanR	Kanamycin resistance
kb	Kilobase
kDa	Kilodalton
KIV	2-ketoisovalerate
KO	Knock-out
KOAc	Potassium acetate
L	Liter
L-CF	Lipid cell-free expression mode
LB	Luria Bertani
LILBID	Laser induced liquid bead ion desorption
M	Molar
Mg <sup>2+</sup>	Magnesium ion
min	Minute
mL	Milliliter
mM	Millimolar
MP	Membrane protein
ms	Millisecond
MS	Methionine sulfoximine
MSP	Membrane scaffold protein
MOV	4-methyl-2-oxovalerate
MWCO	Molecular weight cutoff
N-Acetylglucosamine	GlcNAc
ND	Nanodisc
Ni <sup>2+</sup>	Nickel ion
nm	Nanometer
ns	Nanosecond
NMR	Nuclear magnetic resonance
NP	Nanoparticle
NTP	Nucleoside triphosphate
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
P-CF	Precipitate cell-free
PBS-T	Phosphate-buffered saline Tween 20
PEP	Phosphoenolpyruvate
PIC	Protease inhibitor cocktail
PK	Pyruvate kinase
PLP	Pyridoxal phosphate
POPG	1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylglycerol
PURE	Protein synthesis using recombinant elements
ppm	Parts per million
PR	Proteorhodopsin
PCR	Polymerase chain reaction
PLP	Pyridoxal phosphate
PTM	Post translational modification
PVDF	Polyvinylidene fluoride
rev	Reverse
RM	Reaction mixture
RNA	Ribonucleic acid
rpm	Rotations per minute
RT	Room temperature
S30	Lysates centrifuged at 30,000 x g
SAIL	Stereo-array isotope labeling
SAP	Sphingolipid activator protein
SapA	Saposin A
SapNP	Saposin A based nanoparticle
SDS	Sodium dodecylsulfate
SEC	Size exclusion chromatography
SMALP	Styrene maleic acid lipid particle
SMCS	S methionine-L-cysteine sulfoximine
SOFAST	Band-Selective Optimized-Flip-Angle Short-Transient
ssNMR	Solid state NMR



t	Time
Tts $\beta_1$ AR	Thermostabilized $\beta_1$ adrenergic receptor from turkey
T7RNAP	T7 RNA polymerase
tRNA	Transfer ribonucleic acid
TRIS	Tris(hydroxymethyl)aminomethane
TROSY	Transverse relaxation optimized spectroscopy
TEV	Tobacco etch virus
U	Unit
UTP	Uridine 5'-triphosphate
V	Volt
W	Wash
WT	Wild type



## Summary

Cell-free protein synthesis (CFPS) has greatly evolved over the last decades and became a frequently used system for expression of membrane proteins (MPs). Thereby, MPs are expressed directly and translocon-independently inside provided lipid membranes, such as liposomes or nanodiscs (NDs). This technique is unique since it does not require prior solubilization of MPs in detergent for subsequent reconstitution. Due to the open nature of CF systems, the experimenter is free to adjust the environment of the expressed target. Options include expression conditions, choice of nanoparticles (NPs) and their lipid composition, redox conditions, addition of stabilizing molecules like ligands or cofactors and many more. Best established so far is the CF system based on *E. coli* lysates due to cost-effective production, easy handling, and high product yield. Most frequently, membrane scaffold protein (MSP) based NDs are used for cotranslational insertion of MPs. In the last years, the system was particularly useful for biochemical characterization of MPs, and even structural studies using nuclear magnetic resonance spectroscopy (NMR) and cryogenic electron microscopy (cryo-EM) were possible. However, new alternative systems such as saposin A based nanoparticles (SapNPs) are emerging, and their potential for CF MP production was in detail characterized in the first part of this study. Also, CF synthesis of protein samples for NMR analysis benefits from truncated metabolic pathways of amino acids (AA) in the lysates. This enables selective labeling schemes otherwise impossible in cellular systems due to spectral ambiguity. Nevertheless, few residual enzymatic AA conversions persist, leading to label dilution and ambiguous NMR spectra. The second part of this study was therefore devoted to eliminating these residual AA scrambling activities in the lysate source strain *E. coli* A19 by directed strain engineering. As a result, a new master strain producing optimized CF lysates for NMR applications was obtained.

Similar to MSP, lipid binding properties of saposin A (SapA) were adapted to establish a new NP system for the solubilization of MPs. Thus, SapNPs emerged as a widely used alternative to MSP-based NDs in the realm of MP reconstitution. In contrast to MSP NDs, always composed of 2 MSP entities, SapNPs are more modular with commonly 2-4 SapA entities and fewer lipids than NDs. While MSP ND size is largely determined by the length of the used MSP derivative, SapNP size is dictated by the number of SapA entities in the scaffold. SapNPs were applied in the CF system by using three different strategies: i. preassembly, ii. coassembly and iii. coexpression. By preassembly, SapNPs were formed prior to CF expression and added at defined concentrations into the CF reaction. By coassembly, SapA and lipid are added to the CF reaction for spontaneous assembly with the synthesized MP. By coexpression, lipid is added to the CF reaction and SapA is coexpressed together with the MP target. The different approaches were used to cotranslationally solubilize the light driven proton pump proteorhodopsin (PR) as model protein. Expression was evaluated by mainly analyzing PR yield, particle quality via size exclusion chromatography (SEC) and PR oligomerization using laser induced liquid bead ion desorption mass spectrometry (LILBID-MS).

SEC profiles of purified SapA indicated that ~50 % of expressed SapA had already formed particles with lipids from *E. coli* during purification. This observation was supported by the successful cotranslational expression of ~80  $\mu$ M PR in the CF mix using these particles without further lipid supply. Using LILBID-MS SapA dimers were detected and signals indicating presence of lipids. However, to generate SapNP with

a defined lipid composition it was necessary to use free SapA. This could be achieved, for example, through a preparative SEC purification step. However, it was also shown that during SapA purification, the use of Triton X-100 or DTT during cell disruption prevented the spontaneous assembly with lipids yielding mainly free monomeric SapA. As SapNPs has never been used in combination with CF systems, it was important to establish an efficient method for rapid and high-quality particle generation. A high yield of SapNPs was essential, since the cotranslational expression of MPs usually requires large amounts of NPs. Other particle assembly methods than detergent assisted reconstitution were therefore sought. During the preassembly of free SapA with DOPG, it was found that no detergent was required for the assembly of the particles, since a pH shift to the mildly acidic range already induced SapNP formation. This was advantageous because preassembly without detergent was not only time saving, but it also provided homogeneous particle populations and higher particle yields. The particle diameter increased from 10 nm to 12 nm with higher lipid to SapA ratios during preassembly. The particles used in this study were preassembled at a stoichiometric ratio of 1:24 SapA:DOPG, since no further size increase was observed when higher lipid ratios were used. Overall, SapNP (DOPG) were larger than the SapNPs, which spontaneously assembled with *E. coli* lipids during purification.

Next, preassembled SapNPs were tested regarding their suitability for cotranslational MP insertion. PR synthesized in presence of preassembled SapNPs yielded ~150  $\mu\text{M}$  PR in SapNPs of ~12 nm diameter, showing homogeneous, almost aggregate free SEC profiles. The particle size increased slightly upon PR insertion. The characteristic red color of PR due to bound cofactor all-trans retinal in combination with detected pentameric and hexameric PR species indicated a native-like fold of the PR in the SapNs. PR oligomerization could be, to some extent, modulated by the concentration of the supplied preassembled SapNPs. Particle quality was monitored via SEC and showed good particle quality for all samples with SapNP concentration >200  $\mu\text{M}$  in the CF reaction. Negative stain electron microscopy also showed aggregate-free PR/SapNPs. By coassembly, PR was expressed in presence of lipids and purified SapA. SapA and lipids spontaneously coassembled with expressed PR to PR/SapNPs and similar yields were obtained as with the preassembly strategy. The results even indicated cooperativity of particle formation upon PR synthesis. However, only lower PR template concentrations resulted in aggregate-free samples. This was probably because the particles did not provide enough membrane surface during coassembly and the too few assembled SapNPs became overcrowded by the expressed PR, leading to increased precipitation. A lower template concentration ensured slower expression and gave the SapNPs enough time to assemble so that the SEC profiles drastically improved. In coassembly mode, PR species were only detected up to tetramers and no hexamers were observed. Coexpression of PR and SapA yielded only 15-20  $\mu\text{M}$  solubilized PR and was considered non-suitable for further experiments. Only monomers and dimers of PR could be detected, which is due to the fact that the event of higher oligomerization is unlikely at low expression levels.

Beyond PR, CF expression of further MP targets was tested. Best expression efficiencies were obtained with the bacterial multidrug efflux proteins EmrE and SugE. The proteins showed homogeneous SEC profiles and EmrE was found in its native dimeric state. Cotranslational insertion of other transporters or G-protein coupled receptors (GPCRs) into provided SapNPs was overall less successful and frequently resulted in strong particle aggregation. However, turkey and human  $\beta_1$  adrenergic receptors were both inserted into preassembled SapNPs showing a ligand binding

fraction of 34 % and 21 %, respectively. Both GPCR/SapNP complexes could be purified and analyzed by SEC despite sub-micromolar yields. Currently, the overall performance of SapNPs appears not to be competitive with the well-established MSP ND system. However, SapNPs are clearly suitable as additive for CF systems and future applications with distinct MPs might be considered.

NMR studies with CF generated protein samples are still limited by significant label scrambling. S30 lysates are composed of around 800 proteins with varying abundance. These include several enzymes that are involved in the metabolism of amino acids. Despite the disruption of metabolic routes of AAs during lysate processing, L-Glu, L-Gln, L-Asp, L-Asn and L-Ala are still heavily affected by enzymatic conversions or “scrambling”. Such scrambling is a major problem for selective labeling of proteins for NMR studies, especially for assignment of resonances. Available single strain mutations or the usage of inhibitor cocktails are either only of limited value or too costly. The approach of this study was therefore to generate a cumulative mutant of the *E. coli* strain A19 that combines gene deletions of the most problematic scrambling enzymes. The residual scrambling enzymes in the CF S30 lysate were previously identified by a proteomics study. By using homologous recombination, these identified genes of interest (GOI) were successively deleted. S30 lysates from selected single or multiple mutants were prepared and characterized regarding performance and stability of AAs using NMR and *in vitro* expression assays. Some combinations of mutations were found to strongly affect viability of *E. coli* A19. Other strains required supplementation of culture medium for growth. Strains with a modification of the *glnA* gene were auxotrophic for L-Gln. *glmS* mutants, in turn, required either D-glucosamine or N-Acetyl-D-glucosamine for growth and were sensitive to D-glucose in the culture medium. The final strain designated “Stablelabel” contained deletions and modifications in the genes *asnA*, *ansA*, *ansB*, *glnA*, *aspC* and *ilvE*. Background activities of L-Asn and L-Asp were eliminated, as well as conversions of L-Glu to L-Asp and L-Gln. The only residual AA conversion of L-Gln to L-Glu was due to an array of enzymes with glutaminase activity in the lysates. It was found that the genes of two enzymes with glutaminase activity *glmS* and *gltB* cannot be inactivated simultaneously. Treatment of CF extracts with the glutaminase inhibitor 5 mM 6 diazo-5-oxo-L-norleucine (DON) eliminated L-Gln conversion. This is a significant improvement over previous studies where 20 mM DON was used in the entire CF reaction. A weak residual conversion of L-Glu to L-Ala was also observed, which is probably due to the aminotransferases AlaA/C and TyrB. This conversion was eliminated using the cheap and promiscuous transaminase inhibitor aminooxyacetate (AOA). The mutations in Stablelabel made the use of chemical inhibitors of glutamine synthetase, asparaginase and aspartate ammonia ligase obsolete. In particular, the latter two inhibitors 5 diazo-4-oxo norvaline (DONV) and S methionine-L-cysteine sulfoximine (SMCS) are difficult to obtain commercially and are therefore expensive. Stablelabel showed a slightly slower growth than A19, and an overall good performance with 2.7 mg/mL GFP expressed in the reaction mixture (RM) compared to the parental A19 strain with 3.5 mg/mL.

Furthermore, Stablelabel carried a deletion in the gene *ilvE* to eliminate activity of branched chain amino acid aminotransferase (BCAAT). The mutation was introduced to stabilize labeled L-Val, L-Leu and L-Ile which are commonly used for site-specific methyl labeling of MPs. Application of the AAs directly is rather expensive and therefore, precursors of these AAs are frequently applied. In this study, a new strategy was established by preconversion of the precursors 2-ketoisovalerate (KIV) and 4-

methyl-2-oxovalerate (MOV) to their respective amino acids L-Val and L-Leu using recombinantly expressed BCAAT and L-Glu as amino group donor. Preconversion was carried out in deuterated environment, which alleviates deuterium-proton replacement at C $\alpha$  and C $\beta$  positions during transamination that later can cause isotopic shifts in NMR spectra. Since the deuterium replacement can also happen due to innate BCAAT activity in the lysate, transaminase gene *ilvE* was also deleted in Stablelabel. Finally, Stablelabel was successfully applied for methyl labeling of PR at L-Val and L-Leu residues, preconverted from their respective precursors. [ $^{13}\text{C}$ ,  $^1\text{H}$ ] XL-ALSOFAST-HMQC spectra of L-Val and L-Leu showed no overlapping, indicating no scrambling. The signals were reasonably resolved for a non-deuterated MP inside a detergent micelle. The controlled preconversion of  $\alpha$ -keto acids to their respective AAs for use in CF systems represents a cost-effective alternative to methyl group labeling in live cells. Generally, the consumption of the AA precursor is only about a tenth of what is required for labeling in live cells for efficient labeling. Finally, first attempts were made to label the methyl groups of L-Val in PR or free fatty acid receptor 2 (FFAR2) in MSP1D1 $\Delta$ H5 NDs. Despite good SEC profiles and even deuterated protein in the case of FFAR $_2$ , both HMQC spectra showed broad and poorly resolved signals. This suggests further optimizations such as the use of deuterated MSP and lipid, but also ligand affinity purification of GPCRs to obtain homogeneous samples of functional target proteins.

Conclusively, Stablelabel is a valuable addition to the repertoire of the CF system since it provides a near scrambling free CF lysate minimizing the necessity for inhibitors. Stablelabel was found to be the best compromise between having cumulative mutations to effectively address label scrambling while still retaining sufficient protein synthesis productivity. Future applications in combination with optimized ND systems might be a step towards NMR analysis of CF synthesized GPCR probes and other complex targets.

## Zusammenfassung

Die zellfreie Proteinsynthese (CFPS) hat sich in den letzten Jahrzehnten stark weiterentwickelt und ist zu einem häufig verwendeten System zur Expression von Membranproteinen (MP) geworden. Dabei werden MP kotranslational und ohne die Hilfe des Translokonsystems in die bereitgestellten Lipiddoppelschichten, wie z.B. Liposomen oder Nanodisks inseriert. Dadurch werden die Engpässe der herkömmlichen zellulären Expression wie Produkttoxizität oder aufwendige MP-Rekonstitution umgangen. Nur mit CF-Systemen ist die kotranslationale Expression von MPs ohne die sonst notwendige Solubilisierung von MP und Lipid in Detergenz und Rekonstitution möglich. Aufgrund der offenen Natur von CF-Systemen kann der Experimentator die Reaktionsbedingungen für das Zielprotein frei anpassen. Zu den Optionen gehören Expressionsbedingungen, Auswahl der Nanopartikel (NP) und ihrer Lipidzusammensetzung, Redox-Bedingungen, Zugabe stabilisierender Moleküle wie Liganden oder Cofaktoren und vieles mehr. Am besten etabliert ist bisher das CF-System auf Basis von *E. coli*-Lysaten aufgrund der kosteneffizienten Produktion, der einfachen Handhabung und der hohen Produktausbeute. Am häufigsten werden NDs auf der Basis von Membrangerüstproteinen (MSP) für die kotranslationale Expression von MP verwendet. In den letzten Jahren erwies sich das System als besonders nützlich für die biochemische Charakterisierung von MP, und sogar Strukturstudien mittels Nuklearmagnetresonanz Spektroskopie (NMR) und Kryoelektronenmikroskopie (Cryo-EM) waren möglich. Es sind jedoch auch neue NP-Systeme, wie Saposin A-basierte Nanopartikel (SapNP) auf dem Vormarsch, deren erstmalige Anwendung in Kombination mit dem CF-System im ersten Teil dieser Studie eingehend charakterisiert wurde. Darüber hinaus profitiert die CF Synthese von Proteinproben für die NMR-Analyse von unterbrochenen Stoffwechselwegen der Aminosäuren (AA) in den Lysaten, was Markierungsschemata ermöglicht, die in zellulären Systemen aufgrund der spektralen Mehrdeutigkeit unmöglich wären. Trotz der unterbrochenen Stoffwechselwege werden dennoch einige AA enzymatisch umgesetzt und führen zu Isotopenverdünnung und mehrdeutigen Spektren. Die enzymatische Umwandlung der AA wird als „Scrambling“ bezeichnet. Der zweite Teil dieser Studie war daher der Eliminierung dieser restlichen AA-Scrambling-Aktivitäten im Lysat-Quellstamm *E. coli* A19 durch gezielte Genmanipulationen gewidmet. Resultierend wurde ein neuer Master-Stamm erhalten, der optimierte CF-Lysate für NMR-Anwendungen produzierte.

Ähnlich wie bei MSP, wurde auch die Eigenschaft von Saposin A (SapA), Lipide zu binden, biotechnologisch genutzt, um ein neues Nanopartikelsystem für die Solubilisierung von MP zu etablieren. Heute sind SapNP eine weit verbreitete Alternative zu MSP-basierten NDs im Bereich der MP-Rekonstitution. Im Gegensatz zu MSP-NDs, die immer aus 2 MSP-Einheiten bestehen, sind SapNPs modularer und bestehen üblicherweise aus 2-4 SapA-Einheiten und weniger Lipiden als NDs. Während die Größe der MSP-NDs durch die Länge des verwendeten MSP-Derivats bestimmt wird, wird die SapNP Größe weitgehend durch die Anzahl der SapA-Einheiten im Nanopartikelgerüst bestimmt.

Die SapNP wurden anhand von drei Strategien im CF System eingesetzt: i. Vorassemblierung, ii. Koassemblierung und iii. Koexpression. Durch Vorassemblierung wurden SapNP vor der CF-Expression assembliert und in definierten Konzentrationen der Reaktion zugesetzt. Im Koassemblierungsmodus

wurden SapA und Lipid zur CF-Reaktion für die spontane Assemblierung mit dem MP hinzugegeben. Im Koexpressionsmodus wurde der CF-Reaktion Lipid hinzugefügt und SapA mit dem MP koexprimiert. Die verschiedenen Ansätze wurden verwendet, um die lichtgetriebene Protonenpumpe Proteorhodopsin (PR) als Modellprotein kotranslational in SapNP zu inserieren. Die Expression wurde hauptsächlich durch Analyse der PR-Ausbeute, der Partikelqualität mittels Größenausschlusschromatographie (SEC) und der PR-Oligomerisierung mittels „Laser Induced Liquid Bead Ion Desorption“-Massenspektrometrie (LILBID-MS) bewertet.

Die SEC-Analyse des gereinigten SapA deutete darauf hin, dass SapA mit Lipiden aus *E. coli* schon während der Reinigung Partikel geformt haben muss. Diese Annahme wurde durch die erfolgreiche kotranslationale Expression von ~80 µM PR mit Hilfe dieser Partikel ohne externe Zugabe von Lipid untermauert. Auch die LILBID-MS deutete darauf hin, dass SapNP aus SapA Dimeren bestanden, und zeigte Signale, die auf Lipide hindeuteten. Um SapNP mit einer definierten Lipidkomposition zu erzeugen, war es jedoch notwendig, freies SapA zu verwenden. Dies konnte z.B. durch einen präparativen SEC-Reinigungsschritt erreicht werden. Es wurde jedoch auch festgestellt, dass bei der SapA Reinigung der Einsatz von Triton X-100 oder DTT während dem Zellaufschluss diese spontane Assemblierung mit Lipiden weitestgehend eindämmte. Da SapNP noch nie in Kombination mit CF-Systemen verwendet wurden, war es zunächst wichtig, effiziente Methoden für eine schnelle und qualitativ hochwertige Partikelerzeugung zu etablieren. Da die CF Expression von MP i.d.R. große Mengen an Nanopartikeln erfordert, war eine gute Nanopartikelausbeute unabdingbar. Bei der Vorassemblierung von freiem SapA mit DOPG wurde dabei festgestellt, dass kein Detergenz für die Assemblierung der Partikel erforderlich ist, da ein pH Shift in den mild-sauren Bereich die SapNP-Formation induzierte. Dies war vorteilhaft, da die Vorassemblierung ohne Detergenz nicht nur Zeit sparte und homogene Partikelpopulationen lieferte, sondern auch höhere Partikelausbeuten. Der Partikeldurchmesser nahm von 10 nm mit höheren Lipid zu SapA Verhältnissen auf 12-13 nm zu. Die in dieser Studie verwendeten Partikel wurden in einem stöchiometrischen Verhältnis von 1:24 SapA:DOPG vorassembliert, da höhere Lipid zu SapA Verhältnisse keinen weiteren Größenzuwachs der Partikel aufwiesen. Die Partikel waren insgesamt größer als die SapNP, die während der Reinigung spontan mit *E. coli* Lipiden assemblierten.

Als nächstes wurden die vorassemblierten SapNP auf ihre Eignung für die kotranslationale MP-Expression getestet. PR wurde mit vorassemblierten Partikeln exprimiert und die Expression ergab ~150 µM solubilisiertes MP im CF Ansatz. Die PR/SapNP hatten einen Durchmesser von 12-13 nm und nahezu aggregatfreie SEC-Profile. Die Partikelgröße nahm bei PR-Insertion leicht zu. Die charakteristische rote Farbe der PR-Proben aufgrund des gebundenen Cofaktors all-trans-Retinal sowie die detektierte pentamere und hexamere PR-Spezies wiesen auf eine native Faltung von PR in SapNP hin. Es wurde ferner gezeigt, dass die Oligomerisierung von PR abhängig von der Expressionsstärke oder der SapNP Konzentration im System moduliert werden kann. Die Partikelqualität wurde anhand von SEC-Profilen verfolgt und blieb über nahezu alle Bedingungen hinweg gut. Auch die Elektronenmikroskopie zeigte aggregatfreie PR/SapNP Partikel. Im Koassemblierungsmodus wurde PR in Gegenwart von Lipiden und gereinigtem SapA exprimiert. SapA und Lipide assemblierten spontan mit exprimiertem PR und es wurden ähnliche Ausbeuten wie im Vorassemblierungsmodus erreicht. Die Ergebnisse wiesen sogar auf eine



Kooperativität der Partikelbildung bei der Expression von PR hin. Jedoch führten im Koassemblierungsmodus nur niedrigere PR-Templatkonzentrationen zu aggregatfreien Proben. Dies lag vermutlich daran, dass Partikel während der Koassemblierung nicht schnell genug Membranfläche zur Verfügung stellten, und schnell durch das exprimierte Protein überfüllt wurden, was vermehrt zu Präzipitation führte. Eine niedrigere Templatkonzentration sorgte dabei für eine langsamere Expression, und gab der Assemblierung der SapNP genug Vorsprung, sodass es nicht zu einer Überfüllung der Partikel kam. Im Koassemblierungsmodus wurden jedoch PR-Spezies nur bis zu Tetrameren und keine Hexamere nachgewiesen. Die Koexpression von PR und SapA ergab nur 15–20  $\mu\text{M}$  gelöstes PR und wurde für weitere Experimente als ungeeignet erachtet. Es ließen sich lediglich Dimere von PR nachweisen, was darauf zurückzuführen ist, dass höhere Oligomerisierung bei derart schwacher Expression unwahrscheinlich ist.

Außer PR wurde auch die Insertion anderer MP in SapNP getestet. Die beste Expressionseffizienz wurde mit den bakteriellen Multidrug-Efflux-Proteinen EmrE und SugE erzielt. In SapNP zeigten die Proteine homogene SEC-Profile und EmrE wurde in seinem nativen dimeren Zustand detektiert. Die kotranslationale Expression anderer MP wie z.B. Transporter oder G-Protein gekoppelter Rezeptoren (GPCR) in vorassemblierte SapNP war insgesamt weniger erfolgreich und resultierte häufiger in aggregierten Proben. Allerdings wurden die adrenergen  $\beta_1$ -Rezeptoren aus Truthahn und Mensch, beide in vorgeformte SapNP inseriert, und zeigten einen Ligandenbindungsanteil von jeweils 34 % und 21 %. Beide GPCR/SapNP Komplexe konnten gereinigt und trotz submikromolarer Ausbeute mittels SEC analysiert werden. Derzeit scheint die Gesamtleistung von SapNP nicht mit dem etablierten MSP ND-System mithalten zu können. Allerdings sind SapNP eindeutig als Additiv für CF-Systeme geeignet und zukünftige Anwendungen mit bestimmten MP könnten in Erwägung gezogen werden.

NMR-Studien mit CF synthetisiertem Protein sind weiterhin beeinträchtigt durch signifikante enzymatische Umwandlungen oder „Scrambling“ von markierten AAs. S30-Lysate bestehen aus etwa 800 Proteinen, die unterschiedlich häufig vertreten sind. Darunter sind auch mehrere Enzyme, die in metabolische Umwandlungen von Aminosäuren involviert sind. Trotz der bei der Lysat Herstellung unterbrochenen Stoffwechselwege sind L-Glu, L-Gln, L-Asp, L-Asn und L-Ala immer noch stark von enzymatischem Scrambling betroffen. Scrambling ist ein großes Problem bei der selektiven Markierung von Proteinen für die NMR-Analyse, insbesondere bei der Zuordnung von Resonanzen. Scrambling Enzyme können entweder durch gezielte Gendeletion oder durch Inhibitoren inaktiviert werden. Leider sind verfügbare Einzelstamm Mutationen oder der Einsatz von Inhibitor-Cocktails entweder von begrenztem Nutzen oder zu kostspielig. Der Ansatz dieser Studie bestand daher darin, eine kumulative Mutante des *E. coli* Stammes A19 zu erzeugen, die Gendeletionen der problematischsten Scrambling-Enzyme kombiniert.

Zunächst wurden die Scrambling-Enzyme im S30-Lysat durch das Heranziehen von Daten aus einer vorherigen Proteomik-Studie identifiziert. Mithilfe der homologen Rekombination wurden diese identifizierten Gene von Interesse (GOI) sukzessive gelöscht. Die erhaltenen Stammvarianten wurden kultiviert, S30-Lysate hergestellt und hinsichtlich der Leistung und Stabilität von AAs mithilfe von NMR- und *in vitro* Expressionstests charakterisiert. Es wurde festgestellt, dass einige Mutationskombinationen die Viabilität von *E. coli* A19 stark beeinträchtigten, andere

Stämme erforderten für das Wachstum eine Ergänzung des Kulturmediums. So erforderten alle Stämme mit einer Modifikation des Gens *glnA*, die Zugabe von L-Gln zum Kulturmedium. *glmS* Mutanten wiederum benötigten entweder D-Glucosamin oder N-Acetyl-D-Glucosamin zum Wachstum und waren empfindlich gegenüber D-Glucose. Der endgültige Stamm mit der Bezeichnung „Stablelabel“ enthielt Deletionen und Modifikationen in den Genen *asnA*, *ansA*, *ansB*, *glnA*, *aspC* und *ilvE*. Hintergrundaktivitäten von L-Asn und L-Asp sowie Umwandlungen von L-Glu in L-Asp und L-Gln wurden eliminiert. Lediglich wurde L-Gln weiterhin zu L-Glu umgewandelt, was auf eine Reihe von Enzymen mit Glutaminase-Aktivität in den Lysaten zurückzuführen ist. Es wurde festgestellt, dass die Gene der beiden Enzyme mit Glutaminaseaktivität *glmS* und *gltB* nicht gleichzeitig inaktiviert werden können. Die Behandlung von CF-Extrakten mit dem Glutaminase Inhibitor, 5 mM 6 Diazo-5-oxo-L-norleucin (DON), eliminierte die L-Gln Umwandlung. Dies ist eine deutliche Verbesserung gegenüber früheren Studien, bei denen 20 mM DON in der gesamten Zellfrei Reaktion verwendet wurde. Es wurde auch eine schwache Restumwandlung von L-Glu zu L-Ala beobachtet, die vermutlich auf die Aminotransferasen AlaA/C und TyrB zurückzuführen ist. Diese Umwandlung konnte mit dem billigen und promiskuitiven Aminotransferase-Inhibitor, Aminoxyacetat (AOA), eliminiert werden. Die reduzierten Scrambling Aktivitäten machten den Einsatz chemischer Inhibitoren der Glutamin Synthetase, Asparaginase und Aspartat Ammonia-Ligase obsolet. Vor allem die letzteren beiden Inhibitoren 5-diazo-4-oxo norvaline (DONV) und S-methionine-L-cysteine sulfoximine (SMCS) sind kommerziell nur schwer erhältlich und dementsprechend teuer. Insgesamt zeigte Stablelabel ein etwas langsames Wachstum als A19 und eine insgesamt gute Expressionseffizienz mit 2.7 mg/mL GFP im Reaktionsgemisch (RM) verglichen mit 3.5 mg/mL bei dem Wildtyp Stamm.

Stablelabel wurde außerdem mit einer Deletion des Gens *ilvE* ausgestattet, um die Aktivität der verzweigt-kettigen Aminosäure-Aminotransferase (BCAAT) zu eliminieren. Die Mutation wurde eingeführt, um L-Val und L-Leu zu stabilisieren, die üblicherweise für die selektive <sup>13</sup>C Methylgruppenmarkierung von MP verwendet werden. Die direkte Anwendung der AAs ist teuer und daher werden häufig Vorläufer dieser AAs eingesetzt. In dieser Studie wurden die Vorläufer 2-Ketoisovalerat (KIV) und 4-Methyl-2-oxovalerat (MOV) unter Verwendung von rekombinant exprimierter BCAAT und L-Glu als Aminogruppendonor in ihre jeweiligen Aminosäuren L-Val und L-Leu umgewandelt. Die Vorkonvertierung wurde in einer deuterierten Umgebung durchgeführt. Dies verhinderte den Austausch von Deuterium-Protonen an den Positionen C $\alpha$  und C $\beta$ , was später zu Isotopenverschiebungen in den Spektren führen kann. Da der Deuteriumaustausch auch aufgrund der inhärenten BCAAT-Aktivität im Lysat erfolgen kann, wurde das Transaminase-Gen *ilvE* in Stablelabel deletiert. Schließlich wurde Stablelabel erfolgreich zur Markierung von PR an den Methylgruppen von L-Val und L-Leu eingesetzt. Die Aminosäuren wurden zuvor von ihren jeweiligen Vorläufermolekülen konvertiert. [<sup>13</sup>C, <sup>1</sup>H] XL-ALSOFAST-HMQC Spektren von L-Val und L-Leu zeigten keine Überlappung, was darauf hindeutet, dass kein Scrambling stattfand. Die Signale waren für ein nicht-deuteriertes MP in einer Detergenzmizelle gut aufgelöst. Die kontrollierte Vorkonvertierung von  $\alpha$ -Ketosäuren zu ihren jeweiligen AA für den Einsatz in CF-Systemen stellt eine kostengünstige Alternative zur Methylgruppenmarkierung in lebenden Zellen dar. Der Verbrauch der Aminosäurevorstufe entspricht dabei je nach Zielprotein nur etwa einem zehntel dessen, was für eine Markierung in lebenden Zellen für eine NMR-Probe notwendig wäre. Abschließend wurden die ersten Versuche der Markierung der Methylgruppen von L-Val in PR oder dem Fettsäurerezeptor 2 (FFAR<sub>2</sub>) in MSP1D1 $\Delta$ H5 NDs

unternommen. Trotz guter SEC-Profile und sogar deuteriertem Protein im Fall von FFAR<sub>2</sub>, zeigten beide HMQC-Spektren breite und schlecht definierte Signale. Dies legt weitere Optimierungen nahe, wie z.B. die Verwendung von deuteriertem MSP und Lipid sowie Liganden-oder G-Protein-Affinitätsreinigung von GPCR, um homogene Proben von funktionellem Protein zu erhalten.

Zusammenfassend lässt sich sagen, dass Stablelabel eine wertvolle Ergänzung des Repertoires des CF-Systems darstellt, da es einen nahezu Scrambling-freien CF-Stamm bietet und den Einsatz von Inhibitoren minimiert. Die Kombination der Mutationen im Stablelabel Stamm erwies sich als der beste Kompromiss der wirksamen Bekämpfung von Label-Scrambling und der Beibehaltung einer Proteinausbeute, die den Ansprüchen der NMR genügt. Zukünftige Anwendungen von Stablelabel in Kombination mit optimierten ND-Systemen könnten ein Schritt in Richtung der NMR-Analyse von GPCR und anderen komplexen Zielproteinen sein.



# 1. Introduction

## 1.1 The cell-free system

### 1.1.1 Discovery and development

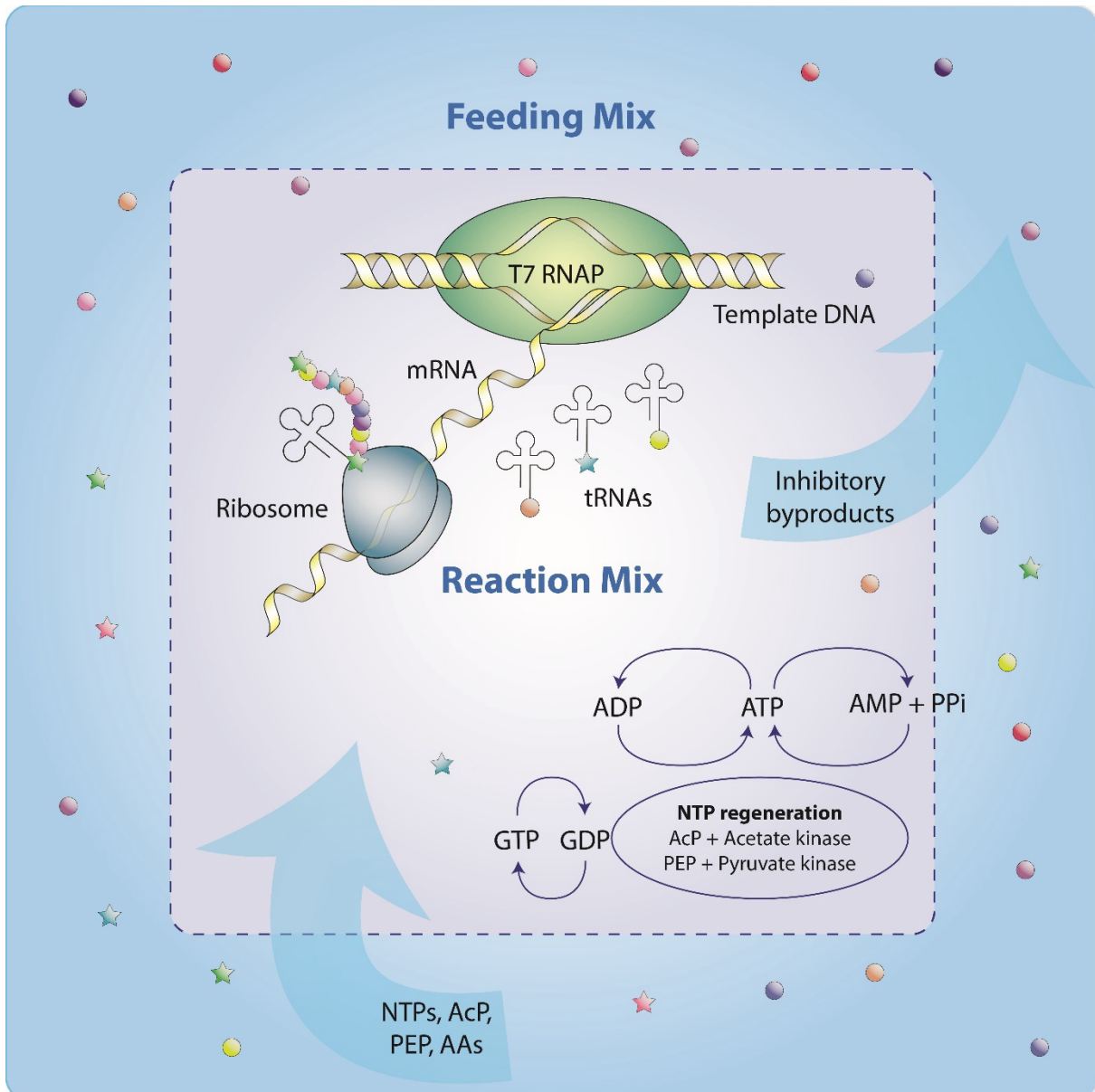
Cell-free (CF) expression of proteins refers to the production of proteins outside of living cells, using only cell components required for *in vitro* transcription-translation (IVTT) systems. In 1897 Eduard Buchner discovered fermentation in yeast cell extracts and received the Nobel Prize in Chemistry ten years later for his pioneering work<sup>1,2</sup>. The earliest study on cell-free protein synthesis (CFPS) using *Staphylococci* lysate was conducted in 1954 by Joan Folkes and Ernest Gale who were both subsequently nominated for the Nobel Prize in Chemistry in 1956<sup>3</sup>. In 1960 the lab of Paul Zamecnik developed a CF extract based on *E. coli* that was used by Marshall Nirenberg and Johann Matthaei to deduce the genetic code<sup>4</sup>. Later Mashall Nirenberg was also rewarded with the Nobel Prize for this achievement<sup>5</sup>. This laid the foundation for subsequent studies that aimed to develop sophisticated CF protein expression systems. A CF expression system can be produced either from lysed cells or by reconstitution of all separately purified transcription and translation components separately such as ribosomes, RNAP, energy regenerating enzymes, transcription factors, tRNA synthetases, etc.<sup>6-8</sup>. Lysates can be also produced from prokaryotic sources, mainly *E. coli*, but also eukaryotic organisms such as yeast<sup>9</sup>, wheat germ<sup>10</sup>, insect cells<sup>11</sup> and rabbit reticulocytes<sup>12</sup>, Chinese hamster ovaries<sup>13</sup> and even human cells, such as HeLa cells<sup>14</sup>. While having higher capabilities of protein modification and expression of larger and complex protein targets, eukaryotic systems mostly suffer from higher cost and low protein yields usually only sufficient for biochemical investigations, but not structural studies or extensive characterization<sup>5</sup>. For post translational modifications (PTMs), ER derived membrane structures termed microsomes containing all required components for PTMs need to be added to the eukaryotic lysates or microsomes need to be preserved during lysate preparation<sup>15-17</sup>. Naturally lacking glycosylation capabilities, *E. coli* lysates can be modified with reconstituted *N*-linked glycosylation pathways<sup>17</sup>.

### 1.1.2 *E. coli* lysates

Most widely used is the bacterial CF system which is commonly produced from *E. coli* derivatives, typically BL21 (DE3), and K12 A19 strains<sup>18</sup>. These systems are relatively easy to produce, efficient and cheap when compared with eukaryotic or reconstituted systems such as PURE<sup>6</sup>. There are different ways to produce *E. coli* lysates. Lysate preparation starts with cultivation of *E. coli* cells, mostly on rich media to ensure high expression capacity, e.g., through high ribosome density<sup>19,20</sup>. Cells are best harvested in mid log phase where these requirements are fulfilled<sup>21</sup>. If desired, cells can be subjected to a heat or cold shock, but also chemical stress, e.g., through addition of ethanol to enhance production of chaperones to assist protein folding and stability, especially of difficult targets<sup>22,23</sup>. After harvesting cells are washed and lysed. During centrifugation, macromolecular structures such as cell membranes, genomic DNA, polysomes and large complexes are removed. However, dependent on centrifugation speed some of these structures may remain in the lysate. Most common are S12 and S30 lysates with varying productivities dependent on the used strain<sup>19,20</sup>. After centrifugation a ribosome runoff step can be included by incubating the lysate in presence of a high-salt buffer<sup>23</sup>. Subsequently, the lysate is dialyzed, aliquoted and

flash frozen. Lysates can be stored for years at -80 °C without significant productivity loss<sup>24</sup>.

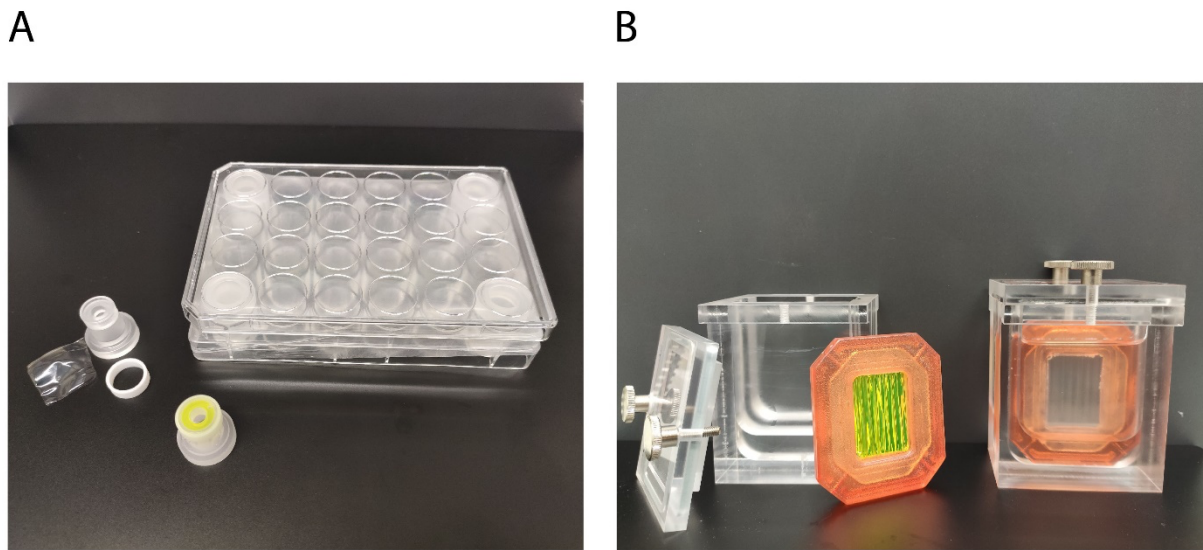
To set up a CF expression it is important to include all compounds in the reaction that are required for transcription and translation. This includes low molecular weight components such as amino acids (AAs), nucleotide triphosphates (NTPs), a redox system, buffer and molecules for energy regeneration such as acetyl phosphate (AcP), phosphoenolpyruvate (PEP) and RNase and protease inhibitors<sup>25</sup>. To produce RNA from DNA template, RNA polymerase (usually T7 RNAP) is required. It can be either expressed during lysate production or expressed separately and added to the reaction<sup>20,25</sup>. For energy regeneration of ATP, pyruvate kinase or creatine kinase need to be included while acetate kinase inherently is present at sufficient levels in lysates<sup>23</sup>. However, excessive phosphate production in turn has negative effects on the CFPS system<sup>26</sup>. In our hands, most *E. coli* lysates contain enough tRNAs for efficient expression, however external tRNA supply might increase protein yield<sup>8</sup>. All other high molecular weight components, such as ribosomes, or tRNA synthetases, initiation, elongation and release factors are provided by addition of lysate. There are two ways of CF protein expression: i. batch configuration and ii. two-compartment configuration. Batch configuration includes all reaction components in one compartment. In a two-compartment configuration the expression lumen, also referred to as the reaction mix (RM), is separated from an additional reservoir with only low molecular weight compounds, also referred to as the feeding mixture (FM). Thereby, FM and RM are separated by a dialysis membrane (Figure 1). This way the reaction lumen is constantly provided with amino acids, NTPs and energy source precursors while toxic expression byproducts, such as phosphate are dialyzed into the feeding mix<sup>26</sup>. With this approach protein yields up to 5 mg/mL can be obtained in the reaction mix, while same amount of high molecular weight components is used as in batch reactions<sup>24</sup>.



**Figure 1.** Cell-free protein synthesis in a two-compartment system. The target protein is expressed in the reaction mixture that holds high and low molecular weight components. The feeding mixture supplies the reaction with low molecular weight components such as NTPs and amino acids through a dialysis membrane (dotted line). Unmodified amino acids are depicted as balls and modified amino acids are depicted as star symbols.

Continuous exchange cell-free (CECF) expression setups can be assembled using either customized or commercially available components as comprehensively illustrated by Schwarz et al<sup>25</sup>. For analytical scale reaction volumes customized reaction containers with a dialysis membrane pinched to the vessel by a PTFE ring can be used while the FM is placed in the wells of cell culture plates (Figure 2A)<sup>25</sup>. This technique is commonly used by our group for quick screening reactions. It has the advantage that reaction containers and the cell culture plate can be reused for further expressions. For large scale expression, commercially available dialysis cassettes can be used in combination with tailored containers for the feeding mixture (Figure 2B)<sup>25</sup>. This technique is also commonly employed by our group for preparative protein expression. However, in contrast to the customized FM chambers, commercial dialysis cassettes cannot be reused. In theory, CECF expression possibilities are as versatile

as dialysis possibilities and good results can be obtained with both, commercial kits and customized devices.



**Figure 2.** Examples of devices for CECF expression. **A:** Analytical scale reaction vessels for up to 60  $\mu\text{L}$  reaction volume are comprised of a container with a dialysis membrane that is fixed by a PTFE ring. The assembled container holding the RM is placed inside a commercially available cell culture plate well holding up to 825  $\mu\text{L}$  of the FM. **B:** The preparative scale and commercially available dialysis cassette with a 3 mL reaction lumen holding the RM is placed inside a customized chamber holding up to 60 mL FM. Larger dialysis cassettes with larger FM chambers can be used. Both, the analytical and preparative scale devices are placed on axial shakers for constant mixing during expression. Modified figure taken from Levin et al<sup>24</sup>.

Since its application for groundbreaking discoveries like deciphering the genetic code, CF systems were used for a variety of applications, including protein production for structural and functional studies, high-throughput screenings, diagnostics, and the production of fine-chemicals, biofuels and biopharmaceuticals<sup>27–36</sup>. CF systems are capable of producing an extensive variety of proteins, including enzymes, antibodies, membrane proteins and toxic or misfolding proteins, that, if expressed in live cells, might be a burden or even lethal to the organism<sup>2,22,37–40</sup>. The open nature of CF systems allows to adjust a wide range of conditions to meet the requirements for the protein of interest<sup>23</sup>. This includes chemical properties such as buffer and redox potential, detergents, membrane mimetics but also site-specific incorporation of non-canonical amino acids, e.g., with isotopic labels and other chemical modifications<sup>22,41–46</sup>. Also, different proteins can be coexpressed simultaneously without the necessity to consider plasmid compatibility<sup>45,47</sup>.

### 1.2 Expression and purification of Membrane proteins

Membrane proteins (MPs) play a vital role in all living domains. Embedded or anchored in the hydrophobic environment of cellular membranes, they function as molecular transporters, channels, receptors, enzymes, and entities ensuring structural integrity of the membrane. Up to 30 % of all human genes are estimated to be encoding for MPs<sup>48</sup>. Their countless roles in cell signaling, mediation of molecular flow and signal transduction made them exceptionally relevant pharmaceutical targets, with G-protein coupled receptors (GPCRs) being the most attractive drug target class being targeted



by ~40 % of all marketed drugs<sup>49</sup>. Residing in a highly specific amphipathic membrane environment, MPs interact with distinct lipids, but also proteins and other carbohydrates. MPs are particularly difficult to study due to their hydrophobic nature. Over the past decades detergents played a crucial role in MP studies<sup>50</sup>. Traditionally, the target protein is extracted from the cell membrane of the expressing host organism using detergents, which are amphipathic molecules that form a hydrophobic core around the MP holding it in a solubilized state. However, the complex physical nature of detergents and sometimes unpredictable behavior can lead to very time consuming screening procedures to find the right conditions to solubilize the protein in its active form, and in a state suitable for the measurement technique<sup>50,51</sup>. Furthermore, even if the right detergent is found, once removed from its native environment, there might be ramifications regarding structural and functional integrity of MPs, as well as stability issues<sup>52,53</sup>. For example, interaction of  $\beta_2$  adrenergic receptor with G-protein was more than tenfold increased in the lipid environment of a ND compared to detergent solubilized receptor<sup>54</sup>.

