Investigation of co-translational protein folding using cryo-EM and solid-state NMR enhanced by DNP

Dissertation
zur
Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt am Fachbereich Biochemie, Chemie und Pharmazie
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von

Linda Schulte
geboren in Darmstadt

Frankfurt am Main 2020
(D30)

Für meine Familie

Contents

Summary

Protein folding occurs in a co-translational manner while the peptide chain passes through the ribosomal exit tunnel. The electrostatic potential of the exit tunnel is mainly negative [155] and positively charged amino acids, including arginine and lysine can cause translational pausing due to increasing interactions with the tunnel surface [154]. However, not only the surface of the ribosomal exit tunnel modulates the translational speed. Translation is a non-uniform process and the genetic code is degenerated with up to six codons encoding a single amino acid. The distribution of these synonymous codons is not random and they are used with varying frequencies within an open reading frame. Codons with higher tRNA abundance are translated faster than those infrequently used [129]. Rare codons are often located between domain boundaries or secondary structural elements leading to the separation of folding events [129]. That the exchange of synonymous codons are not silent mutations was addressed in several studies [5, 22, 126, 282]. Buhr *et al.* [22] demonstrated that synonymous codon mutations effected the translation kinetics and unexpectedly influenced the protein conformation of the bovine eye lens protein γB crystallin (GBC). GBC protein (U) expressed from the unmodified sequence of *B. taurus* in *E. coli* was fully reduced, while the faster translated harmonized GBC protein (H) adapted to the codon usage of *E. coli* was partly (58%) oxidized in the N-terminal domain (NTD) [22]. This observation was the basis for this thesis. We wondered if disulfide bond formation in GBC can occur co-translationally and likely even within the ribosomal exit tunnel.

Disulfide bond formation is considered a post-translational modification and usually restricted to distinct compartments within the cell. The *E. coli* cytosol is a highly reductive environment due to the tripeptide glutathione (GSH) and glutathione disulfide (GSSG) and the oxidoreductases glutaredoxin (Grx) and thioredoxin (Trx). Thus, heterologously expressed proteins in the *E. coli* cytosol are typically found in the reduced state, although exceptions have been reported [8, 230, 111, 10]. Furthermore, it was shown that disulfide bond formation in proteins can indeed occur during protein translation, while the protein is still bound to the ribosome [90, 212, 211].

The observed differences in the oxidation state of GBC are the result of variations in the translation kinetic. The three cysteines Cys18, Cys22 and Cys78 build a cluster and are located within a 5.4-6.4 Å distance of each other (Figure Ia). A disulfide bond between Cys18 and Cys22 was previously reported [177]. However, the torsion angle (χ3) between Cys22 and Cys78, which is defined by the rotation of the two Cβ atoms about the disulfide bond is around ±90°, a value commonly obtained in native disulfide bonds [36]. In this thesis the following questions were addressed:

I) Is disulfide bond formation already observed in the NTD alone?

II) Does the ribosomal tunnel provide sufficient space for disulfide bond formation in GBC?

III) Is oxidation of the nascent polypeptide chain (NC) detectable in ribosome nascent chain complexes (RNCs)?

IV) Can disulfide bond formation be visualized inside the ribosomal exit tunnel?
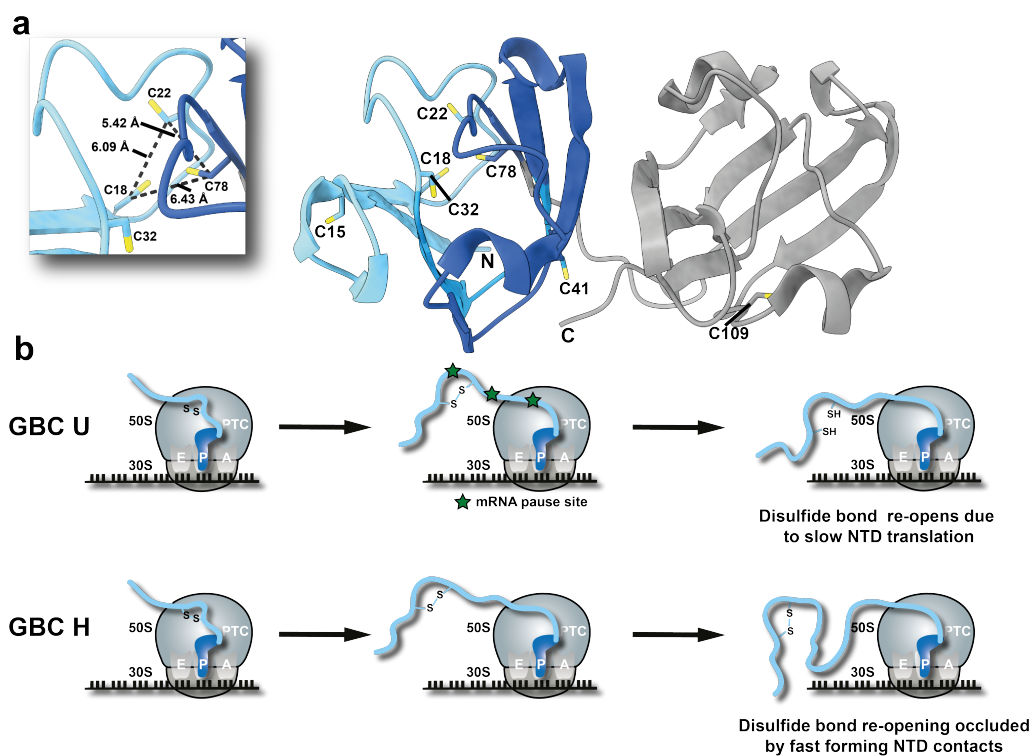
Figure I: *Structure of GBC and hypothesis of oxidation occurrence in H. a)*
GBC structure with detailed view on Cys18, Cy22 and Cys78 and distances
between these Cys residues. Color according to RNC chain length. Number-
ing excludes the N-terminal methionine of GBC to be consistent with other
published studies. Residues 1-32 (light blue), 33-43 (blue), 44-78 (dark blue)
and 79-174 (grey) (PDB: 4W9B). **b)** *Hypothesis explaining the conforma-*
tional differences in gene variants U and H of GBC. A disulfide bond within
the ribosomal tunnel or at an early state of translation is formed in both gene
variants. The disulfide bond is reduced by the E. coli cytosol in the slow trans-
lated version U, while this bond is protected by fast-forming NTD contacts or
the CTD in the gene variant H.

To study whether the cysteine oxidation exists during translation of the NTD, a single domain construct (amino acids 1-82) was studied. This construct did not contain the six-residue peptide linker between the two domains, but the most translational pausing was observed in the translation of the first 70 amino acids [22]. The 2D ^{1}H-^{15}N BEST-TROSY spectrum displayed well-dispersed signals

indicating a folded protein structure. Thus, the CTD was not required for the folding of the NTD. Addition of DTT to both gene variants did not show a severe effect. In contrast to the full-length protein, where the H variant was partly oxidized, the H variant of the NTD was completely reduced. Only in the U protein, dimer formation was observed twice and could be removed by addition of DTT. The dimer was formed including Cys41, Met43, Phe56 and Ile81, the latter three are residues of the hydrophobic patch and are responsible for inter-domain interactions [240]. The cysteine oxidation in the NTD of the full-length GBC might therefore occur at a later state of translation after the expression of the inter-domain peptide linker or parts of the CTD. In addition, we wondered if disulfide bond formation could also happen inside the ribosomal exit tunnel or at an early state of translation when the NC is still attached to the ribosome. Both gene variants might form an initial disulfide bond and in the slow translated U variant this disulfide bond gets reduced when exposed to the reductive environment of the *E. coli* cytosol, whereas in the faster translated H variant this disulfide bond is protected by the upcoming CTD (Figure Ib). To verify whether, disulfide bond formation is indeed possible inside the ribosomal exit tunnel, RNCs with arrested GBC fragments were investigated using various techniques, including theoretical simulations, solid-state NMR enhanced by DNP, mass spectrometry and cryo-EM.

Theoretical simulation using flexible-meccano [14] demonstrated that the ribosomal tunnel is indeed large enough for different disulfide bond formations. In an U32SecM construct, including the first four cysteine residues of GBC and the SecM stalling sequence all possible disulfide bonds were detected. This included a disulfide bridge between Cys15 and Cys18, which are part of the common CXXC motif found in oxidoreductases and the disulfide bridge between Cys18 and Cys22, which was observed in a X-ray structure [177]. But also, a disulfide bond between Cys22 and Cys32 can easily be formed inside the ribosomal exit tunnel.