### 1.2.1 MP solubilization in mimetic lipid bilayers

Indeed, the invention of nanodiscs (NDs) revolutionized membrane protein research by providing a powerful tool for solubilization of MPs in a functional form in a native like environment<sup>55,56</sup>. NDs are small, disc-shaped particles composed of lipids and a scaffold that can be either a scaffold protein such as apolipoprotein A-1 (ApoA-1)<sup>57</sup> or Saposin A (SapA)<sup>58</sup>, short peptides<sup>59</sup>, amphiphatic polymers such as SMALPs<sup>60</sup> or even large DNA-ApoA-1 hybrid-scaffolds<sup>61</sup>. NDs offer several advantages over detergents since they provide a lipid bilayer that comes closer resembling cell membranes than a detergent micelle<sup>62</sup>. This allows for the study of membrane protein structure, function, and interactions using a range of biophysical techniques, including X-ray crystallography, NMR spectroscopy, and especially cryo-EM<sup>63</sup>. To transfer MPs inside such artificial bilayers, MPs are commonly expressed in cellular systems and are inserted inside cellular membranes of the host organism first. After expression, the membranes are harvested and dissolved in detergent to release the target protein<sup>64</sup>. Subsequently, the target protein is purified by affinity chromatography and size exclusion chromatography followed by a reconstitution procedure<sup>53</sup>. Therefore, the purified MP is incubated with the scaffold and lipid in a certain stoichiometric ratio and detergent. Subsequent detergent removal by use of biobeads or chromatography methods ideally leads to the assembly of NDs with the target MP embedded in the lipid bilayer<sup>56</sup>.

### 1.2.2 Apolipoprotein A-1 derivative based nanoparticles

ApoA-1 is a 243-residue multifunctional protein that is a major constituent of high density lipoprotein complexes (HDL) which play a crucial role in cholesterol traffic<sup>65</sup>, but also in immune and inflammation response<sup>66</sup>. The lipid binding capability of ApoA-1 and cholesterol efflux from cells was extensively studied and characterized due to the significance in coronary artery disease caused by atherosclerosis<sup>65,67</sup>. Structural studies of ApoA-1 revealed class A amphiphatic  $\alpha$ -helical repeats that form a horseshoe structure of a seemingly continuous, amphipathic  $\alpha$ -helix that is "punctuated" by regularly spaced kinks at proline residues<sup>68</sup>. The antiparallel dimer structure has a continuous strip of hydrophobic residues capable of lipid binding at the inward face of the ring, while salt bridge forming residues L-Lys, L-Arg, L-Glu and L-Asp resemble a charged, solvent accessible backbone<sup>68</sup>. This property was successfully exploited to form synthetic lipid bilayer structures composed of a dimer of

two truncated ApoA-1 MSPs wrapped around a lipid bilayer of ~150 phospholipid molecules<sup>55</sup>. The resulting particles have a lipid bilayer area of ~45 nm<sup>2</sup> and a diameter of ~9-10 nm<sup>55</sup>. In further studies the scaffold protein was engineered in length to allow larger and smaller particles<sup>57,69</sup>. For example, MSP1D1ΔH5 and MSP2N1-3 variants constitute particles with a diameter of ~8 nm and ~17 nm, respectively, with a plentitude of MSP derivatives covering nanodisc sizes in between<sup>69,70</sup>. Larger NDs provide an increased lipid area and allow the insertion of larger complexes or multimeric protein assemblies, but also enabled studies on virus-membrane interaction or vesicle fusion<sup>71,72</sup>. Reconstitution using scaffold proteins has become a standard technique for MP solubilization. However, being more or less restricted by the size of used MSP derivative, NDs were shown to still have a certain flexibility depending on the size of the incorporated MP. For increased stability and disc homogeneity, sortase-circularized NDs (cNDs) were proposed<sup>72</sup>. One the contrary, less rigid modular systems with a higher flexibility, such as peptidiscs or Salipro, also showed to be suitable for MP solubilization<sup>58,59</sup>.

### 1.2.3 SapA based nanoparticles for cotranslational MP solubilization

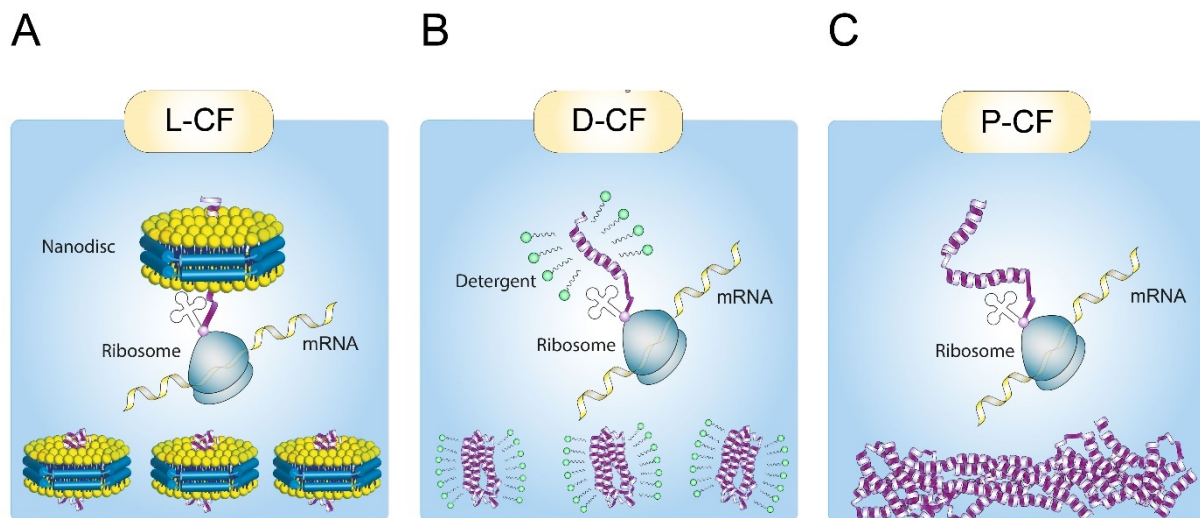
As an alternative to MSP, NDs which are more or less strictly defined in their size by the choice of MSP derivative, a more flexible system was adapted recently. The system is based on SapA, an  $\alpha$ -helical ~9 kDa protein that is part of the highly diverse saposin family with members usually below identity levels of 25-30 %<sup>73</sup>. Physiologically, sphingolipid activator proteins (SAPs) act as cofactors of sphingolipid degrading enzymes<sup>74</sup>. For instance, binding of SapA to sphingomyelin disrupts membrane organization and increases accessibility of sphingomyelin for the catalytic site of sphingomyelinase to create ceramide and phosphocholine<sup>73</sup>. Similarly as for ApoA-1, lipid binding capability of SapA was used to create NPs with SapA as a scaffold protein<sup>75</sup>. The procedure of MP reconstitution using SapNPs is very similar to MSP NDs. SapNPs are more flexible in size and produce a tighter belt around the target protein containing fewer lipids inside the SapNP. This characteristic was claimed advantageous for MP NMR<sup>76</sup>. However, it has to be noted that particle characteristics can vary heavily dependent on the stoichiometric ratios applied for reconstitution, since limiting lipid availability may result in particles with varying numbers of SapA entities<sup>76-78</sup>. To produce SapNPs, SapA needs to be incubated with lipids in presence of detergent. Furthermore, pH was reported to play a role for particle generation, since SapA has an increased lipid binding capability at lysosomal pH range of 4.5-5.5<sup>76,78</sup>. While this is less relevant for reconstitution experiments which are commonly conducted at neutral to slightly basic pH, it's an important consideration regarding NP generation for subsequent cotranslational MP solubilization using the CF system. A systematic comparison of sample quality of an ABC transporter MsbA reconstituted in NDs or SapNPs was already carried out previously<sup>79</sup>. In this work, applicability of SapNPs for cotranslational MP solubilization was systematically studied.

### 1.2.4 CF expression of MPs

To date there are only two routes to obtain MPs incorporated inside lipid bilayer carrying NPs in their functional form that omit the use of detergent: i. MP extraction with SMALPs and ii. cotranslational insertion of MPs inside provided lipid bilayers using CF expression. SMALPs are capable to extract MPs from their native environment directly, without detergent solubilization of the membrane. The advantage of SMALPs is that the native environment of an MP is at least partially co-extracted by the polymer, potentially providing a more native like environment than NDs. Usually NDs are

assembled with an artificial lipid composition that does not fully represent the complexity of a physiological cell membrane<sup>60</sup>. However, SMALPs have also some drawbacks, such as inhomogeneous size distribution of particles, making them less favorable for some applications like NMR.

Another route for detergent free MP solubilization is provided by the CF expression and cotranslational insertion of MPs inside provided lipid bilayers. As early as in 1990, the  $\beta_2$  adrenergic receptor was CF expressed using a rabbit reticulocyte lysate and cotranslationally inserted into microsomal membranes obtained from *Xenopus laevis* oocytes<sup>80</sup>. Later it was found, that membrane insertion can occur as a solely thermodynamically driven process without the necessity of the translocon machinery<sup>81,82</sup>. The translocon independent MP insertion and folding is not a physiological, but a thermodynamically driven process. Nonetheless, a wide array of functional MP probes was since then produced. Cotranslational insertion into MSP based NDs was first carried out by the group of Matthew Coleman, where an ApoA-1 derivative and bacteriorhodopsin were coexpressed using an *E. coli*-based CF system in presence of lipid. As a result, spontaneously formed NDs with inserted functional bacteriorhodopsin were obtained<sup>47</sup>. It soon became clear that the system is suitable to produce a variety of proteins in milligram scale and that eukaryotic lysates are also suitable for cotranslational MP solubilization, albeit yielding far less MP product<sup>83</sup>. While during first trials ApoA-1 was coexpressed with bacteriorhodopsin, a significant expression boost for the bacterial system appeared to be the abundance of preassembled NDs during expression<sup>84</sup>. Ever since, this technique was used to study biochemical properties of MPs, the process of folding, transfer and oligomerization of MPs in membranes itself<sup>45,84–88</sup>, vaccine development<sup>89</sup>, but especially appeared to be suitable for biochemical and structural investigations of GPCR<sup>27,36,90–94</sup>. In CF terminology, cotranslational insertion of MPs into provided membrane structures is referred to as lipid-based CF (L-CF) mode (Figure 3A). Over the last decade, great emphasis was laid on the synthesis of MPs inside provided NDs, which is due to easier probe handling and greater degrees of freedom regarding probe analysis compared to liposomes<sup>95</sup>. Apart from L-CF mode, there is also detergent CF (D-CF) mode, where MPs are solubilized directly during expression by provided detergent micelles in the reaction<sup>44,96</sup> (Figure 3B). This mode was adapted years before the broad use of NDs and was also applied for solving the NMR structure of proteorhodopsin<sup>28</sup>. However, it needs to be considered that the CF system is not compatible with all detergents leading to severe performance ramifications<sup>97</sup>. Also, not all MPs can be obtained in their native-like state in detergent environment<sup>64</sup>. Another way of CF MP expression is closest to conventional reconstitution – the precipitate CF (P-CF) mode. In the P-CF mode no membranes or amphiphatic molecules are provided and the target MP is left to precipitate during expression<sup>98</sup> (Figure 3C). Subsequently, the pellet is washed, solubilized, usually affinity purified and, if required, reconstituted inside NDs or proteoliposomes<sup>88,96</sup>. P-CF with subsequent solubilization in detergent was the most frequently used method for structural analysis of less challenging MP targets<sup>99,100</sup>. However, in some cases P-CF expressed protein precipitate already resembles its native-like form, e.g., amyloid prion structures<sup>101</sup>.



**Figure 3.** Different CF expression modes for MP synthesis. **A:** Lipid mode (L-CF): Lipid membranes, here: NDs, are provided in the CF reaction and the synthesized hydrophobic helices enter the provided membrane without support of the translocon machinery. **B:** Detergent mode (D-CF): Amphipathic detergent molecules are used to solubilize the freshly synthesized MP. **C:** Precipitate mode (P-CF): The freshly synthesized MP is left to precipitate due to lack of hydrophobic environment.

Facing an unknown MP target, it is advisable to conduct initial screening tests, e.g., expression yield of a new target can be estimated in the P-CF mode. Subsequently, expression inside a set of NDs with different lipid composition can be carried out. After purification, nanoparticle quality can be checked by size exclusion chromatography (SEC). This indicates if the protein is prone to produce aggregates with the provided NPs or detergent. If available, activity tests can be carried out to find a favorable hydrophobic environment for the target. This can be done either by ligand binding studies or, as was demonstrated with GPCRs, via interaction partner binding with G-proteins<sup>27,90</sup>. In this study, the SapNP system was tested regarding its suitability for cotranslational solubilization of MPs<sup>45</sup>. Furthermore, different NP assembly modes were studied, such as the use of preassembled NPs, but also the capability to form MP-SapA-lipid particles from separate components. This included addition of purified SapA with lipids or coexpression of both, scaffold protein and MP.

### 1.3 Membrane protein NMR

#### 1.3.1 State of the art and challenges

More than five decades were dedicated to MP investigations using NMR. To date, NMR methods contributed 14,099 protein data bank (PDB) entries which is about 7 % of the overall contributions with X-Ray crystallography dominating the field of structural biology with 178,796 entries. However, over the past decade cryo-EM paved its way to become the dominant technique with an exponentially growing amount of deposited PDB structures (<https://www.rcsb.org/stats/> as of 10/2023). This is due to recent breakthroughs in EM resolution allowing high resolution imaging of very large protein complexes<sup>102</sup>. So far, MPs account for <1 % of all deposited PDB structures and given the advantages of cryo-EM, one might wonder if NMR has a future in structural biology of MPs<sup>102</sup>. Indeed, studies of MPs are particularly challenging due to slow tumbling rates of large molecules inside membrane mimetics that result in faster relaxation of the transverse magnetization, and consequently, decreased sensitivity and broad linewidths<sup>103</sup>. Even the smallest stable MSP1D1ΔH5 NDs will contribute another

~90 kDa to the MP/ND complex<sup>70</sup>. A possibility to maintain a lipid bilayer in a leaner way is the embedding of MPs in bicelles where the MP is inside a lipid bilayer, which itself is kept soluble by a detergent micelle<sup>104</sup>. However, as previously discussed, detergent contact might not be desired for some targets. If however, a functional sample was obtained, NMR can provide information on MP structure and dynamics both in liquid, solid, viscous or fluid samples<sup>105</sup>. To obtain the most information from a sample, <sup>13</sup>C, <sup>15</sup>N and <sup>2</sup>H labeling is highly attractive. Perdeuteration of large samples for liquid state NMR reduces the proton density, slows down relaxation and thus, improves spectral quality<sup>106</sup>. To obtain protonated amide groups in the protein backbone, often the target has to be denatured to allow protons to access the hidden amide groups to replace the deuterons<sup>107</sup>. This approach is convenient when using reconstitution or refolding techniques, where the target is in a more dynamic detergent environment or denatured by chaotropic salt anyway<sup>103</sup>. However, incomplete back-exchange in detergent environment can lead to missing resonances which in turn might require further protein destabilization by higher temperature, denaturants or pressure<sup>108</sup>. Another obstacle for NMR analysis of large targets is spectral overlap and spectral ambiguities due to repetitiveness of hydrophobic amino acids or dominance of a single type of a secondary structure, e.g.,  $\alpha$ -helices in MPs or  $\beta$ -sheets in amyloids<sup>109</sup>. To overcome such ambiguities, selective labeling approaches can be employed<sup>110</sup>.

Also, NMR methodology itself was improved by the advances of ultra-high-field NMR instruments, cryogenic probes and transverse relaxation-optimized spectroscopy (TROSY), where only slowly decaying nuclear magnetization components are contributing to the final signal<sup>111–113</sup>. Indeed, the combination of the latter strategies allowed to obtain structural information and conformational changes of large proteins<sup>111</sup>, protein complexes<sup>114</sup> and MPs<sup>28,104</sup>. Especially in GPCR molecular recognition is often analyzed by NMR, especially in detergent environment<sup>115–119</sup>. However, NMR analysis of GPCR in membrane mimetics, e.g., in bicelles or NDs or SapNPs is also possible<sup>72,76,120</sup>. To date, selective labeling schemes in live *E. coli* cells are restricted as most AAs are prone to metabolic conversion which is referred to as “scrambling”. Complex labeling schemes or unique pair labeling remain restricted as spectra get too ambiguous, and signal intensity is lost due to label dilution<sup>121</sup>. To overcome these problems in live cell expression, auxotrophic strains often carrying sets of mutations for unambiguous labeling of a single AA were employed<sup>121–124</sup>. For L-Arg, L-Cys, L-Gly, L-Ile and L-Thr a single gene deletion respectively eliminates scrambling, which enabled construction of useful cumulative mutant strains<sup>122,125,126</sup>. No misincorporation of <sup>15</sup>N labels is usually observed for L-Met, L-Pro, L-His and L-Lys<sup>126</sup>. Also, unlabeled strategy can be applied, where all AAs except the AAs of interest are provided together with <sup>15</sup>N NH<sub>4</sub>Cl to the medium. The *E. coli* cells will incorporate the available unlabeled AAs for protein synthesis, but will synthesize the missing AA from <sup>15</sup>N NH<sub>4</sub>Cl<sup>123,125</sup>. Furthermore, the presence of certain AAs at a high concentration can also reduce scrambling caused by transamination by feedback inhibition<sup>127</sup>. Another alternative is obtaining selectively labeled methyl groups via relatively cheap AA precursors to incorporate <sup>13</sup>C labeled methyl groups in methyl containing AAs such as L-Val, L-Ile, L-Leu, L-Thr, L-Met and L-Ala. Especially, labeling of the hydrophobic AAs is highly convenient for MPs as these AAs are highly abundant in transmembrane domains (TMDs)<sup>128</sup>. Precursor labeling will be in detail discussed in chapter 1.3.5. Due to disruption of metabolic pathways during lysate processing, CF lysates display a far lower metabolic conversion of amino acids compared to live cells, and, due to the open nature of the system, allow more complex labeling schemes<sup>129</sup>.

### 1.3.2 Advances in NMR studies of CF expressed MPs

Over the last decades CF expression proved to be advantageous for production of labeled MPs, due to its open nature, high productivity, reduced level of AA metabolism and overall system robustness<sup>42</sup>. The different expression modes discussed in 1.2.4 further provide the experimenter with tools to tailor expression conditions for the protein of interest. Some examples of MPs, CF-expressed in P-CF, D-CF or L-CF mode, and subsequently analyzed by NMR are listed in table 1. It is notable that most solved NMR structures of MPs are obtained from proteins not larger than 20 kDa, which is due to restrictions that arise with larger proteins discussed in the previous chapter. Out of 150 human MP targets sized 10-30 kDa, Klammt et al found 32 MPs suitable for NMR studies according to preliminary TROSY-HSQC data<sup>99</sup>. The structure of CF expressed proteorhodopsin (PR) is the second structure of a 7 TMD MP solved by solution NMR which demonstrates the potential of the CF system to produce MPs in D-CF mode suitable for NMR<sup>28</sup>. It is indeed notable that most MP spectra are recorded in detergent environment which is mainly due to molecular size increase if MPs are embedded in NDs<sup>104</sup>. The study by Laguerre and colleagues vividly shows spectral improvement of LspA or EmrE if MSP is stripped off the nanodisc while remaining DMPC bilayer with incorporated MP is stabilized by DH7PC forming a bicelle instead<sup>104</sup>. Thus far, it is the only study of an NMR analysis of a MP that was cotranslationally inserted into the ND. Also, it is the first study where the ND is later converted into a DMPC/DH7PC bicelle<sup>104</sup>. Regarding CF expressed GPCR, so far the only published NMR analysis was carried out with the growth hormone secretagogue receptor (GHSR)<sup>130</sup> and neuropeptide Y receptor type 2 (Y2E)<sup>120</sup>.

**Table 1.** CF expressed MPs for NMR analysis. Name of MP, expression mode, labeling scheme, and purpose of study as well as literature reference are stated.

MP	Mw [kD]	CF mode	Reconstitution	Labeling scheme	Purpose of study	ref
bArcB	13.3	P-CF	LMPG	Uniform/selective	Structure	100
bQseC	21.3	P-CF	LMPG	Uniform/selective	Structure	100
bKdpD	11.2	P-CF	LMPG	Uniform/selective	Structure	100
hHIGD1A	10.1	P-CF	LMPG	Uniform/selective	Structure	99
hHIGD1B	11.0	P-CF	LMPG	Uniform/selective	Structure	99
hTMEM14A	10.1	P-CF	LMPG	Uniform/selective	Structure	99
hTMEM14B	10.6	P-CF	LMPG	Uniform/selective	Structure	99
hTMEM141	11.9	P-CF	LMPG	Uniform/selective	Structure	99
hTMEM14C	11.6	P-CF	LMPG	Uniform/selective	Structure	99
bPR	27.3	D-CF	DH7PC	Uniform/selective/ Methyl labeling	Structure	28
hTMD0	16.1	P-CF	DHPC	Uniform	Method optimization	131
$\Delta$ bTetH	24.0	P-CF	LMPG	Uniform/selective	Method establishment	132
hSREBP	23.6	P-CF	LMPG	Methyl labeling	Method establishment	133
hGHSR <sup>a</sup>	41.9	P-CF	DMPC	Selective	Method establishment	130
hY2R <sup>a</sup>	48.9	P-CF	DMPC/DH7PC	Selective	Assignment/ conformational changes	120
LspA	18.2	L-CF	DMPC/DH7PC	Uniform/selective/ Methyl labeling	Assignment/partial structural characterization/Method establishment	104
EmrE	12.0	L-CF	DMPC/DH7PC	Uniform	Method establishment	104
MscL <sup>a</sup>	75.0	P-CF	DOPC	Selective	Method establishment	134

<sup>a</sup>Analysis by ssNMR

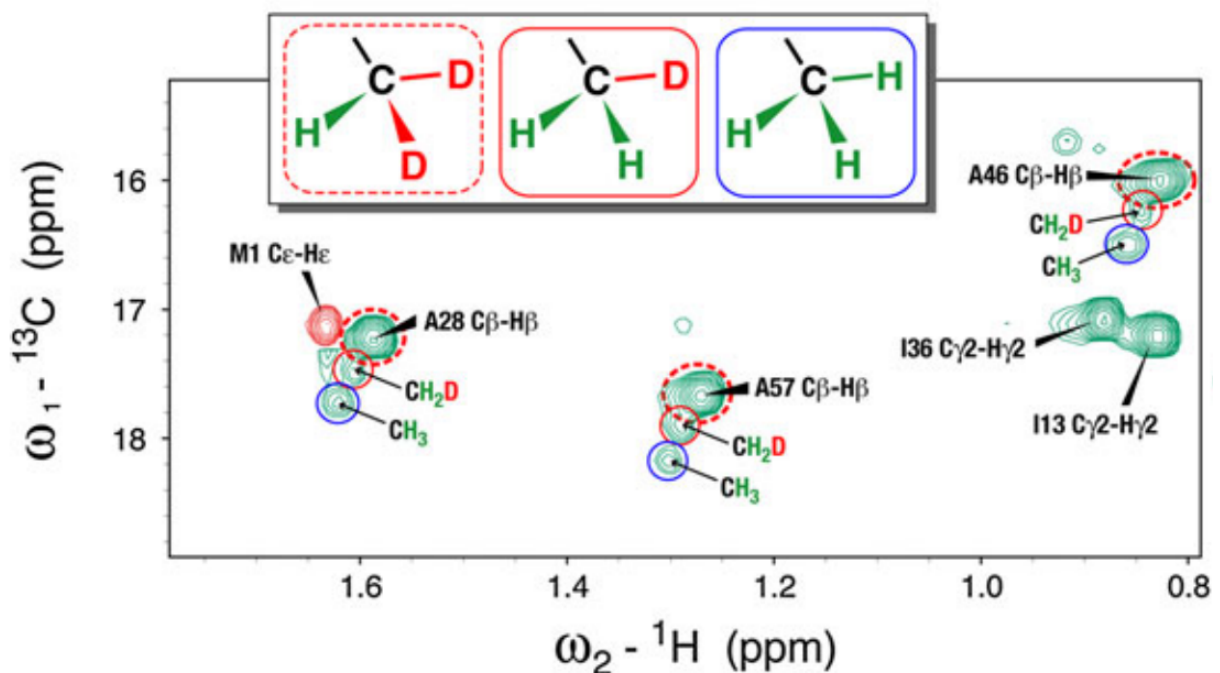
The few studies show that despite the numerous advantages taken together, only few groups seem to use the CF system for NMR analysis of MPs. Partially, the interest in such studies seemed to wane off due to rapid advances in the field of electron microscopy<sup>135</sup> or structure predicting algorithms like alphafold<sup>136</sup>. NMR is a rather insensitive technique and thus, it requires high-field spectrometers and usually numerous highly concentrated samples for assignment of challenging targets, which can lead to escalating costs<sup>42,105</sup>. Additionally, the still rather niche cotranslational MP expression technique is not well established in many groups compared to gold standard live host expression workhorses like *E. coli*, *S. cerevisiae* or *P. pastoris*. Nonetheless, analysis of dynamics at atomic resolution and behavior of proteins in solution will not be replaced so easily, especially regarding the steady use of NMR for GPCR analysis<sup>115,117</sup>. Furthermore, recent advances in CF expression of functional GPCR probes in membrane mimetics suitable for structural studies with cryo-EM could be utilized to produce sample yields suitable for NMR<sup>27,36</sup>. High yield of the CF system paired with selective labeling schemes can help to reduce spectral complexity of large complexes and minimize the number of samples<sup>107</sup>. In the next chapter, CF labeling possibilities will be explored in more detail.

### 1.3.3 Metabolic background activities in *E. coli* lysates

Despite truncated metabolic pathways, CF systems still display an array of metabolic conversions of amino acids. A standard S30 lysate was shown to be composed of ~800 proteins according to a previously conducted proteome study<sup>23</sup>. Differences were found among S12, S30 and heat shock lysates probably due to gene regulation and lysate processing conditions<sup>23</sup>. In our hands, in S30 lysates <sup>15</sup>N label scrambling most heavily affects L-Asn, L-Asp, L-Ala, L-Gln and L-Glu. Other AAs, especially L-Ser, L-Tyr, L-Trp, L-Cys, L-Phe and L-Thr are, to a far lesser extent, affected by an array of aminotransferases. Broadly, these observations are in agreement with other *E. coli* CF lysates that all show strong L-Asn, L-Asp, L-Glu and L-Gln scrambling, while some also report severe L-Ser, L-Ala and L-Gly conversions<sup>137-139</sup>. It should be noted that use of different strains, cultivation and lysate processing conditions might lead to different outcomes regarding residual scrambling activity. To gain control over these reactions, several solutions were proposed, e.g., the use of chemical inhibitors<sup>137-139</sup>. Six inhibitors are employed to inhibit scrambling in a commercially available CF expression kit (<https://www.cortecnet.com>): D-malate (DM), aminooxyacetate (AOA), Methionine sulfoximine (MS), S-methionine-L-cysteine sulfoximine (SMCS), 6-Diazo-5-oxo-norleucine (DON) and 5-diazo-4-oxo-norvaline (DONV)<sup>137</sup>. The price of a "large expression" CF kit with a RM of 750 µL and a FM of 8.5 mL is 1500 €. The scale might be sufficient for analysis of easy to express proteins like ubiquitin, however for complex targets or MPs, the experimenter would need 3-5 such kits for one sample. Unfortunately, the acquisition of glutaminase (DON), asparaginase (DONV), and asparagine synthetase (SMCS) inhibitors is not straightforward either, as the inhibitors are commercially scarce and expensive.

Less sophisticated ways of lysate treatment with NaBH<sub>4</sub> were also suggested to suppress transamination<sup>139</sup>. However, such treatment reduces lysate productivity by ~50 % and only affects PLP dependent aminotransferases<sup>139</sup>. Also, handling of toxic NaBH<sub>4</sub> is rather inconvenient due to excessive lysate foaming upon addition and general precaution measures. For this purpose, the suppression of aminotransferases by AOA, also an inexpensive inhibitor of PLP dependent aminotransferases, is more straightforward<sup>138</sup>. PLP dependent aminotransferases transfer amino groups from the

donor L-Glu to oxoacid precursors of AAs. This reaction not only causes  $^{15}\text{N}$  label scrambling, but also causes peak doublets and triplets of carbon signals due to deuterium-proton back-exchange at  $\text{C}\alpha$  and  $\text{C}\beta$  positions of deuterated AAs (Figure 4)<sup>138</sup>. Strongest back-protonation was observed at  $\text{C}\alpha$  residues of L-Asp, L-Asn, L-Glu and L-Gln, median levels of back-protonation were found in L-Gly, L-Ser, L-Cys, , L-Phe, L-Trp and L-Tyr and low degree of back-protonation in L-Val, L-Ile, L-Leu, L-Lys and L-Arg<sup>140</sup>. Highest  $\text{C}\beta$  back-protonation was observed in L-Asp, L-Asn, L-Gln, L-Glu and moderate levels were found in L-Ala, L-Phe and L-Tyr<sup>140</sup>. However, not all scrambling reactions are catalyzed by PLP dependent transaminases, and consequently, their inhibition will not fully solve the problem.



**Figure 4.** Deuterium-hydrogen back exchange at the methyl group of L-Ala. L-Ala methyl region peaks of 2D  $^1\text{H}$ ,  $^{13}\text{C}$ -HSQC spectrum of SAIL chlorella ubiquitin CF expressed using *E. coli* lysate. Red dashed circles represent the original  $^{13}\text{CD}_2\text{H}$  signals, red circles show  $^{13}\text{CH}_2\text{D}$  and blue circles show  $^{13}\text{CH}_3$  signals, where deuterium proton exchange was complete. Figure was taken from Tonelli et al<sup>138</sup>.

Another way to obtain CF lysates devoid of any side activities is the use of PURE lysates which consist of separately purified components necessary for protein expression and no other proteins<sup>6</sup>. However, PURE systems are expensive and less productive than S30 lysates, and thus, their use in NMR applications is limited<sup>103,141</sup>. To introduce NMR active probes at specific protein sites, incorporation of non-canonical AAs via amber codon use can be considered<sup>142–144</sup>. While this method has the highest labeling accuracy, its use for combinatorial labeling schemes would not be feasible.

Suppression of isotope scrambling in live *E. coli* cells by mutagenesis of genes encoding for scrambling enzymes has already been proved useful for selective isotope labeling<sup>121–124,126</sup>. This strategy appears to be even more convenient in *E. coli* lysates, since some metabolic pathways of AAs are already truncated, and fewer mutations should provide greater effects. Indeed, a plentitude of strains used for live cell



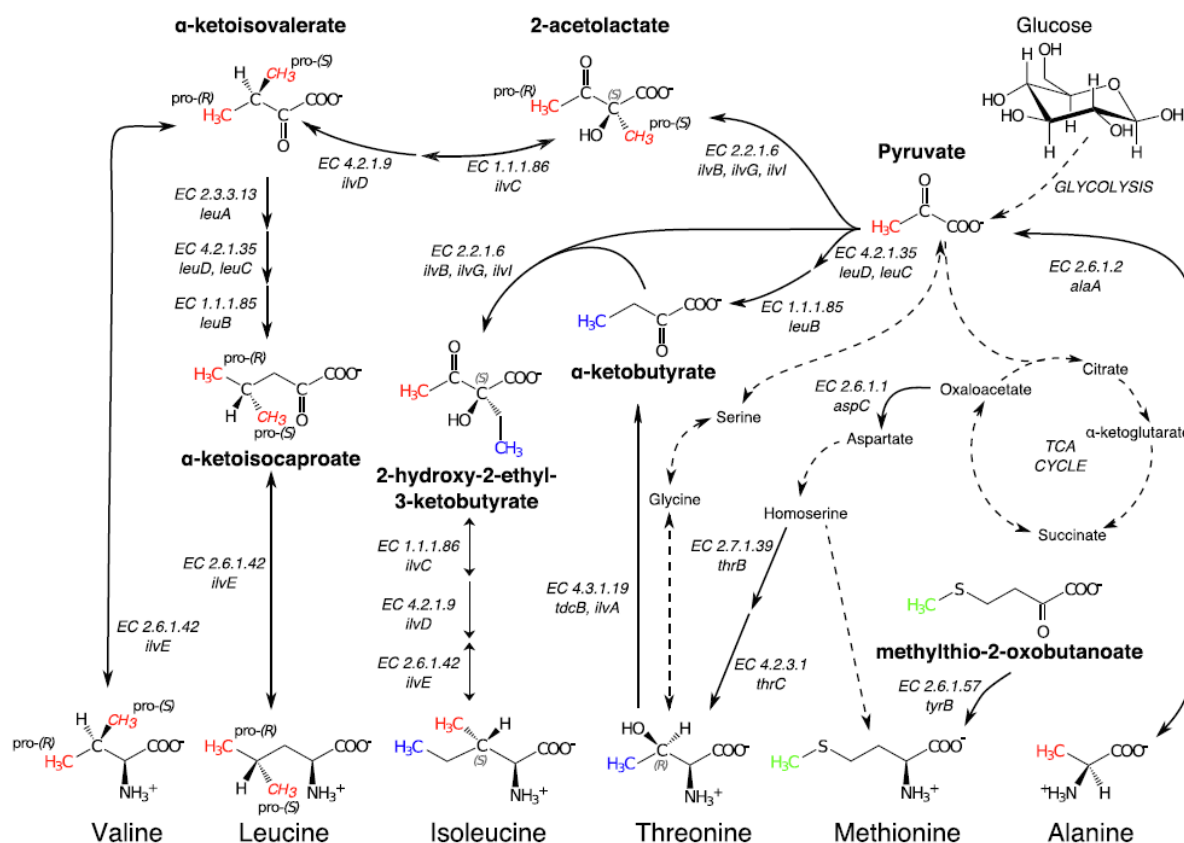
expression is already available. For example, *E. coli* DL30 carries deletions in the major aminotransferase genes *tyrB* (aromatic amino acid aminotransferase), *ilvE* (branched chain amino acid aminotransferase, short BCAAT) and *aspC* (aspartate aminotransferase)<sup>124</sup>. Further strains are engineered for only single AA labeling or carry sets of mutations in AA pathways that are no longer relevant in S30 *E. coli* lysates as they are disrupted during lysate processing anyway<sup>23,122,124</sup>. Strains with cumulative deletions to stabilize L-Trp, L-Ser and L-Arg in CF lysates were also designed<sup>145</sup>. It should also be considered that growth problems of polyauxotrophic mutants can be expected due to severe impact on AA metabolism by multiple deletions<sup>124</sup>. This is to be avoided, since poor growth will affect CF lysate productivity, and thus, product yield, which is absolutely crucial for NMR due to reasons mentioned before (1.3.2). An engineered *E. coli* A19 strain “Stablelabel” tailored for stable isotope labeling of AA by CF expression was recently published by our group and will be in detail discussed in this work<sup>146</sup>.

### 1.3.4 Precursor based methyl labeling of proteins for NMR studies

While metabolic conversions are a serious problem for selective labeling of proteins, they can also be used for other labeling applications, e.g., precursor labeling for methyl TROSY NMR, as extensively reviewed elsewhere<sup>128</sup>. Methyl labeling of MPs is highly attractive due to high abundance of hydrophobic AAs carrying methyl groups. Due to their fast rotation, methyl groups are partially decoupled from the slow overall tumbling of the large protein molecule<sup>147</sup>. Furthermore, due to the threefold protein multiplicity signal sensitivity is enhanced<sup>147</sup>. These properties are advantageous, especially in a deuterated surrounding, and provided insight into protein complexes of up to 1 MDa<sup>147,148</sup>.

There are several metabolic pathways that can be utilized to incorporate a labeled methyl group in the desired AA of a protein using *E. coli* (Figure 5). A straight forward way to obtain characteristic cross-peak patterns is the use of labeled carbon sources, e.g., 1,3-<sup>13</sup>C glycerol or 2-<sup>13</sup>C glycerol<sup>149,150</sup>, 2-<sup>13</sup>C pyruvate or 3-<sup>13</sup>C pyruvate<sup>151</sup>, or isotopologues of D-glucose<sup>152</sup> as carbon source. However, the use of early metabolic precursors on nutrient level increases the undesired effects of unspecific position labeling<sup>153</sup>. Labeling of methyl groups of single amino acids can be challenging in live cells, however certain metabolic tricks can be applied. For instance, labeling of methyl groups of L-Ala is possible by adding the labeled L-Ala directly to the culture medium. However, L-Ala may become converted to pyruvate by transaminases AlaA/AlaC<sup>154</sup>. Secondary conversions to  $\alpha$ -ketoprecursors of L-Val, L-Leu and L-Ile and tertiary conversions to the respective AAs will result in ambiguous spectra<sup>154</sup>. Thus, L-Ala labeling is only possible in combination with saturation of biosynthesis pathways of L-Val, L-Leu and L-Ile, i.e. addition of 2-ketoisovalerate (KIV) and 2-ketobutyrate<sup>154</sup>. Analogously, KIV can be applied to label L-Val only if L-Leu pathway is saturated with 2-ketoisocaproate<sup>155</sup>. This implies that the saturation precursor is deuterated as in most cases protein deuteration is a goal. Only L-Ile- $\delta$ 1 precursor labeling is more convenient as no scrambling of this methyl group occurs. L-Ile- $\gamma$ 2 labeling however, is achieved by using the precursor 2-hydroxy-2-ethyl-3-ketobutyrate in combination with unlabeled and/or deuterated KIV to prevent scrambling to L-Val and L-Leu<sup>156</sup>. <sup>13</sup>C methyl labeled oxoprecursor of L-Met 4-Methylthio-2-oxobutanoate was suggested to be converted to L-Met by the promiscuous aminotransferase TyrB<sup>128,157</sup>.

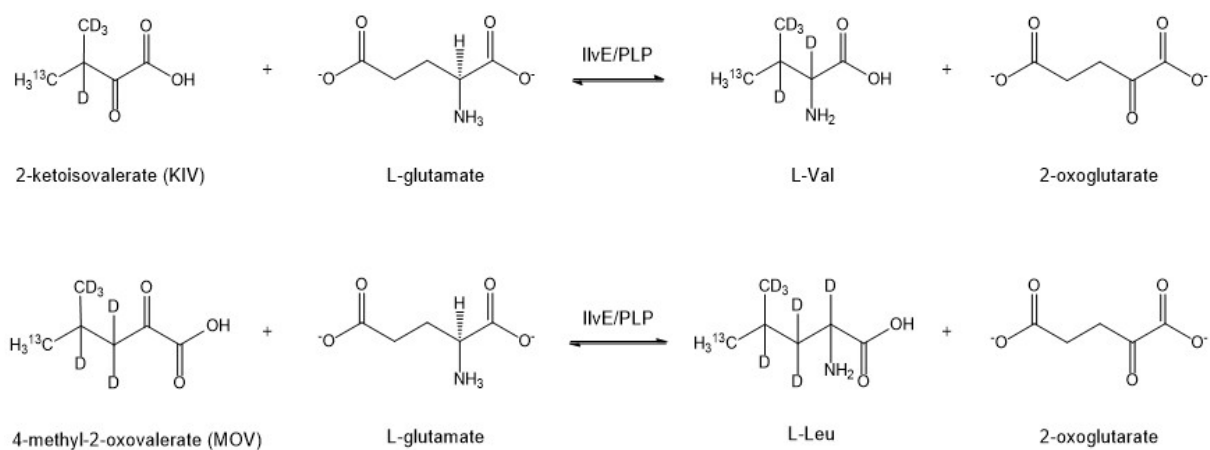
To obtain more precise residue labeling, late  $\alpha$ -ketoacid metabolic precursors of L-Val, L-Leu, L-Ile can be used<sup>158,159</sup>. This is especially attractive for analysis of MPs as transmembrane helices are to 30 % composed of these AAs<sup>160</sup>. However, commercially available KIV is a racemic mixture of pro-*S* and pro-*R* compounds resulting in two signals per L-Val or L-Leu residue<sup>161</sup>. Specific labeling of pro-*S* methyl groups is achievable using 2-acetolactate based on stereospecific rearrangement of methyl groups in 2-acetolactate. While the methyl group at the C2 carbon becomes the pro-*S* methyl group, the pro-*R* methyl group is derived from the C4 methyl moiety of 2-acetolactate<sup>161</sup>. This labeling strategy indeed results in only pro-*S* labeled methyl residues in L-Val and L-Leu, which significantly reduces spectral complexity and shows no scrambling to other AAs<sup>161</sup>. Nonetheless, the strategy does not allow single residue labeling of L-Val or L-Leu or differential pro-*S*/pro-*R* labeling of L-Val and L-Leu, which can be achieved by directly adding these stereo specifically labeled AAs to minimal growth medium<sup>162</sup>. However, stereo-array isotope labeling (SAIL) is expensive and expression protocols should be thoughtfully optimized to obtain as much labeled protein as possible from as few as possible labeled AA<sup>162,163</sup>. Furthermore, scrambling of L-Val residues remains a problem, unless auxotrophic strains are used<sup>164</sup>. Usually, SAIL AAs are applied in CF systems as in live cells the required amount for labeling is too high and thus, in most cases not feasible<sup>163,165</sup>. The catalogue of labeling schemes of L-Val, L-Leu and L-Ile is indeed very versatile and dependent on the labeling requirements, and certain criteria must be considered, such as pathway saturation or use of auxotrophic strains. Schütz and Sprangers provide a comprehensive overview of the studies on methyl labeling schemes (Figure 5)<sup>128</sup>. Despite numerous published protocols protein labeling will require tedious protocol optimization, and despite availability of relatively cheap precursors, labeling in live cells is still quite expensive considering required amounts of AAs, precursors or deuterated “pathway saturators” and deuterated media.



**Figure 5.** Metabolic pathways in *E. coli* for AA synthesis from precursors. Metabolic fate of methyl groups from the respective precursors is color coded in red, green or blue, respectively. Figure adapted from Schütz and Sprangers<sup>128</sup>.

### 1.3.5 Precursor based methyl labeling using the CF system

An elegant way of L-Val labeling via precursor KIV was presented by Lazarova and coworkers<sup>166</sup>. BCAAT (*IlvE*) was shown to be present in *E. coli* A19 lysates to an extent that would allow conversion of oxoacid precursors of L-Leu and L-Val directly in the CF reaction by transfer of the amino group from L-Glutamate (Figure 6)<sup>166</sup>. However, *IlvE* concentration in A19 lysates is still rather low<sup>23</sup> and thus, it is reasonable to either supply external *IlvE* to the CF reaction, or carry out a precursor preconversion before application of the synthesized amino acid in the CF mix<sup>146,166</sup>. The concept of precursor preconversion is not particularly new as any synthesis of labeled precursors or AAs can be considered as such<sup>159,162,167</sup>. Enzymatic *in vitro* preconversion was previously carried out to deuterate <sup>13</sup>C labeled L-Ala<sup>154</sup>. Moreover, a whole enzymatic cascade was employed to produce deuterated and methyl labeled L-Thr starting from <sup>13</sup>C formaldehyde<sup>165</sup>. Although, preconversion adds additional steps to the labeling procedure, it is advantageous because enzymatic preconversion can be better controlled, and thus, precursor amount saved. Additionally, no deuterium is lost during transamination if preconversion is carried out in a deuterated environment<sup>146</sup>. CF expression can be carried out with an *IlvE* deficient lysate to further stabilize the AA. Alternatively, transaminase inhibitors such as AOA can prevent undesired conversions between the AA and the ketoacid precursor<sup>146</sup>. This strategy was extensively characterized in this study<sup>146</sup>.



**Figure 6.** IlvE catalyzed conversion of ketoacid precursors. Precursors KIV (4- $^{13}C$ , 3, 5, 5, 5-D4) and MOV (5- $^{13}C$ , 3, 3, 4, 5, 5, 5-D6) used in this study are converted to the respective AAs L-Val and L-Leu by PLP dependent amino group transfer from L-Glutamate.

## 2 Materials

### 2.1 Equipment

#### 2.1.1 General equipment

---

500 mL/2 L baffled flasks	Schott Duran (Mainz, Germany)
100 mL, 250 mL, 500 mL, 1000 mL, 2000 mL bottles	Schott Duran (Mainz, Germany)
Agarose gel electrophoresis system	Peqlab (Erlangen, Germany)
Autoclave	Integra (Gießen, Germany)
Centrifuge 5810 R	Eppendorf (Hamburg, Germany)
Centrifuge 5920 R	Eppendorf (Hamburg, Germany)
Centrifuge Heraeus Megafuge 16R	Thermo Scientific (Langenselbold, Germany)
Centrifuge Sorvall RC 5C	Thermo Scientific (Langenselbold, Germany)
Customized CF expression reactors	Goethe University (Frankfurt, Germany)
Electroporation cuvettes 1 mm	Thermo Scientific (Langenselbold, Germany)
Electroporator 2510	Eppendorf (Hamburg, Germany)
Fermenter Biostat ED, 10 L	B. Braun Biotech (Braunschweig, Germany)
French pressure cell disruptor	SLM Aminco Instruments (Irvine, USA)
Heated waterbath	Julabo (Seebach, Germany)
Immobilon-P PVDF membrane	Millipore, Merck (Darmstadt, Germany)
Incubator shaker Infors HT	Infors (Bottmingen, Switzerland)
MilliQ water purification system	Millipore, Merck (Darmstadt, Germany)
Mini centrifuge C-1200	National Labnet Co. (Woolbridge, USA)
Mini extruder	Avanti Polar Lipids (Alabaster, USA)
NanoDrop 1000	Peqlab (Erlangen, Germany)
Pipetus	Hirschmann (Eberstadt, Germany)
PHM210 Standard pH-meter	Radiometer Copenhagen (Bronshoj, Denmark)
Rotary evaporator Rotavapor RE12	Buechi (Essen, Germany)
Sonifier Labsonic U	B. Braun Biotech (Braunschweig, Germany)
Table Top Centrifuge Mikro 22	Hettich (Tuttlingen, Germany)
Thermocycler	Peqlab (Erlangen, Germany)
Thermomixer 5436	Eppendorf (Hamburg, Germany)
Ultra-Clear centrifuge tubes	Beckman Coulter (Brea, USA)
Waterbath MW-4	Julabo (Seelbach, Germany)
Xtra TT elmasonic	Elma (Singen, Germany)

#### 2.1.2 Equipment for protein purification and analysis

---

Amicon (3/10/30/50 kDa MWCO)	Millipore Merck (Darmstadt, Germany)
ÅKTA purifier FPLC system	GE healthcare (Munich, Germany)
Dialysis tubes type 27/30 (12-14 kDa MWCO)	Spectrum Labs (Rancho Dominguez, USA)
Gravity Flow Columns	Bio-Rad (Munich, Germany)
HiTrap® desalting columns	(Cytiva, Malborough, Massachusetts, USA)
HiTrap® IMAC Fast Flow columns	(Cytiva, Malborough, Massachusetts, USA)
Immobilon-P PVDF membrane	Millipore Merck (Darmstadt, Germany)
Lumi Imager F1	Roche (Penzberg, Germany)
Mini-Protean electrophoresis system	Bio-Rad (Munich, Germany)
Multimode microplate reader Spark	Tecan (Meannedorf, Switzerland)
Ni-NTA superflow resin	Qiagen (Hilden, Germany)
NMR spectrometer 600 MHz (Avance II)	Bruker (Billerica, Massachusetts, USA)
NMR spectrometer 700 MHz (Avance III HD)	Bruker (Billerica, Massachusetts, USA)

## Materials

NMR spectrometer 800 MHz (Avance III)	Bruker (Billerica, Massachusetts, USA)
NMR spectrometer 900 MHz (Neo)	Bruker (Billerica, Massachusetts, USA)
NMR spectrometer 950 MHz (Avance III)	Bruker (Billerica, Massachusetts, USA)
Nunc™ F96 MicroWell™ Black polystyrene plate	Thermo Scientific (Langensfeld, Germany)
PD Mditrap™ G-25	(Cytiva, Malborough, Massachusetts, USA)
Peristaltic pump	Ismatec (Wertheim, Germany)
Slide-A-Lyzer dialysis cassette 0.5-3 mL and 3-12 mL (10 kDa MWCO)	Thermo Scientific (Langensfeld, Germany)
Sterican® disposable drain tubes	B. Braun (Melsungen, Germany)
StreptII-Tactin resin	Iba (Goettingen, Germany)
Superdex 75 16/600	GE Healthcare (Munich, Germany)
Superdex 200 3.2/300	GE Healthcare (Munich, Germany)
Superdex 200 10/300	GE Healthcare (Munich, Germany)
Superose 6 3.2/300	GE Healthcare (Munich, Germany)
Superose 6 10/300	GE Healthcare (Munich, Germany)
Q-Sepharose	(Cytiva, Malborough, Massachusetts, USA)
Ultrafiltration devices Amicon® Ultra 0.5, 4 and 15 (3 kD, 10, kD, 30 kD, 50 kD)	Millipore Merck (Darmstadt, Germany)
Western blot apparatus Mini-Protean	Bio-Rad (Munich-Germany)

## 2.2 Strains, cloning and DNA preparation

### 2.2.1 *E. coli* strains

strain	genotype
<i>E. coli</i> A19 K12	[ <i>rna-19 gdhA2 his-95, relA1, spoT1, metB1</i> ]
<i>E. coli</i> A19 K12 derivatives	
M1	<i>asnA</i> <sup>-</sup>
M1-6	<i>asnA</i> <sup>-</sup> , <i>glnA-his<sub>6</sub></i>
M1-6-7	<i>asnA</i> <sup>-</sup> , <i>glnA-his<sub>6</sub>, glsA1</i> <sup>-</sup>
M5	<i>gltB</i> <sup>-</sup>
M1-3-4-6-7	<i>asnA</i> <sup>-</sup> , <i>ansA</i> <sup>-</sup> , <i>ansB</i> <sup>-</sup> , <i>glnA-his<sub>6</sub>, glmS</i> <sup>-</sup>
M1-3-4-6-7-9	<i>asnA</i> <sup>-</sup> , <i>ansA</i> <sup>-</sup> , <i>ansB</i> <sup>-</sup> , <i>glnA-his<sub>6</sub>, glmS</i> <sup>-</sup> , <i>aspC</i> <sup>-</sup>
M1-3-4-6-9-10 “Stablelabel”	<i>asnA</i> <sup>-</sup> , <i>ansA</i> <sup>-</sup> , <i>ansB</i> <sup>-</sup> , <i>glnA-his<sub>6</sub>, aspC</i> <sup>-</sup> , <i>ilvE</i> <sup>-</sup>
<i>E. coli</i> DH5α	F <sup>-</sup> endA1 <i>glnV44 thi-1 recA1 relA1</i> <i>gyrA96 deoR nupG purB20 f80dlacZDM15</i> D( <i>lacZYA-argF</i> )U169, <i>hsdR17(rK-mK+)</i> , Δ( <i>ara-leu</i> )7697 Δ <i>lacX74 ΔphoA PvuII phoR</i> <i>araD139 ahpC galE galK rpsL</i> (DE3) F' <i>[lac<sup>+</sup> lacI<sup>q</sup></i> <i>pro] gor522::Tn10 trxB pRARE2</i> (Cam <sup>R</sup> , Str <sup>R</sup> , Tet <sup>R</sup> )
<i>E. coli</i> Rosetta-gami-2 DE3	<i>fhuA2 lacZ::T7 gene1 [lon] ompT gal</i> <i>sulA11 R(mcr-73::miniTn10-TetS)2 [dcm]</i> <i>R(zgb-210::Tn10-TetS) endA1 D(mcrCmrr)</i> 114::IS10
<i>E. coli</i> T7 Express	

### 2.2.2 Knock-out and screening Primers

<i>asnA</i> -KO-fw	GCATGTGTGTCGGTTTTTGTGCTTAATCATAAGCAACAGGACGCAGGAGA ATTAACCCTCACTAAAGGGCG
<i>asnA</i> -KO-rev	GATACCGGATGCGAAGCCGCCTGCTCAGACGCTGGCGGCGATAAATTAT AATACGACTCACTATAGGGCTC
<i>asnA</i> -genscr-fw	GCCAAATTGTTTCGCCAG
<i>asnA</i> -genscr-rev	GGTTTGATGCCGATGCG
<i>ansA</i> -KO-fw	CTGCCTCACGTATATACTTTTGCTCTTTCGATATCATTCAATATCAAT TAACCCTCACTAAAGGGCG
<i>ansA</i> -KO-rev	CATTTTGTAATCGACCAGTAACAGGGCGCGAGGGGGCATTACAGTCTCCT AATACGACTCACTATAGGGCTC
<i>ansA</i> -genscr-fw	CGTTGATGAACCGACCTTC

ansA-genscr-fw CACGCAATAGTCAGTAGC  
 ansB-KO-fw GCGATTAGTACTGATTGAAGATCTGCTGGATCTGCTGCGGATCTTTGGTTT  
 G AATTAACCCTCACTAAAGGGCG  
 ansB-KO-rev CAGAGCTAAGGGATAATGCGTAGCGTTCACGTAAGTGGAGGAATGAAATGT  
 AATACGACTCACTATAGGGCTC  
 ansB-genscr-fw CGTTCCTTTACTGGTTAAC  
 ansB-genscr-rev CAAACTTGATGCGCAGC  
 gltB-KO-fw ATGACACGCAAACCCCGTCGCCACGCTCTTTCTGTGCCCGTGCGCAGCGG  
 AATTAACCCTCACTAAAGGGCG  
 gltB-KO-rev TTACTGCGCTGCACGCGCAACTCTGCTGCGCTACGACTACGGTGACCCA  
 TAATACGACTCACTATAGGGCTC  
 gltB-genscr-fw CCTTTTGCCATAACGAC  
 gltB-genscr-rev GGATGTAGTTGTGTACCGG  
 glnA-His-fw GCGTGCGTATGACTCCGCATCCGGTAGAGTTTGAGCTGTACTACAGCGTCC  
 ATCACCATCACCATCACTAAAATTAACCCTCACTAAAGGGCG  
 glnA-His-rev GATCCCAGGCCTGCCAGAGACAGGCGAAAAGTTTCCACGGCAACTAAAAC  
 ACTAATACGACTCACTATAGGGCTC  
 glnA-genscr-fw GTGATCGCTTTACGGAG  
 glnA-genscr-rev GTCATCTTCTTCGCGACG  
 glmS-KO-fw CGTCTGGAATACCGCGGATATGACTCTGCCGGTCTGGCCGTTGTTGATGCA  
 ATTAACCCTCACTAAAGGGC  
 glmS-KO-rev CACGCGCGCGAACTTCTTCAATGTTGGATTTTCAGTTTTTCCAGCAATTCTAA  
 TACGACTCACTATAGGGCTC  
 glmS-genscr-fw GCAACTGCTGGCTG  
 glmS-genscr-rev GCGTCAACAGGTACAG  
 glsA1-KO-fw GCACCGGGTGAGGAATTACCTCCCGCATCTATAAAAAGGAGTTAACAAAA  
 GAATTAACCCTCACTAAAGGGCG  
 glsA1-KO-rev GTTCCATAGACCGAGAGGTTTGTTACCGTATTACCTTCCGTGTTTCATCATG  
 TAATACGACTCACTATAGGGCTC  
 glsA1-genscr-fw GCAACTGCTGGCTG  
 glsA1-genscr-rev GCGTCAACAGGTACAG  
 aspC-KO-fw GTTACCCTGATAGCGGACTTCCCTTCTGTAACCATAATGGAACCTCGTCATA  
 ATTAACCCTCACTAAAGGGCG  
 aspC-KO-rev GCTTTTCAGCGGGCTTCATTGTTTTAATGCTTACAGCACTGCCACAATCGT  
 AATACGACTCACTATAGGGCTC  
 aspC-genscr-fw GCTGTGGGTATCGTTTACCAG  
 aspC-genscr-rev GGATTTCTGGCAAAGTGCG  
 tyrB-KO-fw CATGAGTCTCACTCTGTTGCTAATTGCCGTTTCGCTCCTGAACATCCACTCGA  
 ATTAACCCTCACTAAAGGGCG  
 tyrB-KO-rev GCTCCAGCCTGCTTTCCTGCATTACATCACCAGCAAAACGCCTTTGCCAC  
 ATAATACGACTCACTATAGGGCTC  
 tyrB-genscr-fw CTGCGTAGACTTAGGTCC  
 tyrB-genscr-rev CGCTTTGCTGTTTTGCC  
 ilvE-KO-fw GAGCACAACCACATCACAACAAATCCGCGCCTGAGCGCAAAGGAATATAA  
 AAAATTAACCCTCACTAAAGGGCG  
 ilvE-KO-rev GTATTTATTGATTAACCTGATCTAACCAGCCCCATTTATCTTCGGTTTTGCCT  
 AATACGACTCACTATAGGGCTC  
 ilvE-genscr-fw CTTCCGGCATCCATGCG  
 ilvE-genscr-rev CGCAGTTAGAGATGCAGAC  
 KanR-screen-fw GTATCCATCATGGCTGATG

### 2.2.3 Equipment for cloning and DNA preparation

---

Mini DNA preparation kit	Macherey-Nagel (Dueren, Germany)
Midi DNA preparation kit	Qiagen (Hilden, Germany)
QIAquick gel extraction kit	Qiagen (Hilden, Germany)
QIA quick PCR purification kit	Qiagen (Hilden, Germany)
Quick and Easy <i>E. coli</i> gene deletion kit	Gene Bridges (Heidelberg, Germany)

## 2.3 Chemicals and reagents

### 2.3.1 General chemicals

---

2-ketoisovaleric acid (KIV)	Sigma-Aldrich (Taufkirchen, Germany)
2,2-Bis(hydroxymethyl)-2,2',2''-nitrilotriethanol	Carl Roth (Karlsruhe, Germany)
2-mercaptoethanol	Carl Roth (Karlsruhe, Germany)
4-methyl-2-oxovaleric acid	Sigma-Aldrich (Taufkirchen, Germany)
6-diazo-5-oxo-norleucine	Sigma-Aldrich (Taufkirchen, Germany)
Acetyl phosphate (AcP)	
Agarose	Roche (Penzberg, Germany)
Alprenolol hydrochloride	Tocris Bioscience (Bristol, UK)
Ammonium persulfate (APS)	Carl Roth (Karlsruhe, Germany)
Ampicillin	Carl Roth (Karlsruhe, Germany)
Antifoam Y-30 Emulsion	Sigma-Aldrich (Taufkirchen, Germany)
Arabinose	Carl Roth (Karlsruhe, Germany)
Bromphenolblue	Carl Roth (Karlsruhe, Germany)
L-alanine (A)	Carl Roth (Karlsruhe, Germany)
L-arginine	Carl Roth (Karlsruhe, Germany)
L-asparagine monohydrate (N)	Carl Roth (Karlsruhe, Germany)
L-aspartic acid (D)	Carl Roth (Karlsruhe, Germany)
Adenosine 5'-triphosphate (ATP)	Roche (Penzberg, Germany)
Chloramphenicol	Carl Roth (Karlsruhe, Germany)
Citric acid	Carl Roth (Karlsruhe, Germany)
cOmplete protease inhibitor cocktail	Roche (Penzberg, Germany)
L-Cysteine	Carl Roth (Karlsruhe, Germany)
Cytidine 5'-triphosphate (CTP)	Biomol (Hamburg, Germany)
d-Desthiobiotin	Sigma-Aldrich (Taufkirchen, Germany)
D <sub>2</sub> O	Sigma-Aldrich (Taufkirchen, Germany)
Dithiothreitol (DTT)	Carl Roth (Karlsruhe, Germany)
Di-potassium hydrogen phosphate (K <sub>2</sub> HPO <sub>4</sub> )	Carl Roth (Karlsruhe, Germany)
Folinic acid calcium salt	Sigma-Aldrich (Taufkirchen, Germany)
Gel loading dye, purple (6x)	NEB (Frankfurt, Germany)
Glucose monohydrate	Carl Roth (Karlsruhe, Germany)
L-glutamine	Carl Roth (Karlsruhe, Germany)
L-glutamic acid	Carl Roth (Karlsruhe, Germany)
Gluthatione oxidized (GSSG)	Carl Roth (Karlsruhe, Germany)
Gluthatione reduced (GSH)	Carl Roth (Karlsruhe, Germany)
Glycine	Sigma-Aldrich (Taufkirchen, Germany)
Guanidine-5'-triphosphate (GTP)	Thermo Scientific (Langensfeld, Germany)
Guanidinium hydrochloride (GDN)	Carl Roth (Karlsruhe, Germany)
L-histidine (H)	Carl Roth (Karlsruhe, Germany)
L-isoleucine (I)	Carl Roth (Karlsruhe, Germany)
Kanamycin sulfate	Carl Roth (Karlsruhe, Germany)
L-leucine (L)	Carl Roth (Karlsruhe, Germany)
L-lysine (K)	Carl Roth (Karlsruhe, Germany)
L-methionine (M)	Carl Roth (Karlsruhe, Germany)
Magnesium acetate tetrahydrate (Mg(OAc) <sub>2</sub> )	Carl Roth (Karlsruhe, Germany)
Magnesium chloride (MgCl <sub>2</sub> )	Carl Roth (Karlsruhe, Germany)
Peptone from casein	Carl Roth (Karlsruhe, Germany)
L-phenylalanine (F)	Carl Roth (Karlsruhe, Germany)
L-phosphoenolpyruvic acid (PEP)	Sigma-Aldrich (Taufkirchen, Germany)



L-proline (P)	Carl Roth (Karlsruhe, Germany)
Potassium dihydrogen phosphate (PH <sub>2</sub> PO <sub>4</sub> )	Carl Roth (Karlsruhe, Germany)
Potassium hydroxide (KOH)	Carl Roth (Karlsruhe, Germany)
RiboLock RNase inhibitor	Thermo Scientific (Langensfeld, Germany)
Rotiphorese 30	Carl Roth (Karlsruhe, Germany)
Rotiszint eco plus	Carl Roth (Karlsruhe, Germany)
L-serine	Carl Roth (Karlsruhe, Germany)
Skim milk powder	Sigma-Aldrich (Taufkirchen, Germany)
Sodium acetate (NaH <sub>2</sub> PO <sub>4</sub> )	Carl Roth (Karlsruhe, Germany)
Sodium chloride (NaCl)	Carl Roth (Karlsruhe, Germany)
Sodium citrate	Carl Roth (Karlsruhe, Germany)
Sodium hydroxide (NaOH)	Carl Roth (Karlsruhe, Germany)
Streptomycin sulfate	Carl Roth (Karlsruhe, Germany)
SYBR green	Bio-Rad (Munich, Germany)
Tetracycline	Carl Roth (Karlsruhe, Germany)
L-Threonine	Sigma-Aldrich (Taufkirchen, Germany)
Tris(hydroxymethyl)aminomethane (TRIS)	Carl Roth (Karlsruhe, Germany)
tRNA <i>E. coli</i> MRE 600	Roche (Penzberg, Germany)
L-Tryptophane	Carl Roth (Karlsruhe, Germany)
Urea	Carl Roth (Karlsruhe, Germany)
Uridine triphosphate (UTP)	Sigma-Aldrich (Taufkirchen, Germany)
L-valine	Carl Roth (Karlsruhe, Germany)
Yeast extract	Carl Roth (Karlsruhe, Germany)

### 2.3.2 Isotopes

[ <sup>3</sup> H]-Dihydroalprenolol	Biotrend (Cologne, Germany)
L-asparagine- <sup>15</sup> N	Cambridge Isotope Laboratories Inc. (Tewksbury, Massachusetts, USA)
L-aspartic acid- <sup>15</sup> N	Cambridge Isotope Laboratories Inc. (Tewksbury, Massachusetts, USA)
L-glutamine- <sup>15</sup> N	Cambridge Isotope Laboratories Inc. (Tewksbury, Massachusetts, USA)
L-glutamic acid- <sup>15</sup> N	Cambridge Isotope Laboratories Inc. (Tewksbury, Massachusetts, USA)
2-ketoisovaleric acid (4- <sup>13</sup> C, 3, 5, 5, 5-D4)	Cambridge Isotope Laboratories Inc. (Tewksbury, Massachusetts, USA)
4-methyl-2-oxovaleric acid (5-methyl <sup>13</sup> C, 3, 3, 4, 6, 6, 6-D6)	Provided by Roman Lichtenecker Institute of Organic Chemistry, University of Vienna
Cell free amino acid mix (20 AA) (U-d)	Cambridge Isotope Laboratories Inc. (Tewksbury, Massachusetts, USA)
Cell free amino acid mix (20 AA) (U- <sup>15</sup> N)	Cambridge Isotope Laboratories Inc. (Tewksbury, Massachusetts, USA)

### 2.3.3 Detergents and lipids

1,2-Dielaidoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DEPG)	Avanti Polar Lipids (Alabaster, USA)
1,2-diheptanoyl-sn-glycero-3-phosphocholine (DH7PC)	Avanti Polar Lipids (Alabaster, USA)
1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC)	Avanti Polar Lipids (Alabaster, USA)
1,2-Dimyristol-sn-glycero-3-phospho-(1'-rac-glycerol) (DMPG)	Avanti Polar Lipids (Alabaster, USA)
1,2-Dioleoyl-sn-glycero-3-phosphocholine (DOPC)	Avanti Polar Lipids (Alabaster, USA)
1,2-Dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DOPG)	Avanti Polar Lipids (Alabaster, USA)
1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (POPG)	Avanti Polar Lipids (Alabaster, USA)
n-Dodecylphosphocholine, Fos-12 (DPC)	Anatrace (Santa Clara, USA)
Polyoxoethylene-sorbitan-monolaurate (Tween 20)	20 Carl Roth (Karlsruhe, Germany)
Polyethylene-glycol-p-1,2,3,3-tetramethyl-butylphenyl-ether (Triton X-100)	Carl Roth (Karlsruhe, Germany)
Sodium cholate	Carl Roth (Karlsruhe, Germany)
Sodium dodecylsulfate	Carl Roth (Karlsruhe, Germany)

### 2.3.4 Enzymes, protein and antibodies

Anti-mouse IgG-Fab HRP conjugate from goat	Sigma-Aldrich (Taufkirchen, Germany)
Anti-penta His IgG from mouse	Qiagen (Hilden, Germany)
Bovine serum albumin (BSA)	Carl Roth (Karlsruhe, Germany)
Gel filtration calibration kit (HMW/LMW)	(Cytiva, Malborough, Massachusetts, USA)
Precision Plus Protein™ Unstained Standards	Bio-Rad (Munich, Germany)
Precision Protein™ Strep-Tactin-HRP	Bio-Rad (Munich, Germany)
Pyruvate kinase	Roche (Penzberg, Germany)
Restriction enzymes	NEB (Frankfurt, Germany)
T4 DNA ligase	NEB (Frankfurt, Germany)
T7 RNA polymerase	in house prepared

## 2.4 Culture media and antibiotics

### 2.4.1 Culture media and antibiotic stocks

Ampicillin stock solution (1000 x)	100 mg/mL Sodium ampicillin in 70 % (v/v) ethanol
Chloramphenicol	34 mg/mL chloramphenicol in 70 % (v/v) ethanol
Double bile (DB) salt solution	5.22 mM NH <sub>4</sub> H <sub>2</sub> PO <sub>4</sub> 0.05 mM MgSO <sub>4</sub> 0.2 mM KCl 44 % (w/v) glycerol
IPTG stock solution (1000 x)	1 M IPTG in H <sub>2</sub> O
Kanamycin stock solution (1000 x)	30-100 mg/mL in Milli-Q
Luria-bertani (LB) broth	10 g/L Peptone from casein 5 g/L Yeast extract 10 g/L NaCl (+ 15 g/L agar-agar for agar plates)
SOC medium	20 g/L peptone from casein 5 g/L yeast extract 0.5 g/L NaCl

Terrific broth (TB)	25 mM KCl 10 mM MgCl <sub>2</sub> 20 mM glucose 12 g/L peptone from casein 24 g/L yeast extract 0.5 % (w/v) glycerol 100 mM glucose (sterile filtered) 1.7 mM KH <sub>2</sub> PO <sub>4</sub> 7.2 mM K <sub>2</sub> HPO <sub>4</sub>
Tetracycline stock solution (1000 x) YPTG broth	5 mg/mL tetracycline in 70 % (v/v) ethanol 16 g/L peptone from casein 10 g/L yeast extract 5 g/L NaCl 100 mM glucose (sterile filtered) 22 mM KH <sub>2</sub> PO <sub>4</sub> 40 mM K <sub>2</sub> HPO <sub>4</sub>

## 2.5 Buffers and stock solutions

### 2.5.1 Cell-free expression and lysate processing

Acetyl phosphate AcP stock (50 x)	1 M AcP-KOH pH 7.0
Amino acid stock (25 x)	25 mM of each natural amino acid in Milli-Q
Bis-Tris buffer stock (21 x)	2.1 M, pH 8.0 in milliQ
Complete protease inhibitor stock (50x)	1 tablet dissolved in 1 mL Milli-Q
DTT stock	500 mM DTT in Milli-Q
Folinic acid stock (100 x)	10 mg/mL in Milli-Q
Glutathione oxidized (GSSG)	100 mM in Milli-Q
Glutathione reduced (GSH)	100 mM in Milli-Q
HEPES buffer stock (25 x)	2.5 M HEPES-KOH, pH 8.0
KOAc stock	0.02 M EDTA 10 M in Milli-Q
Mg(OAc) <sub>2</sub> stock	1 M in Milli-Q
NTP stock (75 x)	90 mM ATP 60 mM CTP, GTP, UTP pH 7.0 (adjusted with NaOH) in Milli-Q
PEP stock	1 M PEP-KOH, pH 7.0
Pyruvate kinase (PK) stock	10 mg/mL PK
RCWMDE stock (33 x)	33 mM stock of the amino acids R, C, W, M, D and E in Milli-Q
Ribolock RNase inhibitor	40 U/μL inhibitor
S30A buffer	10 mM Tris-acetate, pH 8.2 14 mM MgCl <sub>2</sub> 60 mM KOAc
S30B buffer	10 mM Tris-acetate, pH 8.2 14 mM Mg(OAc) <sub>2</sub> 60 mM KOAc
S30C buffer	10 mM Tris-acetate, pH 8.2 14 mM Mg(OAc) <sub>2</sub> 60 mM KOAc
tRNA <i>E. coli</i> MRE 600	40 mg/mL tRNA in Milli-Q

### 2.5.2 Protein purification and analysis

Ammonium persulfate (APS)	10 % (w/v) in Milli-Q
Anode buffer (10 x)	1 M Tris-HCl
BCAAT buffer	100 mM Tris-HCl, pH 8.0 150 mM NaCl
BCAAT wash buffer	100 mM Tris-HCl, pH 8.0

## Materials

---

BCAAT elution buffer	150 mM NaCl 20 mM imidazole 100 mM Tris-HCl, pH 8.0 150 mM NaCl
Cathode buffer (10 x)	300 mM imidazole 1 M Tris-HCl 1 M tricine 1 % (w/v) SDS <i>PH should be 8.3 without adjustment</i>
Staining solution A	0.05 % (w/v) Coomassie Brilliant Blue R-250 25 % (v/v) isopropyl alcohol 10 % (v/v) acetic acid
Staining solution B	0.002 % (w/v) Coomassie Brilliant Blue R-250 10 % (v/v) acetic acid
Staining solution C	10 % (v/v) acetic acid
CypD wash buffer	50 mM HEPES, pH 8 300 mM NaCl
CypD elution buffer	20 mM imidazole 50 mM HEPES, pH 8 300 mM NaCl
CypD dialysis buffer	300 mM imidazole 50 mM Sodium phosphate, pH 7.0 1 mM DTT
Disc formation buffer	10 mM Tris-HCl, pH 8.0 4 M NaCl
ECL 1	100 mM Tris-HCl, pH 7.5 2.9 mM luminol 0.4 % (w/v) <i>p</i> -coumaric acid
ECL 2	100 mM Tris-HCl, pH 7.5 0.06 % (v/v) H <sub>2</sub> O <sub>2</sub>
GFP buffer	25 mM Tris-HCl, pH 7.5 150 mM NaCl
FFAR <sub>2</sub> NMR buffer	40 mM Bis Tris, pH 6.5 40 mM NaCl
MSP buffer 1	40 mM Tris-HCl, pH 8.0 300 mM NaCl
MSP buffer 2	1 % (v/v) Triton X-100 40 mM Tris-HCl, pH 8.9 300 mM NaCl
MSP buffer 3	50 mM cholic acid 40 mM Tris-HCl, pH 8.0 300 mM NaCl
MSP buffer 3B	40 mM Tris-HCl, pH 8.9 300 mM NaCl
MSP buffer 4	40 mM Tris-HCl, pH 8.0 300 mM NaCl 50 mM imidazole
MSP dialysis buffer	40 mM Tris-HCl, pH 8.0 300 mM NaCl
MSP elution buffer	10 % (v/v) glycerol 40 mM Tris-HCl, pH 8.0 300 mM NaCl 300 mM imidazole
LILBID analysis buffer	25 mM Sodium acetate, pH 7.5 10 mM NaCl
Phosphate-buffered saline (PBS)	2.6 mM KCl 1.8 mM KH <sub>2</sub> PO <sub>4</sub> 137 mM NaCl 10 mM Na <sub>2</sub> HPO <sub>4</sub>
Phosphate-buffered saline + Tween20 (PBS-T)	2.6 mM KCl

	1.8 mM KH <sub>2</sub> PO <sub>4</sub>
	137 mM NaCl
	10 mM Na <sub>2</sub> HPO <sub>4</sub>
Radioassay binding buffer	0.005 % (v/v) Tween20
	50 mM HEPES-NaOH, pH 7.5
	5 mM MgCl <sub>2</sub>
	1 mM CaCl <sub>2</sub>
Radioassay filter wash buffer	0.2 % (w/v) BSA
	50 mM HEPES-NaOH, pH 7.5
	500 mM NaCl
SapA disruption buffer	0.1 % (w/v) BSA
	25 mM HEPES-NaOH, pH 7.5
	150 mM NaCl
	2 mM DTT
SapA wash buffer	2 % (w/v) Triton X-100
	25 mM HEPES-NaOH, pH 7.5
	150 mM NaCl
SapA dialysis buffer	20 mM imidazole
	25 mM HEPES-NaOH, pH 7.5
	150 mM NaCl
SapA elution buffer	25 mM HEPES-NaOH, pH 7.5
	150 mM NaCl
	400 mM imidazole
Strep buffer	100 mM Tris-HCl, pH 8.0
	100 mM NaCl
Strep elution buffer	100 mM Tris-HCl, pH 8.0
	100 mM NaCl
SDS-PAGE loading buffer (2 x)	25 mM d-desthiobiotin
	50 mM Tris-HCl, pH 6.8
	4 M urea
	10 % (w/v) SDS
	7.5 % (w/v) glycerol
	10 % (w/v) β-mercaptoethanol
T7 RNAP buffer A	0.12 % (w/v) Bromphenolblue
	30 mM Tris-HCl, pH 8.0
	50 mM NaCl
	1 mM EDTA
	10 mM β-mercaptoethanol
T7 RNAP buffer B	5 % (v/v) glycerol
	30 mM Tris-HCl, pH 8.0
	1 M NaCl
	1 mM EDTA
	10 mM β-mercaptoethanol
T7 RNAP dialysis buffer	5 % (v/v) glycerol
	10 mM K <sub>2</sub> PO <sub>4</sub> , pH 8.0
	100 mM NaCl
	0.5 mM EDTA
	1 mM DTT
T7 RNAP disruption buffer	5 % (v/v) glycerol
	30 mM Tris-HCl, pH 8.0
	50 mM NaCl
	10 mM EDTA
	10 mM β-mercaptoethanol
	5 % (v/v) glycerol
Towbin buffer	25 mM Tris-HCl
	192 mM glycine
	15 % (v/v) MeOH
Tricine gel buffer (3 x)	<i>pH should be 8.3 without adjustment</i>
	3 M Tris-HCl, pH 8.45
	0.3 % (w/v) SDS

### 2.6 Software

---

Adobe Illustrator CS6	Adobe (San José, USA)
Adobe Photoshp CS6	Adobe (San José, USA)
Biorender	Biorender (Toronto, Canada)
ImageJ	National Institute of Health (Maryland, USA)
Mendeley	Mendeley (London, UK)
Microsoft Office	Microsoft (Washington, USA)
Prism 8	GraphPad (Washington, USA)
PyMol	DeLano Scientific LLC, Schrödinger (New York, USA)
Serial Cloner	University of Cambridge (Cambridge, UK)
SPARKY	NMRFAM University of Wisconsin-Madison (Madison, USA)
TopSpin	Bruker (Billerica, USA)

## 3 Methods

### 3.1 Molecular biology

#### 3.1.1 Agarose gel electrophoresis

Agarose gel electrophoresis was carried by melting 1 % (w/v) agarose in TAE buffer in a microwave at 800 W. When agarose was dissolved, the solution was cooled down to 50-60 °C and 1.5 µL SYBR green master mix were added per 20 mL of gel solution and mixed before the suspension was poured into trays of the casting system. A comb was placed into the gel creating the desired amount of gel pockets with the desired volume. The gel was left to polymerize for 30 min at RT before being placed into the electrophoresis system and equilibrated with TAE buffer. DNA samples were mixed in a 1 to 6 ratio with 6 x purple loading dye and loaded into the gel pockets. As marker 6 µL DNA ladder 1 kb was used. Electrophoresis was carried out at 120 V for 20-30 min at RT. DNA bands were visualized on a UV table.

#### 3.1.2 DNA transformation by heat shock

Chemical competent cells stored at -80 °C were thawed on ice for 20 min before 1 µL plasmid DNA was added to the cells, gently mixed, and incubated for another 20 min on ice. In case of previous site directed mutagenesis PCR and subsequent DpnI digest, 10 µL template DNA was added. Then the cells were incubated at 42 °C for 45 s before being resuspended in SOC medium and incubated at 37 °C for 40 min at 200 rpm. Finally, 100 µL of the cell suspension was spread plated on 1.5 % LB agar plates containing 100 µg/mL ampicillin or 30 µg/mL kanamycin and incubated at 37 °C for 16 h.

#### 3.1.3 DNA transformation by electroporation

Competent *E. coli* A19 cells were prepared growing cells either in culture tubes or in centrifuge tubes with punctured lids to an optical density (OD<sub>600nm</sub>) of 0.6-0.9 before the cells were placed on ice for 10 min and subsequently harvested by centrifugation at 10,000 x g, 1 min, 4 °C. The supernatant medium was discarded, and the cell pellets were washed with ice cold 10 % (w/v) glycerol before being centrifuged again at 10,000 x g, 1 min, 4 °C. The supernatant was again carefully discarded, and the washing procedure was repeated another two times in the same manner. The washed pellet was resuspended in 40 µL ice cold 10 % (w/v) glycerol and template DNA was added. Depending on plasmid or linear DNA material, 10 ng or > 1 µg DNA material were added, respectively. Subsequently, the cell suspension was carefully mixed and placed into precooled electroporation cuvettes with a 1 mm slit. Electroporation was carried out at 1450 V with a ~ 5 ms pulse duration, and cells were immediately resuspended in room temperature LB medium. In some cases, LB medium was supplemented with additives such as L-glutamine if genes of L-glutamine metabolism were targeted. Incubation in LB medium was carried out for 1 h before the cells were spread plated on agar plates with appropriate antibiotics.

#### 3.1.4 DNA preparation

Plasmid DNA was propagated in *E. coli* DH5α. For small scale DNA preparation, 2-6 mL LB medium with either 100 µg/mL ampicillin or 30 µg/mL kanamycin were inoculated with single colonies of the *E. coli* strain freshly transformed with the target

plasmid and incubated in culture tubes at 37 °C at 180 rpm for at least 6 h before being harvested by centrifugation at 4500 x g, 1 min, 4 °C. DNA isolation was carried out according to the kits manual. For medium scale DNA preparation, 200 mL of LB medium with stated antibiotic concentrations was inoculated with a single colony of *E. coli* cells freshly transformed with the target plasmid. Cell culture was carried out in baffled shaking flasks at 37 °C, 180 rpm for 16 h, before the cells were harvested by centrifugation at 4500 x g, 10 min, 4 °C. Subsequent DNA purification steps were carried out according to the manufacturer's manual.

### 3.1.5 Restriction and ligation cloning

PCR fragments of the gene of interest were generated using PCR in a 50 µL reaction volume containing 100 ng plasmid DNA, 200 µM dNTPs, 1 x Phusion GC buffer, 2 U Phusion HF polymerase and 500 nM of forward and reverse primers, respectively. Initial denaturation was carried out at 98 °C for 30 s followed by 35 cycles of i. denaturation at 98 °C for 30 s, ii. annealing at 54 °C for 2 min and iii. elongation at 72 °C (1kb/min). Final elongation was carried out at 72 °C for 10 min before storage at 4 °C. Primers were designed to create overhangs with the desired restriction sites, e.g., NdeI and KpnI, followed by the sequence GGTGGT to ensure efficient restriction. Agarose gel electrophoresis (3.1.1) was carried out to separate the amplified PCR fragment from template DNA. The fragment was excised from the gel and DNA was extracted using a gel extraction kit according to the manufacturer's manual. For restriction digestion of 3 µg amplified fragment and the target plasmid, respectively, 50 U of each restriction enzyme and 1 x CutSmart buffer were supplied and incubated at 37 °C for 3 h. Both fragments were then separately purified using a PCR purification kit and 200 ng of digested vector were incubated with a 3-fold molar excess of the in presence of 10 U of T4 DNA ligase in 1 x ligation buffer at 16 °C for 16 h. After incubation, chemically competent *E. coli* DH5α were transformed with 10 µL of the ligation mixture (3.1.1). Single clones were picked and 5 mL LB medium with appropriate antibiotic were inoculated with appropriate antibiotic for DNA preparation (3.1.3). The isolated plasmid DNA was sequenced with T7<sub>fw</sub> or T7<sub>term</sub> primers if not stated otherwise.

### 3.1.6 Genome engineering in *E. coli*

To knock out genes or insert fragments at specific gene loci in the *E. coli* genome, the λ-phage recombination system previously established by Kiril Datsenko and Barry Wanner was used<sup>168</sup>. Knock-out cassettes were amplified by PCR from FRT-PGK-gb2-neo-FRT carrying plasmid using primers with defined annealing regions 5'-N<sub>50</sub>-AATTAACCCTCACTAAAGGGCGG-3' for the forward primer and 5'-N<sub>50</sub>-TAATACGACTCACTATAGGGCTCG-3' for the reverse primer, respectively, where "N<sub>50</sub>" represents a ~ 50 bp sequence complementary to the recombination region (2.2.2). A 50 µL PCR reaction contained 1 x Phusion GC buffer, 0.2 mM dNTPs, 200 nM forward and reverse primer, respectively, 50 ng FRT-PGK-gb2-neo FRT template, and 5 U/µL Phusion polymerase. PCR was carried out with an initial denaturation step at 98 °C for 30 s followed by 35 cycles of i. denaturation at 98 °C for 30 s, ii. annealing at 60 °C for 2 min and iii. elongation at 72 °C (1 kb/min). Final elongation was carried out at 72 °C for 10 min before storage at 4 °C. Amplification was evaluated on Agarose gels by appearance of a band at ca. 1.7 kb. Separation from the plasmid carrying the resistance cassette was not necessary as the plasmid carried a R6K origin that is not propagated by common *E. coli* strains as they lack the pir replication initiator.



First, *E. coli* A19 K12 were inoculated in 5 mL LB medium and incubated at 37 °C until an OD<sub>600 nm</sub> of 0.6-1.0 was reached. Then, the cell culture was harvested and washed and transformed with 10 ng pRed/ET plasmid carrying genes for expression of the  $\lambda$ -phage Red $\alpha$  exonuclease and Red $\beta$  annealing protein by electroporation (3.1.3). After electroporation, cells were suspended in room temperature LB medium and incubated at 30 °C for 90 min at 180 rpm before plating on LB agar containing 4  $\mu$ g/mL tetracycline and incubating at 30 °C for 1-2 d dependent on the colony size. When colonies reached a size of > 1 mm  $\varnothing$ , a single colony was picked and a culture tube containing LB medium with 4  $\mu$ g/mL tetracycline was inoculated and incubated at 30 °C at 180 rpm over night. Next day, 1.4 mL LB medium with 4  $\mu$ g/mL tetracycline were inoculated with 100  $\mu$ L of the pre culture and incubated in a centrifuge tube with a punctured lid at 30 °C at 700 rpm for 3 h in a thermos-mixing device. When cells reached an OD<sub>600 nm</sub> of 0.6-1.0, L-arabinose was added to a final concentration of 0.4 % (w/v) and the culture temperature was adjusted to 37 °C for another 45 min. This induced expression of Red $\alpha$  and Red $\beta$  and simultaneously cause loss of pRed/ET plasmid copies from 5 to 2-3. Non-induced negative controls were included in every knock-out experiment. Cells were harvested and transformed with linear FRT-PGK-gb2-neo-FRT cassettes by electroporation. After electroporation, cells were resuspended in room temperature LB medium without antibiotic and incubated at 37 °C 180 rpm for 3 h. This is the step, where homologous recombination should occur. Subsequently, the cell suspension was plated on LB agar with 50  $\mu$ g/mL kanamycin and incubated at 37 °C until the colonies reached an appropriate size for analysis by Colony PCR (> 2 mm  $\varnothing$ ).

Plates were evaluated for the number of colonies on the agar plates and compared with the negative controls that if any should contain only small pinpoint colonies. Single colonies were picked and resuspended in 50  $\mu$ L Milli-Q and incubated at 98 °C for 5 min to lyse bacteria. A negative control of WT *E. coli* K12 A19 DNA was included in all screening experiments. 1  $\mu$ L of the lysed cell suspension was used as template in 50  $\mu$ L colony PCR reaction that contained 1 x Phusion GC buffer, 0.2 mM dNTPs, 200 nM forward and reverse primer, respectively, and 5 U/ $\mu$ L Phusion polymerase. PCR screening primers were designed in a way that the forward primer specifically anneals to the kanamycin cassette and the reverse primer anneals ~ 300-500 bp downstream of the deleted gene. PCR was carried out with an initial denaturation step at 98 °C for 30 s followed by 35 cycles of i. denaturation at 98 °C for 30 s, ii. annealing at 54 °C for 2 min and iii. elongation at 72 °C (1kb/min). Final elongation was carried out at 72 °C for 10 min before storage at 4 °C. Amplification was evaluated on Agarose gels by appearance of a band at ~ 1 kb. The amplified fragment was then purified by PCR purification kit according to the manufacturers manual and sequenced using the downstream primer to verify gene knockout.

To delete further genes in *E. coli*, the kanamycin cassette had to be removed from *E. coli* genome. Therefore, every cassette contained flippase recognition target (FRT) sites that are recognized by FLP recombinase. To recovery the resistance marker, a single colony of the deletion mutants was picked and inoculated in 5 mL LB medium with 50  $\mu$ g/mL kanamycin and incubated at 37 °C, 180 rpm for 5 h. When OD<sub>600nm</sub> reached 0.6-1.0 cells were harvested and transformed by electroporation with 10 ng of the flippase carrying plasmid (FLP-plasmid) (3.1.3). After electroporation, cells were suspended in room temperature LB medium without antibiotic and incubated at 30 °C, 180 rpm for 3 h before plating on plates containing 30  $\mu$ g/mL kanamycin and 34  $\mu$ g/mL chloramphenicol and incubated at 30 °C until colonies reach an appropriate size of >

1 mm Ø. Then 2 mL LB medium without antibiotics were inoculated with a single colony from the plate and incubated at 30 °C, 180 rpm for 16 h. Then 40 µL of the pre-culture were used to inoculate 1.4 mL LB medium without antibiotics in a 1.5 microfuge tube with a punctured lid and incubated at 30 °C, 700 rpm for 3 h until an OD<sub>600nm</sub> of 0.6-1.0 was reached. Then temperature was adjusted to 37 °C and the culture was incubated overnight. This induced expression of the flippase by the temperature sensitive λPR promoter. The flippase now can excise the antibiotic marker from the *E. coli* genome recognizing the FRT sites up-and downstream of the cassette. Also, this lead to FLP plasmid loss due to temperature sensitive pSC101 origin on the FLP-plasmid. Next day, the culture was diluted 1 to 10<sup>6</sup> with sterile Milli-Q and 100 µL of the diluted suspension was plated on LB agar without antibiotics and incubated at 37 °C overnight. Next day, two LB agar plates, one with 50 µg/mL kanamycine, one without were subdivided into an enumerated grid of ~ 50 fields. To identify which clone lost its resistance, an autoclaved toothpick was used to pick a single colony and to streak the clone on the field with the identical number. After ~ 50 clones (from usually 100-300) were streaked and the two plates were incubated at 37 °C until streaks reached a width of > 1 mm. The clone number that grew on LB agar but not on LB agar with kanamycin, was then streaked out on LB agar so single colonies were obtained after incubation at 37 °C overnight. Usually, at least 4 clones were picked this way. One colony of every streaked-out candidate was then subjected analyzed by colony PCR that was conducted as described above, but this time using primers annealing 500 bp up-and downstream of the target gene respectively. Clones that lost the resistance marker now displayed a DNA band at ~ 1 kb. The fragment was then purified by PCR purification kit and sequenced using the same upstream or downstream primers used for amplification. Verified knock-out mutants were then streaked out on LB agar and incubated at 37 °C overnight and bacterial lawn was then harvested and resuspended in DB salt solution for storage at -80 °C until further use.

### **3.2 Protein expression purification from *E. coli***

#### **3.2.1 Expression and purification of T7 RNAP**

For pre-culture, 50 mL LB medium with 100 µg/mL ampicillin were inoculated with a single culture of *E. coli* T7 freshly transformed with the pAR1219 plasmid. The culture was incubated at 37 °C, 180 rpm and 16 h. Then 1 L Terrific broth in a 2 L baffled shaking flask was inoculated 1:100 with the preculture and incubated at 37 °C, 180 rpm until OD<sub>600nm</sub> reached 0.6-0.9 and 1 mM IPTG was added. The induced culture was incubated for another 4 h at 37 °C, 180 rpm, before cells were harvested at 4500 x g, 4°C for 15 min, washed with T7RNAP buffer A and centrifuged again under the same conditions. Cell pellets were shock-frozen in liquid nitrogen and stored at -80 °C until further use.

Purification was carried out by resuspending the pellet of the harvested culture in T7 RNAP buffer A and adding 1x protease inhibitor cocktail (PIC). Bacterial cells were disrupted by 6 cycles of 1 min sonication with 1 min pause intervals between the cycles. Cell debris was removed at 30,000 x g, 4 °C for 30 min. The supernatant was incubated on ice under gentle agitation while streptomycin was added to a final concentration of 2 % (w/v) to precipitate DNA material in the crude extract. Subsequently the

centrifugation step was repeated to remove insoluble material. The supernatant was then filtered through a 0.45  $\mu\text{m}$  filter and loaded on a Q Sepharose anion exchange column (1 CV = 75 mL) equilibrated with T7 RNAP buffer A using a peristaltic pump. All further chromatography steps were carried out at 16 °C. Then the column was connected to an ÄKTA purifier system and washed with T7 RNAP buffer A at 4 mL/min until the 280 nm absorption reached a baseline. Then a gradient from T7 RNAP buffer A to 50 % T7 RNAP buffer B was applied to reach a final NaCl concentration of 500 mM within 90 min at 4 mL/min flow rate. 1 mL fractions were collected and every 12<sup>th</sup> fraction was analyzed via SDS PAGE and Coomassie staining (3.3.3) to analyze appearance of a prominent band at 90 kDa. These fractions were then pooled and dialyzed in T7 RNAP dialysis buffer at 4 °C for 12-16 h before being concentrated to ~ 12 mg/mL protein concentration monitored via 280 nm UV absorption. Subsequently, the solution was adjusted to 50 % (v/v) glycerol, gently mixed and the protein concentration was measured again before 500  $\mu\text{L}$  aliquots were shock frozen in liquid nitrogen and stored at -80 °C until further use. To determine appropriate T7 RNAP concentration during CF expression, GFP was CF expressed at varying concentrations (3.3.4) and fluorescence was measured to determine GFP yield (3.3.1).

### 3.2.2 Expression and purification of membrane scaffold proteins

Expression and purification of membrane scaffold proteins was carried out as described previously with minor modifications<sup>55</sup>. *E. coli* T7 express cells were freshly transformed with the pET28 plasmid carrying the gene for expression of his-tagged membrane scaffold protein (MSP) variants MSP1E3D1 or MSP1D1 $\Delta$ H5 (3.1.2). 200 mL LB medium with 30  $\mu\text{g}/\text{mL}$  kanamycin were inoculated with a single clone of the transformed cells and incubated at 37 °C, 180 rpm for 12-16 h. Then 10 mL sterile terrific broth (TB) with 30  $\mu\text{g}/\text{mL}$  kanamycin in a 20 L stirred tank reactor was inoculated 1:100 with the preculture and cultivation was carried out at 37 °C, 500 rpm and an air flow rate of >2 bioreactor volumes per minute. 1-2 mL Antifoam Y30 was added to prevent excessive foaming. Cultivation was carried out until an OD of 7-10 was reached, MSP expression was induced with 1 mM IPTG and continued for another 1 h before harvesting at 4,500 x g, 4 °C for 15 min. Cells were washed with MSP buffer 3 and centrifuged again under the same conditions. Then cell pellets were shock frozen in liquid nitrogen and stored at -20 °C until further use.

Purification of MSP was carried out by resuspending the pellets in MSP buffer 1 with 1x PIC and 1 mg/mL DNase. The suspension was sonified in 6 cycles, 1 min each, with 1 min pauses between sonification intervals. Then cell debris was removed by centrifugation at 30,000 x g, 4 °C for 30 min. Subsequently, the crude extract was filtered through 0.45  $\mu\text{m}$  filters and loaded on 4 IMAC HiTrap columns with a total CV of 20 mL at 3 mL per minute. Flow through was collected and loaded on the columns a second time. Then the columns were connected to an ÄKTA purifier system and further purification steps were carried out at 16 °C and 3 mL/min flow rate. Bound MSP was washed with 5 CV MSP buffer 1, 5 CV MSP buffer 2, 2 CV MSP buffer 3B, 5 CV MSP buffer 3 and 5 CV of MSP buffer 4. Elution was carried out with 5 CV MSP elution buffer and fractions of 2 mL were collected. Fractions that showed absorption at 280 nm were pooled and dialyzed in 12-14 kDa MWCO dialysis tubes in MSP dialysis buffer at 4 °C for 12-16 h. Then the protein was concentrated in 10 kDa MWCO ultrafiltration devices until a concentration of at least 300  $\mu\text{M}$  and protein was measured via UV absorption at 280 nm. Protein concentration was calculated using the extinction

coefficient  $\epsilon_{\text{MSP1E3D1}} = 27.31 \text{ M}^{-1} \text{ cm}^{-1}$  or  $\epsilon_{\text{MSP1D1}\Delta\text{H5}} = 19.94$ . MSP was then flash frozen in liquid nitrogen and stored at  $-80 \text{ }^\circ\text{C}$  until further use.

Cleavage of the 6x polyhistidine tag (6x his-tag) from MSP was carried out using tobacco etch virus (TEV) protease that recognizes the amino acid sequence Glu-Asn-Leu-Tyr-Phe-Gln-(Gly/Ser) and cleaves between the Gln and Gly/Ser residues. For cleavage, TEV protease was added in a ratio of 1:25 TEV to MSP. The solution was adjusted to 1 mM DTT and dialyzed against 1 L TEV cleavage buffer in 10 kDa MWCO dialysis devices with gentle stirring at  $4 \text{ }^\circ\text{C}$  for 12-16 h. Afterwards, precipitate was removed by centrifugation at  $20,000 \times g$ ,  $4 \text{ }^\circ\text{C}$  for 30 min and the supernatant was loaded on an IMAC column (CV = 5 mL) equilibrated with TEV equilibration buffer. The flow through was collected. Subsequently, the column was washed with another 10 CV of equilibration buffer. The fractions were collected and concentrated via 10 kDa MWCO ultrafiltration until the final MSP concentration was at least  $300 \text{ }\mu\text{M}$ . The concentration was measured using via UV absorption, this time using the extinction coefficient  $\epsilon = 26.93 \text{ M}^{-1} \text{ cm}^{-1}$  and  $18.45 \text{ M}^{-1} \text{ cm}^{-1}$  for TEV cleaved MSP1D1E3 and MSP1D1 $\Delta$ H5 variants, respectively. TEV cleaved MSP was then flash frozen in liquid nitrogen and stored at  $-80 \text{ }^\circ\text{C}$  until further use. Uncleaved MSP was subsequently washed off the IMAC column with equilibration buffer with 300 mM imidazole.

### 3.2.3 Expression and purification of SapA

Expression of SapA was carried out in *E. coli* Rosetta-gami-2 DE3. 200 mL LB medium with  $30 \text{ }\mu\text{g/mL}$  kanamycin and  $15 \text{ }\mu\text{g/mL}$  chloramphenicol. A 500 mL baffled shaking flask was inoculated with a single colony of cells freshly transformed with pNIC28-BSA4-SapA construct and cultivated at  $37 \text{ }^\circ\text{C}$ , 180 rpm for 12-16 h. Then 10 L of TB medium with  $30 \text{ }\mu\text{g/mL}$  kanamycin and  $15 \text{ }\mu\text{g/mL}$  chloramphenicol in a 20 L stirred tank reactor were inoculated 1:100 with the preculture and cultivation was carried out at  $37 \text{ }^\circ\text{C}$ , 500 rpm and an air flow rate of  $>2$  bioreactor volumes per minute. 1-2 mL Antifoam Y30 was added to prevent excessive foaming. Cultivation was carried out until an OD of 7-10 was reached. SapA expression was induced with 1 mM IPTG and continued for another 1 h before harvesting at  $4,500 \times g$ ,  $4 \text{ }^\circ\text{C}$  for 15 min. Cells were washed with SapA dialysis buffer and centrifuged again under the same conditions. Then cell pellets were shock frozen in liquid nitrogen and stored at  $-20 \text{ }^\circ\text{C}$  until further use.