Using selectively ^{13}C, ^{15}N cysteine labeled RNCs, the chemical shift of cysteine residues of the NC was detected by solid-state NMR enhanced by DNP. Selectively labeled RNCs were studied for the first time using solid-state NMR. We could show that no isotopic scrambling into ribosomal proteins occurred. Three different lengths of RNCs, containing four (U32SecM), five (U43SecM) or six (U78SecM) cysteine residues were investigated. The four-cysteine RNC construct was further mutated to contain either only one cysteine residues (Cys18) or the two cysteines Cys15 and Cys18 or Cys18 and Cys22. In all RNCs, even in the single cysteine construct, Cβ chemical shifts above 35 ppm were detected, indicating oxidation of the NC (Figure IIa). Hence, an inter-molecular disulfide bond was formed.

In addition, an α-helix was observed in the shortest construct U32SecM using solid-state NMR (Figure IIb). LC-MS/MS revealed the occurrence of S-glutathionylated and S-nitrosylated cysteine (Figure IIIa). Both modifications
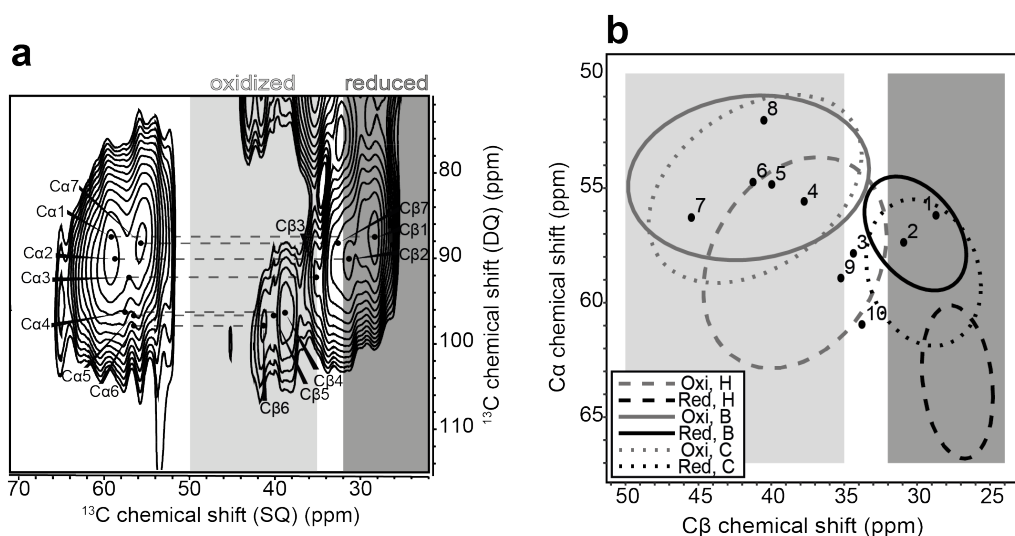
Figure II: *Spectrum of U32SecM C18 and structural analysis of U32SecM. a)*
Cysteine Cα-Cβ cross-peak region of the ^{13}C-^{13}C 2D spectrum of U32SecM
C18. Cβ chemical shifts above 35 ppm are oxidized (light grey) and below
32 ppm reduced (dark grey). b) Cα and Cβ chemcial shifts from U32SecM
cryo-EM sample compared to values derived from a chemical shift database
[261].

are considered as post-translational modifications and can be mediated by GSH
or S-nitrosoglutathione (GSNO) [76, 2]. To the best of our knowledge, these
modifications have not previously been detected in NCs inside or otutside of
the ribosomal tunnel. Although, S-glutathionylated GBC was formerly reported
[148, 277, 248]. Disulfide bond formation was also observed using LC-MS/MS,
including an intra-molecular disulfide bridge between Cys32 and Cys41 in the
U43SecM and U78SecM construct. In the five cysteine construct (U43SecM) this
disulfide bond is located inside and in the U78SecM outside of the ribosomal
exit tunnel. Furthermore, a disulfide bond between Cys15 and Cys32 as well as
Cys22 and Cys32 were detected by LC-MS/MS in the U32SecM construct. Here,
a trypsin cleavage site between the peptides made it impossible to distinguish
between an intra- and an inter-molecular disulfide bond, which could be formed
after release of the ribosome.

The shortest construct U32SecM was further investigated by cryo-EM. We were
able to obtain a 2.53 Å structure with the Gly-tRNA located inside the P-site
and a 3.19 Å structure with two tRNAs bound. 35 amino acids were modeled
into the electron density of the NC. The distance between the first and the last
amino acid of the tunnel was 65 Å resulting in 1.87 Å per residue and indicating
a compact conformation.

An α-helix was formed between Phe150 of SecM and Tyr28 of GBC including one
cysteine residue Cys32. This observation was in line with solid-state NMR results,
showing cysteine signals in an α-helical conformation (Figure IIb). Furthermore,
a disulfide bridge between Cys22 and Cys32 was visible. The non-native α-helix
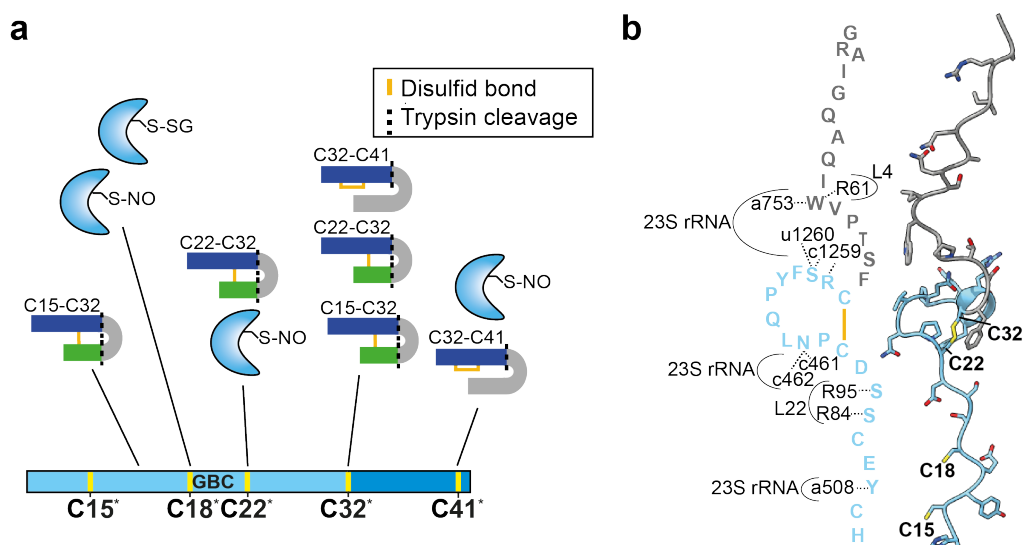
Figure III: *Cysteine modifications detected by LC-MS/MS and model of NC. a)* *Overview of cysteine oxidation states (S-nitrosylation (S-NO), S-glutathionylation (S-SG), disulfide bonding) of the NCs in RNCs as observed by LC-MS/MS.* *b)* *Model of NC with interactions to ribosomal proteins and 23S rRNA. SecM in grey, GBC in light blue.*

is stabilized by interactions to the ribosomal protein L22 and the 23S rRNA, bringing the cysteine residues in close proximity to form a disulfide bond (Figure IIIb). Not only the interactions with the ribosomal tunnel surface, but also the limited access of Trx and Grx to the NC might favor disulfide bond formation inside the ribosomal exit tunnel.

The coupling of disulfide bond formation and protein folding is described by two models: in the folded precursor model, cysteine residues are brought into proximity after the structure is fomed. In the quasi-stochastic model cysteines pair in an unfolded conformation and influence subsequent protein folding [265]. Our observation is in favor of the quasi-stochastic model, since non-native disulfide bonds were detected. The disulfide bond between Cys22 and Cys32 might be shuffeled to Cys32 and Cys41 with further elongation of the NC leading to the detected disulfide bridge in U43SecM and U78SecM with LC-MS/MS.

Some proteins could use early formed disulfide bonds in the ribosomal exit tunnel to escape the highly reductive environment of the cytoplasm. Non-native disulfide bonds, however, need to be reduced and refolded to obtain their native structure. Oxidation of the isolated, full-length GBC with Cu(II) revealed that most likely a disulfide bond between Cys22 and Cys78 is formed in this protein. Hence, the observed difference in the cysteine oxidation state might be either the result of differences in the translation kinetics when translating the CTD or the observed non-native disulfide bond in the exit tunnel is reshuffeled and then reduced to a different extent depending on the translation kinetics. To clarify this question further experiments with longer RNC constructs and isolated proteins including

also residues of the CTD will be required.