For disruption, the pellet was resuspended in SapA disruption buffer with 1 x PIC and 1 mg/mL DNase. The suspension was subjected to 6 cycles of sonification, 1 min each, with 1 min pauses between sonification intervals. Then cell debris was removed by centrifugation at  $30,000 \times g$ ,  $4 \text{ }^\circ\text{C}$  for 30 min. The supernatant was incubated at  $85 \text{ }^\circ\text{C}$  for 10 min and the centrifugation step was repeated to remove the precipitate. The supernatant was filtered through  $0.45 \text{ }\mu\text{m}$  filters and loaded on 20 mL IMAC resin equilibrated with SapA washing buffer in gravity flow columns. Flow through was collected and loaded on the columns a second time. The IMAC beads were washed with 15 CV SapA washing buffer and eluted with 4 CV SapA elution buffer. The eluate was dialyzed against 5 L of SapA dialysis buffer, concentrated to a sample volume of  $< 2 \text{ mL}$  and further purified via size exclusion chromatography (SEC) and concentrated via ultrafiltration with 10 kDa MWCO to a protein concentration of  $\sim 20 \text{ mg/mL}$  before. Protein concentration was measured via UV absorption at 280 nm using the extinction coefficient  $\epsilon = 10.345 \text{ M}^{-1} \text{ cm}^{-1}$ .

### 3.2.4 Expression and purification of BCAAT

Branched chain amino acid aminotransferase (BCAAT) with an N-terminal 6xhis-tag was expressed in *E. coli* BL21 Star express cells as described previously<sup>166</sup>. Briefly, 100 mL LB medium with 100 µg/mL ampicillin in a 0.5 L baffled shaking flask was inoculated with freshly transformed cells and incubated at 37 °C, 180 rpm for 16 h. Then 1 L LB medium with 100 µg/mL ampicillin in a 2 L baffled shaking flask was inoculated with 20 mL of the preculture and incubated at 37 °C, 180 rpm until OD 0.6-0.9. To induce expression, 1 mM IPTG was added. The cells were cultivated for another 4 h at 37 °C and 180 rpm before they were harvested, washed with Strep buffer (2.5.2), resuspended in the same buffer with protease inhibitor cocktail (PIC) and disrupted by sonification. The crude lysate was centrifuged at 30,000 x g, 4 °C for 30 min, passed through a 0.45 µm filter and applied on a 5 mL HiTrap IMAC column previously equilibrated with Strep buffer with 20 mM imidazole. The bound protein was then washed with another 15 CV of Strep buffer with 20 mM imidazole and eluted with 4 CV Strep buffer with 300 mM imidazole. The eluate was dialyzed against the 100-fold volume of BCAAT buffer, concentrated using 10 kDa MWCO ultrafiltration devices and glycerol was added to a final concentration of 50 % (v/v). The concentration of the protein stock was 10 mg/mL. The aliquoted protein solution was flash frozen in liquid nitrogen until further use.

### 3.3 CF expression of proteins

#### 3.3.1 S30 extract preparation

S30 extract was prepared from *E. coli* A19 K12 and its derivatives. First, the cells were streaked out from a cryo stock on a LB agar plate and incubated for 16 h. Then 200 mL LB medium in a 0.5 L baffled flask were inoculated with a single cell and incubated at 37 °C and 180 rpm for 12-16 h. Next day sterile 10 L YPTG medium in a stirred-tank bioreactor (2.4.1) were inoculated with 100 mL of the preculture and incubated at 37 °C, 500 rpm, 2 reactor volumes aeration per minute (vvm). 1-2 mL Y30 antifoam was added to prevent excessive foaming. Culture medium was additionally supplied with L-Gln for all *glnA* mutants. For *glmS* mutants GlcNac was added to all culture media and glucose left out. When OD reached 3.5-4.5, the culture was cooled down to 20 °C within 15-20 min via passive cooling with water circuits in the fermenter. Subsequently, the culture was harvested at 4,500 x g and 4 °C for 15 min and the pellets were resuspended with 300 mL S30A (2.5.1) buffer and centrifuged again at 8,000 x g at 4 °C for 10 min. This washing process was repeated 3 times. Then the pellet was weighed and 1.1 x of the final pellets weight in grams was added in milliliters of buffer S30B (2.5.1). The cells were then filled into a pre-cooled French press and cells were disrupted at 1,000 psig with one passage. Cell debris and cell membrane fragments were then removed at 30,000 x g at 4 °C for 30 min. The pellets were discarded, and the supernatant was centrifuged again under the same conditions. Subsequently, NaCl was added dropwise to the crude lysate to a final concentration of 0.4 M and the solution was incubated at 42 °C for 45 min under gentle agitation. Finally, the lysate with precipitated protein material was placed into 10 MWCO dialysis bags and dialyzed against pre-cooled 5 L S30C (2.5.1) for 2.5 h at 4 °C under agitation and then against fresh pre-cooled 5 L S30C buffer for 12-16 h at 4 °C under agitation. Next day the suspensions were centrifuged again to remove precipitated material at 30,000 x g at 4 °C for 30 min. The supernatant was then thoroughly mixed by inverting and aliquoted in 0.5-1.5 mL volumes in centrifuge tubes and flash frozen in liquid nitrogen and stored at -80 °C until further use.

### 3.3.2 Preassembly of MSP NDs

For ND preassembly lipid stocks of 50 mM lipid in 100-150 mM Na-cholate were prepared. If the lipid did not fully dissolve, further Na-cholate was added from a 1 M stock. Then the molar ratio of lipid to MSP was calculated. The exact MSP concentration was determined for every preassembly setup. In this study, MSP derivatives MSP1E3D1 and MSP1D1ΔH5 were used. The molar ratios for the MSP1E3D1 derivative were 1:80 and 1:85 for DOPG and POPG, respectively. For MSP1D1ΔH5 a 35-fold POPG excess was used for ND formation. Preassembly was initiated by mixing MSP and lipid in the stated ratios. The solution was filled up to 11.88 mL with disc formation (DF) buffer (2.5.1) and n-dodecylphosphocholine (DPC) was added to a final concentration of 0.1 % in a 12 mL preassembly volume. The mixture was incubated at RT under gentle shaking for 1 h and then transferred into 10 MWCO dialyzing cassettes and dialyzed 3 times against 5 L DF buffer at RT under gentle agitation. The buffer exchange was carried out after 12-16 h, respectively. Finally, the dialyzed solution was transferred into centrifuge tubes and centrifuged at 22,000 x g and 4 °C for 20 min to remove insoluble material. The concentration of the supernatant was determined via UV absorption, considering specific molecular masses and extinction coefficients (table 2). The calculated concentration was then divided by 2 since one nanodisc is always composed of two MSP entities. The ND solution was then aliquoted in 50-200 μL batches, flash-frozen in liquid nitrogen and stored at -80 °C until further use.

**Table 2:** Molar masses and extinction coefficients of different MSP derivatives.

MSP	Molar mass [kD]	Extinction coefficient [ $M^{-1} \text{ cm}^{-1}$ ]
MSP1E3D1	31.96	27.14
MSP1E3D1(TEV)	29.98	26.93
MSP1D1ΔH5	21.47	19.94
MSP1D1ΔH5 (TEV)	19.49	18.45

### 3.3.3 Preassembly of SapNPs

Preassembly of SapNPs was carried out by first determining the exact concentration of the his-tagged 11.96 kDa SapA via UV absorption and extinction coefficient of  $10.35 M^{-1} \text{ cm}^{-1}$ . Lipids were prepared by dissolving 60 mg of the target lipid in chloroform in a 2 mL centrifuge tube. The solution was then transferred into a round flask and chloroform was evaporated in a rotary evaporator. The lipid film in the round flask was then resuspended stepwise in 2 mL Strep buffer (2.5.2) and transferred into a 2 mL microfuge tube to yield a 30 mg/mL suspension. Then SapA and lipid were mixed in the molar ratios between 1:15 and 1:30 in a 15 mL falcon tube and pH was adjusted to 5.0 by slow addition of 5 % (v/v) acetic acid, solution mixing and instant pH monitoring. Upon acidification, the solution became clear and was then rapidly diluted with 10 mL of Strep buffer (2.5.2) to neutralize pH. Subsequently, the solution was concentrated using 50 kDa MWCO ultrafiltration devices and centrifuged at 22,000 x g and 4 °C for 10 min to remove insoluble material. Protein concentration was again determined as stated above. Since SapA is able to form NPs of a different number of SapA monomers, the concentration of SapNPs itself was not calculated as was done for MSP (3.3.2). All SapA concentrations used for CF expression were according to SapA concentration itself regardless of the nature of the NPs.

### 3.3.4 Preconversion of methyl labeled precursors

Preconversion of sodium salts of KIV and MOV was carried out for obtaining the respective branched chain amino acids for protein labeling. Thus, unlabeled and labeled KIV (3-methyl  $^{13}\text{C}$ , 3, 4, 4, 4-D4) and MOV (4-methyl  $^{13}\text{C}$ , 3, 3, 4, 5, 5, 5-D6) were used. First, Tris base and NaCl were weighed to yield a final concentration of 100 mM in the conversion mix. L-Glu and KIV/MOV were weighed to yield 350 mM and 70 mM in the final volume of 10 mL conversion buffer, respectively. The components were dissolved together in  $\text{D}_2\text{O}$  and pH was adjusted to 7.5 with 2 M NaOH in  $\text{D}_2\text{O}$ . Then 9 mL of the mixes were filled up with 1 mL of 10 mg/mL BCAT solution to yield a final concentration of 1 mg/mL BCAT. The mixes were incubated at 30 °C, 180 rpm for 16 h. Then, the conversion mixture was passed through 10 kDa MWCO ultrafiltration devices to remove BCAT that remained in the retentate. KIV or MOV permeates of the conversion mixture were used instead of L-Val or L-Leu for CF expression, respectively.

### 3.3.5 CF expression setup

CF expression was carried out using a two-compartment system where the reaction mixture (RM), where protein synthesis occurs, is separated by a 10 or 12-14 kDa MWCO dialysis membrane from the feeding mixture (FM). Both compartments contained the same concentration of low molecular weight components which are 1 mM of all amino acids except L-Arg, L-Cys, L-Trp, L-Met, L-Asp and L-Glu which are added at 2 mM final concentration, 20 mM acetyl phosphate, 20 mM phosphoenolpyruvate, 3.2 mM CTP, GTP and UTP, respectively, 4.5 mM ATP, 0.1 mg/mL folinic acid, 100 mM HEPES-KOH, pH 8.0, 0.8 mM EDTA, 18-20 mM  $\text{MgOAc}_2$  and 170 mM KOAc. Initially, the low molecular weight components were mixed to produce a 3x master mix of the FM. A part of the 3x master mix was then transferred to the RM. Subsequently the FM was filled up to 1x concentration with Milli-Q. Then high molecular weight components and Milli-Q were supplied to the RM to yield a final concentration of 0.35 % (v/v) *E. coli* lysate, 0.3 U/ $\mu\text{L}$  RNasin, 0.04 mg/mL pyruvate kinase, 0.5 mg/mL tRNA from *E. coli*, 1-15 ng/ $\mu\text{L}$  plasmid DNA template and if stated preassembled NPs, scaffold proteins and/or liposomes or DNA template of scaffold proteins.

### 3.3.6 Strep purification of CF expressed MPs

Purification of membrane proteins with strep-tag II was carried out using gravity flow columns after harvesting the proteins from the reaction mix after CF expression (3.3.5). The proteins carried the 8 AA tag (YSH PQFEK\*) at the C-terminus with 4 aa (GGGS) linker between construct and strep-tag II. If possible, concentration was measured previously either via GFP fluorescence (3.4.1) or via absorption of proteorhodopsin (3.4.2). Otherwise, GFP chimeras of the target were expressed to estimate concentration of the non-chimeric proteins. For binding ~1 mL of Streptavidin beads equilibrated with strep buffer was used for binding of 0.5 mg of the MP. Harvested reaction mix was diluted 1 to 3 with strep buffer (2.5.2) and then solution was loaded on the beads. The flow through of the solution was collected and loaded on the beads a second time. Subsequently, the bound protein was washed with 20 CV strep-buffer and eluted with 4 CV strep elution buffer (2.5.2). The eluate was then concentrated using 50 kDa ultrafiltration devices via centrifugation at 4,000 x g and 4 °C. The duration of the centrifugation was adjusted according to the sample volume and concentration.

### 3.3.7 Expression and purification of cyclophilin D

Truncated peptidyl-prolyl cis-trans-isomerase CypD ( $\Delta 43-207$ ) with N-terminal 6x his-tag was CF expressed from pIVEX2.3d plasmid and IMAC purified as described previously<sup>169</sup>. Briefly, 5 mL IMAC Sepharose 6 Fast Flow resin in gravity flow columns was equilibrated with CypD buffer (2.5.2). The crude RM was diluted 1:3 with CypD wash buffer before being applied on the column. The column was washed with 15 CV CypD wash buffer and eluted with 4 CV of CypD elution buffer, dialyzed against 2 x 5 L CypD dialysis buffer and concentrated using 10 kDa MWCO ultrafiltration devices. The final NMR samples were supplied with 100  $\mu\text{g}/\text{mL}$  streptomycin, 1 mM imidazole 1 x complete protease inhibitor, 0.15 mM Sodium trimethylsilylpropanesulfonate and 5%  $\text{D}_2\text{O}$  and had a final protein concentration of 100-150  $\mu\text{M}$ . The final sample volume was 600  $\mu\text{L}$ . Labeling was carried out using single  $^{15}\text{N}$  labeled amino acids during CF expression L-Glu, L-Gln, L-Asp and L-Asn, respectively, while the other 19 amino acids were non-labeled. Inhibitors AOA and DON were applied in some reactions at concentrations 20 mM and 5 mM, respectively.

### 3.3.8 Expression and purification of PR

Proteorhodopsin (PR) with a C-terminal 6xhis-tag was expressed and purified using a modified protocol previously published<sup>28</sup>. Briefly, PR was CF expressed from pIVEX2.3d with 0.4 % (w/v) GDN and 0.1 % (w/v) 07:0 PC in the RM and FM, respectively. Two different methyl labeling schemes were tested: i. expression was carried out presence of 1 mM L-Val (3-methyl  $^{13}\text{C}$ , 3, 4, 4, 4-D4) converted from KIV (3-methyl  $^{13}\text{C}$ , 3, 4, 4, 4-D4) and the other unlabeled 19 amino acids. ii. PR was expressed with 0.5 mM L-Leu (4-methyl  $^{13}\text{C}$ , 3, 3, 4, 5, 5, 5-D6) preconverted from MOV (4-methyl  $^{13}\text{C}$ , 3, 3, 4, 5, 5, 5-D6) with the other 19 aa unlabeled. After harvesting at 20,000 x  $g$  and 4  $^\circ\text{C}$  for 10 min the probes were applied on an IMAC column equilibrated with PR buffer (Na-acetate, pH 5.0 buffer with 0.1 % (w/v) DH7PC). The bound protein was then washed with 15 CV PR buffer with 20 mM imidazole to remove impurities and GDN. The sample was eluted with PR buffer with 300 mM imidazole. The eluate was then concentrated to a volume of 1 mL using 30 MWCO centriprep ultrafiltration devices and the buffer was exchanged on a PD Mditrap G25 column to imidazole free PR buffer and concentrated again using 30 MWCO centriprep ultrafiltration devices. The final NMR samples were supplied with 100  $\mu\text{g}/\text{mL}$  streptomycin, 1 x complete protease inhibitor, 0.15 mM Sodium trimethylsilylpropanesulfonate and 5%  $\text{D}_2\text{O}$ . The final sample volume was 600  $\mu\text{L}$ .

### 3.3.9 FFAR<sub>2</sub> expression and purification for NMR analysis

FFAR<sub>2</sub> was expressed in 2.4 mL RM and 40 mL FM in presence of 150  $\mu\text{M}$  MSP1D1 $\Delta$ H5 POPG NDs. A deuterated AA mix extract from algae was used at a concentration of 3 mg per mL of the CF reaction instead of unlabeled AAs. Additionally, 2 mM L-Val (4- $^{13}\text{C}$ , 2, 3, 5, 5, 5-D5) preconverted from the labeled precursor KIV was included in the CF reaction. Cotranslationally solubilized FFAR<sub>2</sub> in NDs was strep-purified as described (3.3.6) and subjected to SEC purification using a Superose 6 (10/300) column previously equilibrated with FFAR<sub>2</sub> NMR buffer (2.5.2), and pooled main peak fractions concentrated using 50 MWCO ultrafiltration devices. The final NMR samples were supplied with 100  $\mu\text{g}/\text{mL}$  streptomycin, 1 x complete protease inhibitor, 0.15 mM Sodium trimethylsilylpropanesulfonate and 5%  $\text{D}_2\text{O}$ .



### 3.4 Protein analysis

#### 3.4.1 GFP fluorescence assay

GFP fluorescence assay was carried out by mixing 297  $\mu\text{L}$  GFP buffer (2.5.2) with 3  $\mu\text{L}$  of the probe which usually was either GFP or GFP chimera proteins measured either directly in the RM after harvesting or after purification. Protein concentration was calculated from a standard curve that was made by referencing GFP fluorescence to protein concentration of the same sample determined with BCA assay. The calibration curve resulted in the equation  $y = 8125x$  where  $y$  is the fluorescence intensity in arbitrary units and  $x$  is the concentration of GFP in  $\text{mg/mL}$ .

#### 3.4.2 Quantification of PR

PR concentration was measured using vis-absorption at 530 nm. Therefore, PR solutions were measured in 1 cm cuvettes. If absorption was higher than 1, samples were diluted accordingly. Concentration was calculated using the extinction coefficient of  $\epsilon = 44,000 \text{ M}^{-1} \text{ cm}^{-1}$ . With  $c = A/\epsilon \cdot 1000,000$  PR concentration was calculated in  $\mu\text{M}$ .

#### 3.4.3 SDS-PAGE

SDS-polyacrylamide gel electrophoresis (PAGE) was carried out using self-prepared discontinuous separation gels with 11 % (v/v) and stacking gels with 4 % (v/v) acrylamide, respectively. MP samples were mixed 1:1 with 2x SDS loading buffer (2.5.2) and incubated at RT for 30 min while soluble proteins were incubated at 95 °C for 5 min before samples were loaded on the gel. PAGE was carried out using cathode buffer inside the electrode and companion gel assembly and anode buffer inside the mini tank (2.5.2). Protein stacking was carried out at 120 V for 5 min. For protein separation voltage was increased to 150 V.

#### 3.4.4 Coomassie staining

Rapid Coomassie staining was carried as previously described with modifications<sup>170</sup>. After PAGE (3.4.3) gels were immersed in ca. 20 mL staining solution A (2.5.2) and boiled for a few seconds in a microwave at 900 W. The gel was then left to cool down at RT under gentle agitation. The staining solution was then discarded, and the same procedure was repeated with staining solution B and two times with staining solution C (2.5.2). For final destaining, a piece of tissue was added to staining solution C for faster destaining of the gel.

#### 3.4.5 Western blot

Upon completion of SDS-PAGE, proteins were analyzed via western blot. Therefore, proteins were transferred from SDS gels onto 0.45  $\mu\text{m}$  PVDF membranes previously activated with methanol and equilibrated with Towbin buffer (2.5.2). The gel and the membrane were additionally sandwiched by three sheets of Whatman papers equilibrated with Towbin buffer to ensure seamless contact between the membrane and the gel. Blotting was carried out at 340 mA with variable voltage for 35 min under gentle agitation and cooling with an icepack. Subsequently, the membrane was blocked with 4 % (w/v) skimmed milk in PBS-T (2.5.2) and incubated at RT under gentle agitation for 1 h. Then, the blocking solution was discarded. For detection of his-tagged proteins, 1:5000 diluted  $\alpha$ -his antibody from mouse (2.3.4) in 10 mL PBS-T with 0.05 % (w/v) skimmed milk was added and incubated for 1.5 h at RT or at 4 °C over night under gentle agitation. Unbound  $\alpha$ -his antibody was removed by washing the

membrane 3 times with 20 mL PBS-T buffer for 10 min, respectively. Then  $\alpha$ -mouse antibody HRP conjugate in 10 mL PBS-T with 0.05 % (w/v) skimmed milk was added and the membrane was incubated at RT under gentle agitation for 1 h. The second antibody was removed by washing the membrane as was done with the first antibody. After washing, the membrane was rinsed with 0.75 mL of ECL1 and ECL2 solution, respectively and chemiluminescence was detected.

For strep-tag II detection, 1:7500 diluted HRP conjugated  $\alpha$ -strep antibody (2.3.4) in 10 mL PBS-T with 0.05 % (w/v) skimmed milk was added directly after blocking of the membrane and incubated at RT under gentle agitation for 1 h before washing and detection.

### 3.4.6 Radioligand filter binding assays

Radioligand filter binding assays were carried out to determine the percentual fraction of active turkey  $\beta_1$ AR-GFP chimeras, where GFP allowed overall quantification of expressed protein. First, expressed protein was quantified in the harvested reaction mix using GFP fluorescence assay (3.4.1). Then the receptor solution volume was calculated to incubate 10 nM receptor with 50 nM [ $^3$ H]-dihydroalprenolol in 30  $\mu$ L radioassay binding buffer at RT under gentle agitation for 1 h. Unspecific binding was determined by incubating the filters with 40  $\mu$ M unlabeled alprenolol for 1 h before addition of the radioactive ligand.

Then GF/B glass fiber filters were equilibrated with 0.3 (v/v) polyethyleneimine for 30 min and subsequently rinsed with 150  $\mu$ L radioassay filter wash buffer (2.5.2) before the sample was applied and subsequently washed 7 x with 150  $\mu$ L radioassay wash buffer (2.5.2). All washing steps were carried out using a vacuum manifold for 96 well filter plates. Subsequently, filters were solubilized in 2 mL Rotiszint® and scintillation was measured. Radioligand assays were carried out and measured by Zoe Koeck, a member of the Prof. Dötsch/Dr. Bernhard group at the Goethe University of Frankfurt.

### 3.4.7 Size exclusion chromatography

Size exclusion chromatography (SEC) was carried with purified protein samples and nanoparticles. Solid material was removed by centrifugation at 22,000 x g at 4 °C for 10 min prior to application on the column using a gastight syringe. All buffers were filtrated through 0.2  $\mu$ m membranes and degassed prior to use. All columns used for this study are listed with the corresponding working conditions (table 3).

**Table 3.** SEC columns and operational parameters.

Column	Flow rate [mL/min]	V <sub>sample</sub> [mL]	V <sub>loop</sub> [mL]	V <sub>Maxload</sub> [mg]
Superdex 75 16/600	1	1.0-1.8	2	80
Superdex 75 10/300	0.5	0.2-0.45	0.5	8
Superdex 200 3.2/300	0.05	0.02-0.045	0.05	0.8
Superose6 3.2/300	0.05	0.02-0.045	0.05	0.8
Superose6 10/300	0.3	0.2-0.45	0.5	8

### 3.4.8 LILBID native-MS

Laser-induced liquid bead ion desorption mass spectrometry (LILBID-MS) is an ionization technique which releases ions from aqueous microdroplets by irradiation

with an IR laser<sup>171</sup>. The piezo driven droplet generator MD-K-130 with a 50  $\mu\text{m}$  capillary was used throughout the experiments. Droplets were generated at 10 Hz at 100 mbar before entering high-vacuum where irradiation took place. IR laser wavelength was 2.94  $\mu\text{m}$  and the pulse length was 6 ns. Laser power was varied between 12 mJ (soft mode) and 23 mJ (harsh mode) with a laser delay of 1  $\mu\text{s}$ . Irradiation leads to immediate droplet disruption and release of ion material that was accelerated in a home-built time-of-flight (ToF) analyzer equipped with a reflectron to reduce signal broadening. Acquired data was analyzed with the software „Massign“. All spectra were visualized and normed to 1 using Origin 2018. All LILBID-MS experiments were conducted by René Zangl and Janosch Martin both members of Prof. Dr. Nina Morgners group at the Goethe University Frankfurt.

### 3.4.9 Negative stain electron microscopy

For negative stain imaging, G400-C3 grids (Gilder Grids) were carbon-coated by evaporating a carbon thread in a Leica EM ACE600 carbon coater. For each grid, 3  $\mu\text{l}$  of the sample were applied onto the grid and then stained by adding 2 % uranyl formate on to the grid 3 times. After each application, the liquid was blotted away from the side with a filter paper (Whatman 4, GE Health care)<sup>172,173</sup>. TEM images were manually recorded on a Tecnai G<sub>2</sub> Spirit TWIN TEM (FEI), operated at 120 kV at a nominal magnification of 42,000, corresponding to a pixel size of 1.5 Å (nominal defocus - 1.5  $\mu\text{m}$ ) and using a Gatan Rio Camera. Negative stain experiments were carried out by Theresa Gewering, a member of Dr. Arne Moellers group at the Max-Planck Institute of Biophysics.

### 3.4.10 CypD NMR experiments

All [<sup>15</sup>N, <sup>1</sup>H] TROSY spectra of truncated CypD ( $\Delta$ 43-207) were acquired at sample temperatures of 303 K on Bruker NMR spectrometers 600 MHz (Avance II), 700 MHz (Avance III HD), 800 MHz (Avance III), 900 MHz (Neo) and 950 MHz (Avance III) equipped with cryogenic <sup>1</sup>H{<sup>13</sup>C/<sup>15</sup>N} triple-resonance probes. To take advantage of longitudinal <sup>1</sup>H relaxation enhancement between scans, the Band-Selective Excitation Short-Transient (BEST) method was applied with proton pulses with a bandwidth of 4.8 ppm centered at 8.7 ppm<sup>174</sup>. The interscan delay was set to 0.3 s. All CypD samples had concentrations of 100-150  $\mu\text{M}$  in a 5 mm NMR tubes with 600  $\mu\text{L}$  sample volumes.

### 3.4.11 PR NMR experiments

All PR spectra were acquired on a Bruker Avance III 950 MHz spectrometer equipped with a cryogenic probe at a sample temperature of 313 K. An XL-ALSOFAST-HMQC pulse sequence with the <sup>13</sup>C-<sup>1</sup>H back-transfer period shortened to 2 ms combined with delayed decoupling was employed<sup>175</sup>. The delay between scans was set to 0.7 s. Using gradient coherence selection, the sequence was suitable to eliminate otherwise strong *t*<sub>1</sub>-noise from detergent and acetate signals. The L-Val (4-<sup>13</sup>C, 2, 3, 4, 4, 4, -D5) labeled and the L-Leu (5-<sup>13</sup>C, 2, 3, 3, 4, 5, 5, 5-D7) labeled samples had concentrations of 80  $\mu\text{M}$  and 120  $\mu\text{M}$ , respectively, in 5 mm NMR tubes with 600  $\mu\text{L}$  sample volume.

### 3.4.12 FFAR<sub>2</sub> NMR experiments

All free fatty acid receptor 2 (FFAR<sub>2</sub>) spectra were acquired on a Bruker Avance III 950 MHz spectrometer equipped with a cryogenic probe at a sample temperature of 298

and 308 K for [ $^{13}\text{C}$ , $^1\text{H}$ ]-ALSOFAST-HMQC and [ $^{15}\text{N}$ , $^1\text{H}$ ]-IPAP-sfHMQC, respectively. [ $^{13}\text{C}$ , $^1\text{H}$ ]-ALSOFAST-HMQC pulse sequence with the  $^{13}\text{C}$ - $^1\text{H}$  back-transfer period shortened to 2 ms combined with delayed decoupling was employed<sup>175</sup>. The delay between scans was set to 0.7 s. Using gradient coherence selection, the sequence was suitable to eliminate otherwise strong  $t_1$ -noise from detergent and acetate signals.

[ $^2\text{H}/^{15}\text{N}$ ] and L-Val (4- $^{13}\text{C}$ , 2, 3, 4, 4, 4,-D5) labeled FFAR<sub>2</sub> sample contained 40 mM Bis Tris, 40 mM NaCl, pH 6.5, 1x cOmplete protease inhibitor, 5 % D<sub>2</sub>O and 0.15 mM DSS, 100  $\mu\text{g}/\text{mL}$  streptomycin and 100  $\mu\text{M}$  high affinity propionate antagonist TUG1609. FFAR<sub>2</sub> sample was concentrated to 60  $\mu\text{M}$  and measured in a Shigemi NMR tube with 350  $\mu\text{L}$  sample volume.

## 4. Results

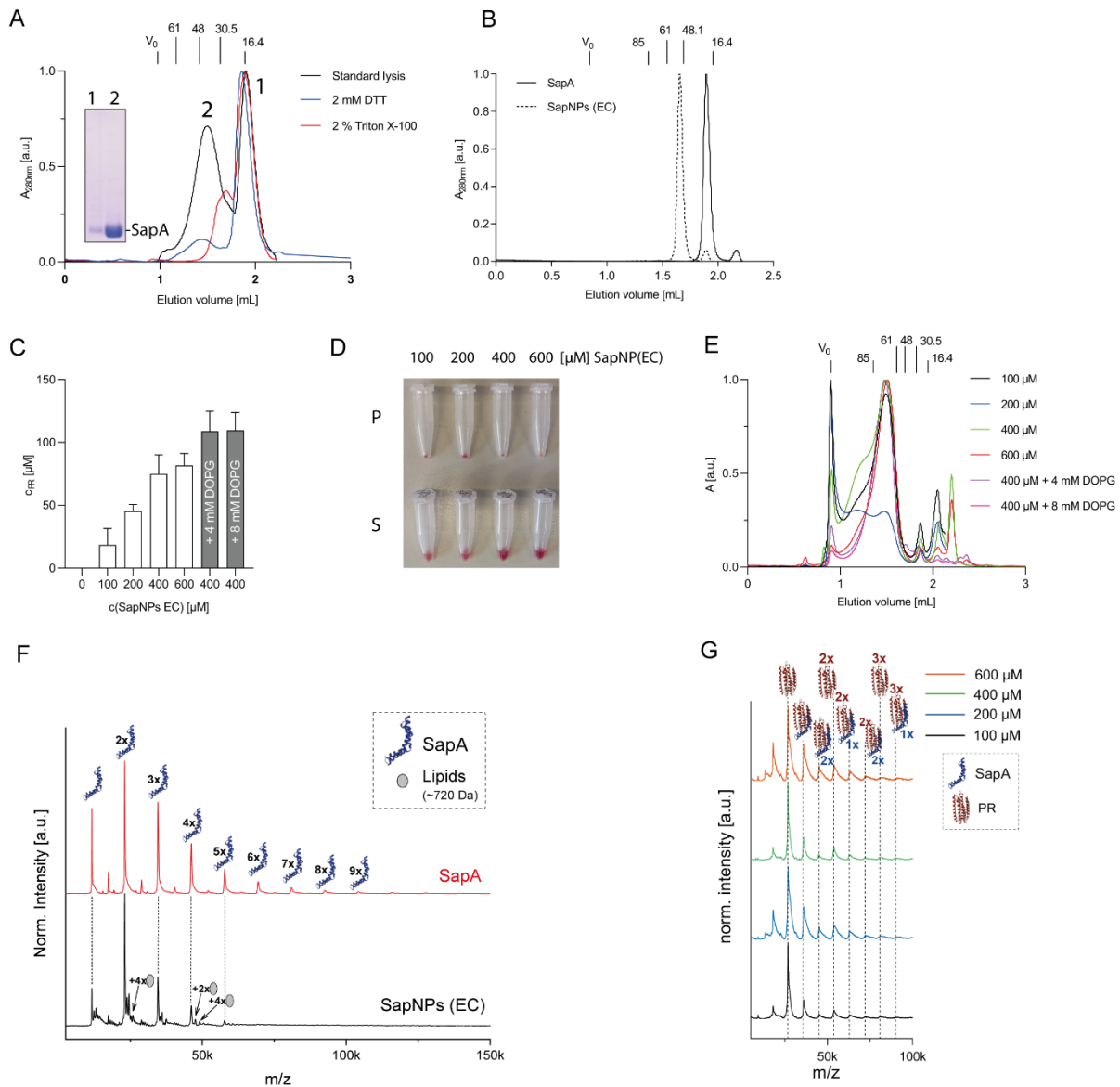
### 4.1 SapNPs for cotranslational stabilization of MPs

#### 4.1.1 Spontaneously formed SapNPs during purification

SapNPs were extensively studied for solubilization of membrane proteins via reconstitution approach and comparisons with MSP nanodiscs were drawn<sup>58,76,79</sup>. However, a systematic approach of SapA application in the CF system has not been studied yet. While NDs are composed of two MSP belts, SapA produces NPs consisting of different numbers of SapA units making it particularly interesting for applications such as NMR where smaller particles are desired.

To produce large amounts suitable for CF expression we chose to produce the protein in a bioreactor at high cell densities ( $OD > 7$  at induction). This approach yielded two fractions of SapA that were initially observed during SEC. (Figure 7A). The fractions were separated by SEC and remained stable after separation which was confirmed by SEC analysis of the separated fractions (Figure 7B). We concluded that the fraction with the retention volume of 1.9 mL corresponds to monomeric SapA while the second fraction was assumed to be SapA in a complex. Indeed, LILBID-MS showed that the particles already contained lipid. That means that SapA captured lipids during purification process, most likely during cell disruption (Figure 7F). Interestingly, monomeric SapA showed oligomerization behavior in LILBID-MS up to nonamers which also might be due to intermolecular disulfide bond formation or hydrophobic interactions. The NP fraction that contained *E. coli* lipids (SapNPs (EC)) was then applied in the CF reaction at different concentrations to solubilize PR and indeed solubilization increased with increasing SapNPs (EC) concentration and plateaued at  $\sim 400 \mu\text{M}$  SapNP (EC) yielding  $\sim 90 \mu\text{M}$  soluble PR (Figure 7C). All expressions showed minor aggregation of PR visible as red precipitate which decreased upon increasing SapNP (EC) concentration (Figure 7D). This is most likely due to particle destabilization upon PR insertion and overcrowding of particles with PR, especially if lower SapNP (EC) concentrations are provided. As SapA is reported to be a system with a certain degree of flexibility, we tested if further lipid addition will improve expression. Indeed, addition of 4 and 8 mM DOPG both boosted expression yield to  $\sim 105 \mu\text{M}$  (Figure 7C) and produced SEC profiles that were slightly more homogenous than with  $400 \mu\text{M}$  SapNPs (EC) without lipids (Figure 7E). This indicates that incorporation of lipids into the particles even after their formation can occur. This potentially provided an increased lipid area for PR insertion and thus, lowered the probability of aggregation (Figure 7E). While PR yield did not significantly increase at  $600 \mu\text{M}$  SapNPs (EC), it improved the SEC profile slightly as was observed with DOPG addition, which supports the theory that a larger lipid area will lead to less aggregation and better SEC profiles. LILBID-MS analysis of PR/SapNP (EC) complexes revealed particles containing up to 3 PR and 2 SapA entities (Figure 7G). Probably particle size did not allow for further insertion of PR to reach physiologic pentameric or hexameric assembly<sup>176,177</sup>.

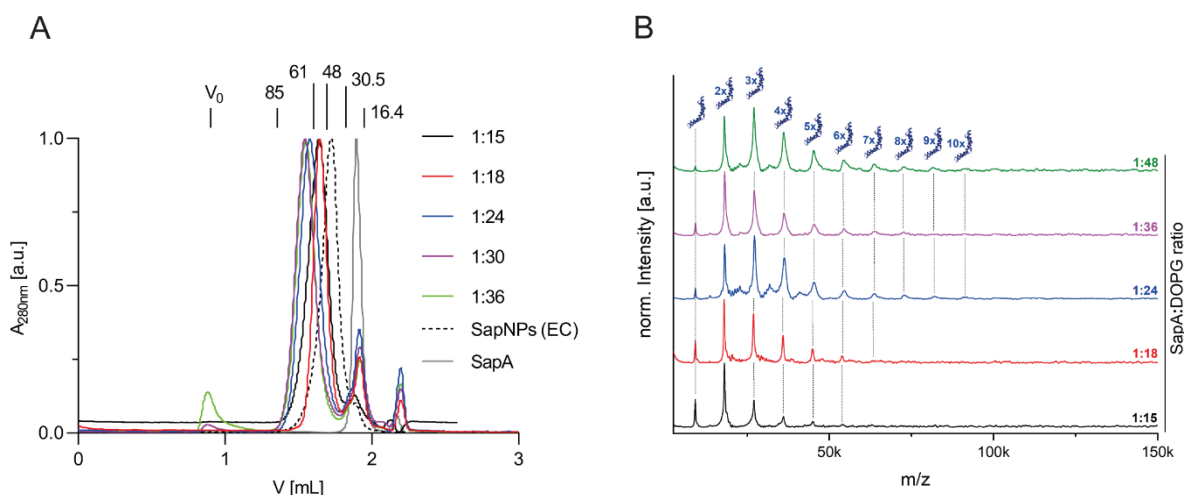
## Results



**Figure 7.** Purification of SapNPs (EC) and application for PR solubilization. **A:** SEC analysis of purified SapA conformers with a Superdex 200 3.2/300 column. Samples were prepared with (red line) or without addition of 2% Triton X-100 (black line). Addition of 2 mM DTT to lysis buffer eliminates larger SapNPs (EC) (blue line). **B:** SEC analysis of separated fractions from A on a Superose 6 (3.2/300) column. **C:** Spontaneously formed SapNP (EC) particles were used for PR solubilization by expression with 15 ng/ $\mu$ L template and increasing concentrations of SapNPs (EC) (white bars) or 400  $\mu$ M SapNPs (EC) with additional DOPG (grey bars). Concentration of PR/SapNPs was determined via absorption at 530 nm measured after three independent expressions. Bars are means of three independent reactions. **D:** Pellet (P) and supernatant (S) fractions of PR samples from C. **E:** Analytical SEC on a Superose 6 (3.2/300) column of samples from C. Radii of calibration proteins are given in  $\text{\AA}$  and  $V_0$  indicates the column void volume of the SEC columns. **F:** LILBID-MS analysis of free SapA and SapNPs (EC) after purification. **G:** LILBID-MS spectra of SapA and SapNPs (EC). In SapNP (EC) spectrum peak signals adjacent to the SapA signals indicate the presence of lipids with a mass of approx. 720 Da. Pictograms in the LILBID-MS spectra illustrate the detected scaffold protein SapA (blue symbol) and detected PR (red symbol). Numbers of detected monomers are indicated. Previously published modified figure taken from Levin et al<sup>45</sup>.

#### 4.1.2 Preassembly of SapNPs for MP expression

To ensure a homogenous composition of SapNPs, only monomeric SapA fractions were used for particle preforming with DOPG. Since SapA was reported to be a NP system with a high adaptability, different SapA to lipid ratios were screened. Unlike other studies, in this study no detergent was used for particle formation and the sole acidification effect was used to induce particle assembly (3.3.3). This approach appeared to be very convenient as detergent removal would require either incubation with SM2 beads, dialysis, or another preparative SEC step. Upon increased lipid ratio the size of SapNPs increased and reached its maximum of ca. 13 nm  $\varnothing$  at SapA:DOPG molar ratios of 1:24 and 1:30. Even at lower ratios, particles were larger than SapNPs (EC) (Figure 8A). LILBID-MS analysis showed higher SapA complexes up to decamers at 1:48 SapA:DOPG ratio when compared to a lower lipid ratio of 1:15 SapA:DOPG (Figure 8B). This indicates that NPs with more than commonly reported 4 SapA can be produced. However, their overall stability, especially during mechanical stress of cotranslational MP insertion, remains questionable. It is also to be noted that LILBID-MS rather gives qualitative than quantitative information about entities in the sample. Thus, the fraction of nonameric SapNPs is not expected to be very large.

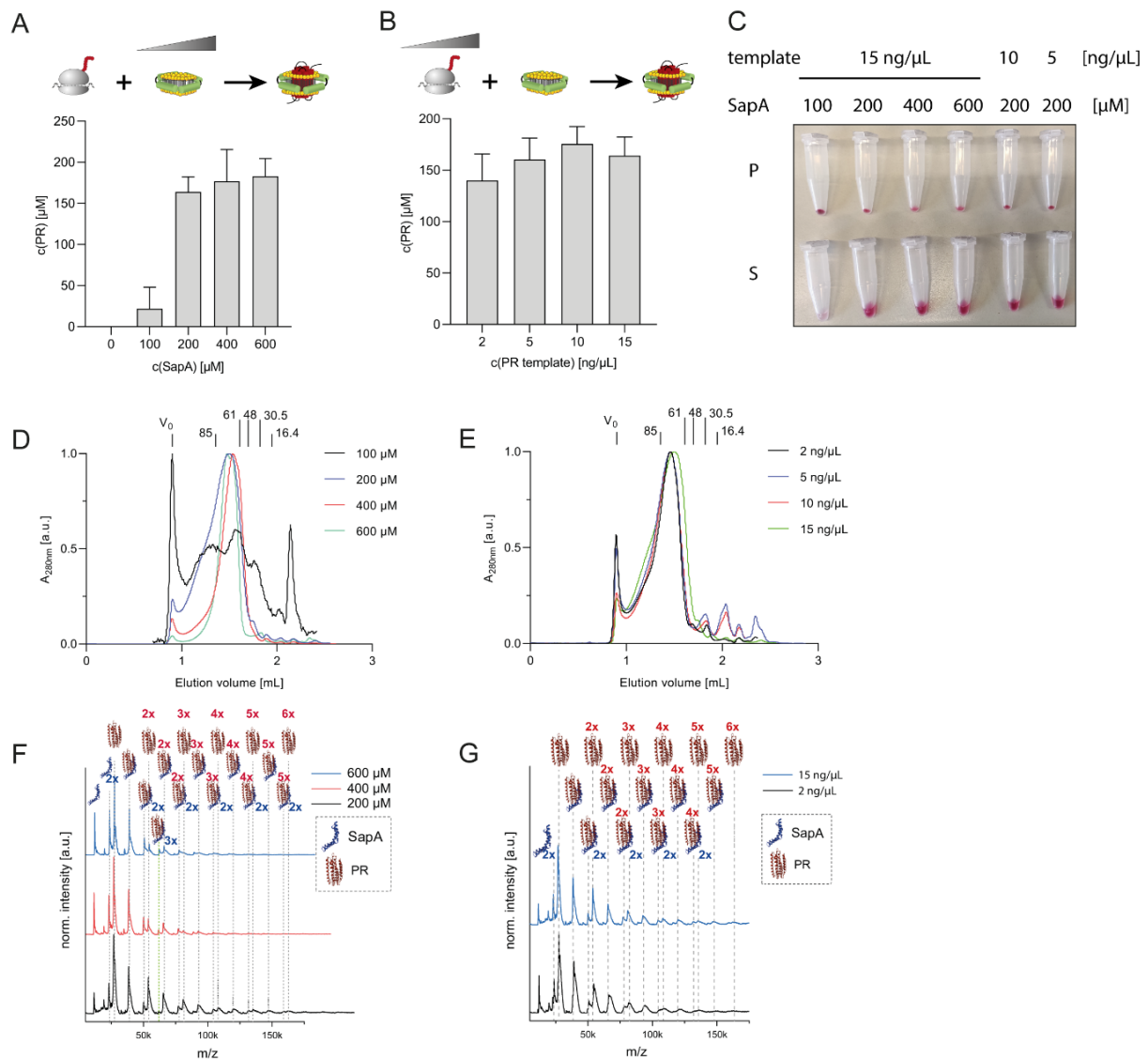


**Figure 8.** Analysis of preassembled SapNPs (DOPG). **A:** SEC analysis with a Superose 6 (3.2/300) column of SapNPs (DOPG) preassembled at different SapA:DOPG ratios. Radii of calibration proteins are given in Å and  $V_0$  indicates the column void volume. **B:** LILBID-MS analysis of SapNPs (DOPG) from A. Blue pictograms in the LILBID-MS spectra illustrate the calculated number of detected SapA entities in the SapA:DOPG particles in the assigned peaks. Figure taken from Levin et al<sup>45</sup>.

Next, the preformed SapNPs were applied to analyze their capability to solubilize PR and other MP targets. For the preassembly strategy, SapNPs with a 1:24 SapA:DOPG ratio were used, since NP size seemed not to further increase at higher ratios. First, PR yields were investigated at different SapNP concentration and different template concentrations. PR expression reached its plateau at ca. 170  $\mu\text{M}$  soluble protein already at 200  $\mu\text{M}$  SapA (Figure 9A). Expression at different template concentrations only slightly varied between 140 and 170  $\mu\text{M}$  at 2 and 15 ng/ $\mu\text{L}$ , respectively (Figure 9B). One must consider that assuming every SapNP is constituted of 4 SapA subunits, this makes up for  $\sim 50$   $\mu\text{M}$  NP concentration. However, as indicated previously, the exact particle composition of SapNPs is unknown. Insoluble red aggregates were observed in all samples, which indicates that at least partially folded PR is present in lipid aggregates. Towards higher concentrations of SapNPs the

amount of red precipitate slightly decreased (Figure 9C). Only at 100  $\mu\text{M}$  concentration of SapA, the pellet size was larger, and the supernatant showed only weak rose color. This indicates that the low concentration of SapNPs was insufficient for this level of expression (Figure 9C). Increase of SapNP concentration in the RM did not improve yield, however, it improved SapNP/PR particle homogeneity as observed using SEC analysis (Figure 9D). LILBID-MS analysis showed that using 200  $\mu\text{M}$  SapA and 15  $\text{ng}/\mu\text{L}$  PR template physiological pentameric and hexameric states of PR in SapNPs can be achieved<sup>176,177</sup>. Increasing SapNP concentration led to improved sample homogeneity, but the oligomerization of PR decreased to trimeric species. This indicates that at higher SapNP concentration, PR gets distributed among the abundant NPs making events of oligomerization less likely (Figure 9F). Variation of PR template concentration between 2 and 15  $\text{ng}/\mu\text{L}$  using 200  $\mu\text{M}$  1:24 SapA:DOPG SapNPs did not show a major effect on particle quality and oligomerization. At 2  $\text{ng}/\mu\text{L}$  template still 140  $\mu\text{M}$  soluble PR were synthesized and the SEC chromatograms (Figure 9E) and LILBID-MS (Figure 9G) data were highly congruent throughout the samples.

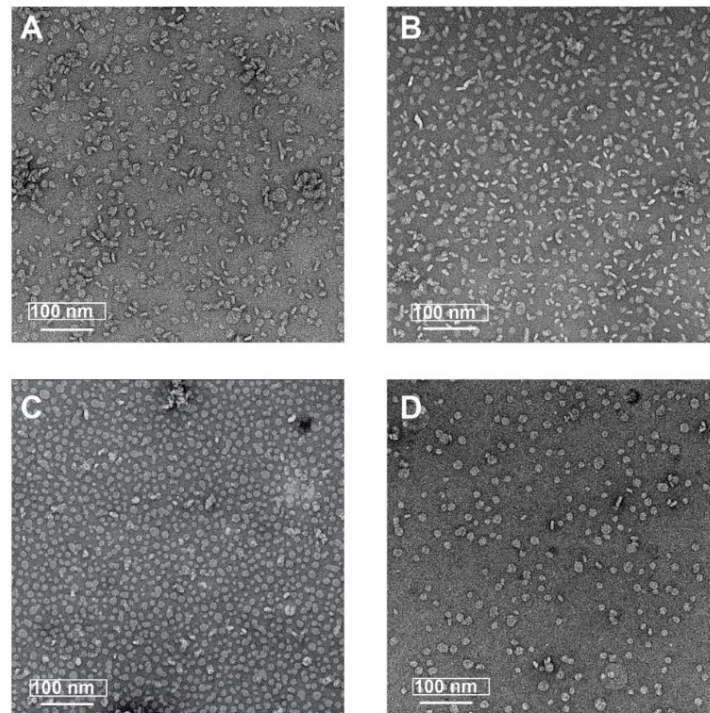




**Figure 9** PR solubilization with preassembled SapNP (DOPG) particles. **A:** PR solubilization with increasing SapNP (DOPG) preassembled at a 1:24 SapA:DOPG ratio. PR template concentration was 15 ng/μL. **B:** PR solubilization with increasing PR template concentrations in presence of 200 μM preassembled SapNPs (DOPG). Concentration of PR/SapNPs were determined via absorption at 530 nm measured after three independent expressions. Bars are means of three independent expressions. **C:** Pellet (P) and supernatant (S) fractions from B obtained at different PR template concentrations. **D:** SEC analysis of purified PR/SapNPs (DOPG) samples from A. **E:** SEC analysis of purified PR/SapNPs (DOPG) complexes from B. All SEC runs were conducted on a Superose 6 (3.2/300) column. Radii of calibration proteins are given in Å and  $V_0$  indicates the column void volume in the respective chromatograms. **F:** LILBID-MS of PR/SapNPs from A. **G:** LILBID-MS of PR/SapNPs from B. Pictograms in the LILBID-MS spectra illustrate the detected scaffold protein SapA (blue symbol) and detected PR (red symbol). Numbers of detected monomers are indicated. Previously published modified figure taken from Levin et al<sup>45</sup>.

Negative stain EM revealed that MSP NDs as well as SapNPs yielded particles that appeared to be different in size, with MSP NDs seeming slightly more homogenous (Figure 10A/B). PR expressed in presence of 20 μM (Figure 10A) NDs (DOPG) showed a larger amount of aggregation clusters compared to PR expressed with 50 μM NDs (Figure 10B). This is in accordance with observations made with preformed SapNPs (Figure 9D). As reported, SapNPs appear to have a larger range of size in the negative stain EM (Figure 10C). This was also observed during preforming studies with

SapA and DOPG (Figure 8A/B). However, even the small SapNPs did not show prominent aggregates and should carry PR as they have been purified via strep tag II fused to the C-terminus of PR (Figure 10D).

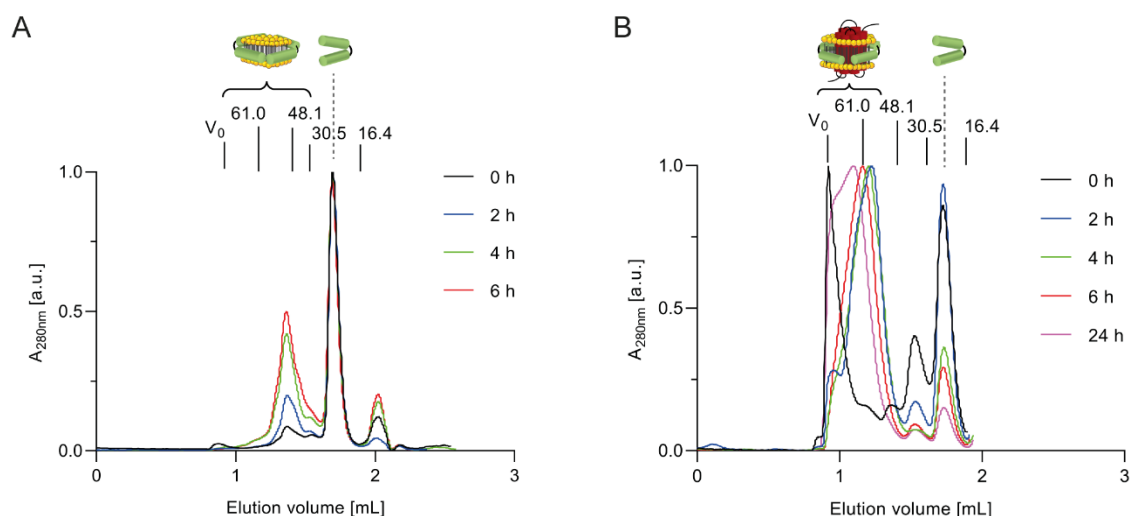


**Figure 10.** Negative stain EM images of NDs, PR/NDs, SapNPs and PR/SapNPs. **A+B:** PR/NDs synthesized with 15 ng/μL PR template and in presence of preassembled MSP1E3D1 NDs (DOPG). Samples were diluted 1:10. **A:** 20 μM NDs added in the RM during expression. **B:** 50 μM NDs added in the RM during expression. **C+D:** Preassembled SapNPs (DOPG) with SapA:DOPG ratio of 1:24. Samples were 1:20 diluted. **C:** SapNPs (DOPG). **D:** PR/SapNP (DOPG) synthesized in presence of 250 μM SapNPs (DOPG) and with 15 ng/μL PR template. Previously published figure taken from Levin et al<sup>45</sup>.

#### 4.1.3 Coassembly of SapA with lipids and MPs

Even though preassembly with SapNPs does not require three days of dialysis as is the case for MSP ND preforming, a faster and more convenient route is the coassembly of scaffold protein, MP and DOPG directly during expression of the MP. After finding key parameters for the implementation of preassembled SapNPs in the CF system, next we investigated parameters important for SapNP/PR/DOPG coassembly. First, we analyzed the capability of SapA and lipid self-assembly in the RM over the course of time, by purifying every SapA product via the his-tag of SapA. Without PR expression, particle assembly was rather slow and resulted in ~30 % of SapA incorporated into SapNPs after 4 h and did not significantly change after 6 h (Figure 11A). This is not surprising as SapA requires low pH or addition of detergent for transition into a state with lipid binding capability which is both not provided in the CF reaction. Interestingly, if PR was expressed, particle assembly was facilitated and after 4 h SapA was largely incorporated into SapA/PR/lipid complexes. The particles became slightly larger after 6 h indicating ongoing PR insertion and after 24 h particles became partially aggregated as they eluted at void volume during SEC (Figure 11B). This indicates a cooperative process of SapNP formation that is probably favored by

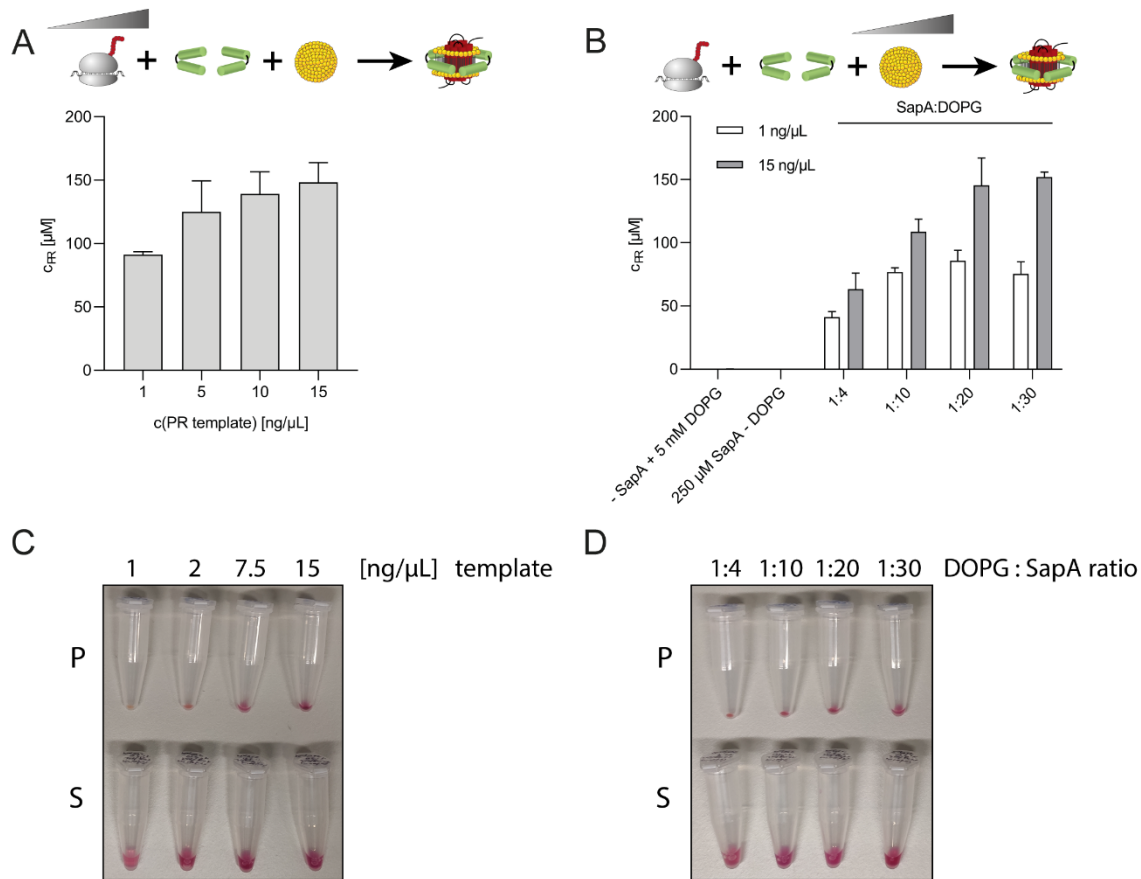
hydrophobic surfaces provided by PR helices. This vividly demonstrates that the kinetics of PR insertion, and thus, expression duration also influence sample quality.



**Figure 11.** Coassembly kinetics of SapA and lipids in the CF mixture with or without PR expression. **A:** Kinetics of spontaneous empty SapNPs (DOPG) formation in CF reaction mixtures during coassembly of 200  $\mu\text{M}$  SapA with 6 mM DOPG. SEC chromatograms show SapNP assembly products obtained after 0 - 6 h incubation in the CF RM. **B:** Kinetics of spontaneous formation of PR/SapNPs over the course of expression with 1 ng/ $\mu\text{L}$  PR template. SEC chromatograms show SapNP assembly products obtained after 0 - 24 h incubation in the CF reaction mixture. SapNPs and SapNPs/PR were recaptured via his-tag of SapA and eluted from IMAC beads. SEC runs were performed with a Superdex 200 (3.2/300) column. Radii of calibration proteins are given in Å and  $V_0$  indicates the column void volume in the respective chromatograms. Figure modified after Levin et al<sup>45</sup>.

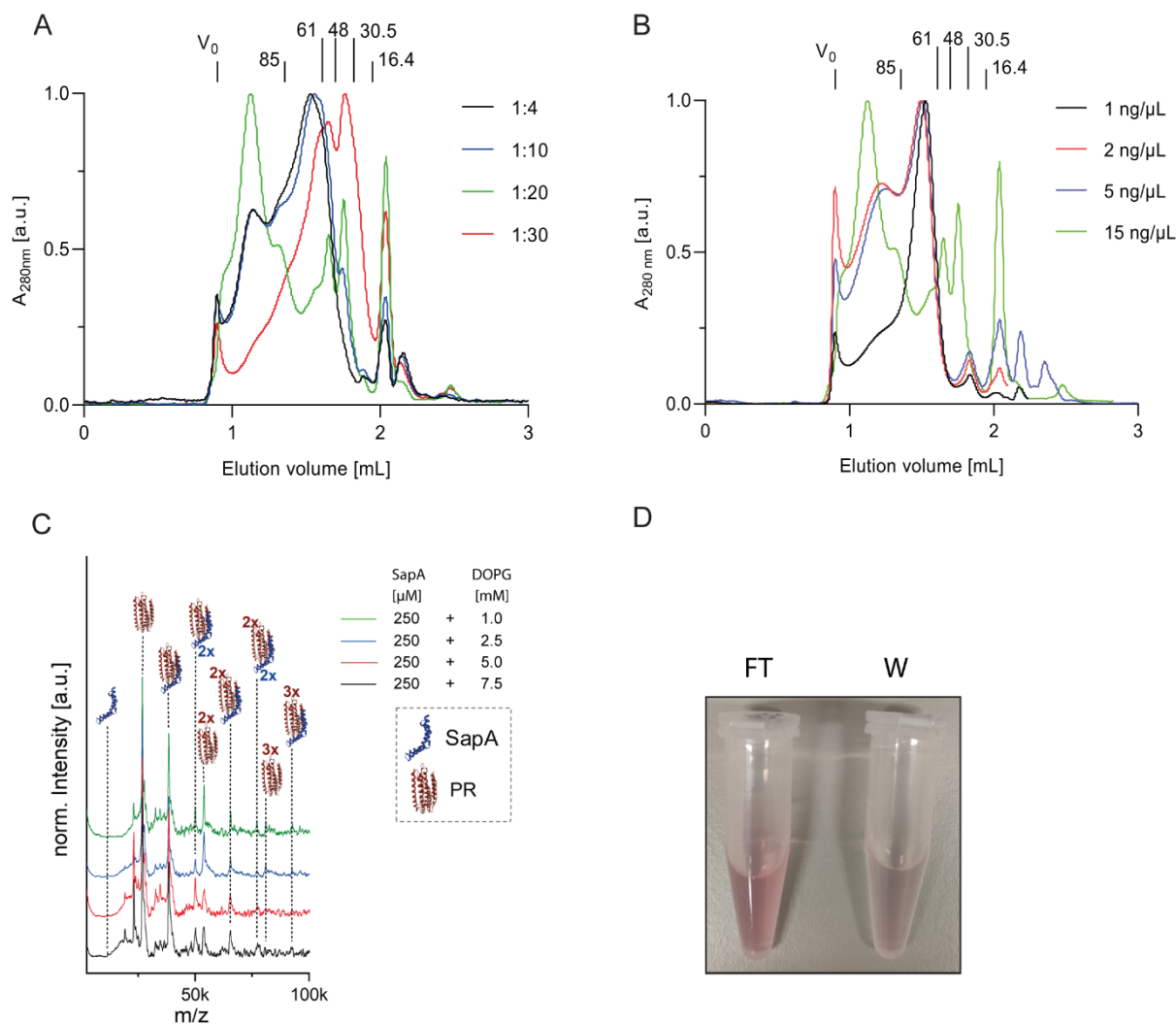
Next, the effect of PR template concentration on PR expression and solubilization in coassembly mode was investigated. Template concentration of PR had only a minor effect on the yield of solubilized product using the coassembly approach. Already at 1 ng/ $\mu\text{L}$  a PR concentration of 90  $\mu\text{M}$  was reached and an increase of template concentration from 5-15 ng/ $\mu\text{L}$  lead to expression yields of 120-150  $\mu\text{M}$  (Figure 12A). More detrimental was the effect of the amount of provided DOPG for particle assembly. While removing either SapA or DOPG was detrimental for PR solubilization, increase of DOPG concentration during expression resulted in PR yield increase plateauing at a ratio of 1:20 SapA:DOPG. Expressions at 1 ng/ $\mu\text{L}$  PR and 15 ng/ $\mu\text{L}$  PR template concentration yielded  $\sim 90 \mu\text{M}$  and  $\sim 140 \mu\text{M}$ , respectively (Figure 12B). Interestingly, the coassembly mode is more sensitive to template concentration during expression as almost no precipitate was formed at 1 ng/ $\mu\text{L}$  compared to 15 ng/ $\mu\text{L}$  PR template. It is possible that lower amount of provided particles tends to result in SapNP overcrowding or insertion into liposomes as not enough lipid area is provided by SapNPs (Figure 12C). The liposomes on the other hand, are destabilized by PR insertion and precipitate resulting in red pellets. Increase of DOPG liposome concentration in the reaction mixture resulted in slightly larger insoluble PR aggregates which supports the assumption that PR also inserts into liposomes (Figure 12D).

## Results



**Figure 12.** Coassembly of PR/SapNPs with nascent PR, purified SapA and DOPG liposomes. **A:** PR solubilization with increasing PR template concentrations in presence of 250  $\mu\text{M}$  SapA and 5 mM DOPG (SapA:DOPG ratio of 1:20). **B:** PR solubilization with increasing DOPG concentrations in presence of 250  $\mu\text{M}$  SapA and 1 ng/ $\mu\text{L}$  (white bars) or 15 ng/ $\mu\text{L}$  (grey bars) PR template. Concentration of PR/SapNPs were determined via absorption at 530 nm measured after three independent expressions. Bars are means of three independent reactions. **C:** Pellet and supernatant fractions from A. Increasing template concentration during expression leads to larger PR pellets indicating insoluble aggregate formation. **D:** Pellet and supernatant fractions of PR samples from B expressed at 15 ng/ $\mu\text{L}$ . Increase of DOPG liposomes leads to increased PR pellet fractions. Previously published modified figure taken from Levin et al<sup>45</sup>.

Next, we analyzed the impact of different SapA to DOPG ratios and template concentrations on the PR/SapNP sample quality by SEC and LILBID-MS. While all samples expressed at 15 ng/ $\mu\text{L}$  resulted in inhomogeneous samples with large aggregate fractions (Figure 13A), reduction of template concentration resulted in a more homogenous PR/SapNP peak (Figure 13B). This is in accordance with the previous observation that larger red pellets are formed at higher template concentration (Figure 12C). It was further apparent that excessive sample aggregation resulted in poor quality LILBID-MS spectra (Figure 13C). Also, aggregation seems to have a negative effect on the overall protein purification yield since PR only poorly bound to strep beads. This was already indicated by red colored flow through (FT) and wash (W) fractions indicating presence of folded but unbound PR (Figure 13D). This is most likely due to poor accessibility of the strep-tag in soluble SapA/membrane aggregates that contained folded protein leading to up to 80 % of sample loss. Such effects were not observed with preassembled particles.

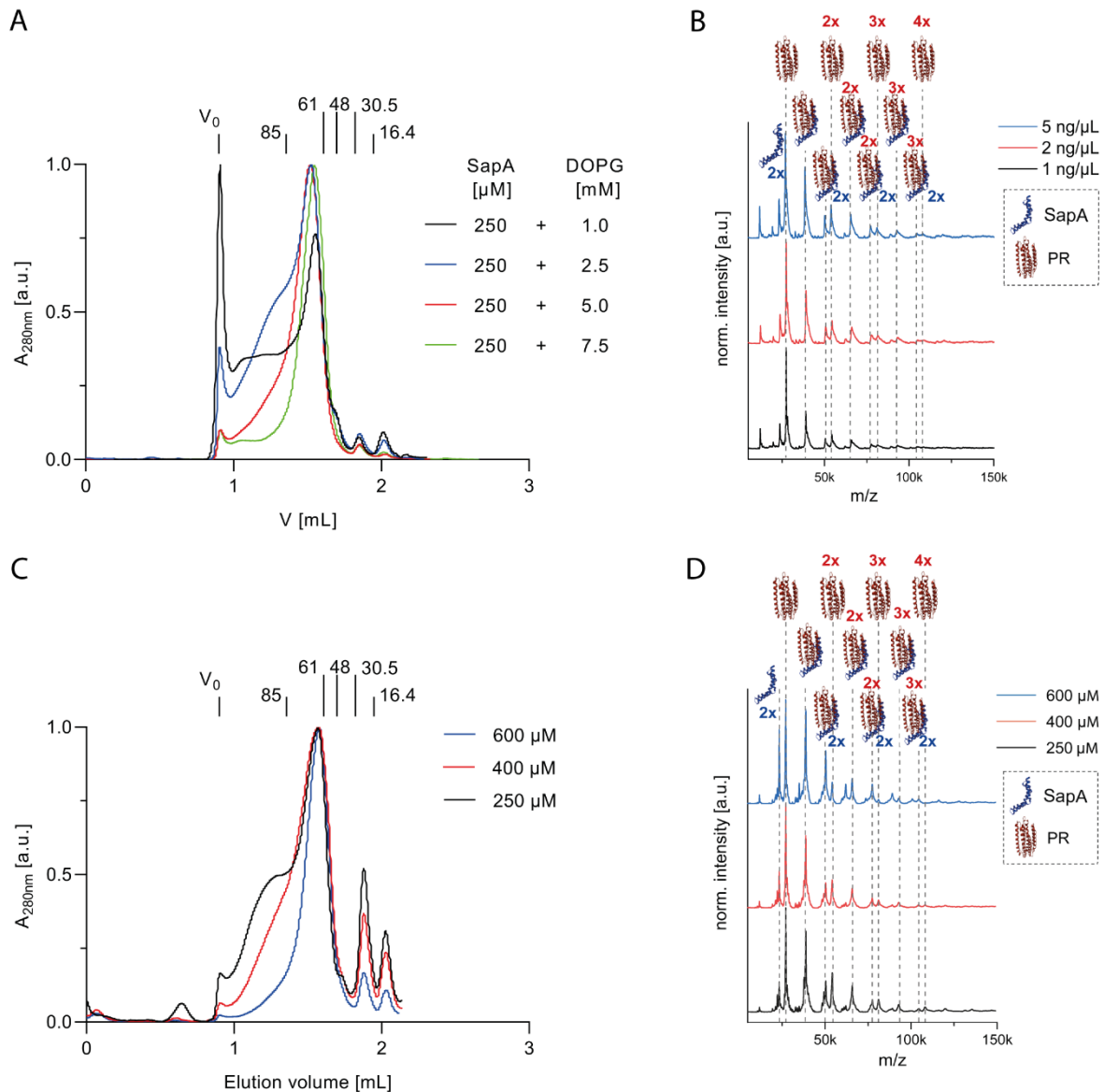


**Figure 13.** Sample quality of coassembled PR/SapNPs at different template and lipid ratios. **A:** SEC profiles of PR expressed in coassembly mode at 250  $\mu\text{M}$  SapA and DOPG concentrations of 1-7.5 mM (1:4-1:30 ratio) and 15 ng/ $\mu\text{L}$  template. **B:** SEC profiles of PR expressed in presence of 250  $\mu\text{M}$  and 5 mM DOPG at template concentrations of 1-15 ng/ $\mu\text{L}$ . SEC runs were carried out with a Superose 6 (3.2/300) column. Radii of calibration proteins are given in  $\text{\AA}$  and  $V_0$  indicates the column void volume in the respective chromatograms. **C:** LILBID-MS analysis of oligomerization of PR samples from A. Pictograms in the LILBID-MS spectra illustrate the detected scaffold protein (blue symbol) and detected PR (red symbol) and numbers of detected monomers are indicated. **D:** Flow through (FT) and wash (W) fractions indicate that PR samples with high number of soluble aggregates poorly bind to Strep-tactin beads. Previously published modified figure taken from Levin et al<sup>45</sup>.

Apparently, the coassembly approach is more sensitive to expression conditions than the use of preformed particles which showed to be far less sensitive to PR template concentration. Further analysis of samples expressed at 1 ng/ $\mu\text{L}$  PR template concentration showed that increase of lipid concentration is beneficial for sample homogeneity while short supply of lipid results in aggregation (Figure 14A). Vice versa, increase of SapA concentration at a fixed concentration of 7.5 mM DOPG was beneficial for sample quality (Figure 14C). Increase of either lipid, SapA or both seems beneficial for particle formation. Reasonably, the higher the concentration of either component, the higher the chance of particle formation. The more particles, and thus the more lipid area is provided, the more PR can distribute across particles. If, however, SapA and/or DOPG is provided in insufficient amounts, then too high expression capacity provided by high plasmid concentration will cause aggregation. Presumably

## Results

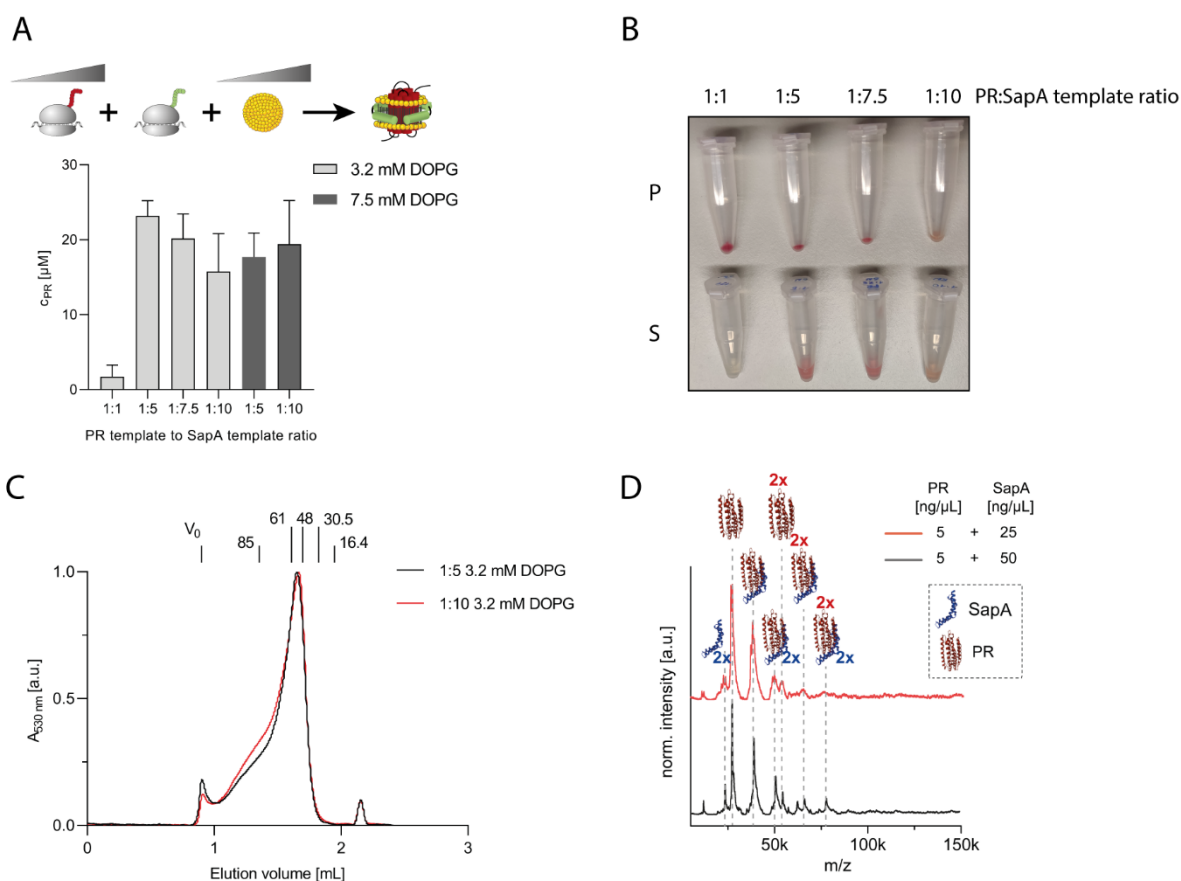
lower template concentration slows down expression, so particles did not get overcrowded by PR in the usual expression time frame of 12-16 h. The coassembly approach did not yield pentameric or hexameric species. Up to tetramers and trimers were found at 5 ng/ $\mu$ L and 1 ng/ $\mu$ L expression, respectively (Figure 14B). This was also the case for experiments at different SapA:DOPG ratios (Figure 14D).



**Figure 14.** PR/SapNP sample quality analysis after synthesis by coassembly. SEC was conducted on a Superose 6 (3.2/300) column. **A:** SEC analysis of coassembled PR/SapNP complexes with 250  $\mu$ M SapA, 1 ng/ $\mu$ L PR template and at SapA:DOPG ratios from 1:4 to 1:30 (= 1, 2.5, 5, 7.5 mM DOPG). **B:** Oligomerization of PR in coassembled SapA:DOPG (250  $\mu$ M SapA + 5 mM DOPG, 1:20 ratio) and 1–5 ng/ $\mu$ L PR template. **C:** SEC analysis of coassembled PR/SapNPs at constant DOPG concentration of 7.5 mM, 250–600  $\mu$ M SapA and 2 ng/ $\mu$ L PR template. **D:** LILBID-MS analysis of samples from C. Pictograms in the LILBID-MS spectra illustrate the detected scaffold protein SapA (blue symbol) and detected PR (red symbol). Numbers of detected monomers are indicated. Radii of calibration proteins are given in Å, and  $V_0$  indicates the column void volume in the respective chromatograms. Previously published figure taken from Levin et al<sup>45</sup>.

#### 4.1.4 Coexpression of SapA with PR

A final simplification of SapA application in the CF system was the simultaneous expression of both components, MP and SapA, in presence of lipids. Since both, PR and SapA compete over the same transcription and translation machinery during CF expression, PR template to SapA template ratios and lipid concentration were screened. While at a ratio of 1:1 no PR was solubilized, all ratios above yielded  $\sim 15\text{--}20\ \mu\text{M}$  PR regardless of the applied lipid concentration (Figure 15A). The amount of red precipitate decreased with increasing SapA template concentration, which took more capacity from the transcription and translation machinery resulting in more SapA and less PR expression. At 1:1 ratio most likely PR preferably inserted into liposomes leading to more precipitation. It's also possible that the 1:1 ratio did not yield enough SapA for particle formation (Figure 15B). The particles appeared to be somewhat smaller than the particles obtained with preassembly or coassembly strategies (Figure 15C). This is most likely due to lower oligomerization of PR reaching only up to dimers with the coexpression approach (Figure 15D). Reasonably, at such low PR yield the particles were not overcrowded by PR, and thus, low oligomeric states were observed. Another explanation is the assembly mode of particles probably consisting of only two SapA units due to low SapA availability. Coexpression yielded significantly less solubilized PR than the previous approaches. This consolidates the assumption that a sufficient amount of SapNPs is important for MP yield.

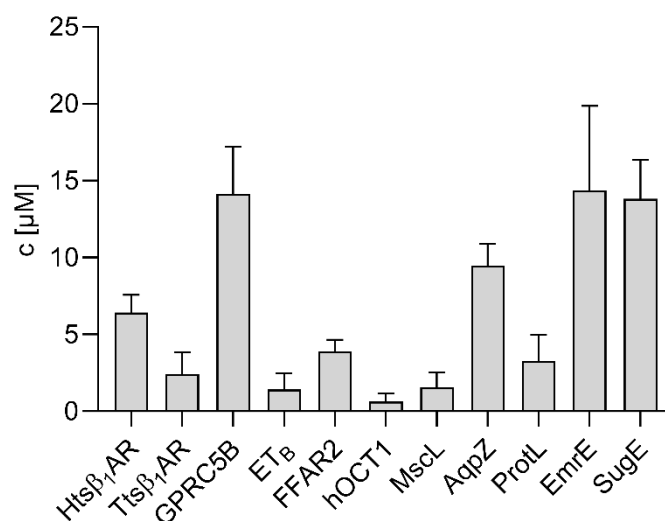


**Figure 15.** Analysis of SapNP/PR particles obtained by coexpression of PR and SapA in the CF reaction. **A:** Yield of PR/SapNPs at varying PR to SapA template ratios. Expression was performed at a fixed PR template concentration of 5 ng/ $\mu\text{L}$  and with 3.2 mM or 7.5 mM DOPG. The template ratios of PR to SapA are indicated. Concentration of PR/SapNPs were determined via absorption at 530 nm measured after three independent expressions. Bars are means of three independent reactions. **B:** SEC profiles of strep

purified PR/SapNPs after coexpression at 3.2 mM DOPG and 1:5 or 1:10 PR:SapA template ratios. SEC analysis was performed with a Superose 6 (3.2/300) column. Radii of calibration proteins are given in Å and  $V_0$  indicates the column void volume in the respective chromatograms. **C**: LILBID-MS analysis of samples from A synthesized with 7.5 mM DOPG. Pictograms in the LILBID-MS spectra illustrate the detected scaffold protein SapA (blue symbol) and detected PR (red symbol). Numbers of detected monomers are indicated. Previously published modified figure taken from Levin et al<sup>45</sup>.

#### 4.1.5 Applicability of SapNPs in the CF system

SapNPs were shown to be capable to solubilize PR which is a suitable model protein for versatile studies on particle quality and MP oligomerization. Next, expression of an array of MP-GFP chimeras was tested regarding yield, behavior during purification and particle quality. Expression included GPCR chimeras of free fatty acid receptor 2 (FFAR<sub>2</sub>), thermostabilized  $\beta_1$  adrenergic receptors from turkey and human (Tts $\beta_1$ AR/Hts $\beta_1$ AR), GPRC5B and endothelin receptor B (ET<sub>B</sub>). Also included were human organic cation transporter 1 (hOCT1), phage protein L (ProtL) and bacterial MPs, such as water channel aquaporin Z (AqpZ), large conductance mechanosensitive channel MscL and multidrug efflux proteins EmrE and SugE. GFP fluorescence was detected in all CF supernatants after expression and removal of insoluble aggregates by centrifugation. Protein fractions remaining in solution after centrifugation at 20,000 x g for 10 min were considered soluble (Figure 16).

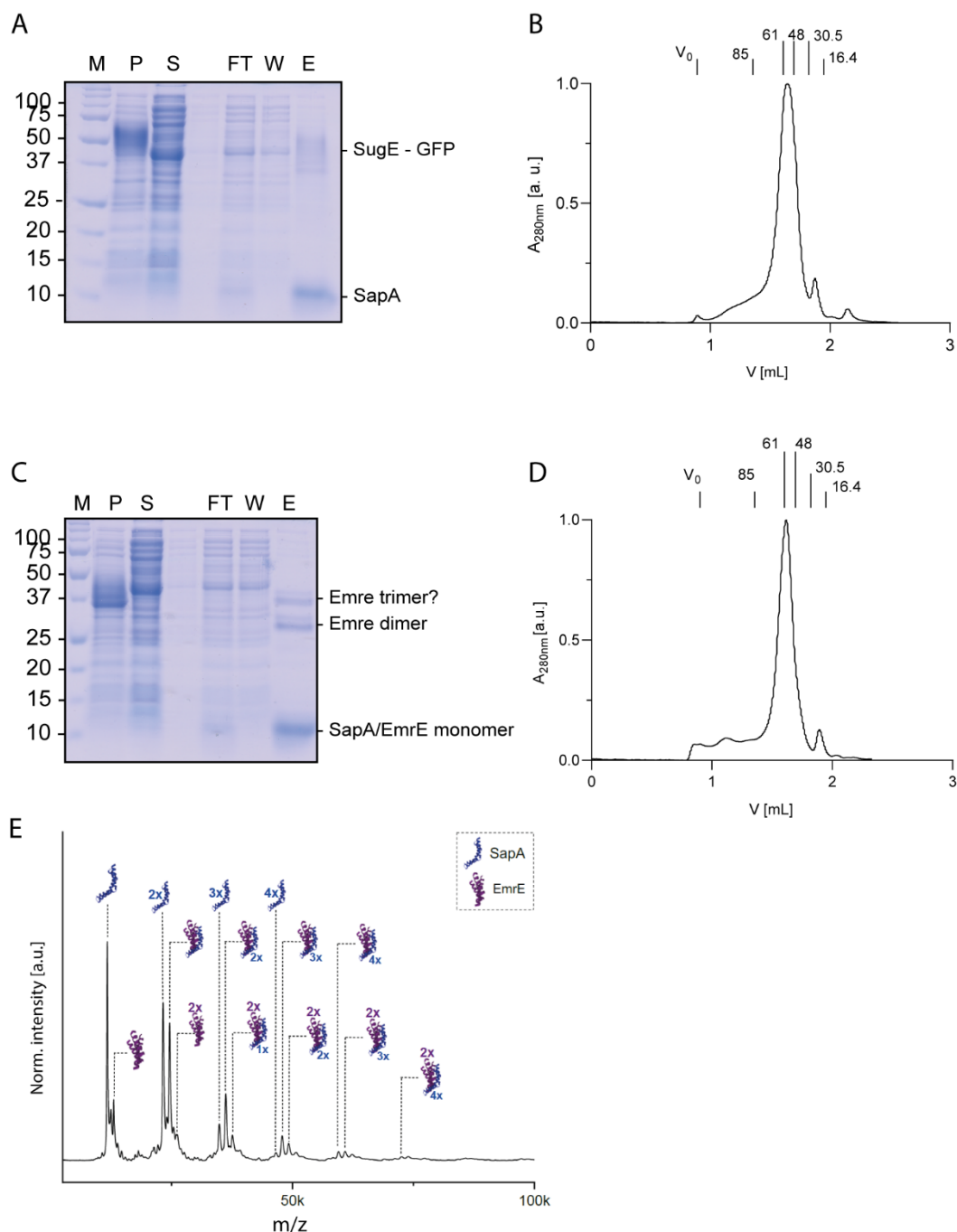


**Figure 16:** Cotranslational solubilization of GFP chimeras of various MPs. 250  $\mu$ M SapNPs preassembled with DOPG in a 24:1 DOPG:SapA ratio were used in the RM. Median values obtained from fluorescence measurement of three independent CF expressions. Error bars indicate standard deviation.

Compared to PR, solubilization of other MPs was far less efficient and reached expression levels comparable with PR coexpressed with SapA (Figure 15A/16). However, this observation was also made with NDs where PR is expressed even up to 300  $\mu$ M while other MPs yields are far below. Indeed, only small bacterial multidrug efflux proteins EmrE and SugE and GPRC5B showed expression levels reaching  $\sim$ 15  $\mu$ M. Other proteins yielded  $\sim$ 10  $\mu$ M (AqpZ) or below. As comparison, except GPRC5B, EmrE and SugE that showed comparable yields, other MPs usually reached



at least 2-fold higher yields using MSP NDs instead of SapNPs. Analyzing particle quality of the bacterial efflux proteins EmrE and SugE, it was found that both proteins can be expressed in coassembly mode at 250  $\mu$ M SapA and 5 mM DOPG. The samples yielded comparable amounts as the preforming mode, sufficient for SEC and LILBID-MS analysis (Figure 17).

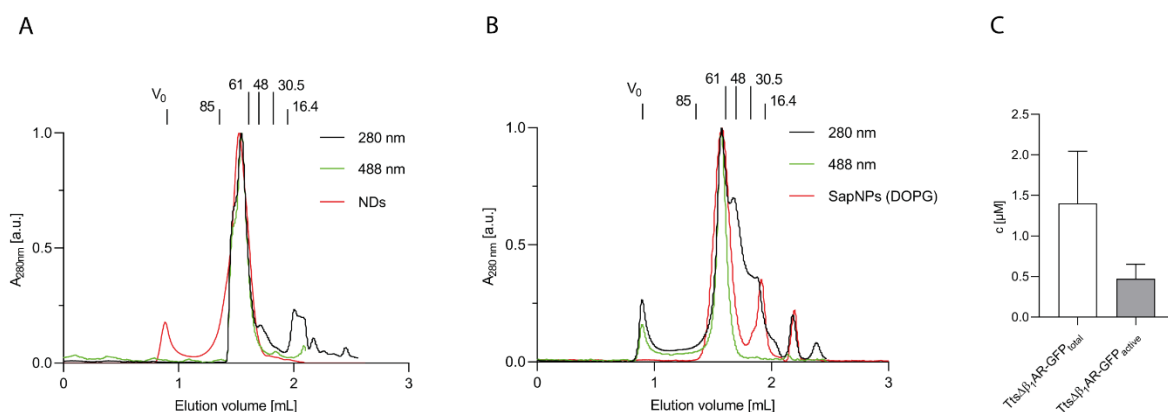


**Figure 17.** Expression and purification of bacterial multidrug efflux proteins SugE-GFP-his and EmrE-strep. Proteins were expressed in coassembly mode with 250  $\mu$ M SapA and 5 mM DOPG. For SugE, TEV cleaved SapA was used to avoid purification of empty particles. **A:** SDS-PAGE of SugE-GFP-his IMAC purification. **B:** SEC of elution fraction of SugE-GFP from A. **C:** SDS-PAGE of strep purification of EmrE-strep. **D:** SEC of elution fraction of EmrE-strep from C. **E:** LILBID-MS analysis of elution fraction of EmrE-strep. Pictograms in the LILBID-MS spectrum illustrate the detected scaffold protein (blue

## Results

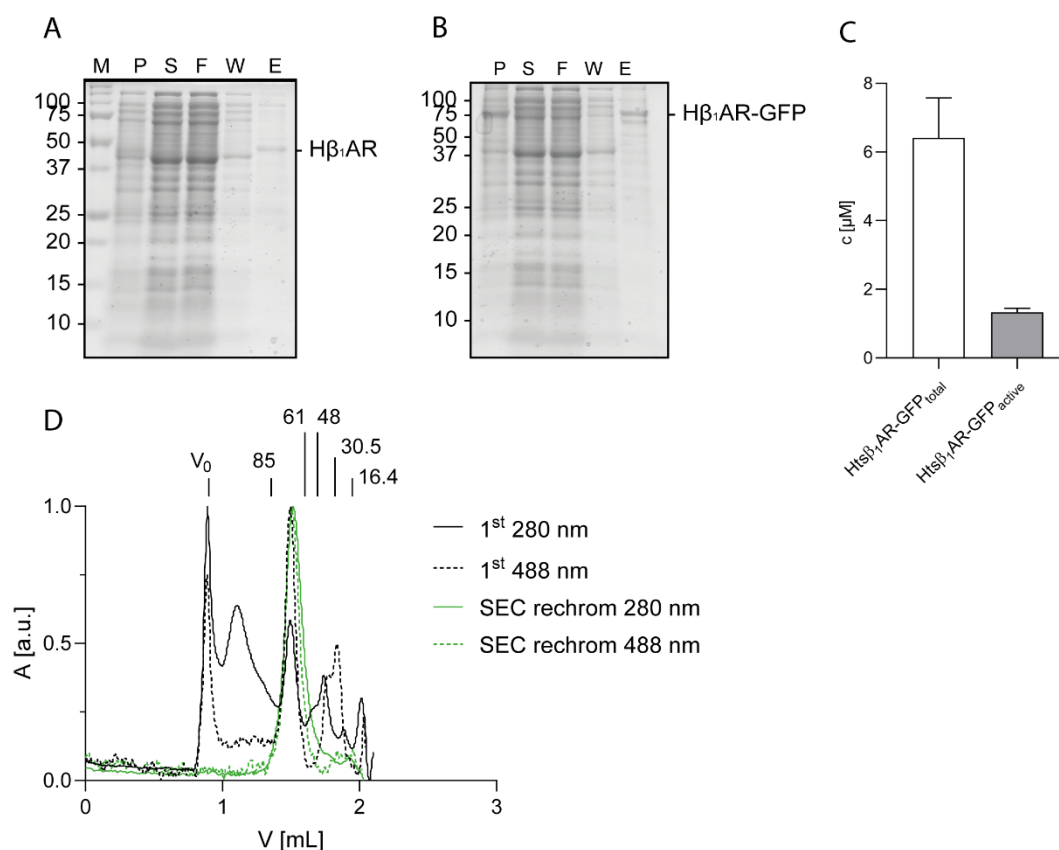
symbol) and detected EmrE-strep (violet symbol). The corresponding numbers of detected EmrE or SapA are indicated, respectively. SDS-PAGE bands of purification fractions are pellet (P), supernatant (S), flow-through (FT), wash (W) and elution (E). SEC analysis was performed with a Superose 6 (3.2/300) column. Radii of calibration proteins are given in Å and  $V_0$  indicates the column void volume in the respective chromatograms.

Expression and purification of small bacterial multidrug efflux proteins SugE and EmrE was proven successful. SDS-PAGE showed a clear band of SugE-GFP and SapA (Figure 17A) while EmrE showed dimers that appeared to remain intact even in SDS loading buffer. Surprisingly, a band that would correspond to an artificial trimer was detected (Figure 17C). SEC profiles indicated that the samples were almost completely free of aggregates, and both showed a diameter of ca. 12 nm which is similar to the usual nanodisc size (Figure 17B/D). LILBID-MS analysis showed prevalently monomer species of EmrE but also dimers. Interestingly, SapA signals corresponding to 2, 3 and 4 SapA units were resolved, indicating that SapNPs consisting of 4 SapA units can be formed even in coassembly mode (Figure 17E). No artificial trimers of EmrE were detected in LILBID-MS analysis, indicating that the SDS band at ca. 37 kDa (Figure 17C) may correspond to a complex of EmrE/SapA. SapNPs combined with CF system appeared to be robust when working with PR or small proteins with low oligomerization states. Next, we analyzed, how the system will behave with more complex targets with flexible regions such as a GPCR. Starting off with GPCRs yielding highest expression levels appeared reasonable. However, FFAR<sub>2</sub> or GPRC5B were not obtained in significant amounts after strep purification. Like PR samples showing presence of soluble aggregates, these GPCRs did not bind to the strep matrix, which was also indicated by green flow-through of the GFP chimeras (data not shown). Therefore, we expressed Ttsβ<sub>1</sub>AR from fused to GFP which allowed to determine yield and ligand binding fraction directly in the unpurified sample. It was also possible to purify the receptor via IMAC to analyze particle quality by SEC and compare it to a sample expressed in MSP1D1E3 NDs (Figure 18A/B).



**Figure 18.** Quality control of Ttsβ<sub>1</sub>AR-GFP-his purified via IMAC. **A:** Analytical SEC of Ttsβ<sub>1</sub>AR-GFP-his expressed in presence of 60 μM MSP1D1E3 (TEV cleaved) NDs. **B:** Ttsβ<sub>1</sub>AR-GFP-his expressed in presence of 400 μM SapNPs preassembled in a ratio of 1:24 with SapA (TEV cleaved):DOPG. Elution was monitored via 280 nm and 488 nm to observe in which fraction the GFP chimera eluted. Radii of calibration proteins are given in Å, and  $V_0$  indicates the column void volume in the respective chromatograms. **C:** Determination of the amount of active GPCR via radioligand binding assay in the unpurified sample. Bars indicate median concentration and activity of the GPCR calculated from three independent expressions.

The yield of expressed Tts $\beta_1$ AR-GFP-his in SapNPs was lower than FFAR $_2$  and GPRC5B. However, it was possible to purify the receptor in contrast to the other two GPCRs. The protein sample seemed to contain aggregates as fluorescence in the elution sample was barely measurable yielding a submicromolar probe. However, the sample was enough to perform analytical size exclusion analysis, which was compared to Tts $\beta_1$ AR-GFP-his expressed with 60  $\mu$ M NDs (DOPG) (Figure 18A). SEC analysis of the Tts $\beta_1$ AR-GFP-his/SapNP sample showed an almost aggregate free peak with minor peaks at the peak tail (Figure 18B). Presumably these peak shoulders correspond to impurities after IMAC purification, which appear very prominent, due to low concentration of the receptor. The GFP absorbance was congruent with the empty SapNP elution peak confirming that the receptor chimera successfully entered the NP. Radioligand assay showed a 34 % activity of unpurified sample which is comparable to activity measured in nanodiscs (Figure 18C). Finally, it was attempted to express the thermostabilized GFP chimera of the human  $\beta_1$ AR (Hts $\beta_1$ AR-GFP-his). Hts $\beta_1$ AR was previously successfully expressed and biochemically and structurally characterized by our group<sup>27</sup>. Now the receptor was expressed with 400  $\mu$ M preformed SapNPs and particle quality and activity in the crude CF mixture were investigated (Figure 19).



**Figure 19.** Expression and purification of Hts $\beta_1$ AR variants. Expression was carried out at 10 ng/ $\mu$ L template concentration and 400  $\mu$ M preformed SapNP/DOPG (1:24) particles. SDS-PAGE of strep purification of Hts $\beta_1$ AR-GFP (**A**) and IMAC purification of Hts $\beta_1$ AR-GFP (**B**). Purification fractions of pellet (P), supernatant (S), flow-through (F), wash (W), elution (E) are indicated as well as marker lane (M). **C:** Total receptor yield (Hts $\beta_1$ AR-GFP<sub>total</sub>) determined by GFP fluorescence and amount of active fraction (Hts $\beta_1$ AR-GFP<sub>active</sub>) determined radioligand assay measured in the crude CF supernatant. **D:** SEC analysis of strep purified elution fraction (black chromatograms). To analyze particle stability, the main peak fraction collected at ca. 1.5. mL elution volume was again concentrated and subjected to a second round of SEC (green chromatograms). Absorption was monitored at 280 nm and 488 nm to

## Results

---

detect fractions containing GFP. Radii of calibration proteins are given in Å and  $V_0$  indicates the column void volume in the respective chromatograms.