Zusammenfassung

Proteinfaltung findet kotranslational statt, während die Peptidkette durch den ribosomalen Exit-Tunnel geschleust wird. Das elektrostatische Potenzial des Tunnels ist hauptsächlich negativ [155] und positiv geladene Aminosäuren wie Arginine und Lysine interagieren verstärkt mit der Tunneloberfläche, was zu einer Verlangsamung der Translationsgeschwindigkeit führt [154]. Aber nicht nur die Interaktion mit der Tunneloberfläche ist entscheidend für die Translationsgeschwindigkeit. Die Translation ist ein nicht-uniformer Prozess und der genetische Code degeneriert mit bis zu sechs Codons, die eine einzelne Aminosäure kodieren. Die Verteilung dieser synonymen Codons ist nicht zufällig und diese werden mit verschiedenen Frequenzen innerhalb eines offenen Leserahmens verwendet. Codons mit einer höheren tRNA-Häufigkeit werden schneller eingebaut als seltene Codons [129]. Diese seltenen Codons sind häufig zwischen Proteindomänen oder Sekundärstrukturelementen platziert und könnten daher zur Separierung von Faltungsevents dienen [129].

Dass der Austausch von synonymen Codons nicht ohne Folgen ist, zeigten verschiedene Studien [5, 22, 126, 282]. Buhr *et al.* [22] zeigte, dass der synonyme Austausch die Translationsgeschwindigkeit, aber auch die Proteinkonformation des bovinen Augenlinsenproteins γB crystallin (GBC) beeinflusst. Während die unmodifizierte Gensequenz aus *B. taurus* in *E. coli* langsamer translatiert wurde und zu einem vollständig reduzierten GBC Protein (U) führte, wurde die harmonisierte Genvariante, die der Codon-Verwendung in *E. coli* angepasst war, schneller exprimiert und resultierte in einem teilweise oxidierten (58%) GBC Protein (H) [22]. Dieser Befund war der Ausgangspunkt für die vorliegende Doktorarbeit in der untersucht werden soll, ob Disulfidbrücken kotranslational und/oder womöglich sogar im ribosomalen Exit-Tunnel entstehen können.

Disulfidbrücken sind post-translationale Modifikationen, die normalerweise in bestimmten Kompartimenten der Zelle gebildet werden. Das *E. coli* Zytoplasma, in dem die Proteinbiosynthese stattfindet, ist eine extrem reduzierende Umgebung mit hohen Konzentrationen des Tripeptids Glutathion (GSH) und Glutathiondisulfid (GSSG) sowie der Oxidoreduktasen Glutaredoxin (Grx) und Thioredoxin (Trx). Somit sind heterolog exprimierte Proteine normalerweise vollständig reduziert und verfügen über keine Disulfidbrücken. Allerdings sind auch Ausnahmen zu dieser Regel bekannt [8, 10, 111, 230]. Ebenfalls konnten Disulfidbrücken innerhalb der gebundenen Proteinkette am Ribosom detektiert werden, also während der Translation des Proteins [90, 211, 212].

Die gemessenen Oxidationsunterschiede basieren auf der unterschiedlichen Translationsgeschwindigkeit der beiden Gensequenzen. Die N-terminale Domäne (NTD) des Zweidomänen-Proteins GBC enthält sechs der insgesamt sieben Cysteinreste. Nur in dieser Domäne wurde Oxidation detektiert und die drei Cysteine Cys18,

Cys22 und Cys78 weisen einem Abstand von 5.4-6.4 Å zueinander auf (Abbildung IVa). Eine Disulfidbrücke zwischen Cys18 und Cys22 wurde in einer Kristallstruktur gemessen [177]. Der Torsionswinkel ($\chi3$) zwischen den zwei Cβ-Atomen von Cys22 und Cys78 liegt jedoch im Bereich von $\pm 90°$, der in vielen nativen Disulfidbrücke angenommen wird [36]. In der vorliegenden Doktorarbeit wurden die folgenden Fragen untersucht:

I) Werden Disulfidbrücken bereits in der eigenständigen NTD gebildet?

II) Ist der ribosomale Exit-Tunnel groß genug, um die Ausbildung möglicher Disulfidbrücken in GBC zu ermöglichen?

III) Kann Oxidation in der naszierenden Polypeptidkette am Ribosom detektiert werden?

IV) Kann eine Disulfidbrücke innerhalb des ribosomalen Exit-Tunnels visualisiert werden?
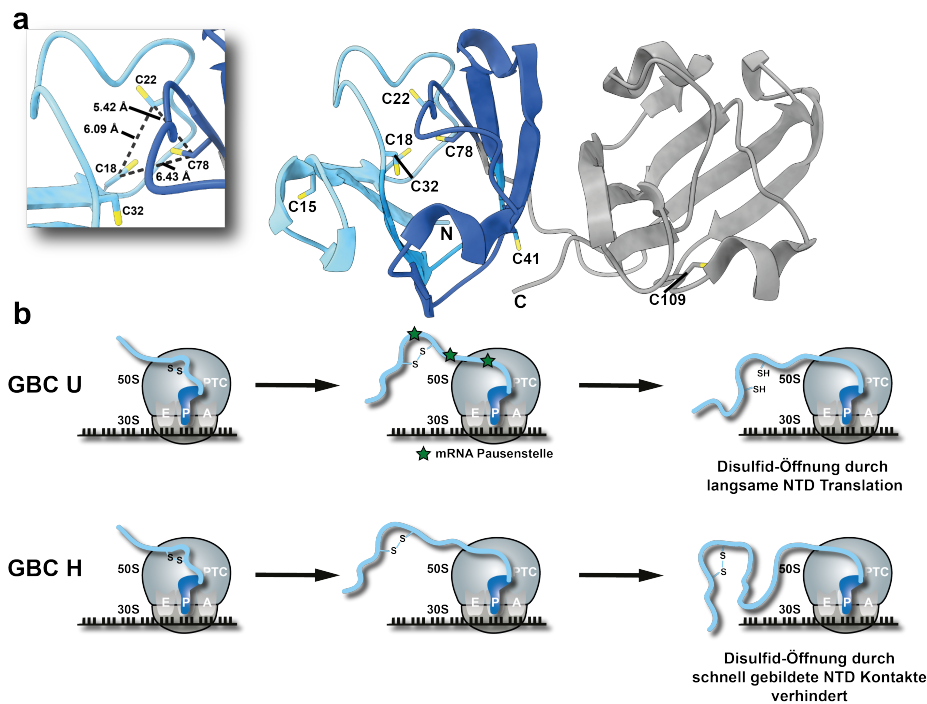
Abbildung IV: *GBC Struktur und Hypothese zur Oxidation in GBC Variante H. a) GBC Struktur des Volllängen-Proteins mit Detailansicht auf Cys18, Cys22 und Cys78 sowie der Abstände zwischen diesen. Die Farbgebung entspricht der unterschiedlichen RNC-Konstrukte. Reste 1-32 (hellblau), 33-43 (blau), 44-78 (dunkelblau) und 79-174 (grau). b) Hypothese zur Erklärung der konformationellen Unterschiede zwischen den Genvarianten U und H von GBC. Eine Disulfidbrücke kann im ribosomalen Exit-Tunnel oder früh während der Translation gebildet werden. Im langsam translatierenden U wird die Disulfidbrücke reduziert, während diese in der schnell translatierenden Variante H durch N-terminalen Kontakte geschützt wird.*

Um zu untersuchen, ob die Unterschiede bereits während der Translation der NTD ausgebildet werden, wurde ein Ein-Domänen-Konstrukt hergestellt. Dieses Konstrukt beinhaltete die Aminosäuren 1-82, aber nicht den Peptidlinker zwischen den beide Domänen. Allerdings wurden die meisten Translationspausen bei der Translation der ersten 70 Aminosäuren detektiert [22]. Das 2D ^{1}H-^{15}N BEST-TROSY-Spektrum wies anhand der unterschiedlichen chemischen Verschiebung der Signale auf eine gefaltete Proteinstruktur hin. Daher konnte sich die NTD ohne Beteiligung der CTD eigenständig falten. Die Zugabe von DTT zu beiden Proteinvarianten U und H führte zu keinem messbaren Effekt. Im Gegensatz zu dem Volllängen-Protein, indem die Variante H teilweise oxidiert war, war die NTD der Variante H vollständig reduziert. Nur in der U-Variante konnte zweimal Dimerausbildung detektiert werden, die durch die Zugabe von DTT teilweise aufgehoben werden konnte. Entscheidend an der Dimerbildung beteiligt waren die Aminosäuren Cys42, Met43, Phe56 und Ile81. Die letzteren drei sind ebenfalls an der hydrophoben Interaktion mit der CTD involviert [240]. Dadurch, dass kein Unterschied zwischen den beiden Proteinen festgestellt werden konnte, könnte der Peptidlinker sowie Teile der CTD entscheidend für die Cysteinoxidation im Volllängen-Protein sein.