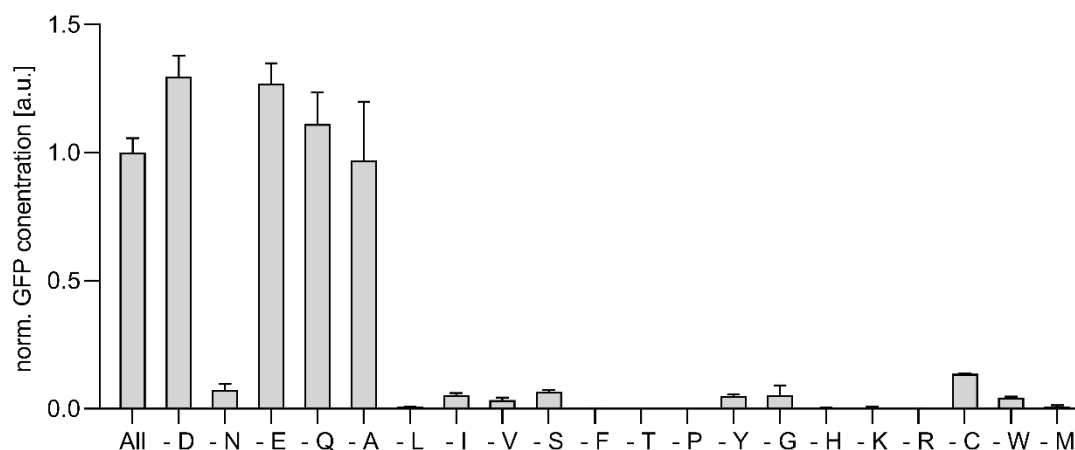
Compared to purifications of smaller MPs, no SapA band was detected on SDS-PAGE and only weak bands of GPCR were visible (Figure 19A/B). This again indicates formation of soluble aggregates that prevent protein probe from binding. Despite ~21 % ligand binding fraction in the supernatant of the crude CF reaction (Figure 19C), the CF reaction volume had to be increased to 600  $\mu$ L to obtain a probe sufficient for analytical SEC analysis (Figure 19D). The Hts $\beta_1$ AR/SapNP complexes eluted at around 1.5 mL elution volume which roughly corresponds to particles observed with Tts $\beta_1$ AR-GFP. Compared to Tts $\beta_1$ AR-GFP, this sample still contained soluble aggregates. To investigate if these are formed due to particle instability, the main peak fraction was collected, concentrated, and applied on SEC for a second time. The peak eluted at the same height as previously and no new aggregates or low molecular degradation products were observed, indicating that a certain amount of the particles was stable. Both chromatograms showed GFP absorption in the main elution fraction. Aggregate fraction without GFP absorption might be due to disintegrated H $\beta_1$ AR/SapNP complexes that did not contain folded GFP. (Figure 19D).

The extensive study on SapNPs revealed that SapNPs can be applied to stabilize some MPs by cotranslational insertion. It revealed valuable insights into cotranslational MP solubilization and pointed out factors and parameters to be considered when working with other NP scaffolds. Although, at this point of research, SapNPs cannot compete with MSP NDs in the CF system regarding target yield and particle quality, it was still shown that the system is capable of stabilizing active GPCRs and MPs of lower complexity such as PR, EmrE or SugE. However, further optimization might yield cryo-EM grade GPCR samples in the future. To achieve this goal, SapNP stability would have to be drastically improved by either fine tuning reaction conditions or even engineering of SapA. It is possible, that despite the extensive work, some details and optimization possibilities were not yet considered. However, thus far, the system has also proven to be compatible with LILBID-MS and according to negative stain EM, single particles of CF synthesized MP/SapNP complexes can be obtained for cryo-EM studies.

## 4.2 Lysate engineering for reduced $^{15}\text{N}$ label scrambling

### 4.2.1 Identification of mutation targets

The CF expression of proteins for NMR studies is widely applied since the CF system allows more precise amino acid labeling due to truncated metabolic pathways of AA when compared to *in vivo* systems. While some internal conversions can be beneficial e. g. for relatively cheap labeling via late AA precursors, such as  $\alpha$ -ketoacids, generally stability of labeled AAs is highly desired as otherwise ambiguous NMR spectra hamper analysis and assignment of proteins. Although in CF lysates metabolic pathways of AAs are disrupted to a certain degree, some pathways are still intact causing label scrambling<sup>42,137,139,146</sup>. To tackle this problem, an array of inhibitors was proposed previously<sup>137</sup> which are also sold commercially (<https://www.cortecnet.com>). In this study the approach was to create an engineered *E. coli* A19 strain with least possible AA conversion in the lysate by cumulative gene deletion. To identify AAs most susceptible to internal conversions, GFP was expressed using AA mixtures lacking one AA, respectively (Figure 20). If GFP was still expressed in the absence of a particular AA, the AA must be metabolically provided to the reaction by enzymatic conversion. In this regard, GFP levels were a helpful measure of the degree of AA scrambling during CF expression.



**Figure 20:** Investigation of AA conversions in the CF reaction. GFP was expressed using AA mixtures lacking one AA indicated on the X-axis, respectively. Fluorescence was measured in the crude CF reaction to investigate residual GFP expression yield in the absence of the respective AA and compared to a control with all AAs included during CF expression. Expression levels were measured in independent triplicates and were standardized to the control experiment (3.5 mg/mL GFP). Error bars indicate standard deviation. Previously published figure taken from Levin et al<sup>146</sup>.

In the absence of L-Ala, L-Asp, L-Glu and L-Gln GFP levels were at standard levels or even higher than the control, which is unexpected. L-Cys, L-Gln, L-Ser, L-Tyr, L-Trp, L-Ile, L-Gly showed only minor scrambling. In contrast to a previous study, L-Asn also showed low conversion<sup>137</sup>. Subsequently, metabolic pathways of scrambling AAs were reviewed using the KEGG PATHWAY database<sup>178</sup>. The abundance and thus, likelihood of involvement in AA conversion was then estimated with the help of a proteomics study carried out previously<sup>23</sup>. Put together these tools helped to confine enzymes most likely involved in scrambling (table 4). The other, less prominent

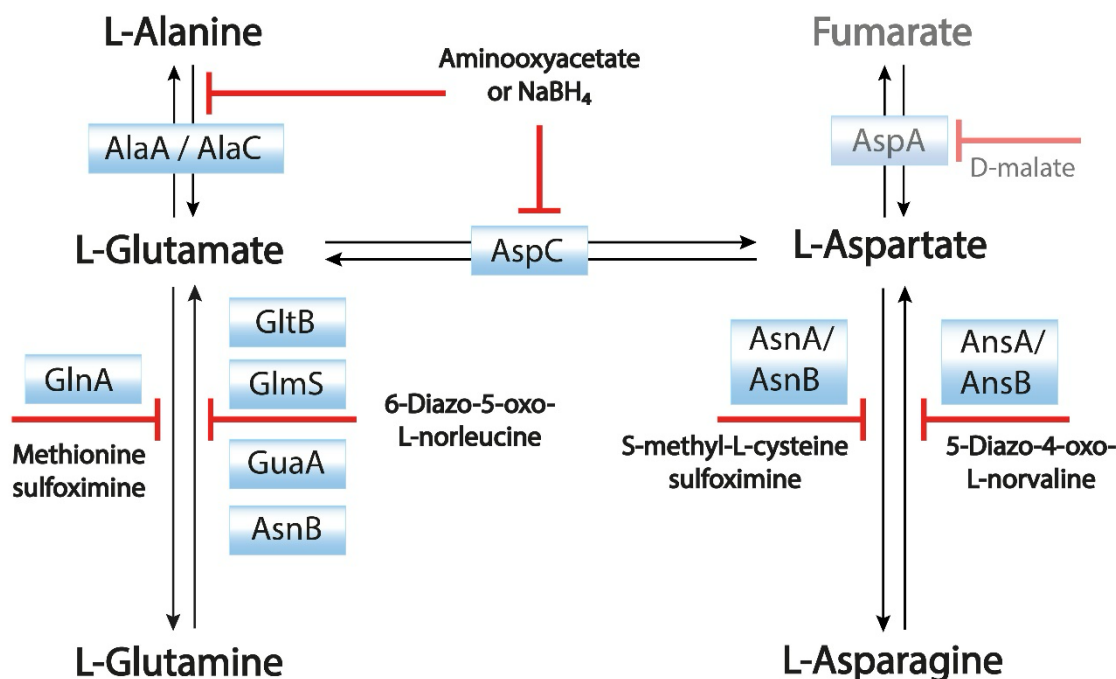
conversions observed in the GFP expression assay were attributed to an array of enzymes, especially PLP dependent enzymes such as the highly promiscuous aromatic AA aminotransferase (TyrB), glutamate-pyruvate aminotransferases (AlaA/C) or tryptophanase (TnaA)<sup>23,122,145,179</sup>. It was considered that inactivation of all genes will not be possible and that it is reasonable to inhibit PLP dependent aminotransferases by addition of unselective inhibitor AOA which irreversibly inhibits PLP-dependent enzymes<sup>180</sup>. For simplification of mutant nomenclature, target genes were assigned a number (Table 4).

**Table 4.** Enzymes with potential AA scrambling activity in standard A19 S30 lysates. Enzymes with EC number gene name are shown. The emPAI indicates the degree of the abundance of the enzymes in the CF lysate. Previously published table taken from Levin et al<sup>146</sup>.

Nr.	Enzyme	EC	Gene	emPAI <sup>a</sup>
1	Asparagine synthase	6.3.5.4	<i>asnA</i>	0.4 ± 0.2
2	Asparagine synthase (L-Gln hydrolyzing)	5.3.5.4	<i>asnB</i>	0.4 ± 0.3
3	Asparaginase	3.5.1.1	<i>ansA</i>	0.4 ± 0.2
4	Asparaginase II	3.5.1.1	<i>ansB</i>	-
5	Glutamate synthase [NADPH]	1.4.1.13	<i>gltB</i>	2.9 ± 1.1
6	Glutamine synthetase	6.3.1.2	<i>glnA</i>	6.2 ± 1.6
7	Glutamine-fructose-6-phosphate transaminase (GFAT)	2.6.1.16	<i>glmS</i>	5.9 ± 0.9
8	Aspartate aminotransferase	2.6.1.1	<i>aspC</i>	2.6 ± 0.6
9	Branched-chain-amino-acid transaminase	2.6.1.42	<i>ilvE</i>	0.4 ± 0.2
10	GMP synthetase (glutamine-hydrolyzing)	6.3.5.2	<i>guaA</i>	7.9 ± 1.0
11	Glutamate-pyruvate aminotransferase	2.6.1.2	<i>alaC</i>	0.1 ± 0.0
12	Aromatic AA aminotransferase	2.6.1.57	<i>tyrB</i>	0.1 ± 0.0

<sup>a</sup> emPAI values as determined by Foshag et al<sup>23</sup>.

Although, asparaginase II (AnsB) is an extracellular enzyme, and thus, should not be present in CF lysates, it was still included in the list since high substrate affinity was reported compared to its intracellular isoform (AnsA)<sup>181</sup>. A metabolic scheme illustrates relevant scrambling enzymes and inhibitors reported previously<sup>137</sup> (Figure 21).

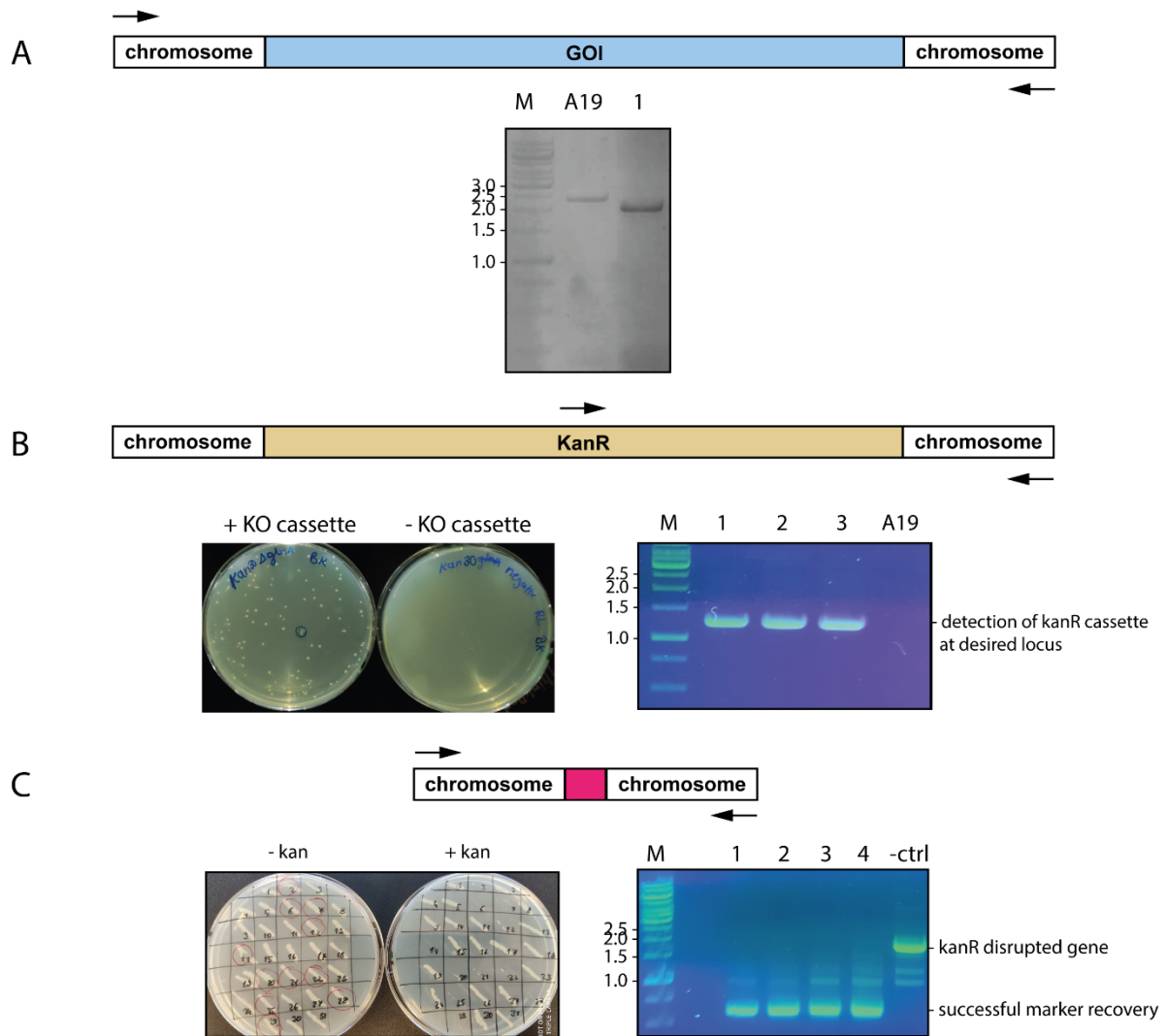


**Figure 21.** Scheme of enzymes involved in most prominent <sup>15</sup>N scrambling in CF reactions. The scheme is based on data shown in figure 20 combined with proteomics data published previously<sup>23</sup> and KEGG PATHWAY database. Previously published figure taken from Levin et al<sup>146</sup>.

#### 4.2.2 Engineering of *E. coli* A19 and characterization of strains

To perform alterations in the *E. coli* genome,  $\lambda$ -phage recombination system established by Kirill Datsenko and Barry Wanner<sup>168</sup> was applied. Briefly, the *E. coli* cells were first transformed with a plasmid containing all  $\lambda$ -phage genes required for recombination. Upon induction, the genes were expressed, and cells were then transformed with a linear DNA cassette carrying a kanamycin resistance (kanR) with up-and downstream regions of ~50 kb that are homologous to target sequences in the genome of the *E. coli* strain. Colonies on kanR plates were screened for cassette integration in two different variations of colony PCR: i. Primer's annealing ~200-500 bp up-and downstream of the target gene, respectively (Figure 22A). ii. Forward primer annealing at the kanR cassette and the reverse primer annealing 200-500 bp downstream of the gene target (Figure 22B). Mostly, the second colony PCR method was preferred as it confirmed locus specific presence of kanR and was favorable, especially when the deleted fragment had a similar size as the inserted resistance cassette. It was usually sufficient to screen 2-8 clones to confirm gene deletion. At least that was the case when growth on kanR plates was unambiguous (Figure 22B). To be able to continue mutagenesis using the same system, marker recovery was necessary. Therefore, cells were transformed with FLP plasmid carrying a flippase that removed the marker from the genome upon induction of its expression. Cells were then plated on LB agar without antibiotic and then single colonies were patched on a plate with kanamycin and a plate without. Clones that showed no growth on the plate with kanamycin were subjected to a final colony PCR with primers annealing up and downstream of the target locus. The fragments were purified and sequenced to confirm marker recovery (Figure 22C).

## Results



**Figure 22.** Screening for gene deletion mutants of *E. coli* A19 and recovery of kanR marker. **A:** Colony PCR screening with primers annealing 200-500 bp up-and downstream of the GOI, respectively. **B:** Clones obtained after successful deletion of GOI by integration of a kanR cassette. Clones transformed with linear DNA fragment (+KO cassette) containing kanR resistance flanked by 50 bp homology regions complementary to GOI regions showed growth on LB plates with 30  $\mu$ g/mL kanamycin, while non-transformed cells were not able to proliferate (-KO cassette). Colony PCR was carried out with forward primer annealing to kanR cassette and reverse primer annealing 200-500 bp downstream of the GOI. In general, this method was preferred. **C:** Screening for kanR cassette removal after expression. After transient flippase expression, candidates were patched on a plate with and without kanamycin. Clones that showed no kanR were further analyzed by colony PCR with primers annealing 200-500 bp up-and downstream of the deleted GOI. As -ctrl, colony PCR was carried out with cells carrying kanR. Lane numbers correspond to clone numbers analyzed. Marker descriptions are given in kb.

Using the  $\lambda$ -phage recombination method a set of strains was created. Selected single mutants and combinatorial mutant strains were further characterized (Table 5). Deletion mutants of *alaC* and *guaA* mentioned previously (table 4) were not possible to obtain and were thus excluded from the list of mutations. Furthermore, it was attempted to delete *tyrB*, which was not successful in strains carrying *aspC* and *ilvE* deletions. It was also tested if an introduction of a C-terminal his-tag for removal of scrambling enzymes would be a suitable alternative to gene knockout. This was tested with the gene *glnA*. However, after introduction of the kanR cassette and marker recovery, lysates showed no glutamine synthetase activity in the lysates (Figure 25). Thus, the gene was already inactivated, despite no full gene deletion was obtained.



**Table 5.** Engineered chromosomal mutants in *E. coli* A19. Table taken from Levin and colleagues<sup>146</sup>.

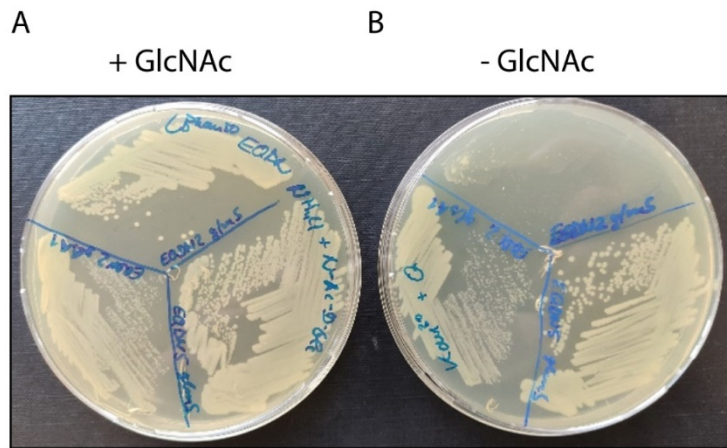
Number Name	1 <i>asnA</i>	2 <i>asnB</i>	3 <i>ansA</i>	4 <i>ansB</i>	5 <i>gltB</i>	6 <i>glnA</i>	7 <i>glmS</i>	8 <i>aspC</i>	9 <i>ilvE</i>
K 3 <sup>a</sup>			+						
K 4 <sup>a</sup>				+					
M 1 <sup>a</sup>	+								
M 5					+				
M 1-2	+	+							
M 1-2-3	+	+	+						
M 1-2-4	+	+		+					
M 1-2-4-5	+	+		+	+				
M 1-3-4-6-8	+		+	+		+		+	
M 1-2-4-5-6	+	+		+	+	+			
M 1-8	+							+	
M 1-2-5-8	+	+			+			+	
M 1-6 <sup>c</sup>	+					+			
M 1-6-7 <sup>c</sup>	+					+	+		
M 1-3-6-7 <sup>c</sup>	+		+			+	+		
M 1-3-4-6-7 <sup>c</sup>	+		+	+		+	+		
M 1-3-4-5-6-7 <sup>b, c</sup>	+		+	+	+	+	+		
M 1-3-4-6-7-8 <sup>c</sup>	+		+	+		+	+	+	
M 1-3-4-6-8-9 <sup>c</sup>	+		+	+		+		+	+
“Stablelabel”									

<sup>a</sup>, Strains designated with a K resemble single knockout strains obtained from the Keio collection strain BW25113.

<sup>b</sup>, Poor or no strain proliferation

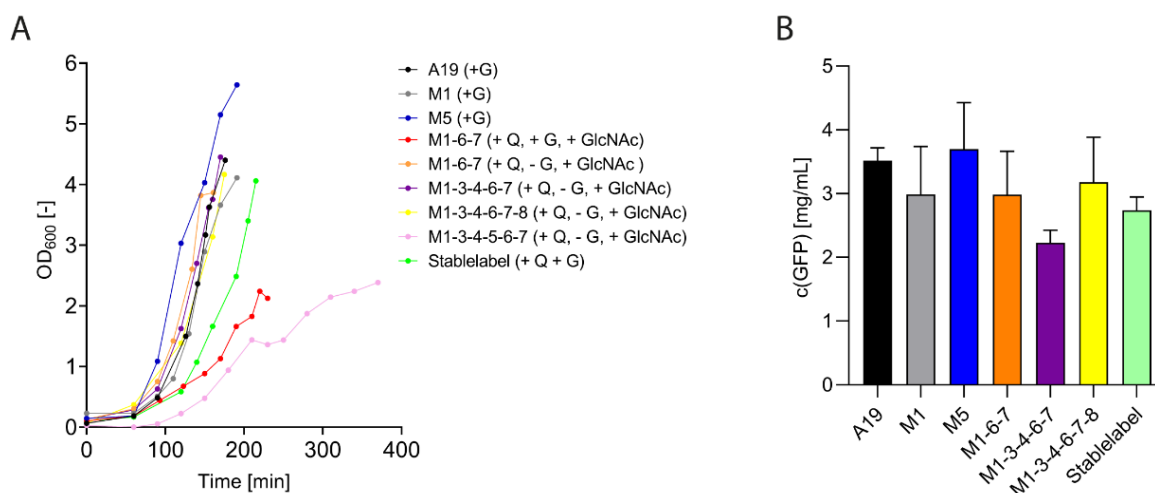
<sup>c</sup>, A *hisx6* tag was initially introduced to remove glutamine synthetase from the lysate. It was observed that introduction of *hisx6* tag was sufficient to remove glutamine synthetase activity without interfering with lysate performance

Especially when working with mutants at later engineering stages, false positive mutants were detected more frequently. Presumably, this is due to 34 bp FRT cloning scars remaining in the genome, which might have served as recombination targets. Furthermore, mutants with a high number of mutations showed poor marker recovery efficiency which might also be due to residual cloning scars occupying flippase capacity. This was the case with mutants already carrying five deletions or more. The highest mutants in this study contained 7 mutations, however marker recovery was not possible at this point. Some mutants showed no viability after deletion of certain genes. This was for example the case for cells that upon *glmS* deletion lost their ability to synthesize D-glucosamine and now required external D-glucosamine or N-acetylglucosamine (GlcNAc) media supplementation for growth (Figure 23A/B).



**Figure 23.** Auxotrophic mutations determining growth requirements of created mutants. Upper parts of LB agar plates show compartment where *glmS*<sup>-</sup> mutants of A19 were streaked out on a plate containing GlcNAc (A) and a plate without GlcNAc (B). The auxotrophic mutants showed no proliferation capability on LB agar without GlcNAc.

Furthermore, it was observed that D-glucose resulted in growth stagnation of *glmS*<sup>-</sup> mutants. This mutation side effect was not observed on LB-agar as it contained no glucose. However, the effect was strongly pronounced in bioreactor cultivations, where it appeared to be best to add 1 g/L GlcNAc and leave out glucose as otherwise growth was hampered (Figure 24A). Auxotrophy was also observed with *glnA* mutants which required L-Gln for growth and thus all culture media of *glnA*<sup>-</sup> or *glnA-his* strains were supplemented with 1 g/L L-Gln. Some combinations of mutations, e.g., *glmS* and *gltB*, both suspected to be most contributing to glutaminase activity, resulted in poor growth (Figure 24A). Generally, all strains entered log-phase around 100-120 min after inoculation of the fermenter, independent of glucose or other supplementation. Subsequently, lysates from bioreactor cultivations were produced and their productivity compared (Figure 24B). One must keep in mind that not all lysates were produced at the same time. The evolution of the mutants followed the results of NMR experiments and GFP expression assays in which scrambling activity was analyzed (Figure 25/26). The final strain carrying a set of most beneficial mutations was M1-3-4-6-8-9, designated “Stablelabel”. It showed a slightly delayed log phase (Figure 24A) and a slightly decreased productivity of ~2.7 mg/mL GFP compared to A19 with ~3.5 mg/mL (Figure 24B).



**Figure 24.** Growth kinetics and CF protein production performance of selected engineered *E. coli* A19 lysates. **A:** Growth curves of *E. coli* A19 and mutants cultivated in a 15 L stirred-tank bioreactor in 10 L YPTG medium at 37 °C, 500 rpm. 1 g/L L-Gln was supplied to the culture medium to restore growth of *glnA* mutants (+Q). Standard YPTG culture medium contained 0.1 M glucose (+G), except for some *glmS* mutants such as M 1-3-4-5-6-7 (-G) where no glucose was added due to growth problems. Furthermore, *glmS* mutant cultures were supplied with additional 1 g/L GlcNAc (+GlcNAc). **B:** Extract performance was evaluated by expression of GFP under standard conditions. Quantification was carried out via GFP fluorescence measurement in the CF reaction mixture. Bars represent means of expression yields obtained from at least three independent CF reactions. Previously published figure taken from Levin et al<sup>146</sup>.

#### 4.2.3 NMR analysis of CypD expressed with engineered lysates

M1-3-4-6-8-9 “Stablelabel” was defined as the final strain for NMR experiments as it combined mutations eliminating L-Asn to L-Asp (*ansA/B*<sup>-</sup>) and L-Asp to L-Asn (*asnA*<sup>-</sup>) scrambling, as well as L-Glu conversions to L-Asp (*aspC*<sup>-</sup>) and L-Gln (*glnA-his*). Furthermore, branched chain amino acid transferase (BCAAT) was expected to have stabilizing effects on methyl group labeling of L-Val and L-Leu (*ilvE*<sup>-</sup>). A previously assigned peptidyl-prolyl cis-trans-isomerase CypD variant was used as model protein for <sup>15</sup>N label scrambling studies of L-Asp, L-Asn, L-Glu, L-Gln and L-Ala (Figure 25). CypD was labeled with <sup>15</sup>N L-Asp, L-Asn, L-Glu or L-Gln using either A19 lysate (red spectra) or Stablelabel lysate (blue spectra).

Compared to A19, L-Asp labeling showed only the 9 expected L-Asp peaks while A19 showed 21 peaks which corresponded to L-Glu, L-Gln and L-Ala signals (Figure 25A). L-Ala is synthesized by amino group transfer from L-Glu to pyruvate catalyzed by AlaA/C aminotransferases. Since the pathway from L-Asp to L-Glu is disrupted by *aspC* deletion in Stablelabel, <sup>15</sup>N labeled amino group is not transferred to L-Glu to be passed on to pyruvate in a subsequent reaction.

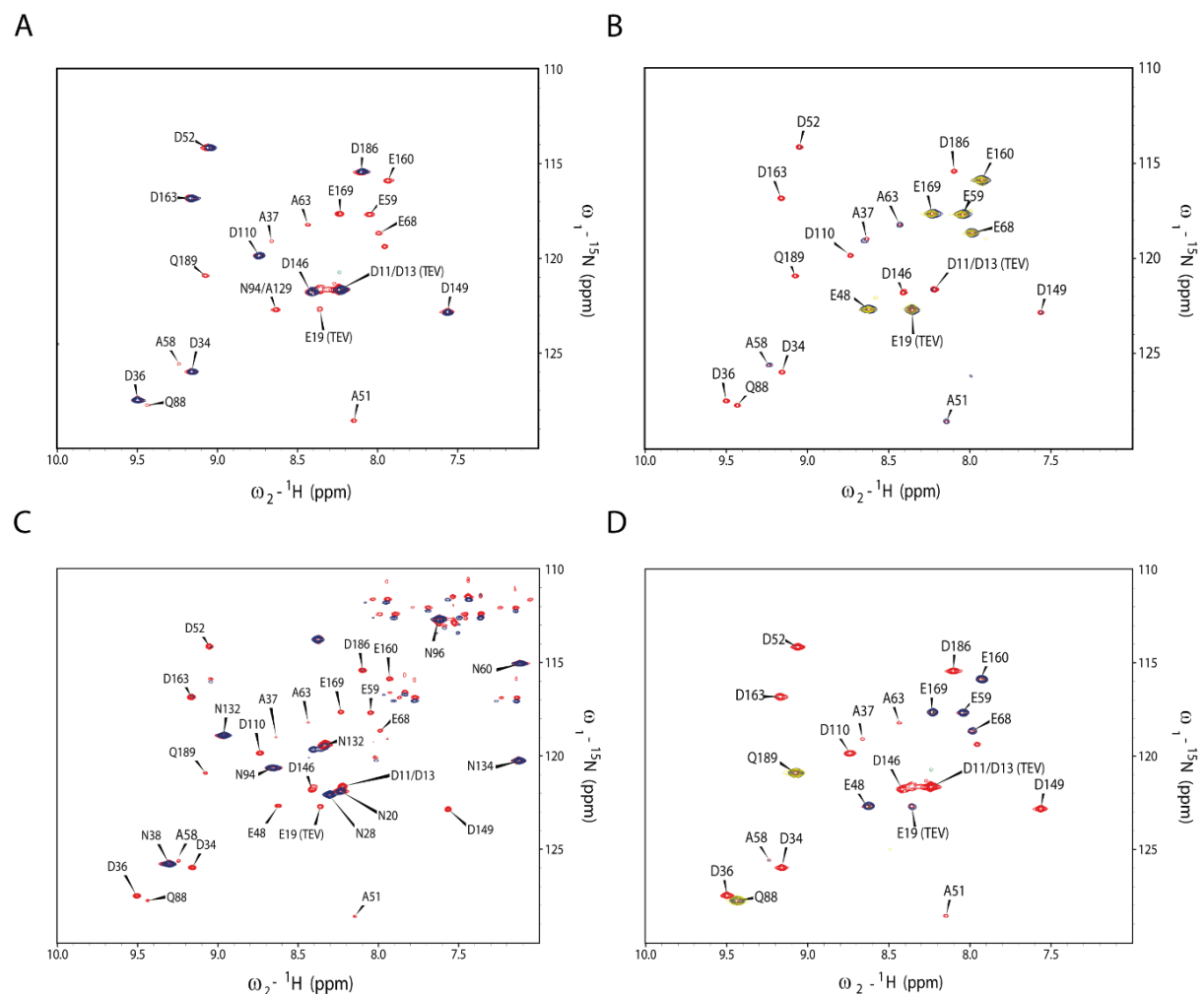
When labeling CypD with <sup>15</sup>N L-Glu with A19 (Figure 25B), the same 21 signals were observed as in Figure 25A, showing L-Glu scrambling to L-Gln, L-Asp and L-Ala. When performing the same experiment with Stablelabel, only L-Glu and weak L-Ala signals were detected. This confirms that *in vivo* C-terminal his-tagging of GlnA completely inactivated either the gene by introduction of a cloning scar in the terminator sequence or a his-tag at the C-terminus (Figure 25B). To suppress AlaC transamination, expression was carried out with Stablelabel and additional broadband transaminase

## Results

inhibitor AOA. As a result, L-Ala signals disappeared in the yellow spectrum (Figure 25B).

$^{15}\text{N}$  L-Asn labeling using A19 resulted in a similar scrambling pattern, only this time L-Asn was converted to L-Asp first, secondarily to L-Glu and lastly to L-Gln and L-Ala (Figure 25C).  $^{15}\text{N}$  labeled side chains of L-Asn further contributed a typical pattern of weak signals to the spectrum. Expression of CypD with Stablelabel showed only L-Asn signals as deletion of asparaginases *ansA/B* inhibited conversion to L-Asp and thus, all secondary conversions (Figure 25C).

$^{15}\text{N}$  L-Gln labeling using A19 resulted in labeled L-Glu, L-Asp and L-Ala, due to conversions mentioned above. Stablelabel showed only scrambling to L-Glu, no scrambling to L-Asp due to *aspC* deletion, and interestingly far weaker L-Glu transamination to L-Ala (Figure 25D). To inhibit residual glutaminase activity, Stablelabel lysate was treated with 5 mM DON, a covalently binding glutaminase inhibitor. As a result, only L-Gln peaks were observed (Figure 25D, yellow spectrum).

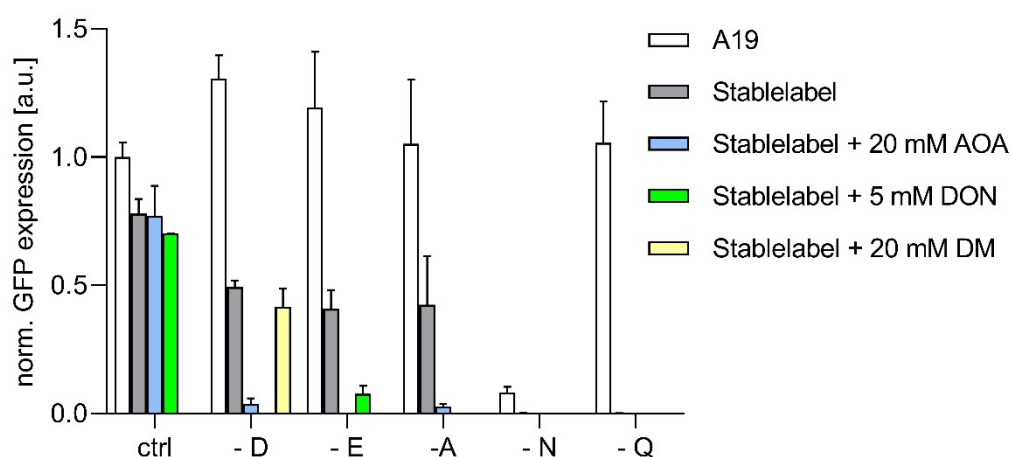


**Figure 25:** Selective labeling of the previously fully assigned CypD (43-207) variant. Spectra overlay depicts different outcomes of CypD selective labeling with  $^{15}\text{N}$  L-Asp (**A**),  $^{15}\text{N}$  L-Glu (**B**),  $^{15}\text{N}$  L-Asn (**C**) or  $^{15}\text{N}$  L-Gln (**D**) using either *E. coli* A19 (red) or engineered strain M1-3-4-6-9-10 “Stablelabel” (blue) for CypD expression. The yellow spectrum in **B** was obtained after addition of 20 mM AOA in the CF reaction to suppress L-Glu to L-Ala label conversion. The yellow spectrum in **D** was obtained after incubation of CF lysate with 5 mM DON at RT for 1 h before CypD expression to suppress L-Gln to L-Glu conversion. [ $^{15}\text{N}$ ,  $^1\text{H}$ ] BEST-TROSY spectra were measured on CypD samples concentrated to 100-150  $\mu\text{M}$  in Na phosphate buffer pH 7.0 at 303 K. All A19 spectra were recorded on 900 or 950 MHz

spectrometers while Stablelabel spectra were recorded on 600 or 700 MHz spectrometers. Previously published figure taken from Levin et al<sup>146</sup>.

#### 4.2.4 Analysis of AA scrambling in A19 Stablelabel by GFP expression

Despite NMR studies gave a clear perspective on amino acids most affected by metabolic conversion, stabilized *E. coli* lysates can be beneficial for other applications, such as quantification of amino acids in body fluids<sup>35</sup>. Therefore, GFP expression was now carried out with Stablelabel in the absence of either L-Asp, L-Asn, L-Glu, L-Gln or L-Ala, respectively, and compared with A19 (Figure 26). Prominently, no GFP expression was observed in absence of L-Gln which is due to eliminated GlnA activity (GlnA-his). Also, weak GFP expression observed with A19 in absence of L-Asn vanished if GFP expression was carried out with Stablelabel. Interestingly, GFP expression in absence of L-Asp was still ~50 % of the A19 control, despite main metabolic routes of L-Asp supply from L-Glu, or L-Asn were eliminated by *aspC* and *ansA/B* deletion. To investigate if L-Asp is supplied from fumarate, D-malate (DM) was included in the CF reaction, to suppress activity of aspartate ammonia lyase (AspA) which did not reduce GFP levels. Finally, AOA was included in the reaction to test if conversion is to be attributed to an aminotransferase, and indeed, AOA almost eliminated GFP expression. This indicates that transamination was due to aromatic AA aminotransferase TyrB, which has a broad substrate spectrum including L-His, L-Asp, L-Glu, L-Val, L-Ile, L-Leu, L-Cys and L-Ala<sup>182</sup>. This is also in agreement with aminotransferase scrambling to L-Ala possibly catalyzed not only by AlaA/C but also by TyrB. Treatment of Stablelabel lysate with 5 mM DON showed that most conversion to L-Glu is due to glutaminase activity since GFP level decreased by ~90 % (Figure 26).



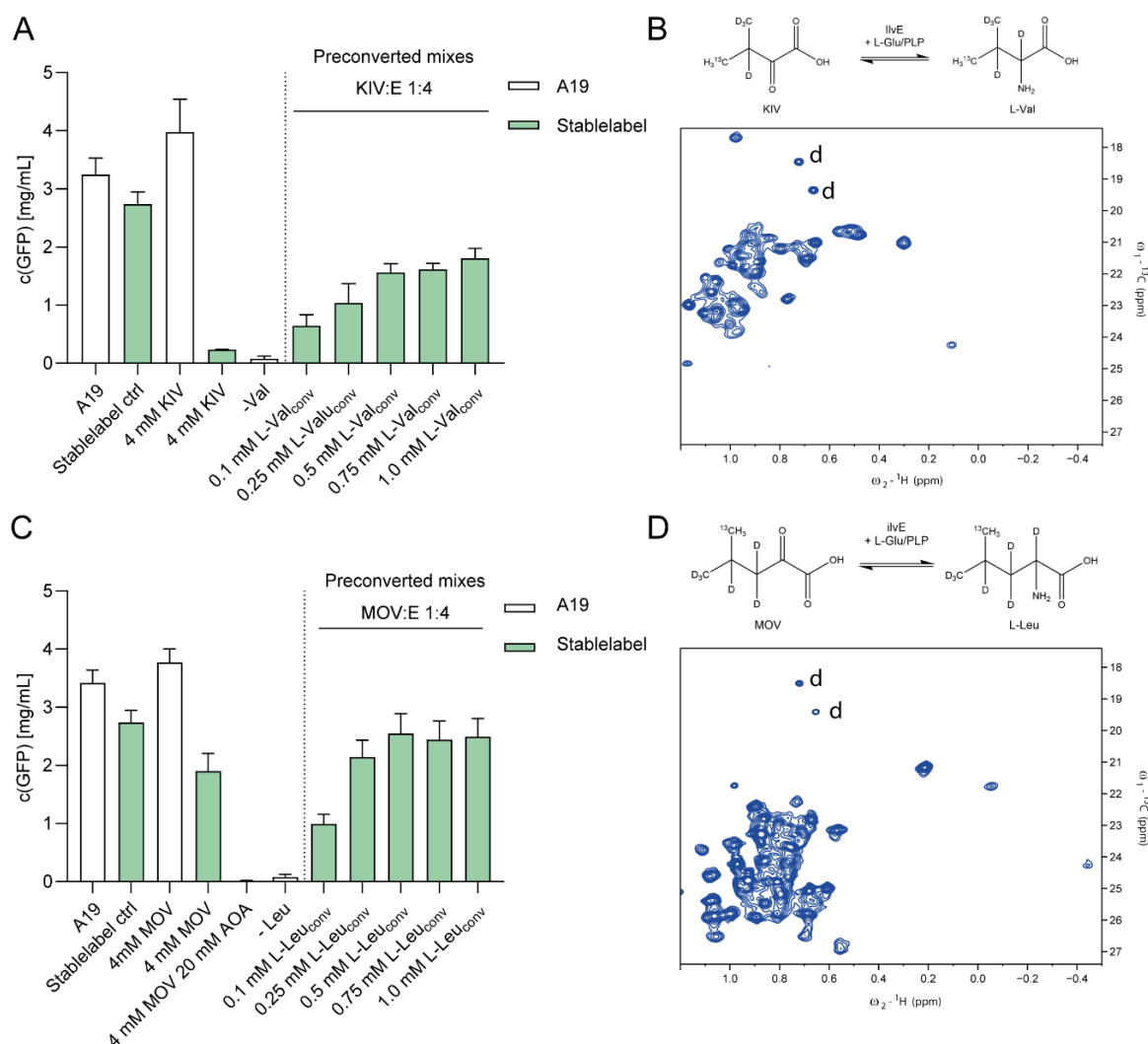
**Figure 26:** Expression of GFP with A19 (white bars) and M1-3-4-6-8-9 “Stablelabel” (grey bars) lysates in absence of L-Asp, L-Asn, L-Glu, L-Gln or L-Ala, respectively. Inhibitors AOA (blue bars), DON (green bars) and DM (yellow bar) were applied to test their influence on L-Asp and L-Ala conversions using Stablelabel lysate. Quantification was carried out via GFP fluorescence measurement in the CF reaction mixture. Bars represent means of expression yields standardized to A19 ctrl GFP yield of 3.5 mg/mL. Median values were obtained from fluorescence measurements of at least three independent CF reactions. Previously published figure taken from Levin et al<sup>146</sup>.

#### 4.2.5 Precursor preconversion for methyl labeling of BCAA

As another application, Stablelabel was optimized for methyl labeling by stabilization of BCAA L-Val and L-Leu by deletion of *IlvE* (gene: *ilvE*). Methyl labeling is beneficial for membrane protein labeling, as methyl groups are distributed across all helices and are mainly clustered in the protein core, where ligand interaction is likely. First,  $\alpha$ -ketoacids KIV (4- $^{13}\text{C}$ , 3, 5, 5, 5-D4) and MOV (5- $^{13}\text{C}$ , 3, 3, 4, 5, 5, 5-D6), designated as KIV and MOV, respectively, were converted to their respective amino acids L-Val (4- $^{13}\text{C}$ , 2, 3, 5, 5, 5-D5) and L-Leu (5- $^{13}\text{C}$ , 2, 3, 3, 4, 5, 5, 5-D7) using recombinantly expressed *IlvE*. Introduction of protons at C $\alpha$  position in the deuterated amino acids was avoided by carrying out enzymatic transamination in ~90 % deuterated environment. Initially, GFP expression trials were carried out to investigate the effect of *IlvE* deletion and get an estimate how much preconverted KIV/MOV is required to obtain sufficient expression levels (Figure 27A/C). Expression levels of A19 and Stablelabel were compared in the first two bars, in the next two bars L-Val (Figure 27A) and L-Leu (Figure 27C) were left out of the reaction and replaced by 4 mM KIV, respectively. This provided information on residual internal conversions of KIV/MOV to L-Val/L-Leu in the lysates. Internal KIV conversion completely disappeared in Stablelabel, while A19 conversion was strong enough to reach GFP expression above basal level (Figure 27A). The same was observed with A19 expressing GFP with no L-Leu but with 4 mM MOV in the reaction. Interestingly, MOV this time was also internally converted to L-Leu and only addition of 20 mM AOA prohibited the internal conversion, again indicating other transaminases than *IlvE* being involved, such as potentially TyrB (Figure 27C). Consequently, expressions with preconverted L-Leu were carried out in presence of 20 mM AOA. As presence of KIV and MOV were prerequisite to these internal conversions, negative controls were included, where L-Val/L-Leu (Figure 27A/C) were omitted from the reaction and not replaced by their respective precursor. In both cases GFP expression was abolished. The next part of the respective bar diagrams shows titration of the preconverted 1:4 KIV/L-Val (Figure 27A) and MOV/L-Leu mixes (Figure 27C) to the reaction. Concentrations of converted mixes are given in mM and represent the theoretical amount of L-Val or L-Leu at 100 % precursor conversion. GFP expression levels plateaued at around 0.5 mM theoretical L-Val/L-Leu preconverted from KIV/MOV.

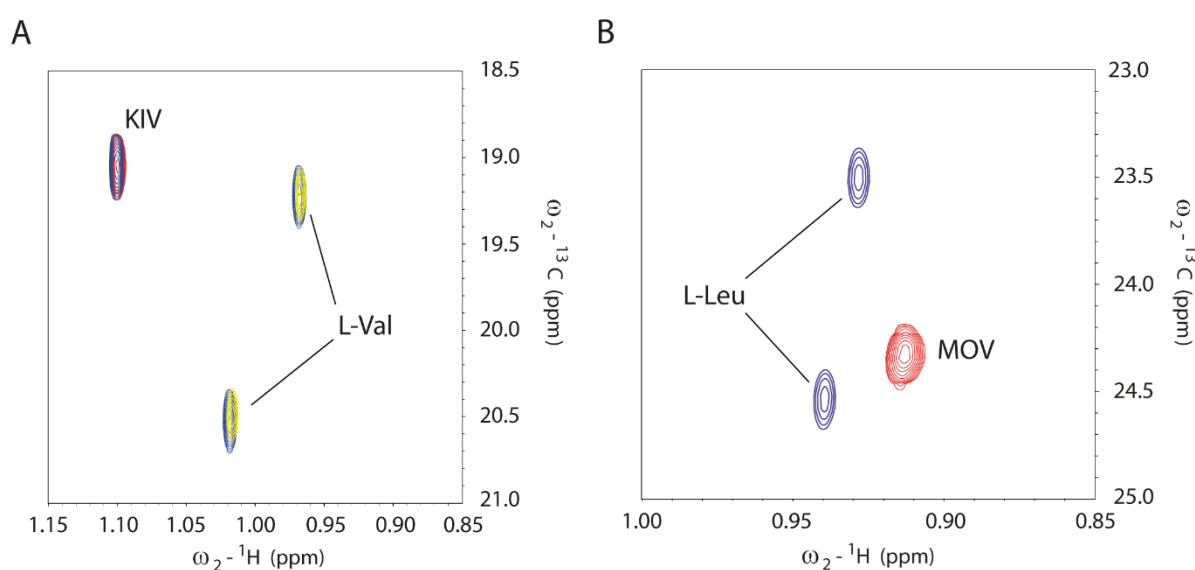
To investigate applicability of the method for MP labeling, his-tagged PR was expressed in GDN/DH7PC in presence of 0.5 mM mixtures of L-Val or L-Leu preconverted from their respective precursors with the rest of amino acids being unlabeled and non-deuterated. After purification, PR samples of 80-100  $\mu\text{M}$  were obtained which was sufficient to record [ $^{13}\text{C}$ ,  $^1\text{H}$ ] XL-ALSOFAST-HMQC spectra of L-Val (4- $^{13}\text{C}$ , 2, 3, 5, 5, 5-D5) and L-Leu (5- $^{13}\text{C}$ , 2, 3, 3, 4, 5, 5, 5-D7) labeled PR (Figure 27B/D). Due to non-stereospecific incorporation of the pro-S or the pro-R [ $^{13}\text{C}$ ,  $^1\text{H}$ ] methyl group in L-Val (4- $^{13}\text{C}$ , 2, 3, 4, 4, 4,-D5), ~37 detected peaks are in the expected range of 20 L-Val (4- $^{13}\text{C}$ , 2, 3, 4, 4, 4,-D5) residues (Figure 27B). The PR spectrum with preconverted L-Val (4- $^{13}\text{C}$ , 2, 3, 5, 5, 5-D5) is in accordance with data previously acquired with KIV (4- $^{13}\text{C}$ , 3, 5, 5, 5-D4) conversion directly in the CF reaction with A19 S30 lysates containing a combination of residual and additional recombinantly expressed *IlvE*<sup>166</sup>. PR labeling with preconverted L-Leu (5- $^{13}\text{C}$ , 2, 3, 3, 4, 5, 5, 5-D7) resulted in a similar picture as KIV. PR contains 27 L-Leu residues and around 40 signals are visible, which however, is difficult to estimate due to strong signal overlap and broad lines. Again, the larger number of signals is due to non-stereospecific methyl groups of the precursors. Detergent signals were visible in both spectra, which showed

no signal overlap indicating no label scrambling of KIV or MOV by IlvE. In general, both spectra yielded broad lines which is due to the use of non-deuterated amino acids apart from L-Val and L-Leu.