Zusätzlich wurde untersucht, ob die Ausbildung von Disulfidbrücken im Inneren des Ribosoms möglich ist. Dann könnte sich in beiden Genvarianten eine anfängliche Disulfidbrücke bilden und durch die unterschiedliche Translationsgeschwindigkeit die Disulfidbrücke in der langsamen Genvariante (U) im *E. coli* Zytosol reduziert werden, während diese in der schneller translatierten Variante H von der CTD geschützt wird (Figure IVb). Um diese Hypothese zu untersuchen, wurden GBC-Fragmente an das Ribosom arretiert und diese Ribosom-naszierenden Polypeptidketten-Komplexe (RNCs) mit verschiedenen Techniken, einschließlich theoretischer Simulation, Festkörper-NMR, Massenspektrometrie und cryo-EM, analysiert. Theoretische Simulation mittels flexible-mecanno [14] zeigten, dass der ribosomale Tunnel groß genug für die Ausbildung verschiedenster Disulfidbrücken ist. In einem U32SecM-Konstrukt, das vier Cysteine und die SecM-Sequenz beinhaltet, konnten alle theoretisch möglichen Disulfidbrücken gebildet werden. Dies beinhaltet eine Disulfidbrücke zwischen Cys15 und Cys18, die beide Teil des CXXC Motifs sind, welches häufig in Oxidoreduktasen vorkommt, sowie die Disulfidbrücke zwischen Cys18 und Cys22, die in einer Kristallstruktur gemessen wurde [177]. Aber auch eine Disulfidbrücke zwischen Cys22 und Cys32, die sich nicht im nativen Volllängen-GBC formen kann, konnte innerhalb des ribosomalen Tunnels gebildet werden.

Mithilfe von selektiv ^{13}C, ^{15}N Cystein markierten RNCs konnte die chemische Verschiebung der naszierenden Peptidkette innerhalb und außerhalb des Ribosoms mittels Festkörper-NMR gemessen werden. Zum ersten Mal wurden selektiv markierte RNCs mittels Festkörper-NMR untersucht. Drei unterschiedlich lange RNCs mit vier (U32SecM), fünf (U43SecM) und sechs Cysteinen (U78SecM)

wurden verwendet. Zusätzlich wurde das Vier-Cystein-Konstrukt (U32SecM) mutiert, sodass es entweder nur ein Cystein (Cys18) oder aber die zwei Cysteine Cys15 und Cys18 oder Cys18 und Cys22 enthielt. In allen RNC-Konstrukten, auch in der Ein-Cystein-Variante wurden Cβ Signale oberhalb von 35 ppm gemessen, was auf Cysteinoxidation und in diesem Fall auf die Ausbildung einer inter-molekularen Disulfidbrücke schließen lässt (Abbildung Va).
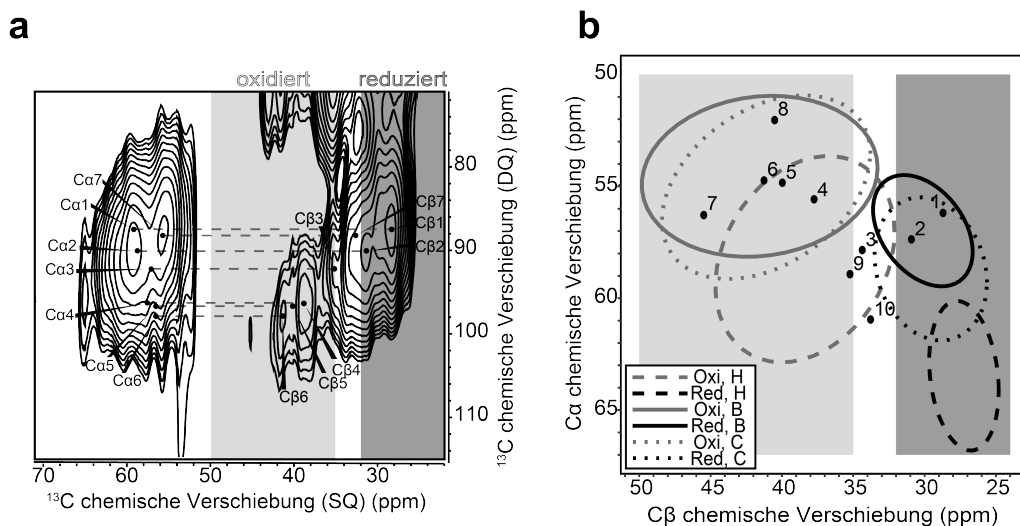
Abbildung V: *Spektrum von U32SecM C18 und strukturelle Analyse der Cystein-Signale von U32SecM. **a)** Cystein Cα-Cβ Kreuzpeak-Region des ^{13}C-^{13}C DQ-SQ-Spektrums von U32SecM C18. Cβ Signale oberhalb von 35 ppm sind oxidiert (hellgrau) und unterhalb von 32 ppm reduziert (dunkelgrau). **b)** Cα und Cβ chemische Verschiebungen der U32SecM cryo-EM Probe verglichen mit Werten der chemischen Verschiebungs-datenbank [261].*

Zudem wurden im kürzesten U32SecM Konstrukt Cysteinsignale innerhalb einer α-helikalen Struktur mittels Festkörper-NMR detektiert (Abbildung Vb). LC-MS/MS-Messungen detektierten S-glutathionylierte und S-nitrosylierte Cysteine (Abbildung VIa). Beide Cystein-Modifikationen sind post-translationale Mod-ifikationen und können durch GSH oder S-Nitrosoglutathione (GSNO) [76, 2] gebildet werden. Uns ist nicht bekannt, dass diese Modifikationen bis dato in naszierenden Polypeptidketten innerhalb des ribosomalen Tunnels oder auch außerhalb gemessen worden sind. An GBC gebundenes GSH wurde allerdings bereits identifiziert [148, 277, 248]. Mittels LC-MS/MS konnten ebenfalls intra-molekulare Disulfidbrücken zwischen Cys32 und Cys41 in dem U43SecM- und U78SecM-Konstrukt detektiert werden. Während die Disulfidbrücke im ersten Konstrukt innerhalb des ribosomalen Tunnels lag, befand sie sich im U78SecM Konstrukt außerhalb des ribosomalen Tunnels. Ebenfalls, wurde in dem kürzesten U32SecM-Konstrukt eine Disulfidbrücke zwischen Cys15 und Cys32 sowie zwis-chen Cys22 und Cys32 gemessen. Da aber zwischen beiden Peptiden eine Trypsin-schnittstelle lag, konnte nicht zwischen inter- und intramolekularer Disulfidbrücke unterschieden werden. Daher wurde das kürzeste Konstrukt ebenfalls mittels

cryo-EM untersucht. Eine 2.53 Å Ribosomenstruktur mit einer gebundenen Gly-tRNA in der P-site und eine 3.19 Å Ribosomenstruktur mit zwei gebunden tR-NAs konnte gemessen werden. 35 Aminosäuren konnten in die Elektronendichte der NC modelliert werden. Der Abstand zwischen der ersten und der letzten Aminosäure betrug 65 Å, was 1.87 Å pro Aminosäure entspricht.
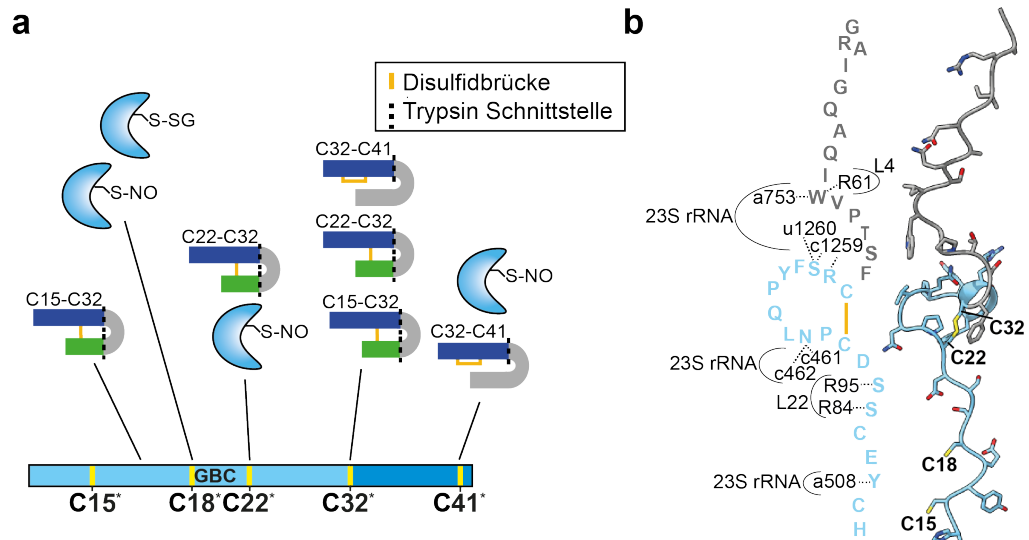
Abbildung VI: *Cysteinmodifikationen identifiziert durch LC-MS/MS und Modell der NC. a) Übersicht der mittels LC-MS/MS detektierten Cysteinoxidationszustände (S-Nitrosylierung (S-NO), S-Glutathionylierung (S-SG), Disulfidbrücke der naszierenden Polypeptidkette im Komplex mit dem Ribosom. b) Modell der naszierenden Polypeptidkette mit Interaktionen zu den ribosomalen Proteinen und der 23S rRNA. SecM in grau und GBC in hellblau.*