**Figure 27.** Preconversion of methyl labeled precursor KIV to L-Val and MOV to L-Leu. GFP expression was carried out either with A19 (white bars) or M 1-3-4-6-8-9 “Stablelabel” (green bars). **A/C:** The first two bars indicate control expressions of both strains where 1 mM L-Val/L-Leu was included in the reaction. The next two bars show GFP expression levels with L-Val/L-Leu subtracted from the reaction and instead 4 mM KIV/MOV was supplied. For MOV the same reaction was carried out in the presence of 20 mM AOA, an unselective transaminase inhibitor. The next bar (-Val/-Leu) indicates expression with A19 lysate without L-Val/L-Ile. The next section of the graph shows GFP expression levels with L-Val/L-Leu subtracted and instead preconverted KIV/L-Glu or MOV/L-Glu mixture supplied. L-Glu was present in 4-fold molar excess of the precursor during preconversion. The molarities represent the theoretical L-Val/L-Leu concentration if KIV/MOV was converted to the respective amino acid completely. Preconversion of both precursors was carried out in ~90 % deuterated buffer to avoid proton incorporation at C $\alpha$  position of the ketoacid during transamination by IlvE. **B:** [<sup>13</sup>C, <sup>1</sup>H] XL-ALSOFAST-HMQC spectrum of proteorhodopsin in non-deuterated DH7PC labeled with L-Val (4-<sup>13</sup>C, 2, 3, 5, 5, 5-D<sub>5</sub>) preconverted from KIV (4-<sup>13</sup>C, 3, 5, 5, 5-D<sub>4</sub>). **D:** [<sup>13</sup>C, <sup>1</sup>H] XL-ALSOFAST-HMQC spectrum of proteorhodopsin in non-deuterated DH7PC labeled with L-leu (5-<sup>13</sup>C, 2, 3, 3, 4, 6, 6, 6-D<sub>8</sub>) preconverted from MOV (5-methyl <sup>13</sup>C, 3, 3, 4, 6, 6, 6-D<sub>7</sub>). Detergent signals in all spectra are labeled with a “d”. Spectra were acquired on a 950 MHz Bruker Avance Neo spectrometer at 313 K sample temperature. Samples contained 5 % D<sub>2</sub>O and 0.15 mM DSS, 1x cComplete protease inhibitor and 100  $\mu$ g/mL streptomycin. Previously published figure taken from Levin et al<sup>146</sup>.

Regarding preconversion, it was found that it is important to use a higher ratio of L-Glu during preconversion of KIV and MOV (Figure 28). In this study a ratio of 4:1 was used however for increased preconversion efficiency an even higher L-Glu to  $\alpha$ -ketoacid precursor ratio can be applied. 1:1 and 1:4 KIV:L-Glu mixtures were analyzed by recording  $[^{13}\text{C}, ^1\text{H}]$ -HSQC spectra after conversion and compared to non-converted KIV ( $4\text{-}^{13}\text{C}, 3, 5, 5, 5\text{-D}_4$ ) (Figure 28A). Using a KIV:L-Glu ratio of 1:1 already converted ~50 % of the precursor to L-Val. If the ratio was increased to 1:4 KIV:L-Glu during preconversion, >90 % of the precursor were converted to L-Val (Figure 28A, yellow spectrum). Due to scarce availability of the MOV ( $5\text{-}^{13}\text{C}, 3, 3, 4, 5, 5, 5\text{-D}_6$ ), only preconversion at 1:4 ratio MOV:L-Glu was carried out yielding similar results as KIV preconversion (Figure 28B, blue spectrum).



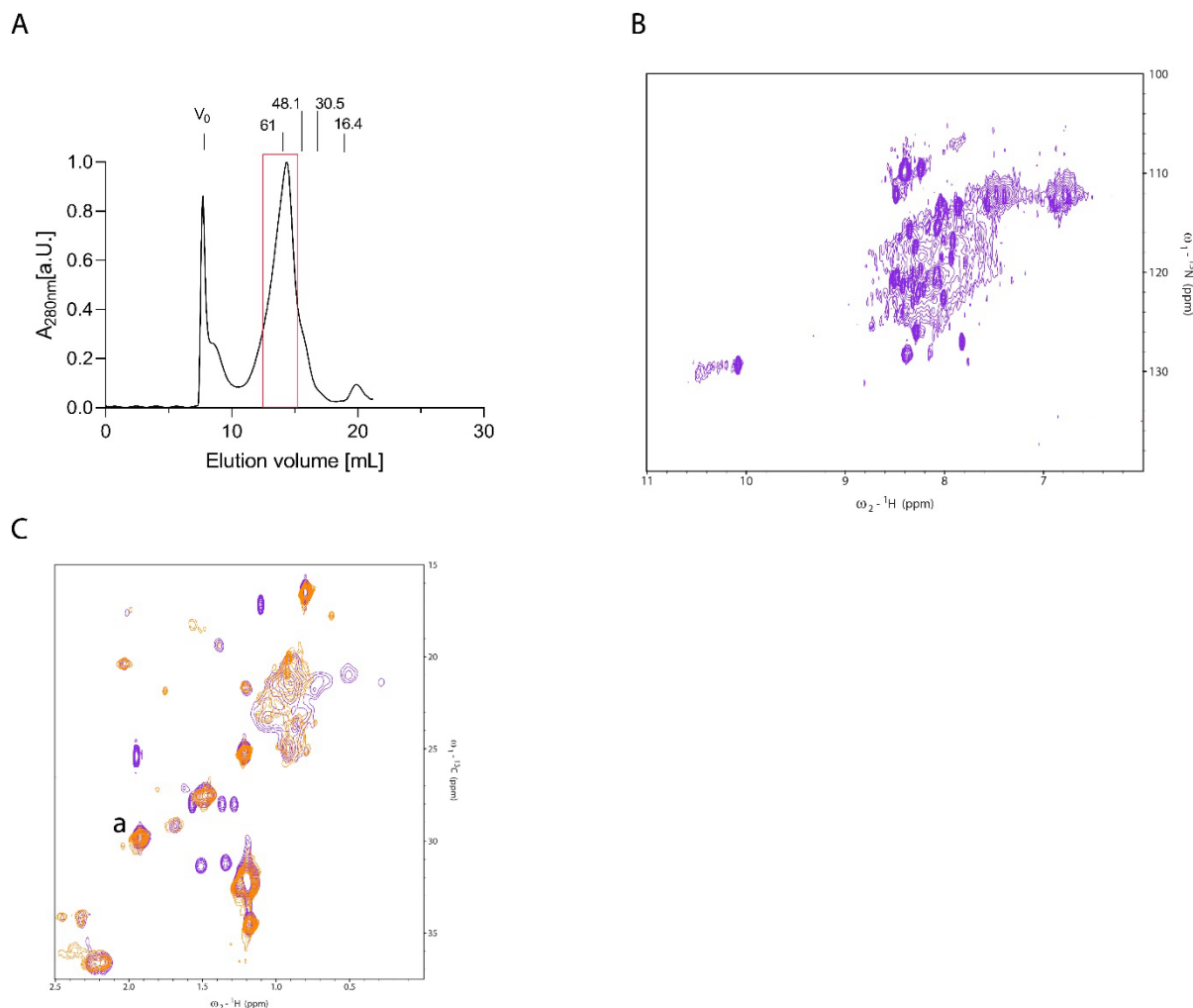
**Figure 28.** Conversion of BCAA precursors prior to labeling. Conversion of KIV ( $4\text{-}^{13}\text{C}, 3, 5, 5, 5\text{-D}_4$ ) to L-Val ( $4\text{-}^{13}\text{C}, 2, 3, 5, 5, 5\text{-D}_5$ ) and MOV ( $5\text{-}^{13}\text{C}, 3, 3, 4, 5, 5, 5\text{-D}_6$ ) to L-Leu ( $5\text{-}^{13}\text{C}, 2, 3, 3, 4, 5, 5, 5\text{-D}_7$ ) was carried out by recombinantly expressed IlvE. **A:** Overlay of non-converted KIV (red), KIV after conversion with 75 mM KIV and 75 mM L-Glu (blue) and KIV after conversion with 75 mM KIV and 300 mM L-Glu (yellow). **B:** Overlay of non-converted MOV (red) and MOV after conversion with 24 mM MOV and 100 mM L-Glu (blue).  $[^{13}\text{C}, ^1\text{H}]$ -HSQC spectra were recorded at 313 K sample temperature on a Bruker Avance II 500 MHz spectrometer, equipped with a room-temperature  $^1\text{H}\{^{13}\text{C}/^{15}\text{N}\}$  three-axes gradient triple-resonance probe. Precursor conversion was performed in 100 mM Tris-HCl, pH 8.0, 100 mM NaCl dissolved in  $\text{D}_2\text{O}$  and 1 mg/mL IlvE. Upon completion of conversion, IlvE was removed by ultrafiltration and the mixtures were diluted 1:50 in 25 mM Na-acetate, pH 5.0 containing 5 %  $\text{D}_2\text{O}$  and 0.15 mM DSS. Non converted precursors were measured under the same conditions. Previously published figure taken from Levin et al<sup>146</sup>.

Precursor conversion appears to be an efficient way for labeling CF expressed MPs, which is cheaper than similar labeling strategies in live cells. Furthermore, the open nature of the CF system allowed to avoid potential MOV label dilution and saturation of the preconversion with L-Glu will even further increase conversion efficiency. Stablelabel does not only display decreased  $^{15}\text{N}$ -label scrambling but has also stabilizing effects on the precursors.



#### 4.2.6 First NMR attempts of CF synthesized GPCR in NDs

The CF system can produce up to 1 mg/mL GPCR of which up to 40 % active GPCR probes without addition of ligands or G-proteins. Since SEC analysis is usually a good first indicator of sample quality, U-D/ $^{15}\text{N}$  and L-Val ( $4\text{-}^{13}\text{C}$ , 2, 3, 5, 5, 5-D5) labeled FFAR<sub>2</sub> was expressed in MSP1D1ΔH5 NDs, affinity purified via strep tag, subsequently via SEC (Figure 29A) and finally analyzed by NMR (Figure 29B/C).



**Figure 29.** Preliminary NMR data of FFAR<sub>2</sub> and PR in MSP1D1ΔH5 NDs (POPG). U-D/ $^{15}\text{N}$  and L-Val ( $4\text{-}^{13}\text{C}$ , 2, 3, 5, 5, 5-D5) labeled FFAR<sub>2</sub> and L-Val ( $4\text{-}^{13}\text{C}$ , 2, 3, 5, 5, 5-D5) labeled PR were CF expressed in presence of 160  $\mu\text{M}$  POPG NDs at 5 ng/ $\mu\text{L}$  template concentration in 2.5 mL RM and 18 mL FM. **A:** SEC purified fraction of FFAR<sub>2</sub> (red square) was used for NMR measurement. **B:** [ $^{15}\text{N}$ ,  $^1\text{H}$ ]-IPAP-sfHMQC of FFAR<sub>2</sub>. **C:** [ $^{13}\text{C}$ / $^1\text{H}$ ]-ALSOFAST-HMQC of FFAR<sub>2</sub> (orange spectrum) and PR (violet spectrum). NMR measurements were carried out with  $\sim 60$   $\mu\text{M}$  FFAR in 40 mM Bis-Tris, 40 mM NaCl, pH 6.5. PR was concentrated to 200  $\mu\text{M}$  in 20 mM NaOAc, pH 5.0. Both samples contained 5 % D<sub>2</sub>O and 0.15 mM DSS, 1x cComplete protease inhibitor, 100  $\mu\text{g}/\text{mL}$  streptomycin and FFAR<sub>2</sub> additionally contained 100  $\mu\text{M}$  high affinity propionate antagonist TUG1609. Spectra were acquired on a 950 MHz Bruker Avance Neo spectrometer at 308 K temperature for [ $^{15}\text{N}$ ,  $^1\text{H}$ ]-IPAP-sfHMQC (**B**) and 298 K or 313 K for [ $^{13}\text{C}$ ,  $^1\text{H}$ ]-ALSOFAST-HMQC of FFAR<sub>2</sub> or PR (**C**).

Despite a SEC profile that showed a distinct peak of particles with  $\sim 11$  nm diameter (Figure 29A), neither U-D/ $^{15}\text{N}$  (Figure 29B) nor methyl labeling of L-Val (Figure 29C) provided good quality spectra. This is surprising as even PR showed poorly defined methyl signals of L-Val. Even though analysis of complexes as large as FFAR<sub>2</sub>/ND with

~140 kDa is highly challenging due to reasons mentioned above (1.3.1), one would expect at least better resolved methyl group signals of L-Val. The [<sup>15</sup>N, <sup>1</sup>H]-IPAP-sfHMQC spectrum of FFAR<sub>2</sub> is overall not well resolved despite some signals with higher intensity. These signals are to be attributed to mobile residues in loops or at the termini regions of the protein (Figure 29B). In large complexes, regions with higher rigidity and possible conformational exchange can further broaden signals. Sample inhomogeneity might be another reason for poorly resolved signals, despite a distinct SEC peak. Possibly, affinity chromatography with subsequent SEC is not sufficient to obtain a homogeneous sample (Figure 29A).

Regarding PR it is possible that higher oligomeric complexes, which were observed up to PR tetramers in MSP1D1ΔH5 NDs, lead to undefined signals due to size increase of the particle to as large as ~200 kD. [<sup>13</sup>C/<sup>1</sup>H]-ALSOFAST-HMQC spectra of PR and FFAR<sub>2</sub> show overlapping carbon signals. The signal at 1.95/29 ppm corresponds to acetate, while the other overlapping signals are most likely POPG signals. Again, labeling of L-Val, mostly present in rigid helices might have resulted in broader signals in large complexes, as well as the non-stereospecific labeling of methyl groups might have played a role. It is also to be noted that proton abundance in non-deuterated lipid and MSP can further worsen spectral quality in both samples. For NMR structure of PR, deuterated DH7PC was used. Also, PR was analyzed in its monomeric state as no oligomerization was observed in DH7PC<sup>183</sup>. Conclusively, the outcome of future NMR experiments with CF expressed GPCR in NDs is strongly dependent on the target, and it is unlikely that SEC alone will provide homogeneous particles suitable for NMR analysis. Ligand or cofactor affinity chromatography will provide better chances to obtain good NMR probes. This will strongly depend on availability of ligands and cofactors, but also on a reasonably high yield of the CF-expressed target.

## 5. Discussion

### 5.1 Introducing SapNPs in the bacterial CF system

NPs composed of defined lipid bilayers embedded in amphipathic scaffolds provide a valuable tool for protein reconstitution and cotranslational protein solubilization using CF systems. While MSP based NDs were widely applied in combination with the CF system, other scaffolds such as SMALPs<sup>184</sup>, peptidiscs (unpublished) and SapA<sup>146</sup> were employed only by our lab so far. To expand the toolbox of NPs inside a CF reaction, an extensive study for SapNPs was carried out. Starting the study with the expression of SapA, cultivation in a fermenter was used right from the start, as it has proven very useful to obtain 15-30 mg MSP per liter culture medium. Bearing in mind that 1 mg of purified MSP is commercially priced at ~100 € (<https://www.sigmaaldrich.com>), this expression strategy at first glance seems to be providing the experimenter with a sheer unexploitable amount of MSP. However, one must keep in mind that in CF systems nanodiscs are usually employed at concentrations  $\geq 60 \mu\text{M}$  (120  $\mu\text{M}$  MSP). This puts cost calculation in perspective, especially when working with small scaffolds that form small NDs such as MSP1D1 $\Delta$ H5 or SapA. Small NPs provide a smaller lipid area, and thus must be supplied at higher concentrations than large MSP1D1E3 NDs. Already preliminary experiments showed that SapNPs are not suitable for MP solubilization at concentrations below 200  $\mu\text{M}$  SapA resulting in aggregated SapNP/MP complexes. These factors considered; good scaffold expression is a necessity. Expression of SapA in a fermenter with induction at OD  $\geq 7$  yielded SapA amounts similar to MSP1D1E3, however, SEC analysis showed that around 50 % of the purified protein already formed particles with *E. coli* lipid. SapA is a protein with a high lipid binding capacity due to its physiological roles<sup>73</sup>. Apparently, disrupted *E. coli* provided SapA with sufficient lipids to produce particles spontaneously. Although this finding might be particularly interesting to investigate prokaryotic proteins in SapNP complexes with native *E. coli* lipids, our interest was rather to have defined lipid compositions in SapNPs. Treatment with Triton X-100 and reducing agent DTT was found suitable to prevent spontaneous particle formation. Although SapA contains three disulfide bonds<sup>185</sup>, its structural integrity seemed to withstand reducing conditions during CF expression. Thus, the model protein PR was expressed at considerable amounts already in the preliminary experiments with spontaneously formed SapNPs (EC).

To efficiently obtain SapNPs with defined lipid composition, it was first considered to take the classical route of detergent assisted particle reconstitution. However, this approach was problematic due to tedious and extensive use of SM-2 bio-beads and often resulted in heterogeneous NP populations. To overcome this problem, it was considered to exploit the capability of SapA to bind lipids at lysosomal pH of ~5 which instantly cleared turbid SapA/lipid mixtures and produced SapA/lipid complexes very efficiently. Titrations of SapA with DOPG showed that the NP size can be modulated and peaked around a diameter of 12-13 nm, similar to MSP1D1E3 based NDs. However, even at this diameter, SapNPs should contain only up to ~65 % of the lipid carriage of MSP1D1E3 NDs, since the highest preforming SapA to DOPG ratio was 1:24. Considering that NPs consist of 2-4 SapA entities, the complexed membrane would contain 50-100 DOPG molecules, in comparison to ~160 DOPG molecules in corresponding MSP1D1E3 NDs<sup>70</sup>. This already explains the higher demand of SapNPs for efficient cotranslational MP solubilization. Fortunately, the preassembly of SapNPs proved to be very simple and time saving. After mixing SapA with lipids, the pH was

adjusted to 4.8-5.0, then diluted with slightly basic buffer, and concentrated again to reach the desired NP concentration.

This established strategy was suitable to solubilize PR, EmrE and SugE, but also to obtain functional GPCR probes such as turkey and human isoforms of  $\beta_1$ AR. It was also demonstrated by our group that GFP tagged PR in SapNPs is transferrable into live cell membranes just as efficiently as in MSP NDs<sup>87</sup>. However, SapNPs were not suitable to solubilize all targets that are compatible with MSP NDs. It is likely that the overall scaffold stability is crucial as the particle undergoes mechanical stress during MP insertion. While the MSP scaffold forms a continuous ring with kinks at L-Pro residues, SapNPs have more gaps between the SapA units. That potentially increases likelihood of particle disintegration. Furthermore, the overall smaller size of SapNPs might be problematic for efficient insertion and folding of some MPs. It also seems that long N-termini as well as long loop regions of MPs cause problems during insertion. These flexible regions are likely to support disintegration of the NPs. Similar problems were already observed by our group, albeit to lesser extent, with MSP NDs.

PR is a commonly used model protein by our group. With seven TMDs and its oligomerization capability it is a good first choice when testing or optimizing NP systems<sup>84,85</sup>. SapNPs successfully accommodated pentameric and hexameric PR still showing good particle dispersion according to SEC. Also, during coassembly, PR expression itself even seemed to support particle formation. However, NPs formed in the coassembly approach were a lot more prone to aggregation. One explanation could be that NP formation could not keep pace with the rate of PR expression. Consequently, insufficient membrane area was available for proper PR insertion. The results also showed vividly that tuning stoichiometry and kinetics is important for obtaining high quality samples. Prolonged expression favored aggregation probably due to NP overcrowding and extended insertion stress. Generally, it appears that NDs are more reliable in both, preassembly and coassembly modes, yielding not only high-quality PR/ND particles but also GPCR samples with a larger active fraction<sup>45</sup>. Unfortunately, the coassembly mode proved not to be reliable for expression of GPCRs with the SapA system. Although relatively high yields of solubilized GFP fusions with the GPCRs FFAR<sub>2</sub> and GPCR5B were synthesized, further analysis revealed that the GFP moiety obviously folded independently, and potentially only soluble aggregates were formed, that did not even bind to the affinity column used for purification. An exception were  $\beta_1$ AR-GFP fusions as samples in low nanomolar concentrations could be purified. In summary, although SapNPs are generally suitable for the cotranslational solubilization of MPs, they are currently not competitive to MSP ND systems and further refinement might be considered.

### **5.2 The future of membrane mimetics in CF systems**

The thesis of Zoe Köck has already demonstrated that GPCR/ND particles obtained during CF expression usually represent an inhomogeneous mixture of functional protein and misfolded soluble aggregates<sup>27</sup>. To obtain high quality samples, correctly folded GPCR/ND particles were isolated by using their interaction with G-proteins as bait. SEC analysis as preliminary quality control assay helps to identify reaction compounds and parameters that support the folding of synthesized MP targets. Overall, compact MPs with a higher rigidity are easier to obtain at high yields and in a

functional state, as was demonstrated with PR, EmrE and several others<sup>45</sup>. In contrast, more complex MPs having extended flexible regions deserve more intense expression protocol refinement.

Since cotranslational MP insertion is not assisted by the translocon machinery, it is only plausible that sole thermodynamics will not always be sufficient to fold a functional protein. Thus, the question arises, how to optimize and to refine MP sample preparation to obtain as few as possible aggregates. Experiments with SapNPs have already proposed that aggregate formation can be reduced by reducing template concentration, increasing availability of SapNPs for insertion or shortening expression time. However, these strategies seem to have only minor effects on the particle quality of some targets. This means that folding within the NP itself might be a problem. For future studies one might think about further optimizing MSP NDs for CF systems. For example, a different way of processing of the eluate of purified MP/ND complexes might already show some effect on the product. After CF expression, samples are usually centrifuged at 20,000 x g and everything that is not precipitated is considered “soluble”. Assuming aggregates have a higher density than intact MP/ND samples, ultracentrifugation methods of either crude CF reactions or eluates might be a way at least to partially remove aggregated material. Secondly, application of covalently circularized MSP might provide particles with a higher stability, able to withstand mechanical stress during MP folding as no gaps will be present between the MSP belts. On one hand this strategy is promising, on the other it was previously postulated that nascent MP peptides might favorably insert at the gaps between the MSP belts. However, so far there is no evidence for that. cNDs were reported to be highly stable, and thus, favorable for a variety of applications<sup>72</sup>. For the CF system particle stability is a great advantage. ND stability will also depend on the lipid composition of the particles, since lipids not only have an influence on MP insertion, but also on the lateral pressure of the membrane<sup>22,186</sup>.

Once a protocol for circularized NDs is established, it would be interesting to use large MSP variants to obtain membrane surfaces that will provide space for large complexes and give flexible targets some “breathing room”. Indeed, not all MSP variants were shown to be forming stable NDs. In our hands, especially large MSP2N2 NDs resulted in very inhomogeneous particle populations in pilot studies. One can imagine that the larger the ND, the more flexible the bilayer will be. It's possible that probes will have to be handled with extreme caution not to be exposed to too harsh mechanical stress. Especially during concentration of the sample, disruption of membranes due to shear forces or particle fusion due to high local concentration might occur. Apart from cNDs, circularized SapNPs might also be an alternative. In theory, this would require cloning of several SapA copies in a row connected with linkers. The amount of SapA copies will be determinant to the particle size and sortase treatment would circularize the protein belt. The same strategy is plausible with peptidiscs which provide the experimenter with an even higher modularity. However, stability of such systems is not guaranteed.

The idea to use different scaffolds is indeed an interesting one. Recently, derivatives of SMALPs and DIBMAs were tested by our group regarding their compatibility with the CF system. Initial trials with typical derivatives failed most likely because SMALPs tend to bind  $Mg^{2+}$  which is essential for CF expression. However, engineered variants, especially chemically altered Sulfo-DIBMAs depleted of this characteristic were successfully applied<sup>184</sup>. Although, still behind the MSP1D1E3 NDs regarding efficiency

and production of high-quality samples, they might open new routes of CF expression of MPs<sup>184</sup>. For example, MPs could be inserted into liposome membranes and then extracted by SMAs. Since liposomes are larger than NDs, size limitations would be avoided this way. Unfortunately, in our hands, liposomes lack stability in the CF reaction which might be due to mechanical stress upon insertion and excessive overcrowding of the membranes, causing collapse or coagulation of the extruded liposomes and precipitation. Decrease of expression strength or higher liposome concentrations might be helpful in this regard. Apart from SMALPs, MSPs or SapA might also be applied for protein extraction since spontaneous formation of particles was successfully demonstrated<sup>45</sup>. The approach to create isotropic bicelles as was successfully demonstrated with MSP1D1ΔH5<sup>104</sup> might be also employed with all of the mentioned scaffolds.

Conclusively, despite most NP systems were already tested in combination with the CF system by our group, one must consider that cotranslational MP insertion is a technique that opens a plenitude of new paths. Even though MSP NDs are a gold standard at this moment, alternatives discussed above are yet to be explored. More effort to engineer NPs might not only improve the current system, but also provide further answers to translocon-unassisted MP insertion.

### 5.3 Engineered lysates for stable isotope labeling and beyond

Engineering of *E. coli* strains for live cell expression or production of improved CF lysates is not particularly a new technique to increase labeling efficiency. However, thus far, no *E. coli* strain is available that combines mutations to minimize label scrambling between L-Glu, L-Asp, L-Gln and L-Asn. Moreover, no study so far combined the use of mutated strains with aminotransferase inhibitors. The engineered strain Stablelabel combined mutations that reduced conversions between L-Asp, L-Asn, L-Glu and L-Gln. Residual PLP dependent aminotransferase reactions were eliminated using broadband aminotransferase inhibitor AOA. The only non PLP dependent AA conversion still present in Stablelabel is L-Gln to L-Glu which is only controllable by treatment of lysate with the glutaminase inhibitor DON. Yokoyama and coworkers proposed to use DON at 20 mM concentration in the CF reaction. However, pre-treatment of Stablelabel lysate alone with 5 mM DON is sufficient to almost completely eliminate conversion, since DON irreversibly binds to the glutaminase active site<sup>137</sup>. There are several proteins with glutaminase activity in *E. coli* lysates. Most prevalent are GltB, GlnS, GlnA and to lesser extent AsnB, Gdh and possibly GlnA1. In our hands, a chromosomal *guaA* deletion was not possible to obtain. Deletion of *gltB* and *glnS* genes separately was possible. However, cumulative deletion severely impacted strain viability and yielded an unproductive lysate. Furthermore, strains with *glnS* deletion were sensitive to glucose in the culture medium and growth stagnated around OD 1.5. It was previously reported that growth of a *glnS*<sup>-</sup> mutant was inhibited by glucose suggesting that it impairs glucosamine uptake via phosphotransferase systems<sup>187,188</sup>. This might indicate that the two enzymes are essential in the L-Glu and L-Gln metabolism, although an array of enzymes with glutaminase activity is present in live *E. coli* cells<sup>189</sup>. In our efforts to find different solutions for strain engineering, a his-tag was fused *in vivo* to glutamine synthetase GlnA. The initial strategy was to test if scrambling enzymes can be affinity tagged and subsequently pulled out of the lysate, e.g., after final clarification steps. However, attachment of the his-tag already completely eliminated glutaminase activity and no his-tagged protein was found in the lysate via western blot. This indicates that the his-tag and the introduced 32 bp cloning

scar might have already abolished GlnA expression on the transcription level. Indeed, one disadvantage of the Datsenko and Wanner gene deletion protocol is the use of flippase recognition sites flanking the resistance marker for recovery with a flippase encoded on a separate plasmid<sup>168</sup>. It was observed that marker recovery became more difficult the higher the strains were evolved. It is possible that the flippase was “distracted” by FRT cloning scars scattered across the genome at the cost of recovery efficiency, also increasing risk of chromosomal rearrangements<sup>190</sup>. However, optimized homologous recombination protocols like FRUIT and also CRISPR-Cas9 assisted methods allow scarless gene deletions<sup>190–192</sup>. In the future this might allow to produce tailored CF lysates, where unwanted enzymes such as AA scrambling enzymes, RNases or proteases can be removed by simply passing the lysate over nickel beads a few times. In preliminary experiments it was already shown that A19 lysates do not lose efficiency due to such treatment.

One step further, one might also imagine knock-in strains, e.g., expressing additional chaperones, cofactors, or transcriptional enhancers for certain genes. However, it is to be tested how such mutations would affect strain viability. Of course, instead of engineering knock-in strains one could simply separately express all proteins desired and supply them in the CF reaction, or as was done here, use for precursor conversion. Nonetheless, considering the plethora of modern cloning methods, engineering of CF lysates was so far highly underappreciated. *In vitro* expression will always demand defined systems such as bottom up constructed PURE lysates since they contain only the proteins required for synthesis<sup>6</sup>. Engineered CF lysates, on the other hand, would rather peel off unnecessary components of the *E. coli* lysate and combine beneficial properties of PURE lysates with a high protein yield contributing a powerful tool for protein expression. As *E. coli* lysates are also applied as biosensors in diagnostics, engineered lysates might further expand their use in this field too<sup>33,35</sup>.

Despite prevalent scrambling in CF lysates, metabolic pathways to employ early-stage precursors, e.g., for precise methyl labeling, are truncated or show only minor activity in the lysates. Late metabolic precursors, however, can be converted into the AA by only one transamination step with L-Glu as amino group donor<sup>166</sup>. This labeling technique was exploited in this study to label PR at L-Val and L-Leu residues. As a result, somewhat crowded, but nonetheless well-resolved spectra of PR in a detergent environment were obtained. The sole drawback of the method is the impossibility to label the pro-*S* or pro-*R* methyl groups selectively. This will always result in a larger number of resonances, due to slight differences of the chemical environment between the two methyl groups. The great advantage of precursor preconversion, however, is the possibility to transform the ketoacid in a deuterated environment. This way deuterons at C $\alpha$  and C $\beta$  are preserved, and peak doubling or tripling is avoided. Furthermore, the amount of applied precursor can be reduced compared to using conversion in the CF reaction directly. This reduces labeling cost to as few as 20-30 €. Furthermore, if the rest of the protein must be deuterated, one would still have to use AOA to prevent deuterium-proton exchange. With AOA in the reaction however, innate IlvE is inhibited too, and no precursor conversion will take place. So far methyl labeled precursors are commercially available only for KIV and cost around 1400 € per gram (<https://www.isotope.com>). Other precursors in this study were either unlabeled or kindly provided by Roman J. Lichtenecker and coworkers. Although, a labeled oxoacid precursor for L-Ile was also provided, in our hands preconversion with IlvE was not successful. Optimization regarding varying concentration of the amino group donor L-Glu might be beneficial for this approach. More recently, Schörghuber and coworkers

added also aromatic ketoacids to the repertoire of oxoacid precursors<sup>153</sup>. This would require heterologous expression and purification of the required enzymes and a characterization for optimized preconversion conditions. However, problems with non-stereoselective methyl group labeling will remain.

### 5.4 Outlook on NMR investigations with cotranslationally solubilized GPCR

To knit together the approaches to produce a functional labeled MP probe, this chapter will try to provide a “recipe” for future applications with emphasis on GPCR. As it was already discussed in 5.2., different NP systems can be tested regarding their usability. However, most of the NP systems are not suitable for NMR studies due to size, stability or inhomogeneity restrictions discussed before. A rather promising ND scaffold for NMR studies is MSP1D1ΔH5 which produces 8-10 nm NDs, commonly used for NMR experiments with MPs<sup>70</sup>. The first step would be of course to test expression capacity and check quality by SEC, and if possible, to perform an activity assay. If no ligand binding assay is available, a pulldown with suitable G-protein as bait can be carried out. The capability to bind G-protein is already a good indicator that the GPCR is functional. To optimize sample quality, G-protein can be added directly to the CF reaction as was recently shown by our group<sup>27,36</sup>. Furthermore, ND and template concentration during CF expression can be modulated to optimize sample quality, as was demonstrated in this study. Smaller NDs provide a smaller lipid surface area and thus, a higher ND concentration will be required.

In our first attempt of GPCR labeling, it was shown that a decent SEC profile alone does not guarantee good sample quality. Apparently, to obtain a homogeneous functional GPCR, ligand or cofactor affinity purification will always be required. A purification strategy like this would require several steps beginning with affinity purification of the tagged GPCR. In this case, it is convenient to perform strep purification, as it separates tagged constructs with high specificity. Once the eluate is obtained, it can be passed over an IMAC resin with G-protein already bound to it. If the GPCR was already expressed in presence of the G-protein, the crude expression mixture can be passed over IMAC beads directly. Once protein that does not interact with the G-protein is washed away, bound complexes can be eluted with imidazole containing buffer. To remove free G-protein that will also elute from the nickel column, a strep purification is required. The sample needs then to be concentrated for a polishing SEC step that will remove potential aggregates or disintegrated G-protein complexes. The SEC fractions with active complexes can then be pooled. An apyrase treatment can be introduced to help stabilize GPCR/G-protein complexes<sup>27</sup>. Once the amount of active probe after purification is known, one can estimate the feasibility to analyze the target with NMR. A recent study of the histamine 2 receptor showed very promising yields of active GPCR<sup>36</sup>. On the other hand, CF expressed endothelin B yielded less than 10 % active GPCR which is not feasible for NMR<sup>41</sup>. One can imagine that 2-3 affinity purifications, a SEC purification and concentration steps in between will lead to protein loss, which is important to consider when scaling up the reaction for NMR.

To overcome issues with large complex size, it's necessary to deuterate at least the MP. Deuterated lipid and MSP would potentially further improve spectra, although this is an extremely expensive approach. One liter of D<sub>2</sub>O costs around 1500 € which amounts to 15,000 € for a 10 L MSP expression, which is a realistic scale required for CF production of NMR probes. On top comes that expression in D<sub>2</sub>O is usually less



efficient than in H<sub>2</sub>O based media. An optimistic estimate is that 10 L culture medium will yield 150 mg MSP1D1ΔH5. 250 mg deuterated POPG additionally costs ~6,700 € (<https://www.cortecnet.com>). Assuming 150 μM MSP1D1ΔH5 NDs in the RM, the expressed MSP and lipid would be sufficient for a RM volume of 24 mL, i.e. ~10 NMR probes, dependent on the target yield after purification. Assuming a GPCR that is expressed at 20 μM in the RM with 50 % active protein and additional loss of 20 % during purification, a 300 μL sample with a concentration of ~60 μM would remain for NMR. In the next step, one can think of possible ways to label an MP.

Methyl side chain labeling from precursors is very attractive for labeling MPs due to reasons discussed above (1.3.4). In this study, emphasis was laid on L-Val and L-Leu precursors KIV and MOV, respectively. At current price of 1400 €/g for KIV (4-<sup>13</sup>C, 3, 5, 5, 5-D4), this corresponds to ~20-30 € per 50 mL FM and 2.5 mL RM expression (<https://www.isotope.com>). To deuterate the protein, all other single deuterated AAs have to be supplied to the reaction separately or mixes with deuterated AA extracted from algae can be used. 1 g of deuterated AAs extracts from algae cost around 1500 € and provide enough material for ~5-7 expressions at 50 mL FM scale. However, composition of such algal AA extracts is highly disbalanced. More than 60 % of the mixtures is composed of 6 AAs: L-Leu (~14 %), L-Glu (~13 %), L-Ala (~10 %), L-Val (9 %), L-Ile (8 %) and L-Phe (8 %), while abundance of L-Trp, L-Cys, L-His, L-Ser, L-Lys, L-Arg and L-Met is 3 % or less, respectively. The other AAs are present at median 4-6 % abundance (<https://shop.isotope.com>). Insufficient AA concentration during CF expression will lead to poor yield and thus, supplementation of some single deuterated AAs might be required. Before any NMR probe, expression should be tested with the labeled AAs in a small scale. For labeling with preconverted KIV or MOV, one has to consider that the algae AA extracts already contain L-Val and L-Leu in overproportionate amounts. Consequently, it is recommended to use higher concentration to outcompete the deuterated AA from the algae extract. Concentration of preconverted L-Val can also be increased to outcompete deuterated, non-methyl labeled L-Val. This will amount to ~60-90 € labeling cost for a 2.5 mL RM and 50 mL FM expression. Unfortunately, MOV (5-<sup>13</sup>C, 3, 3, 4, 5, 5, 5-D6) is not commercially available and thus, one has to consider feasibility of this approach, as well as for the other precursors that are not yet on the market<sup>155,167</sup>.

This study shows vividly all bottlenecks to consider when trying to analyze MPs cotranslationally inserted into NDs by NMR. So far NMR analysis of cotranslationally inserted MPs in NDs was only reported by Aisha Laguerre for LspA and EmrE<sup>104</sup>. Immense spectral improvements were demonstrated upon stripping away MSP from MP/ND complexes to replace the protein scaffold with a lighter lipid-detergent mixed micelle, and thus, reducing correlational tumbling time<sup>104</sup>. GPCR targets are larger and have a higher complexity than the analyzed bacterial proteins like LspA<sup>104</sup>. However, as already discussed above, there are methods to obtain homogeneous samples with active receptors. There is certainly a plentitude of GPCR targets suitable for CF expression. Ligand binding or G-protein interaction studies will help to identify the compatible ones. Some targets might never become accessible to CF expression due to absence of supporting folding mechanisms. For other targets cotranslational expression will dramatically reduce complexity in obtaining functionally folded and labeled GPCR.



## 6. References

- (1) Buchner, E., and Rapp, R. (1897) Alkoholische Gahrung ohne Hefezellen. *Berichte der Dtsch. Chem. Gesellschaft* 30, 2668–2678.
- (2) Khambhati, K., Bhattacharjee, G., Gohil, N., Braddick, D., Kulkarni, V., and Singh, V. (2019) Exploring the potential of cell-free protein synthesis for extending the abilities of biological systems. *Front. Bioeng. Biotechnol.* 7, 1–16.
- (3) Gale, E. F., and Folkes, J. P. (1954) Effect of nucleic acids on protein synthesis and amino acid incorporation in disrupted Staphylococcal cells. *Nature* 173, 1223–1227.
- (4) Nirenberg, M. W., and Matthaei, J. H. (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. U. S. A.* 47, 1588–1602.
- (5) Silverman, A. D., Karim, A. S., and Jewett, M. C. (2020) Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* 21, 151–170.
- (6) Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001) Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* 19, 751–755.
- (7) Spirin, A. S., Baranov, V. I., Ryabova, L. A., Ovodov, S. Y., and Alakhov, Y. B. (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. *Science* (80- ). 242, 1162–1164.
- (8) Kigawa, T., Yabuki, T., Matsuda, N., Matsuda, T., Tanaka, A., and Yokoyama, S. (2004) Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. *J. Struct. Funct. Genomics* 5, 63–68.
- (9) Gan, R., and Jewett, M. C. (2014) A combined cell-free transcription-translation system from *Saccharomyces cerevisiae* for rapid and robust protein synthesis. *Biotechnol. J.* 9, 641–651.
- (10) Roberts, B. E., and Paterson, B. M. (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. U. S. A.* 70, 2330–2334.
- (11) Stech, M., Hust, M., Schulze, C., Dubel, S., and Kubick, S. (2014) Cell-free eukaryotic systems for the production, engineering, and modification of scFv antibody fragments. *Eng. Life Sci.* 14, 387–398.
- (12) Pelham, H. R. B., and Jackson, R. J. (1976) An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur. J. Biochem.* 67, 247–256.
- (13) Brodel, A. K., Sonnabend, A., and Kubick, S. (2014) Cell-free protein expression based on extracts from CHO cells. *Biotechnol. Bioeng.* 111, 25–36.
- (14) Mikami, S., Masutani, M., Sonenberg, N., Yokoyama, S., and Imataka, H. (2006) An efficient mammalian cell-free translation system supplemented with translation factors. *Protein Expr. Purif.* 46, 348–357.
- (15) Zemella, A., Thoring, L., Hoffmeister, C., amalıkova, M., Ehren, P., Wustenhagen, D. A., and Kubick, S. (2018) Cell-free protein synthesis as a novel tool for directed glycoengineering of active erythropoietin. *Sci. Rep.* 8, 1–12.
- (16) Tarui, H., Murata, M., Tani, I., Imanishi, S., Nishikawa, S., and Hara, T. (2001) Establishment and characterization of cell-free translation/glycosylation in insect cell (*Spodoptera frugiperda* 21) extract prepared with high pressure treatment. *Appl. Microbiol. Biotechnol.* 55, 446–453.

## References

---

- (17) Guarino, C., and Delisa, M. P. (2012) A prokaryote-based cell-free translation system that efficiently synthesizes glycoproteins. *Glycobiology* 22, 596–601.
- (18) Cole, S. D., Miklos, A. E., Chiao, A. C., Sun, Z. Z., and Lux, M. W. (2020) Methodologies for preparation of prokaryotic extracts for cell-free expression systems. *Synth. Syst. Biotechnol.* 5, 252–267.
- (19) Kwon, Y. C., and Jewett, M. C. (2015) High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep.* 5, 1–8.
- (20) Kim, T. W., Keum, J. W., Oh, I. S., Choi, C. Y., Park, C. G., and Kim, D. M. (2006) Simple procedures for the construction of a robust and cost-effective cell-free protein synthesis system. *J. Biotechnol.* 126, 554–561.
- (21) Zubay, G. (1973) In vitro synthesis of protein in microbial systems. *Annu. Rev. Genet.* 7, 267–287.
- (22) Henrich, E., Hein, C., Dötsch, V., and Bernhard, F. (2015) Membrane protein production in *Escherichia coli* cell-free lysates. *FEBS Lett.* 589, 1713–1722.
- (23) Foshag, D., Henrich, E., Hiller, E., Schäfer, M., Kerger, C., Burger-Kentischer, A., Diaz-Moreno, I., García-Mauriño, S. M., Dötsch, V., Rupp, S., and Bernhard, F. (2018) The *E. coli* S30 lysate proteome: A prototype for cell-free protein production. *N. Biotechnol.* 40, 245–260.
- (24) Levin, R., Koeck, Z., Dötsch, V., and Bernhard, F. (2020) Co-translational insertion of membrane proteins into preformed nanodiscs. *J. Vis. Exp.* 165, 1–27.
- (25) Schwarz, D., Junge, F., Durst, F., Frölich, N., Schneider, B., Reckel, S., Sobhanifar, S., Dötsch, V., and Bernhard, F. (2007) Preparative scale expression of membrane proteins in *Escherichia coli*-based continuous exchange cell-free systems. *Nat. Protoc.* 2, 2945–2957.
- (26) Jewett, M. C., and Swartz, J. R. (2004) Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.* 86, 19–26.
- (27) Köck, Z., Ermel, U., Martin, J., Morgner, N., Frangakis, A. S., Dötsch, V., Hilger, D., and Bernhard, F. (2022) Biochemical characterization of cell-free synthesized human  $\beta 1$  adrenergic receptor cotranslationally inserted into nanodiscs. *J. Mol. Biol.* 434, 167687.
- (28) Reckel, S., Gottstein, D., Stehle, J., Löhr, F., Verhoefen, M. K., Takeda, M., Silvers, R., Kainosho, M., Glaubitz, C., Wachtveitl, J., Bernhard, F., Schwalbe, H., Güntert, P., and Dötsch, V. (2011) Solution NMR structure of proteorhodopsin. *Angew. Chemie - Int. Ed.* 50, 11942–11946.
- (29) Dondapati, S. K., Stech, M., Zemella, A., and Kubick, S. (2020) Cell-Free Protein Synthesis: A Promising Option for Future Drug Development. *BioDrugs* 34, 327–348.
- (30) Chiba, C. H., Knirsch, M. C., Azzoni, A. R., Moreira, A. R., and Stephano, M. A. (2021) Cell-free protein synthesis: Advances on production process for biopharmaceuticals and immunobiological products. *Biotechniques* 70, 1–8.
- (31) Carlson, E. D., Gan, R., Hodgman, C. E., and Jewett, M. C. (2012) Cell-free protein synthesis: Applications come of age. *Biotechnol. Adv.*
- (32) Pardee, K., Green, A. A., Ferrante, T., Cameron, D. E., Daleykeyser, A., Yin, P., and Collins, J. J. (2014) Paper-based synthetic gene networks. *Cell* 159, 940–954.
- (33) Copeland, C. E., Langlois, A., Kim, J., and Kwon, Y. C. (2021) The cell-free system: A new apparatus for affordable, sensitive, and portable healthcare. *Biochem. Eng. J.* 175, 108124.

- (34) Bergquist, P. L., Siddiqui, S., and Sunna, A. (2020) Cell-free biocatalysis for the production of platform chemicals. *Front. Energy Res.* 8, 1–22.
- (35) Soltani, Mehran; Brigham Young; Hunt, John; Bundy, B. (2021) Rapid RNase inhibitor production to enable low-cost, on demand cell-free protein synthesis biosensor use in human body fluids. *Biotechnol. Bioeng.* 119, 3973–3983.
- (36) Köck, Z., Schnelle, K., Persechino, M., Umbach, S., Schihada, H., Janulienė, D., Parey, K., Pockes, S., Kolb, P., Dötsch, V., Möller, A., Hilger, D., and Bernhard, F. Cryo-EM structure of cell-free synthesized human histamine H2 receptor coupled to heterotrimeric Gs protein in lipid nanodisc environment. Preprint at <https://www.biorxiv.org/content/10.110>.
- (37) Mezhyrova, J., Martin, J., Peetz, O., Dötsch, V., Morgner, N., Ma, Y., and Bernhard, F. (2021) Membrane insertion mechanism and molecular assembly of the bacteriophage lysis toxin  $\Phi$ X174-E. *FEBS J.* 288, 3300–3316.
- (38) Rosenblum, G., and Cooperman, B. S. (2014) Engine out of the chassis: Cell-free protein synthesis and its uses. *FEBS Lett.*
- (39) Ryabova, L. A., Desplancq, D., Spirin, A. S., and Plückthun, A. (1997) Functional antibody production using cell-free translation: Effects of protein disulfide isomerase and chaperones. *Nat. Biotechnol.* 15, 79–84.
- (40) Walter, P., and Blobel, G. (1981) Translocation of proteins across the endoplasmic reticulum. III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *J. Cell Biol.* 91, 557–561.
- (41) Rues, R. B., Dong, F., Dötsch, V., and Bernhard, F. (2018) Systematic optimization of cell-free synthesized human endothelin B receptor folding. *Methods* 147, 73–83.
- (42) Laguerre, A., Löhr, F., Bernhard, F., and Dötsch, V. (2015) Labeling of membrane proteins by cell-free expression. *Methods Enzymol.* 565, 367–388.
- (43) Klammt, C., Schwarz, D., Fendler, K., Haase, W., Dötsch, V., and 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, 6024–6038.
- (44) Haberstock, S., Roos, C., Hoevens, Y., Dötsch, V., Schnapp, G., Pautsch, A., and Bernhard, F. (2012) A systematic approach to increase the efficiency of membrane protein production in cell-free expression systems. *Protein Expr. Purif.* 82, 308–316.
- (45) Levin, R., Köck, Z., Martin, J., Zangl, R., Gewering, T., Schüler, L., Moeller, A., Dötsch, V., Morgner, N., and Bernhard, F. (2022) Cotranslational assembly of membrane protein/nanoparticles in cell-free systems. *Biochim. Biophys. Acta - Biomembr.* 1864, 184017.
- (46) Zimmerman, E. S., Heibeck, T. H., Gill, A., Li, X., Murray, C. J., Madlansacay, M. R., Tran, C., Uter, N. T., Yin, G., Rivers, P. J., Yam, A. Y., Wang, W. D., Steiner, A. R., Bajad, S. U., Penta, K., Yang, W., Hallam, T. J., Thanos, C. D., and Sato, A. K. (2014) Production of site-specific antibody-drug conjugates using optimized non-natural amino acids in a cell-free expression system. *Bioconjug. Chem.* 25, 351–361.
- (47) Cappuccio, J. A., Blanchette, C. D., Sulchek, T. A., Arroyo, E. S., Kralj, J. M., Hinz, A. K., Kuhn, E. A., Chromy, B. A., Segelke, B. W., Rothschild, K. J., Fletcher, J. E., Katzen, F., Peterson, T. C., Kudlicki, W. A., Bench, G., Hoepflich, P. D., and Coleman, M. A. (2008) Cell-free co-expression of functional membrane proteins and apolipoprotein, forming soluble nanolipoprotein particles. *Mol. Cell. Proteomics* 7, 2246–2253.
- (48) Jelokhani-Niaraki, M. (2023) Membrane proteins: Structure, function and motion. *Int. J. Mol. Sci.*

## References

---

24.

(49) Laeremans, T., Sands, Z. A., Claes, P., De Blicke, A., De Cesco, S., Triest, S., Busch, A., Felix, D., Kumar, A., Jaakola, V. P., and Menet, C. (2022) Accelerating GPCR drug discovery with conformation-stabilizing VHHs. *Front. Mol. Biosci.* 9, 1–21.

(50) Garavito, R. M., and Ferguson-Miller, S. (2001) Detergents as Tools in Membrane Biochemistry. *J. Biol. Chem.* 276, 32403–32406.

(51) Le Maire, M., Chempeil, P., and Møller, Jesper, V. (2000) Interaction of membrane proteins and lipids with solubilizing agents. *Biochim. Biophys. Acta* 1508, 86–111.

(52) Thoma, J., and Burmann, B. M. (2021) Fake it 'till you make it - the pursuit of suitable membrane mimetics for membrane protein biophysics. *Int. J. Mol. Sci.* 22, 1–22.