Dies weist auf eine kompakte Struktur der Peptidkette innerhalb des Ribosoms hin. Zwischen den Aminosäuren Phe150 aus SecM und Tyr28 aus GBC wurde eine α-Helix gebildet, die auch das Cys32 von GBC enthält. Außerdem wurde zwischen den beiden Cysteinen Cys22 und Cys32 eine Disulfidbrücke geformt. Die nicht native α-Helix wird durch Interaktionen zum ribosomalen Protein L22 sowie der 23S rRNA stabilisiert (Abbildung VIb). Durch die Ausbildung dieser α-Helix werden die beiden Cysteine Cys22 und Cys32 in räumliche Nähe gebracht und die Disulfidbrücke kann sich ausbilden. Allerdings spielen nicht nur die Interaktionen innerhalb des Ribosoms eine Rolle zur Bildung der Disulfidbrücke. Auch der limitierte Zugang der Oxidoreduktasen Grx und Trx führt dazu, dass die Disulfidbrücke nicht innerhalb des Tunnels reduziert werden kann. Die Kopplung zwischen Proteinfaltung und Bildung von Disulfidbrücken wird durch zwei Modelle beschrieben: Im Gefalteten-Vorgänger Modell bildet sich die Proteinstruktur zuerst aus und bringt die Cysteine in räumliche Nähe, sodass sich eine Disulfidbrücke formen kann. Im Quasi-Stochastischen Modell paaren sich die Cysteine in einer ungefalten Konformation, wodurch die weitere Proteinfaltung beeinträchtigt wird [265]. Die Entdeckung der nicht-nativen Disulfidbrücke im

ribosomalen Exit-Tunnel entspricht dem Quasi-Stochastischen Modell. Mit weiterer Elongation der naszierenden Polypeptidkette könnte die Disulfidbrücke zwischen Cys22 und Cys32 auf Cys32 und Cys41 übertragen worden sein, was somit zur Detektion dieser Disulfidbrücke mittels LC-MS/MS führt.

Einige Proteine könnten diese früh gefalteten Disulfidbrücken innerhalb des ribosomalen Exit-Tunnels nutzen, um dem reduzierenden Zytoplasma zu entkommen. Nicht-native Disulfidbrücken auf der anderen Seite müssen zunächst reduziert werden, damit das Protein seine native Struktur annehmen kann.

Durch Oxidation des isolierten Volllängen-GBC mit Cu(II) wurde sehr wahrscheinlich eine Disulfidbrücke zwischen Cys22 und Cys78 im Protein gebildet. Die Ausbildung dieser Disulfidbrücke, die vermutlich auch in dem teilweise oxidierten GBC H-Protein vorlag, basiert somit entweder auf der unterschiedlichen Translationsgeschwindigkeit, wenn die CTD translatiert wird oder eine nicht-native Disulfidbrücke wird zu einem unterschiedlichen Ausmaß auf andere Cysteine übertagen und reduziert. Um den genauen Prozess zu verstehen, der zur Ausbildung dieser Disulfidbrücken führt, müssen noch weitere längere Konstrukte der Polypeptidkette am Ribosom aber auch isolierte Proteine mit Sequenzen der CTD untersucht werden.

List of Abbreviations

°C	Degree Celsius
Å	Ångström
aa	Amino acid
a.u.	Arbitrary units
BEST	Band selective short transient
β–ME	β––Mercaptoethanol
BMRB	Biological Magnetic Resonance data bank
bp	base pairs
cryo-EM	cryogenic electron microscopy
Da	Dalton
DNA	Deoxyribonucleic acid
DNP	Dynamic nuclear polarization
DSS	Dimethyl-silapetane-sulfonate
ddH$_2$O	Miilipore water
DTT	Dithiothreitol
DQ	Double quantum
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
FAD	flavin adenine dinucleotide
FRET	Förster resonance energy transfer
g	Centrifugal force
g	gram
GBC	γB crystallin
γ–GCS	γ-glutamyl-cysteinesynthetase
GS	GSH-synthetase
Gor	Glutathione reductase
Grx	Glutaredoxin

GSH	Glutathione
GSSG	Glutathione disulfide
h	hour
His-tag	Polyhistidine-tag
hrs	hours
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High pressure liquid chromatography
HSQC	Heteronuclear single quantum coherence spectroscopy
IC	Initiation complex
IF	Initiation factor
IEX	Ion exchange chromatography
IPTG	Isopropyl β-d-1-thiogalactopyranoside
K	Kelvin
kbp	kilo base pairs
kDa	kilo Dalton
L	Liter
LB	Lysogeny broth
LC-MS	Liquid chromatography–mass spectrometry
LMW	Low molecular weight
min	minute
mL	milliliter
mM	millimolar
MWCO	Molecular weight cut-off
mRNA	Messenger RNA
NADPH	Nicotinamide adenine dinucleotide phosphate
Ni-NTA	Nitrilotriacetic acid
nm	nanometer
nmol	nanomol

NMR	Nuclear Magnetic Resonance
NS	Number of scans
NUS	Non-uniform sampling
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
pH	pondus hydrogenii
ppm	parts per million
PSSG	Protein mixed disulfide with glutathione
RBS	Ribosome-binding site
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RT	room temperature
RP	Reversed phase
rpm	revolutions per minute
S	Svedberg
SD	Shine-Dalgarno
SDM	Site directed mutagenesis
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SQ	Single quantum
Tris	Tris(hydroxylmethyl-)aminoethan
tRNA	Transfer RNA
TROSY	Transverse relaxation optimized spectroscopy
Trx	Thioredoxin
UV	Ultraviolet
Vis	Visible light

v/v   Volume per volume

w/v   Weight per volume

1 Introduction

1.1 Translation

Proteins are synthesized from the N- to C-terminus on the ribosome, a ribonucleoprotein complex, universal in all kingdoms of life. The ribosome consists of two subunits, the 30S and 50S subunit in prokaryotes and the 40S and 60S subunit in eukaryotes. Both subunits constitute the 70S prokaryotic ribosome and the 80S eukaryotic ribosome. The prokaryotic ribosome consists of three rRNAs and 54 r-proteins, with 21 r-proteins and the 16S rRNA within the small subunit and 33 r-proteins, the 5S and 23S rRNA in the large subunit [266]. The eukaryotic ribosome consists of 80 proteins and four rRNAs, 28S, 5.8S and 5S rRNA in the large subunit and the 18S rRNA in the small subunit. The translation can be divided into four parts: Initiation, Elongation, Termination and Recycling (Figure 1.1). The focus of this thesis chapter will be set to translation in bacteria, for additional information on the translation mechanism in eukaryotes the review of Wilson and Cate [266] is recommended. For most proteins, the rate limiting step of protein synthesis is the translation initiation.
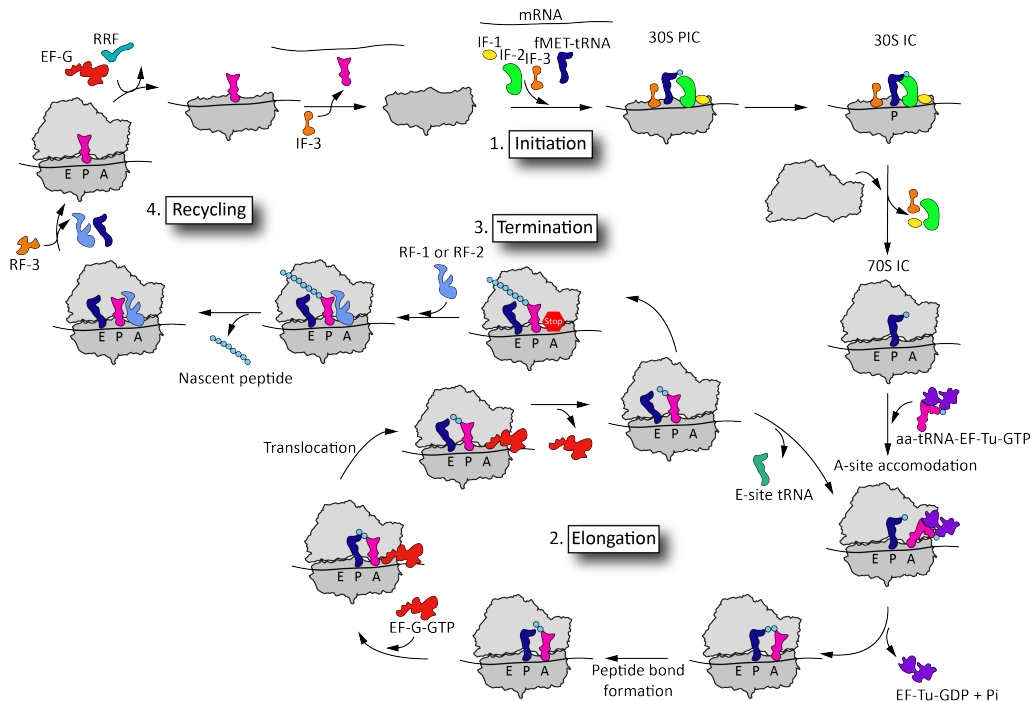
Figure 1.1: *Overview of bacterial protein biosynthesis.* *Translation is divided into Inititation, Elongation, Termination and Recycling. Aminoacyl-tRNA (aa-tRNA), elongation factor (EF), inititation factor (IF), release factor (RF), recycling factor (RRF) are labeled accordingly. Figure adapted from Javed et al. [115], Keiler [121], Schmeing and Ramakrishnan [225].*

1.1.1 Inititation

Briefly, during initiation the three initiation factors IF1-IF3, the initiator formylmethionyl-tRNA (fMet-tRNA) and mRNA bind to the 30S subunit of the ribosome and form the 30S pre-initiation complex (PIC). The mRNA interacts with its Shine-Dalgarno (SD) sequence (if present) with the 3' end of 16S rRNA anti-Shine-Dalgarno sequence. Binding of the initiator tRNA to the start codon, results in a conformational change leading to the 30S initiation complex (IC). Finally, assembly of the 50S subunit forms the 70S IC, which is promoted by the GTPase activity of IF-2 and accompanied by IF-3 release [225]. The elongation process can start after GTP hydrolysis and phosphate release from IF-2 and movement of the fMet-tRNA into the peptidyl transferase center (PTC) [225].

1.1.2 Elongation

Elongation is divided into A-site accommodation or decoding, in peptide bond formation and translocation. The elongation cycle starts with a peptidyl-tRNA in the P-site and an empty accommodation site (A-site). In a first step, the elongation factor (EF)-Tu carrying an aminoacyl-tRNA (aa-tRNA) and GTP binds to the A-site of the ribosome. After proofreading EF-Tu hydrolyzes GTP and the product EF-Tu GDP dissociates from the ribosome. A new peptide bond is formed in the PTC, due to a nucleophilic attack on the ester carbon of the peptidyl-tRNA in the P-site by the α-amino group of the aa-tRNA [213]. This reaction is catalysed within domain V of the 23S rRNA by binding the CCA ends of both aminoacyl- and peptidyl-tRNA [225]. Further, the ribosomal proteins L27 and L16 seem to support the RNA components to perform the peptidyl-transfer activity [225]. Before the A-site tRNA translocates to the P-site, the ribosome fluctuates between different pre-translocation states (PRE). In the classical PRE-state (PRE(C)), the peptidyl-tRNA remains in the A-site and the deacylated tRNA in the P-site. While in the hybrid PRE-state (PRE(H)) the 30S and 50S subunits have rotated by approximately 6 degrees to each other [67], which is called the ratcheted state. This rotation brings the tRNAs in hybrid position with the anticodons still in the A- and P-site on the 30S subunit but the 3' ends shifted towards the P- and E-site on the 50S [213]. EF-G binds to both PRE-states and stabilizes the PRE(H) state and hydrolyzes bound GTP to GDP, thereby bringing the two tRNAs into the A-P and P-E hybrid state. Pi is released and the tRNAs move to the E- and P-site, respectively. The E-site tRNA dissociates and EF-G is released. In contrast to the translation initiation, the elongation proceeds rather fast with an average incorporation of 13-22 aa per second in *E. coli* [21]. For further details of the elongation process, the reviews of Rodnina and Wintermeyer [213] and Schmeing and Ramakrishnan [225] are recommended.

1.1.3   Termination

Elongation proceeds until an mRNA stop codon moves into the A-site. In bacteria, the stop codons UAG and UAA are recognized by release factor 1 (RF-1), while RF-2 recognizes the stop codons UGA and UAA. Both release factors are also known as class I release factors. In eukaryotes, however only a single release factor eRF1 recognizes all three stop codons. The stop codons are recognized by a tripeptide motif PXT in RF-1 and SPF in RF-2 [109] and a conserved GGQ motif in both release factors catalyzes the hydrolysis of the peptidyl-tRNA bond [69, 245]. The class II release factor RF-3 promotes the dissociation of the class I release factors from the ribosome after peptide release by binding GTP which induces a conformational change [74].

1.1.4   Recycling

Before another protein can be synthesized, the ribosomes need to be recycled before reinitiation. RF-3 hydrolyzes GTP and dissociates. The recycling factor RRF and EF-G bind to the ribosome and GTP hydrolysis leads to dissociation of the 30S and 50S subunit [225, 121]. Binding of IF-3 to the 30S subunit removes mRNA and tRNA from the 30S subunit, preparing it for a new round of initiation [225, 121]. There are several mechanisms, which rescue ribosomes that are stalled in the middle of the mRNA, however two examples of ribosome stalling exist, which do not lead to mRNA cleavage and ribosome rescue, namely *secM* and *tnaC* [121]. These two arrest peptides will be explained in section 1.3.2 and 1.3.3.

1.2   Protein folding

After biosynthesis, proteins need to adapt a folded compact structure in order to function. Since most proteins can be unfolded to a random coil structure and refolded again, the information of how a protein finds its unique native structure is encoded within the amino acid sequence. However, a protein cannot simply try out all possible conformations until it obtains the right one. Assuming a protein with n residues and 2n backbone torsion angles, φ and ψ, could only adapt three conformations per torsion angle would already obtain $2^{3 \cdot n}$ conformations. For a small 100 amino acid protein this would result in 10^{89} conformations and if a protein samples one configuration in 10^{-13} seconds it would require 10^{66} years until all conformations were sampled. Since an average human life time in Germany is around 8×10^{1} years, almost no protein would have obtained its native conformations in a human body. These calculations are also known as the Levinthal paradox. Therefore, it became obvious that proteins require a folding pathway to exclude unfavorable conformations. One mechanism explaining the folding is the nucleation-condensation or nucleation-collapse mechanism in which a folding

nucleus forms, with short secondary structure elements like α-helices and β-sheets and the remaining structure then condenses first to domain-like structures and subsequently to the final tertiary structure [63]. Long range contacts and other native hydrophobic interactions stabilize weak secondary structure elements [39]. The folding of the protein from an isolated denatured conformation into its native structure is often described using the energy landscape perspective with a folding funnel (Figure 1.2a). Using the folding funnel, the conformational space accessible to the protein reduces until the native state is approached [269]. On the pathway to the native structure the number of native contacts (Q-values) as well as the number of total contacts (C) increases, while the free energy reduces during protein folding until the global energetic minimum is reached. The vertical axis of the folding funnel corresponds to the free energy, while the horizontal axis represents the conformations. The width of the funnel at a certain height incorporates the entropy of the protein.

1.2.1 Co–translational protein folding

Proteins usually start to fold, while their being synthesized by the ribosome, hence intermediates are populated, which differ from those obtained by *in vitro* experiments [71, 58, 34]. The newly synthesized polypetpide chain does not appear all at once, but in a vectorial manner in the cytoplasm [33] in an extremely crowded and heterogeneous environment. Those conditions cleary differ from those obtained in *in vitro* refolding studies. Nevertheless, proteins typically fold more efficient *in vivo* than *in vitro* [58, 227] and at least 67% of all cytoplasmatic polypeptide chains are capable to fold independently without the requirement of molecular chaperons [23, 70]. The ribosome is able to accommodate at least 33 amino acids in an extended conformation [152], therefore the conformational space of a newly synthesized polypeptide chain is very restricted and the folding funnel will be both narrow and shallow [33]. While the protein elongates and leaves the ribosomal tunnel, more conformational space becomes accessible and the funnel width and depth increases [33] (Figure 1.2b and c). To study co-translational protein folding ribosome nascent chain complexes (RNCs) (see section 1.3) are commonly used.