(53) Chen, A., Majdinasab, E. J., Fiori, M. C., Liang, H., and Altenberg, G. A. (2020) Polymer-encased nanodiscs and polymer nanodiscs: new platforms for membrane protein research and applications. *Front. Bioeng. Biotechnol.* 8, 1–8.

(54) Leitz, A. J., Bayburt, T. H., Barnakov, A. N., Springer, B. A., and Sligar, S. G. (2006) Functional reconstitution of  $\beta$ 2-adrenergic receptors utilizing self-assembling nanodisc technology. *Biotechniques* 40, 601–612.

(55) Bayburt, T. H., Grinkova, Y. V., and Sligar, S. G. (2002) Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano Lett.* 2, 853–856.

(56) Ritchie, T. K., Grinkova, Y. V., Bayburt, T. H., Denisov, I.G., Zolerciks, J. K., and Atkins, W.M., Sligar, S. G. (2009) Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol.* 464, 211–231.

(57) Denisov, I. G., Grinkova, Y. V., Lazarides, A. A., and Sligar, S. G. (2004) Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size. *J. Am. Chem. Soc.* 126, 3477–3487.

(58) Frauenfeld, J., Löving, R., Armache, J.-P., Sonnen, A., Guettou, F., Moberg, P., Zhu, L., Jegerschöld, C., Flayhan, A., Briggs, J. A. G., Garoff, H., Löw, C., Cheng, Y., and Nordlund, P. (2016) A novel lipoprotein nanoparticle system for membrane proteins HHS Public Access. *Nat Methods* 13, 345–351.

(59) Carlson, M. L., Young, J. W., Zhao, Z., Fabre, L., Jun, D., Li, J., Li, J., Dhupar, H. S., Wason, I., Mills, A. T., Beatty, J. T., Klassen, J. S., Rouiller, I., and Duong, F. (2018) The peptidisc, a simple method for stabilizing membrane proteins in detergent-free solution. *Elife* 7, 1–23.

(60) Jamshad, M., Lin, Y., Knowles, T. J., Parslow, R. A., Harris, C., Wheatley, M., Poyner, D. R., Bill, R. M., Thomas, O. R. T., Overduin, M., and Dafforn, T. R. (2011) Surfactant-free purification of membrane proteins with intact native membrane environment. *Biochem. Soc. Trans.* 39, 813–818.

(61) Zhao, Z., Zhang, M., Hogle, J. M., Shih, W. M., Wagner, G., and Nasr, M. L. (2018) DNA-corralled nanodiscs for the structural and functional characterization of membrane proteins and viral entry. *J. Am. Chem. Soc.* 140, 10639–10643.

(62) Denisov, I. G., and Sligar, S. G. (2016) Nanodiscs for structural and functional studies of membrane proteins. *Nat. Struct. Mol. Biol.* 23, 481–486.

(63) Denisov, I. G., and Sligar, S. G. (2017) Nanodiscs in membrane biochemistry and biophysics. *Chem. Rev.* 117, 4669–4713.

(64) Simon, K. S., Pollock, N. L., and Lee, S. C. (2018) Membrane protein nanoparticles: The shape of things to come. *Biochem. Soc. Trans.* 46, 1495–1504.

- (65) Brouillette, C. G., Anantharamaiah, G. M., Engler, J. A., and Borhani, D. W. (2001) Structural models of human apolipoprotein A-I: A critical analysis and review. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* 1531, 4–46.
- (66) Georgila, K., Vyrla, D., and Drakos, E. (2019) Apolipoprotein A-I (ApoA-I), immunity, inflammation and cancer. *Cancers (Basel)*. 11, 1–25.
- (67) Brouillette, C. G., Jones, J. L., Ng, T. C., Kercret, H., Chung, B. H., and Segrest, J. P. (1984) Structural Studies of Apolipoprotein A-I/phosphatidylcholine recombinants by high-field proton NMR, nondenaturing gradient gel electrophoresis, and electron microscopy. *Biochemistry* 23, 359–367.
- (68) Borhani, D. W., Rogers, D. P., Engler, J. A., and Brouillette, C. G. (1997) Crystal structure of truncated human apolipoprotein A-I suggests a lipid-bound conformation. *Proc. Natl. Acad. Sci. U. S. A.* 94, 12291–12296.
- (69) Grinkova, Y. V., Denisov, I. G., and Sligar, S. G. (2010) Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers. *Protein Eng. Des. Sel.* 23, 843–848.
- (70) Hagn, F., Nasr, M. L., and Wagner, G. (2018) Assembly of phospholipid nanodiscs of controlled size for structural studies of membrane proteins by NMR. *Nat. Protoc.* 13, 79–98.
- (71) Padmanabha Das, K. M., Shih, W. M., Wagner, G., and Nasr, M. L. (2020) Large nanodiscs: A potential game changer in structural biology of membrane protein complexes and virus entry. *Front. Bioeng. Biotechnol.* 8, 1–7.
- (72) Nasr, M. L., Baptista, D., Strauss, M., Sun, Z. Y. J., Grigoriu, S., Huser, S., Plückthun, A., Hagn, F., Walz, T., Hogle, J. M., and Wagner, G. (2016) Covalently circularized nanodiscs for studying membrane proteins and viral entry. *Nat. Methods* 14, 49–52.
- (73) Bruhn, H. (2005) A short guided tour through functional and structural features of saposin-like proteins. *Biochem. J.* 289, 249–257.
- (74) Kolter, T., and Sandhoff, K. (2005) Principles of lysosomal membrane digestion: stimulation of sphingolipid degradation by sphingo- lipid activator proteins and anionic lysosomal lipids. *Annu. Rev. Cell Dev. Biol.* 21, 81–103.
- (75) Frauenfeld, J., Löving, R., Armache, J., Sonnen, A. F., Guettou, F., Moberg, P., Zhu, L., Jegerschöld, C., Flayhan, A., Briggs, J. A. G., Garoff, H., and Löw, C. (2016) A saposin-lipoprotein nanoparticle system for membrane proteins. *Nat. Methods* 13, 345–351.
- (76) Chien, C. T. H., Helfinger, L. R., Bostock, M. J., Solt, A., Tan, Y. L., and Nietlispach, D. (2017) An adaptable phospholipid membrane mimetic system for solution NMR studies of membrane proteins. *J. Am. Chem. Soc.* 139, 14829–14832.
- (77) Flayhan, A., Mertens, H. D. T., Ural-Blimke, Y., Martinez Molledo, M., Svergun, D. I., and Löw, C. (2018) Saposin lipid nanoparticles: A highly versatile and modular tool for membrane protein Research. *Structure* 26, 345-355.e5.
- (78) Li, J., Richards, M. R., Bagal, D., Campuzano, I. D. G., Kitova, E. N., Xiong, Z. J., Privé, G. G., and Klassen, J. S. (2016) Characterizing the size and composition of saposin A lipoprotein picodiscs. *Anal. Chem.* 88, 9524–9531.
- (79) Kehlenbeck, D. M., Josts, I., Nitsche, J., Busch, S., Forsyth, T., and Tidow, H. (2019) Comparison of lipidic carrier systems for integral membrane proteins - MsbA as case study. *Biol. Chem.* 400, 1509–1518.
- (80) Kobilka, K. (1990) The Role of cytosolic and membrane factors in processing of the Human B-2 adrenergic receptor following translocation and glycosylation in a cell-free system\*. *J. Biol. Chem.* 265,

## References

---

7610–7618.

- (81) Nomura, S. ichiro M., Kondoh, S., Asayama, W., Asada, A., Nishikawa, S., and Akiyoshi, K. (2008) Direct preparation of giant proteo-liposomes by in vitro membrane protein synthesis. *J. Biotechnol.* *133*, 190–195.
- (82) Kalmbach, R., Chizhov, I., Schumacher, M. C., Friedrich, T., Bamberg, E., and Engelhard, M. (2007) Functional cell-free synthesis of a seven helix membrane protein: In situ insertion of bacteriorhodopsin into liposomes. *J. Mol. Biol.* *371*, 639–648.
- (83) Katzen, F., Fletcher, J. E., Yang, J., Kang, D., Peterson, T. C., Cappuccio, J. A., Blanchette, C. D., Sulchek, T., Chromy, B. A., Hoepflich, P. D., Coleman, M. A., and Kudlicki, W. (2008) Insertion of membrane proteins into discoidal membranes using a cell-free protein expression approach. *J. Proteome Res.* *7*, 3535–3542.
- (84) Roos, C., Zocher, M., Müller, D., Münch, D., Schneider, T., Sahl, H. G., Scholz, F., Wachtveitl, J., Ma, Y., Proverbio, D., Henrich, E., Dötsch, V., and Bernhard, F. (2012) Characterization of co-translationally formed nanodisc complexes with small multidrug transporters, proteorhodopsin and with the *E. coli* MraY translocase. *Biochim. Biophys. Acta - Biomembr.* *1818*, 3098–3106.
- (85) Peetz, O., Henrich, E., Laguerre, A., Löhr, F., Hein, C., Dötsch, V., Bernhard, F., and Morgner, N. (2017) Insights into cotranslational membrane protein insertion by combined LILBID-mass spectrometry and NMR spectroscopy. *Anal. Chem.* *89*, 12314–12318.
- (86) Baumann, A., Kerruth, S., Fitter, J., Büldt, G., Heberle, J., Schlesinger, R., and Ataka, K. (2016) In-situ observation of membrane protein folding during cell-free expression. *PLoS One* *11*, 1–15.
- (87) Umbach, S., Levin, R., Neumann, S., Steinmetzer, T., Dötsch, V., and Bernhard, F. (2022) Transfer mechanism of cell-free synthesized membrane proteins into mammalian cells. *Front. Bioeng. Biotechnol.* *10*, 1–17.
- (88) Keller, T., Gorboulev, V., Mueller, T. D., Dötsch, V., Bernhard, F., and Koepsell, H. (2019) Rat organic cation transporter 1 contains three binding sites for substrate 1-methyl-4-phenylpyridinium per monomer. *Mol. Pharmacol.* *95*, 169–182.
- (89) He, W., Felderman, M., Evans, A. C., Geng, J., Homan, D., Bourguet, F., Fischer, N. O., Li, Y., Lam, K. S., Noy, A., Xing, L., Cheng, R. H., Rasley, A., Blanchette, C. D., Kamrud, K., Wang, N., Gouvis, H., Peterson, T. C., Hubby, B., and Coleman, M. A. (2017) Cell-free production of a functional oligomeric form of a Chlamydia major outer-membrane protein (MOMP) for vaccine development. *J. Biol. Chem.* *292*, 15121–15132.
- (90) Rues, R. B., Dötsch, V., and Bernhard, F. (2016) Co-translational formation and pharmacological characterization of beta1-adrenergic receptor/nanodisc complexes with different lipid environments. *Biochim. Biophys. Acta - Biomembr.* *1858*, 1306–1316.
- (91) Proverbio, D., Roos, C., Beyermann, M., Orbán, E., Dötsch, V., and Bernhard, F. (2013) Functional properties of cell-free expressed human endothelin A and endothelin B receptors in artificial membrane environments. *Biochim. Biophys. Acta - Biomembr.* *1828*, 2182–2192.
- (92) Dong, F., Rues, R. B., Kazemi, S., Dötsch, V., and Bernhard, F. (2018) Molecular determinants for ligand selectivity of the cell-free synthesized human endothelin B receptor. *J. Mol. Biol.* *430*, 5105–5119.
- (93) Yang, J. P., Cirico, T., Katzen, F., Peterson, T. C., and Kudlicki, W. (2011) Cell-free synthesis of a functional G protein-coupled receptor complexed with nanometer scale bilayer discs. *BMC Biotechnol.* *11*.
- (94) Gessesse, B., Nagaike, T., Nagata, K., Shimizu, Y., and Ueda, T. (2018) G-protein coupled receptor protein synthesis on a lipid bilayer using a reconstituted cell-free protein synthesis system. *Life* *8*.



- (95) Manzer, Z. A., Selivanovitch, E., Ostwalt, A. R., and Daniel, S. (2023) Membrane protein synthesis: No cells required. *Trends Biochem. Sci.* 48, 642–654.
- (96) Ma, Y., Münch, D., Schneider, T., Sahl, H. G., Bouhss, A., Ghoshdastider, U., Wang, J., Dötsch, V., Wang, X., and Bernhard, F. (2011) Preparative scale cell-free production and quality optimization of MraY homologues in different expression modes. *J. Biol. Chem.* 286, 38844–38853.
- (97) Hein, C., Henrich, E., Orbán, E., Dötsch, V., and Bernhard, F. (2014) Hydrophobic supplements in cell-free systems: Designing artificial environments for membrane proteins. *Eng. Life Sci.* 14, 365–379.
- (98) Sobhanifar, S., Reckel, S., Junge, F., Schwarz, D., Kai, L., Karbyshev, M., Löhr, F., Bernhard, F., and Dötsch, V. (2010) Cell-free expression and stable isotope labelling strategies for membrane proteins. *J. Biomol. NMR* 46, 33–43.
- (99) Klammt, C., Maslennikov, I., Bayrhuber, M., Eichmann, C., Vajpai, N., Chiu, E. J. C., Blain, K. Y., Esquivies, L., Kwon, J. H. J., Balana, B., Pieper, U., Sali, A., Slesinger, P. A., Kwiatkowski, W., Riek, R., and Choe, S. (2012) Facile backbone structure determination of human membrane proteins by NMR spectroscopy. *Nat. Methods* 9, 834–839.
- (100) Maslennikov, I., Klammt, C., Hwang, E., Kefala, G., Okamura, M., Esquivies, L., Mörs, K., Glaubitz, C., Kwiatkowski, W., Jeon, Y. H., and Choe, S. (2010) Membrane domain structures of three classes of histidine kinase receptors by cell-free expression and rapid NMR analysis. *Proc. Natl. Acad. Sci. U. S. A.* 107, 10902–10907.
- (101) Lends, A., Daskalov, A., Maleckis, A., Delamare, A., Berbon, M., Grélard, A., Morvan, E., Shenoy, J., Dutour, A., Tolchard, J., Noubhani, A., Giraud, M., Sanchez, C., Habenstein, B., Guichard, G., Compain, G., Jaudzems, K., Saupe, S. J., and Loquet, A. (2022) Cell-free synthesis of amyloid fibrils with infectious properties and amenable to sub-milligram magic-angle spinning NMR analysis. *Nat. Commun. Biol.* 5, 12022.
- (102) Danmaliki, G. I., and Hwang, P. M. (2020) Solution NMR spectroscopy of membrane proteins. *Biochim. Biophys. Acta - Biomembr.* 1862, 183356.
- (103) Hoffmann, B., Löhr, F., Laguerre, A., Bernhard, F., and Dötsch, V. (2018) Protein labeling strategies for liquid-state NMR spectroscopy using cell-free synthesis. *Prog. Nucl. Magn. Reson. Spectrosc.* 105, 1–22.
- (104) Laguerre, A., Henrich, E., Hoffmann, B., Abdul-manan, N., Connolly, P. J., Perozo, E., Moore, J. M., and Bernhard, F. (2016) From nanodiscs to isotropic bicelles: a procedure for solution NMR studies of detergent sensitive integral membrane proteins. *Structure* 24, 1830–1841.
- (105) Warschawski, D. E., Arnold, A. A., Beaugrand, M., Gravel, A., Chartrand, É., and Marcotte, I. (2011) Choosing membrane mimetics for NMR structural studies of transmembrane proteins. *Biochim. Biophys. Acta - Biomembr.* 1808, 1957–1974.
- (106) Movellan, K. T., Najbauer, E. E., Pratihari, S., Salvi, M., Giller, K., Becker, S., and Andreas, L. B. (2019) Alpha protons as NMR probes in deuterated proteins. *J. Biomol. NMR* 73, 81–91.
- (107) Löhr, F., Katsemi, V., Hartleib, J., Günther, U., and Rüterjans, H. (2003) A strategy to obtain backbone resonance assignments of deuterated proteins in the presence of incomplete amide  $^2\text{H}/^1\text{H}$  back-exchange. *J. Biomol. NMR* 25, 291–311.
- (108) Abiko, L. A., Rogowski, M., Gautier, A., Schertler, G., and Grzesiek, S. (2021) Efficient production of a functional G protein-coupled receptor in *E. coli* for structural studies. *J. Biomol. NMR* 75, 25–38.
- (109) Kaplan, M., Pinto, C., Houben, K., and Baldus, M. (2016) Nuclear magnetic resonance (NMR) applied to membrane–protein complexes. *Q. Rev. Biophys.* 49, 1–26.

## References

---

- (110) Lewis, B. A., Harbison, G. S., Herzfeld, J., and Griffin, R. G. (1985) NMR structural analysis of a membrane protein: Bacteriorhodopsin peptide backbone orientation and motion. *Biochemistry* 24, 4671–4679.
- (111) Tugarinov, V., Choy, W. Y., Orekhov, V. Y., and Kay, L. E. (2005) Solution NMR-derived global fold of a monomeric 82-kDa enzyme. *Proc. Natl. Acad. Sci. U. S. A.* 102, 622–627.
- (112) Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy. *Proc. Natl. Acad. Sci. U. S. A.* 94, 12366–12371.
- (113) Shimada, I., Ueda, T., Kofuku, Y., Eddy, M. T., and Wüthrich, K. (2021) GPCR drug discovery: Integrating solution NMR data with crystal and cryo-EM structures. *NMR with Biol. Macromol. Solut. A Sel. Pap. Publ. from 1996 to 2020 by Kurt Wüthrich* 18, 197–220.
- (114) Gauto, D. F., Estrozi, L. F., Schwieters, C. D., Effantin, G., Macek, P., Sounier, R., Sivertsen, A. C., Schmidt, E., Kerfah, R., Mas, G., Colletier, J. P., Güntert, P., Favier, A., Schoehn, G., Schanda, P., and Boisbouvier, J. (2019) Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. *Nat. Commun.* 10, 1–12.
- (115) Isogai, S., Deupi, X., Opitz, C., Heydenreich, F. M., Tsai, C. J., Brueckner, F., Schertler, G. F. X., Veprintsev, D. B., and Grzesiek, S. (2016) Backbone NMR reveals allosteric signal transduction networks in the  $\beta$ 1-adrenergic receptor. *Nature* 530, 237–241.
- (116) Grahl, A., Abiko, L. A., Isogai, S., Sharpe, T., and Grzesiek, S. (2020) A high-resolution description of  $\beta$ 1-adrenergic receptor functional dynamics and allosteric coupling from backbone NMR. *Nat. Commun.* 11, 2216.
- (117) Rößler, P., Mayer, D., Tsai, C. J., Veprintsev, D. B., Schertler, G. F. X., and Gossert, A. D. (2020) GPCR activation states induced by nanobodies and mini-G proteins compared by NMR spectroscopy. *Molecules* 25, 1–17.
- (118) Kofuku, Y., Ueda, T., Okude, J., Shiraishi, Y., Kondo, K., Maeda, M., Tsujishita, H., and Shimada, I. (2012) Efficacy of the  $\beta$ 2-adrenergic receptor is determined by conformational equilibrium in the transmembrane region. *Nat. Commun.* 3, 1045–1049.
- (119) Nygaard, R., Zou, Y., Dror, R. O., Mildorf, T. J., Arlow, D. H., Manglik, A., Pan, A. C., Liu, C. W., Fung, J. R., Bokoch, M. P., Thian, F. S., Kobilka, T. S., Shaw, D. E., Mueller, L., Prosser, R. S., and Kobilka, B. K. (2013) The dynamic process of  $\beta$ 2-adrenergic receptor activation. *Cell* 152, 532–542.
- (120) Krug, U., Gloge, A., Schmidt, P., Becker-Baldus, J., Bernhard, F., Kaiser, A., Montag, C., Gauglitz, M., Vishnivetskiy, S. A., Gurevich, V. V., Beck-Sickinger, A. G., Glaubitz, C., and Huster, D. (2020) The conformational equilibrium of the neuropeptide Y2 receptor in bilayer membranes. *Angew. Chemie - Int. Ed.* 59, 23854–23861.
- (121) Waugh, David, S. (1996) Genetic tools for selective labeling of proteins with. *J. Biomol. NMR* 8, 184–192.
- (122) Lin, M. T., Sperling, L. J., Schmidt, H. L. F., Tang, M., Samoilova, R. I., Kumasaka, T., Iwasaki, T., Dikanov, S. A., Rienstra, C. M., and Gennis, R. B. (2011) A rapid and robust method for selective isotope labeling of proteins. *Methods* 55, 370–378.
- (123) Vance, C. K., Kang, Y. M., and Miller, A. F. (1997) Selective  $^{15}\text{N}$  labeling and direct observation by NMR of the active-site glutamine of Fe-containing superoxide dismutase. *J. Biomol. NMR* 9, 201–206.
- (124) LeMaster, D. M., and Richards, F. M. (1988) NMR sequential assignment of *Escherichia coli* thioredoxin utilizing random fractional deuteration. *Biochemistry* 27, 142–150.

- (125) Verardi, R., Traaseth, N. J., Masterson, L. R., Vostrikov, V. V., and Veglia, G. (2012) Isotope labeling for solution and solid-state NMR spectroscopy of membrane proteins. *Adv. Exp. Med. Biol.*
- (126) Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., and Dahlquist, F. W. (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods Enzymol.* 177, 44–73.
- (127) Griffey, R. H., Redfield, A. G., Loomis, R. E., and Dahlquist, F. W. (1985) Nuclear magnetic resonance observation and dynamics of specific amide protons in T4 lysozyme. *Biochemistry* 24, 817–822.
- (128) Schütz, S., and Sprangers, R. (2020) Methyl TROSY spectroscopy: A versatile NMR approach to study challenging biological systems. *Prog. Nucl. Magn. Reson. Spectrosc.* 116, 56–84.
- (129) Matsuda, T., Koshiba, S., Tochio, N., Seki, E., Iwasaki, N., Yabuki, T., Inoue, M., Yokoyama, S., and Kigawa, T. (2007) Improving cell-free protein synthesis for stable-isotope labeling. *J. Biomol. NMR* 37, 225–229.
- (130) Pacull, E. M., Sendker, F., Bernhard, F., Scheidt, H. A., Schmidt, P., Huster, D., and Krug, U. (2020) Integration of cell-free expression and solid-state NMR to investigate the dynamic properties of different sites of the growth hormone secretagogue receptor. *Front. Pharmacol.* 11, 1–13.
- (131) Tumulka, F., Roos, C., Löhr, F., Bock, C., Bernhard, F., Dötsch, V., and Abele, R. (2013) Conformational stabilization of the membrane embedded targeting domain of the lysosomal peptide transporter TAPL for solution NMR. *J. Biomol. NMR* 57, 141–154.
- (132) Schwarz, D., Klammt, C., Koglin, A., Löhr, F., Schneider, B., Dötsch, V., and Bernhard, F. (2007) Preparative scale cell-free expression systems: New tools for the large scale preparation of integral membrane proteins for functional and structural studies. *Methods* 41, 355–369.
- (133) Linser, R., Gelev, V., Hagn, F., Arthanari, H., Hyberts, S. G., and Wagner, G. (2014) Selective methyl labeling of eukaryotic membrane proteins using cell-free expression. *J. Am. Chem. Soc.* 136, 11308–11310.
- (134) Abdine, A., Verhoeven, M. A., Park, K. H., Ghazi, A., Guittet, E., Berrier, C., Van Heijenoort, C., and Warschawski, D. E. (2010) Structural study of the membrane protein MscL using cell-free expression and solid-state NMR. *J. Magn. Reson.* 204, 155–159.
- (135) Kühlbrandt, W. (2014) The resolution revolution. *Science* 343, 1443–1444.
- (136) Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589.
- (137) Yokoyama, J., Matsuda, T., Koshiba, S., Tochio, N., and Kigawa, T. (2011) A practical method for cell-free protein synthesis to avoid stable isotope scrambling and dilution. *Anal. Biochem.* 411, 223–229.
- (138) Tonelli, M., Singarapu, K. K., Makino, S. I., Sahu, S. C., Matsubara, Y., Endo, Y., Kainosho, M., and Markley, J. L. (2011) Hydrogen exchange during cell-free incorporation of deuterated amino acids and an approach to its inhibition. *J. Biomol. NMR* 51, 467–476.
- (139) Su, X. C., Loh, C. T., Qi, R., and Otting, G. (2011) Suppression of isotope scrambling in cell-free protein synthesis by broadband inhibition of PLP enzymes for selective <sup>15</sup>N-labelling and production of perdeuterated proteins in H<sub>2</sub>O. *J. Biomol. NMR* 50, 35–42.

## References

---

- (140) Etezady-Esfarjani, T., Hiller, S., Villalba, C., and Wüthrich, K. (2007) Cell-free protein synthesis of perdeuterated proteins for NMR studies. *J. Biomol. NMR* 39, 229–238.
- (141) Lavickova, B., and Maerkl, S. J. (2019) A simple, robust, and low-cost method to produce the PURE cell-free system. *ACS Synth. Biol.* 8, 455–462.
- (142) Adachi, J., Katsura, K., Seki, E., Takemoto, C., Shirouzu, M., Terada, T., Mukai, T., Sakamoto, K., and Yokoyama, S. (2019) Cell-free protein synthesis using S30 extracts from *Escherichia coli* rzero strains for efficient incorporation of non-natural amino acids into proteins. *Int. J. Mol. Sci.* 20, 1–12.
- (143) Hong, S. H., Ntai, I., Haimovich, A. D., Kelleher, N. L., Isaacs, F. J., and Jewett, M. C. (2014) Cell-free protein synthesis from a release factor 1 deficient *Escherichia coli* activates efficient and multiple site-specific nonstandard amino acid incorporation. *ACS Synth. Biol.* 3, 398–409.
- (144) Jones, D. H., Cellitti, S. E., Hao, X., Zhang, Q., Jahnz, M., Summerer, D., Schultz, P. G., Uno, T., and Geierstanger, B. H. (2010) Site-specific labeling of proteins with NMR-active unnatural amino acids. *J. Biomol. NMR* 46, 89–100.
- (145) Michel-Reydellet, N., Calhoun, K., and Swartz, J. (2004) Amino acid stabilization for cell-free protein synthesis by modification of the *Escherichia coli* genome. *Metab. Eng.* 6, 197–203.
- (146) Levin, R., Löhr, F., Karakoc, B., Lichtenecker, R., Dötsch, V., and Bernhard, F. (2023) *E. coli* “Stablelabel” S30 lysate for optimized cell-free NMR sample preparation. *J. Biomol. NMR* 77, 131–147.
- (147) Kurauskas, V., Schanda, P., and Sounier, R. (2017) Methyl-Specific Isotope Labeling Strategies for NMR Studies of Membrane Proteins. *Membr. Protein Struct. Funct. Character.* 1635, 109–123.
- (148) Sprangers, R., and Kay, L. E. (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* 445, 618–622.
- (149) LeMaster, D. M., and Kushlan, D. M. (1996) Dynamical mapping of *E. coli* thioredoxin via  $^{13}\text{C}$  NMR relaxation analysis. *J. Am. Chem. Soc.* 118, 9255–9264.
- (150) Higman, V. A., Flinders, J., Hiller, M., Jehle, S., Markovic, S., Fiedler, S., van Rossum, B. J., and Oschkinat, H. (2009) Assigning large proteins in the solid state: A MAS NMR resonance assignment strategy using selectively and extensively  $^{13}\text{C}$ -labelled proteins. *J. Biomol. NMR* 44, 245–260.
- (151) Robson, S. A., Takeuchi, K., Boeszoermyeni, A., Coote, P. W., Dubey, A., Hyberts, S., Wagner, G., and Arthanari, H. (2018) Mixed pyruvate labeling enables backbone resonance assignment of large proteins using a single experiment. *Nat. Commun.* 9, 2–3.
- (152) Lundström, P., Teilum, K., Carstensen, T., Bezsonova, I., Wiesner, S., Hansen, D. F., Religa, T. L., Akke, M., and Kay, L. E. (2007) Fractional  $^{13}\text{C}$  enrichment of isolated carbons using  $[1-^{13}\text{C}]$ - or  $[2-^{13}\text{C}]$ -glucose facilitates the accurate measurement of dynamics at backbone  $\text{C}\alpha$  and side-chain methyl positions in proteins. *J. Biomol. NMR* 38, 199–212.
- (153) Schörghuber, J., Geist, L., Platzer, G., Feichtinger, M., Bisaccia, M., Scheibelberger, L., Weber, F., Konrat, R., and Lichtenecker, R. J. (2018) Late metabolic precursors for selective aromatic residue labeling. *J. Biomol. NMR* 71, 129–140.
- (154) Ayala, I., Sounier, R., Usé, N., Gans, P., and Boisbouvier, J. (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *J. Biomol. NMR* 43, 111–119.
- (155) Lichtenecker, R. J., Weinhäupl, K., Reuther, L., Schörghuber, J., Schmid, W., and Konrat, R. (2013) Independent valine and leucine isotope labeling in *Escherichia coli* protein overexpression systems. *J. Biomol. NMR* 57, 205–209.

- (156) Ruschak, A. M., and Kay, L. E. (2010) Methyl groups as probes of supra-molecular structure, dynamics and function. *J. Biomol. NMR* 46, 75–87.
- (157) Fischer, M., Kloiber, K., Häusler, J., Ledolter, K., Konrat, R., and Schmid, W. (2007) Synthesis of a  $^{13}\text{C}$ -methyl-group-labeled methionine precursor as a useful tool for simplifying protein structural analysis by NMR spectroscopy. *ChemBioChem* 8, 610–612.
- (158) Gardner, K. H., and Kay, L. E. (1997) Production and incorporation of  $^{15}\text{N}$ ,  $^{13}\text{C}$ ,  $^2\text{H}$ ( $^1\text{H}$ - $\delta$ 1 Methyl) isoleucine into proteins for multidimensional NMR studies. *J. Am. Chem. Soc.* 7863, 7599–7600.
- (159) Lichtenecker, R., Ludwiczek, M. L., Schmid, W., and Konrat, R. (2004) Simplification of Protein NOESY Spectra Using Bioorganic Precursor Synthesis and NMR Spectral Editing. *J. Am. Chem. Soc.* 126, 5348–5349.
- (160) Deber, C. M., Brandl, C. J., Deber, R. B., Hsu, L. C., and Young, X. K. (1986) Amino acid composition of the membrane and aqueous domains of integral membrane proteins. *Arch. Biochem. Biophys.* 251, 68–76.
- (161) Gans, P., Hamelin, O., Sounier, R., Ayala, I., Durá, M. A., Amero, C. D., Noirclerc-Savoye, M., Franzetti, B., Plevin, M. J., and Boisbouvier, J. (2010) Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew. Chemie - Int. Ed.* 49, 1958–1962.
- (162) Miyanoiri, Y., Takeda, M., Okuma, K., Ono, A. M., Terauchi, T., and Kainosho, M. (2013) Differential isotope-labeling for Leu and Val residues in a protein by *E. coli* cellular expression using stereo-specifically methyl labeled amino acids. *J. Biomol. NMR* 57, 237–249.
- (163) Kainosho, M., Torizawa, T., Iwashita, Y., Terauchi, T., Mei Ono, A., and Güntert, P. (2006) Optimal isotope labelling for NMR protein structure determinations. *Nature* 440, 52–57.
- (164) Miyanoiri, Y., Ishida, Y., Takeda, M., Terauchi, T., Inouye, M., and Kainosho, M. (2016) Highly efficient residue-selective labeling with isotope-labeled Ile, Leu, and Val using a new auxotrophic *E. coli* strain. *J. Biomol. NMR* 65, 109–119.
- (165) Velyvis, A., Ruschak, A. M., and Kay, L. E. (2012) An economical method for production of  $^2\text{H},^{13}\text{C}$ -threonine for solution NMR studies of large protein complexes: Application to the 670 kDa proteasome. *PLoS One* 7, 1–8.
- (166) Lazarova, M., Löhr, F., Rues, R. B., Kleebach, R., Dötsch, V., and Bernhard, F. (2018) Precursor-based selective methyl labeling of cell-free synthesized proteins. *ACS Chem. Biol.* 13, 2170–2178.
- (167) Lichtenecker, R. J., Weinhäupl, K., Schmid, W., and Konrat, R. (2013)  $\alpha$ -Ketoacids as precursors for phenylalanine and tyrosine labelling in cell-based protein overexpression. *J. Biomol. NMR* 57, 327–331.
- (168) Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6640–6645.
- (169) Hein, C., Löhr, F., Schwarz, D., and Dötsch, V. (2017) Acceleration of protein backbone NMR assignment by combinatorial labeling: Application to a small molecule binding study. *Biopolymers* 107, e23013.
- (170) Fairbanks, G., Steck, T. L., and Wallach, D. F. H. (1971) Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. *Biochemistry* 10, 2606–2617.
- (171) Peetz, O., Hellwig, N., Henrich, E., Mezhyrova, J., Dötsch, V., Bernhard, F., and Morgner, N. (2019) LILBID and nESI: different native mass spectrometry techniques as tools in structural biology. *J. Am. Soc. Mass Spectrom.* 30, 181–191.

## References

---

- (172) Gewering, T., Januliene, D., Ries, A. B., and Moeller, A. (2018) Know your detergents: A case study on detergent background in negative stain electron microscopy. *J. Struct. Biol.* *203*, 242–246.
- (173) Januliene, D., and Moeller, A. (2021) Single-particle cryo-EM of membrane proteins, in *Structure and Function of Membrane Proteins* (Schmidt-Krey, I., and Gumbart, J. C., Eds.), pp 153–178. Springer US, New York, NY.
- (174) Farjon, J., Boisbouvier, J. R., Schanda, P., Pardi, A., Simorre, J. P., and Brutscher, B. (2009) Longitudinal-relaxation-enhanced NMR experiments for the study of nucleic acids in solution. *J. Am. Chem. Soc.* *131*, 8571–8577.
- (175) Rößler, P., Mathieu, D., and Gossert, A. D. (2020) Enabling NMR studies of high molecular weight systems without the need for deuteration: The XL-ALSOFAST experiment with delayed decoupling. *Angew. Chemie - Int. Ed.* *59*, 19329–19337.
- (176) Stone, K. M., Voska, J., Kinnebrew, M., Pavlova, A., Junk, M. J. N., and Han, S. (2013) Structural insight into proteorhodopsin oligomers. *Biophys. J.* *104*, 472–481.
- (177) Maciejko, J., Kaur, J., Becker-Baldus, J., and Glaubitz, C. (2019) Photocycle-dependent conformational changes in the proteorhodopsin cross-protomer Asp–His–Trp triad revealed by DNP-enhanced MAS-NMR. *Proc. Natl. Acad. Sci. U. S. A.* *116*, 8342–8349.
- (178) Kanehisa, M., and Goto, S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* *28*, 27–30.
- (179) Kim, S. H., Schneider, B. L., and Reitzer, L. (2010) Genetics and regulation of the major enzymes of alanine synthesis in *Escherichia coli*. *J. Bacteriol.* *192*, 5304–5311.
- (180) John, R. A., Charteris, A., and Fowler, L. J. (1978) The reaction of amino-oxyacetate with pyridoxal phosphate-dependent enzymes. *Biochem. J.* *171*, 771–779.
- (181) Willis, R. C., and Woolfolk, C. A. (1974) Asparagine utilization in *Escherichia coli*. *J. Bacteriol.* *118*, 231–241.
- (182) Hayashi, H., Inoue, K., Nagata, T., Kuramitsu, S., and Kagamiyama, H. (1993) *Escherichia coli* aromatic amino acid aminotransferase: Characterization and comparison with aspartate aminotransferase. *Biochemistry* *32*, 12229–12239.
- (183) Henrich, E., Peetz, O., Hein, C., Laguerre, A., Hoffmann, B., Hoffmann, J., Dötsch, V., Bernhard, F., and Morgner, N. (2017) Analyzing native membrane protein assembly in nanodiscs by combined non-covalent mass spectrometry and synthetic biology. *Elife* *6*, 1–19.
- (184) Glueck, D., Grethen, A., Das, M., Mmeka, O. P., Patallo, E. P., Meister, A., Rajender, R., Kins, S., Räschele, M., Victor, J., Chu, C., Etzkorn, M., Köck, Z., Bernhard, F., Babalola, J. O., Vargas, C., and Keller, S. (2022) Electroneutral polymer nanodiscs enable interference-free probing of membrane proteins in a lipid-bilayer environment. *Small* *18*.
- (185) Vaccaro, A. M., Ciaffoni, F., Tatti, M., Salvioli, R., Barca, A., Tognozzi, D., and Scerch, C. (1995) pH-Dependent conformational properties of saposins and their interactions with phospholipid membranes. *J. Biol. Chem.* *270*, 30576–30580.
- (186) Booth, P. J. (2005) Sane in the membrane: Designing systems to modulate membrane proteins. *Curr. Opin. Struct. Biol.* *15*, 435–440.
- (187) Wu, H. C., and Wu, T. C. (1971) Isolation and characterization of a glucosamine-requiring mutant of *Escherichia coli* K-12 defective in glucosamine-6-phosphate synthetase. *J. Bacteriol.* *105*, 455–466.
- (188) Kundig, W., Ghosh, S., and Roseman, S. (1964) Phosphate bound to histidine in a protein as an

intermediate in a novel phospho-transferase system. *Proc. Natl. Acad. Sci. United States* 52, 1067–1074.

(189) Sakamoto, N., Kotre, A. M., and Savageau, M. A. (1975) Glutamate dehydrogenase from *Escherichia coli*: purification and properties. *J. Bacteriol.* 124, 775–783.

(190) Ronda, C., Pedersen, L. E., Sommer, M. O. A., and Nielsen, A. T. (2016) CRMAGE: CRISPR Optimized MAGE Recombineering. *Sci. Rep.* 6, 1–11.

(191) Stringer, A. M., Singh, N., Yermakova, A., Petrone, B. L., Amarasinghe, J. J., Reyes-Diaz, L., Mantis, N. J., and Wade, J. T. (2012) FRUIT, a scar-free system for targeted chromosomal mutagenesis, epitope tagging, and promoter replacement in *Escherichia coli* and *Salmonella enterica*. *PLoS One* 7.

(192) Reisch, C. R., and Prather, K. L. J. (2015) The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*. *Sci. Rep.* 5, 1–12.





## 7. Appendix

### 7.1 Declaration of contributions

Unless stated otherwise the following work presented in this thesis was carried out by me under the supervision of Dr. Frank Bernhard and Prof. Dr. Volker Dötsch at the Institute of Biophysical Chemistry, Goethe University Frankfurt. All contributions from colleagues are explicitly acknowledged in the following section.

- Figure 7F Experiment design, protein expression, purification, and sample preparation were carried out by me. LILBID-MS measurements and data plotting were carried out by Janosch Martin and (Morgner lab, Institute of Physical and Theoretical Chemistry Goethe University).
- Figure 7G Experiment design and sample preparation was carried out by me. LILBID-MS measurements and data plotting were carried out by René Zangl (Morgner lab, Institute of Physical and Theoretical Chemistry Goethe University).
- Figure 8B Experiment design and detergent-free SapNP formation were carried out by me. LILBID-MS measurements and data plotting were carried out by Janosch Martin (Morgner lab, Institute of Physical and Theoretical Chemistry Goethe University).
- Figure 9F/G Experiment design, sample preparation and SEC analysis were carried out by me. LILBID-MS measurements and data plotting were carried out by René Zangl (Morgner lab, Institute of Physical and Theoretical Chemistry Goethe University).
- Figure 10 Experiment design and sample preparation were carried out by me. Negative stain EM data acquisition was carried out by Theresa Gewering (Möller Lab, Department of Structural Biology at the Max-Planck Institute for Biophysics, Frankfurt)
- Figure 13C Experiment design and sample preparation were carried out by me. LILBID-MS measurements and data plotting were carried out by Janosch Martin (Morgner lab, Institute of Physical and Theoretical Chemistry Goethe University).
- Figure 14D Experiment design and sample preparation were carried out by me. LILBID-MS measurements and data plotting were carried out by René Zangl (Morgner lab, Institute of Physical and Theoretical Chemistry Goethe University).
- Figure 15D Experiment design and sample preparation were carried out by me. LILBID-MS measurements and data plotting were carried out by

- Janosch Martin (Morgner lab, Institute of Physical and Theoretical Chemistry Goethe University).
- Figure 17E Experiment design and sample preparation were carried out by me. LILBID-MS measurements and data plotting were carried out by Janosch Martin (Morgner lab, Institute of Physical and Theoretical Chemistry Goethe University).
- Figure 18C Experiment design and sample preparation were carried out by me. Radioligand binding was measured by Zoe Köck (Dötsch/Bernhard lab, Institute of Biophysical Chemistry, Goethe University)
- Figure 19C Experiment design and sample preparation were carried out by me. Radioligand binding was measured by Zoe Köck (Dötsch/Bernhard lab, Institute of Biophysical Chemistry, Goethe University).
- Figure 22B Experiment design was layed out by me. Deletion of *glnA* was carried out by my bachelor student Betül Karakoc (Dötsch/Bernhard lab, Institute of Biophysical Chemistry, Goethe University).
- Table 5 Experiment design was layed out by me. *In vivo* modification of *glnA* with a C-terminal his-tag in *E. coli* A19 M1 resulting in *E. coli* A19 M1-6 was carried out by my bachelor student Betül Karakoc (Dötsch/Bernhard lab, Institute of Biophysical Chemistry, Goethe University).
- Figure 25A-D Experiment design and labeled CypD samples were prepared by me. NMR measurements were carried out by Frank Löhr (Dötsch/Bernhard lab, Institute of Biophysical Chemistry, Goethe University).
- Figure 25B Experiment design and preparation of labeled PR sample were carried out by me. NMR measurements were carried out by Frank Löhr (Dötsch/Bernhard lab, Institute of Biophysical Chemistry, Goethe University).
- Figure 25C Experiment design and GFP expression with preconverted precursor MOV were carried out by me. Synthesis of labeled precursor MOV was carried out by Roman J. Lichtenecker (Institute of Organic Chemistry, University of Vienna)
- Figure 25D Experiment design and preparation of labeled PR sample were carried out by me. NMR measurements were carried out by Frank Löhr (Dötsch/Bernhard lab, Institute of Biophysical Chemistry, Goethe University). Synthesis of labeled precursor MOV was carried

out by Roman J. Lichtenecker (Institute of Organic Chemistry, University of Vienna)

Figure 28A/B Experiment design and precursor conversion were carried out by me. NMR measurements were carried out by Frank Löhr (Dötsch/Bernhard lab, Institute of Biophysical Chemistry, Goethe University).

Figure 29 Experiment design and protein labeling were carried out by me. NMR samples were measured by Frank Löhr who also gave labeling advice (Dötsch/Bernhard lab, Institute of Biophysical Chemistry, Goethe University).



## 7.2 Curriculum vitae

### Roman Levin

Date of birth:

Place of birth: 16.01.1991

St. Petersburg, Russia



### EDUCATION

Since 12/2018	<p><b>PhD in Biochemistry</b>, Research group of Prof. Dr. Volker Dötsch, Institute of Biophysical Chemistry Goethe University   Frankfurt, Germany <b>Title:</b> Customization of cell-free lysates for membrane protein expression and NMR analysis</p>
10/2014-01/2018	<p><b>M.Sc. in Food Microbiology and Biotechnology</b> in the group of Prof. Dr. Lutz Fischer, department of Biotechnology and Enzyme Science University of Hohenheim   Stuttgart, Germany  <b>Thesis title:</b> Recombinant expression of a cellobiose 2-epimerase in <i>Yarrowia lipolytica</i> PO1f.</p>
10/2011-01/2014	<p><b>B. Sc. in Food Science and Biotechnology</b> in the group of Prof. Dr. Lutz Fischer, department of Biotechnology and Enzyme Science University of Hohenheim   Stuttgart, Germany  <b>Thesis title:</b> Partial purification and characterization of a laccase from <i>Trametes versicolor</i> and application for crosslinking of milk protein.</p>
09/2002-06/2011	<p><b>Secondary School</b> Paracelsus Gymnasium Hohenheim   Stuttgart, Germany German university entrance level</p>

### PROFESSIONAL EXPERIENCE

Since 12/2018	<p><b>Research Scientist</b> in the group of Prof. Dr. Volker Dötsch and Dr. Frank Bernhard, department of Biophysical Chemistry Goethe University   Frankfurt, Germany</p>
06/2017-08/2018	<p><b>Master Thesis</b> in the group of Prof. Dr. Lutz Fischer, department of Biotechnology and Enzyme Science University of Hohenheim   Stuttgart, Germany</p>
08/2017-09/2017	<p><b>Research assistant</b> in the group of Prof. Dr. Lutz Fischer, department of Biotechnology and Enzyme Science University of Hohenheim   Stuttgart, Germany</p>
04/2016-09/2016	<p><b>Research assistant</b> in the group of Prof. Dr. Bo Jiang, State Key Laboratory of Food Science and Technology Jiangnan University   Wuxi, PR China</p>
05/2015-07/2015	<p><b>Tutor of organic chemistry</b> in the group of Prof. Dr. Uwe Beifuß, department of Bioorganic Chemistry University of Hohenheim   Stuttgart, Germany</p>

**PUBLICATIONS**

- 
- 2023 **Levin, R.**, Löhr, F., Karakoc, B., Lichtenecker, R., and Bernhard, F. (2023) *E. coli* “ Stablelabel ” S30 lysate for optimized cell-free NMR sample preparation *J. Biomol. NMR* 77, 131–147.
- 2022 **Levin, R.**, Köck, Z., Martin, J., Zangl, R., Gewering, T., Schüler, L., Moeller, A., Dötsch, V., Morgner, N., and Bernhard, F. (2022) Cotranslational assembly of membrane protein/nanoparticles in cell-free systems. *Biochim. Biophys. Acta - Biomembr.* 1864, 184017.
- Umbach, S., **Levin, R.**, Neumann, S., Steinmetzer, T., Dötsch, V., and Bernhard, F. (2022) Transfer mechanism of cell-free synthesized membrane proteins into mammalian cells. *Front. Bioeng. Biotechnol.* 10, 1–17.
- 2020 **Levin, R.**, Koeck, Z., Dötsch, V., and Bernhard, F. (2020) Co-translational insertion of membrane proteins into preformed nanodiscs. *J. Vis. Exp.* 165, 1–27.
- 2016 Chen, Q., **Levin, R.**, Zhang, W., Zhang, T., Stressler, T., Fischer, L., Mu, W. (2017) Characterization of a novel cellobiose-2-epimerase from thermophilic *Caldicellulosiruptor obsidiansis* for lactulose production. *J. Sci. Food Agric.* 97, 3095-3105

**SKILLS**


---

Biochemistry	Protein expression, cell-free expression, purification, protein refolding, nanoparticles, reconstitution, affinity chromatography, size exclusion chromatography, ion exchange chromatography, hydrophobic interaction chromatography, SDS-PAGE, western blotting, pulldown assays, dialysis
Molecular biology	PCR, cloning, strain engineering, $\lambda$ -bacteriophage based recombination in <i>E. coli</i> and CRISPR-Cas9 assisted recombination in yeast.
Biophysics	Fluorescence anisotropy, Radioligand binding assays
Informatics	MS Office, Adobe Photoshop, Adobe Illustrator, GraphPad prism, Image J, PyMOL, SPARKY, TopSpin, ChemDraw, Serial Cloner, Snappgene
Languages	German (native), Russian (native), English (fluent)

---

**PRESENTATIONS AND SCHOLARSHIPS**

---

2022	International Workshop on Advances Isotopic Labelling Methods. Grenoble, France. Speaker
2022	International Workshop on Advances Isotopic Labelling Methods. Grenoble, France. Poster presenter
2016	Scholarship from the Ministry of Science, Research and Arts Baden Württemberg for a research stay at the University of Jiangnan in Wuxi, PR China.

**ACADEMIC TEACHERS**

---

Bernhard, F., Dötsch V., Seitzl, I., Fischer, L., Glück, C., Swietalski, P., Beifuß, U., Lupo, D. Mu, W., Chen, Q., Stressler, T., Kuhn, A., Conrad, J., Hausmann, R., Schmidt, H., Carle, R. Lutz-Wahl, S., Kölling R. Kuschel, B. Claaßen, W.

**FURTHER OCCUPATIONS**

---

02/2022 -	Volunteer and translator for Ukrainian war refugees
04/2011-06/2018	Nursing assistant at the Association of the Physically Disabled Stuttgart (Körperbehinderten Verein Stuttgart e.V.)



Publiziert unter der Creative Commons-Lizenz Namensnennung (CC BY) 4.0 International.  
Published under a Creative Commons Attribution (CC BY) 4.0 International License.  
<https://creativecommons.org/licenses/by/4.0/>