1.2.2 Synonymous codon usage

The genetic code is degenerated with up to six codons encoding a single amino acid, as in the case of serine, arginine and leucine. These codons are termed "synonymous" or "silent" since the amino acid sequence of the protein remains the same when one codon is replaced by another synonymous codon. However, the distribution of these synonymous codons is not random and they are used with different frequencies, a phenomenon called codon bias. Codon bias exists in any organism [236] and corresponds to the tRNA abundance [104]. Codons with higher tRNA abundance are thus faster translated than infrequently used
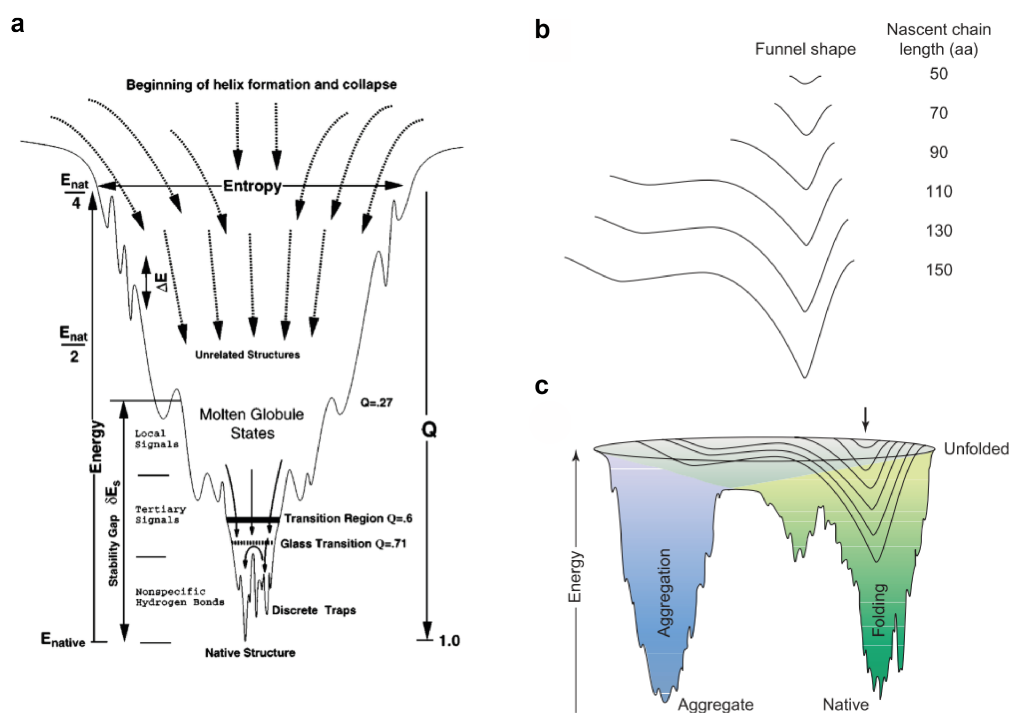
Figure 1.2: *Protein folding landscape. a) Funnel-shaped landscape of a small heli-
cal protein into its unique native structure. The ensemble of conformations
in the upper part of the funnel start to form a helix, lowering the entropy
(funnel diameter) to obtain its native structure, which is the energetic min-
imum. The vertical axis of the funnel contains the solvent-averaged en-
ergy E and the fraction of native-like contacts Q. Figure reprinted with
permission from Onuchic et al. [193]. © 1997, Annual Reviews, Inc. b)
Co-translational folding funnel of a polypeptide chain synthesized in vivo.
Funnel shape as a function of polypeptide chain length. Short chains have
a shallow and narrow funnel shape due to limited conformational entropy
and energetic differences between the conformations. c) Co-translational
funnel shape overlaid on the in vitro folding and aggregation funnel for a
full-length protein. Figure 1.2b and 1.2c reprinted with permission from
Clark [33]. © 2004 Elsevier Ltd.*

codons, which was supported by studies substituting frequently used codons by
infrequently used ones, or vice versa [129]. Therefore, translation is a non-uniform
process and rare codons can lead to translational pausing. In the case of γB crys-
tallin, rare codons are located in the peptide linker between the two domains and
the N-terminal domain is translated faster than the C-terminal domain [131]. Due
to the location of rare codons between domains, the folding of proteins might be
separated by translational pauses.

According to Anfinsen [4], the amino acid sequence contains all necessary infor-
mation for a protein to obtain its three-dimensional structure. However, studies
showed that an additional layer of information might be included in the transla-
tion kinetics and that not only the amino acid sequence is necessary for a protein
to function. A few of these studies are summarized below, giving a glimpse on
the effect of synonymous codon usage on protein folding and function. The re-
views of Komar [129] and Jacobson and Clark [110] are recommended for further

information. Komar *et al.* [134] mutated 16 consecutive rate codons in the *E. coli* chloramphenicol acetyltransferase into frequent ones, resulting in an accelerated protein synthesis rate, but also in a 20% lower protein activity. Angov *et al.* [5] adapted the gene sequence from the codon usage of the initial host to the designated heterologous expression host, a process called codon harmonization. Three harmonized proteins of *P. falciparum* were expressed in *E. coli* and the soluble protein yield was increased by a factor of 4 - to 1000. Replacement of 10 consecutive rare codons to preferred codons in the yeast TRP3 gene encoding a bifunctional protein with anthranilate synthase II and indoleglycerol-phosphate synthase activity resulted in 1.5-fold reduction in indoleglycerol-phosphate synthase activity [38]. Kimchi-Sarfaty *et al.* [126] showed that a few synonymous codon mutations in the multidrug resistance 1 gene correlated to altered function and interactions of the P-glycoprotein with substrate and inhibitors, indicating a conformational change of the protein. Codon optimization in the *Neurospora* circadian clock protein FREQUENCY (FRQ) resulted in increased protein levels, but also changes in conformational changes as an altered phosphorylation profile, stability and impaired functions in the circadian feedback loop [282]. In a follow-up study, Yu *et al.* [276] demonstrated in an *in vitro* expression system derived from *Neurospora* and *Saccharomyces* that the codon sequence rather than changes in the specific RNA structure determines the elongation rate of the firefly luciferase. While most of these findings detected the conformational changes with altered protein activity, Buhr *et al.* [22] showed using NMR spectroscopy that synonymous codon mutations in γB crystallin resulted in a different oxidation state of the harmonized protein. Two synonymous gene variants were investigated, one with the gene sequence from *B. taurus* and one adapted to the codon usage of *E. coli*. The harmonized construct (H) was translated faster, more resistant to protease K digestion and partly oxidized, while the unmodified one was completely reduced.

All the above mentioned studies show that not only the amino acids sequence is responsible for the correct function of a protein, but an additional layer of information might be buried within the translation kinetics of the protein.

1.2.3   The ribosomal exit tunnel

The existence of the ribosomal exit tunnel within the 50S ribosomal subunit was proven in 1995 by cryo-electron microscopy [68]. The ribosomal tunnel spans ∼80-100 Å from the PTC to the exit site and is between 10-20 Å wide, with an average diameter of 15 Å [189]. This provides sufficient space for around 33 amino acids in an fully-extended [152] and approx. 67 residues in an α-helical conformation [152, 153]. The narrowest point with a diameter of about 10 Å is at the beginning and at a position 28 Å from the exit [189]. Beside a bend between 20-35 Å from the PTC, the tunnel is largely straight [189]. The majority of the tunnel surface consists of 23S rRNA domain I till IV (28S rRNA in eukaryotes)

and is mainly hydrophilic including hydrogen bonding groups from bases, back-
bone phosphates and polar protein side chains [189]. Therefore, the electrostatic
potential of the tunnel is mainly negative and varies in magnitude along the tun-
nel length [155]. The proteins L4 and L22 form a constriction point in the mid
tunnel region, with a long β-hairpin of L22 being the largest protein contribution
to the ribosomal tunnel surface [189]. An overview of the tunnel dimensions are
shown in Figure 1.3.
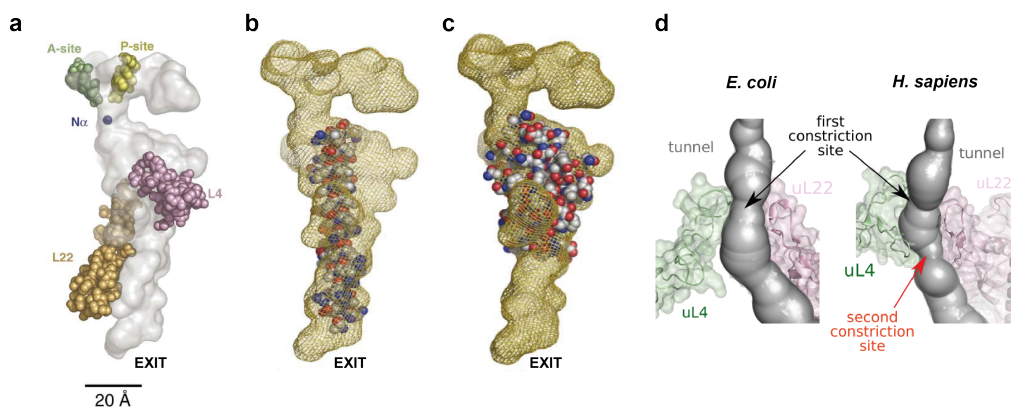
Figure 1.3: *Overview of the ribosomal tunnel dimensions. **a)** Stereo diagramm
of the ribosomal exit tunnel from H. marismortui including the tRNA bind-
ing sites, the PTC and the nascent chain intercalating ribosomal proteins
L4 (pink) and L22 (orange) relative to the tunnel surface (grey). **b)** Space
filling model of an α-helix and **c)** immunoglobulin domain IgG inside the
ribosomal tunnel. **d)** Overview of a prokayotic and eukaryotic tunnel di-
mensions. E. coli (left) and H. sapiens (right). The eukaryotic tunnel
contains a second constriction site 20 Å below the first one. Figure 1.3a-c
reprinted with permission from Voss et al. [260] © 2006 Elsevier Ltd. Fig-
ure 1.3d reprinted with permission from Duc et al. [46] © 2019, Oxford
University Press.*

While in bacteria L23 is located near the tunnel exit, this protein is replaced by
two proteins L23 and L39e in eukaryotes [85, 189], which may provide a better
control of the ribosomal tunnel exit. Until recently, the tunnel dimensions of
cytoplasmic ribosomes were considered to be universal through all kingdoms of
life [266]. However, comparison of 20 ribosome structures revealed a second con-
striction site in eukaryotes and to a lesser extent also in archaea, which is absent
in bacteria [46]. This second constriction site occurs through an additional L4
loop in eukaryotes, which reduces the tunnel radius from 5.0 ± 0.5 Å in bacteria
to 3.9 ± 0.6 Å in eukaryotes at around 50 Å from the tunnel beginning [46]. Addi-
tionally, 31.6 ± 2.3 Å of the ribosomal tunnel are covered by L39e in eukaryotes,
compared to 19.0 ± 2.8 Å by L23 in bacteria, which reduces the average radius of
the tunnel by approximately 1 Å in the last 30 Å in eukaryotes [46]. These mod-
ifications might be useful to reduce the PTC from external threats, like higher
resistance to macrolide binding [278, 255]. Increasing evidence indicates that the
ribosomal tunnel is rather involved in protein folding events than being a pas-
sive, teflon-like channel. This has been shown among others by Lu and Deutsch

[154] by measuring the translation rates of peptides containing each consecutive arginine, glutamine, glutamate, lysine or aspartate residues. While positively charged residues, like arginine and lysine, resulted in reduced or paused translation this effect was not seen with neutral or negatively charged amino acids [154]. Another effect seems to be not only the charge, but also the geometry of the tunnel. Kudva *et al.* [140] showed that the 29-residue zinc-finger domain (ADR1a) folds within different parts of the tunnel, depending on the presence of ribosomal proteins L23 and L24. Altogether, the ribosomal tunnel seems to play an important role in protein folding and several studies address the question to which extent the nascent chain (NC) can be folded within this tunnel. These studies are explained in more detail according to the used method in the following sections.

1.2.4   Co–translational chaperones

Although at least 67% of the proteins can fold independently [23], specialized proteins exist that are called molecular chaperons. They are known to bind to non-native states of a protein and assist folding to their native structure. Among those chaperons there are a few, which can directly interact with the NC. The probably best studied chaperon is the trigger factor (TF) of *E. coli*. The dragon-shaped TF is an ATP-independent chaperone consisting of 432 amino acids with a ribosome-binding, a peptidyl prolyl isomerase domain and in between those two domains a discontinuous C-terminal domain able to bind substrates (residues 113 to 149 and 247 to 432) [220]. The TF binds in a 1:1 stoichiometry to the 50S subunit of both translating and non-translating ribosomes and exists in a 2-fold molar excess over ribosomes in the cytosol [23]. However, the binding affinity to NCs is about 50-fold higher than for non-translating ribosomes with a K_d=1-2 μM [138, 218]. The TF binds to NCs, which are around 110 amino acids long and expose hydrophobic residues as for example a stretch of eight amino acids enriched with aromatic and basic amino acids and a positive net charge [138, 197]. A NMR study revealed four contact sites of TF which are mainly composed of hydrophobic residues [220]. The TF can bind via a charged-loop in its N-terminal doimain (NTD) to the ribosomal protein L23 [137] and is not only able to assist proper protein folding but also unfolds pre-existing folded structures and decelerates disulfide bond formation [90]. Beside TF, the ATP-dependent Hsp70 chaperone DnaK can act with its co-chaperones DnaJ and nucleotide exchange factor GrpE, both co-and postranslationally, to guide protein folding [23, 138]. DnaK is one of the most abundant chaperones within *E. coli* and does not bind to the ribosome but acts downstream of TF. While deletion of the genes encoding TF or DnaK alone, only leads to minor folding problems, simultaneous deletion leads to severe aggregation and is lethal at temperatures above 30°C [43]. In contrast to TF, DnaK with its cofactors cannot postpone or revert co-translational structure formation and neither decelerate disulfide bond formation

[90]. Dnak interacts with at least 700 cytosolic proteins, including approximately 180 relatively aggregation-prone proteins in *E. coli* and with 1000 proteins in ΔTF strains [29]. DnaJ presents non-native substrates to ATP-bound DnaK, hydrolysis of ATP to ADP leads to closing of a α-helical lid and thereby tight binding of the substrate. The peptide-binding domain of DnaK recognizes an extended approximately 7 residue long segment with a hydrophobic core, flanked by positively charged residues [215, 197]. DnaJ dissociates after ATP-hydrolysis and GrpE binds to the ATPase domain of DnaK. GrpE catalyzes the release of ADP and binding of another ATP results in lid opening and substrate release. The substrate is either transferred to other chaperones, as for instance to GroEL, or achieves its native conformation. The GroEL chaperonin with its lid-like cofactor GroES belong to the ATP-dependent chaperones and are found in bacteria, mitochondria, chloroplasts and in eukaryotic organelles originating from bacterial endosymbionts [86]. The complex interacts with at least 250 proteins, corresponding to approximately 10% of the total *E. coli* cytosolic proteins, with around 80 proteins relying fully on GroEL for proper protein folding [138, 86]. It was long assumed that GroEL-GroES act exclusively post-translational, however studies revealed association of GroEL with nascent polypeptides implementing also a co-translational function [275]. Protein substrates for GroEL need to be between 10-55 kDa in size [52]. GroEL is a complex of two heptameric rings with each subunit of around 57 kDa in size and GroES is a dome-shaped heptameric ring with 10 kDa subunits [86]. Each subunit of GroEL provides an apical domain with a hydrophobic cleft competent for binding of short hydrophobic peptides [86]. The exact mechanism of how the GroEL-GroES complex assists in protein folding is still not fully understood, the review of Hayer-Hartl *et al.* [86] provides three models in which the complex act passively by preventing aggregation or as an active cage promoting protein folding. TF, DnaK and GroEL guide protein folding in prokaryotes. In eukaryotes the folding-machinery obtains higher complexity with several ribosome-associated chaperones, like the nascent polypeptide-associated complex (NAC) Hsp70 Ssb and the ribosome-associated complex (RAC) [138]. For additional information on the chaperone interactions in eukaryotes the recent review of Kramer *et al.* [138] is recommended.

1.3   Ribosome nascent chain complexes (RNCs)

Ribosome nascent chain complexes (RNCs) are macromolecular complexes consisting of the 70S ribosome and the elongating polypeptide chain. To keep the NC attached to the ribosome different approaches can be used. Using *in vitro* translation systems, RNCs can be created by either using an mRNA sequence missing a stop codon or missing release factors or with ribosome arrest peptides (RAPs), which are needed for RNCs produced *in vivo*.

1.3.1   Ribosome arrest peptides (RAPs)

Ribosome arrest peptides (RAPs) contain specific amino acids that interact with ribosomal residues inside the ribosomal exit tunnel and thereby keeping the NC attached to it. The big advantage of RAPs is that they can be used to produce RNCs *in vivo*. The translational arrest of those peptides is independent of the codon usage and mRNA secondary structure [107]. Figure 1.4 shows an overview of known arresting peptides. The arrest sequences are usually around 20 amino acids long and contain essential residues around the PTC-proximal element (residues -4 to +1) and in the mid-tunnel region close to the constriction point (residues -9 to -21) [107]. While RAPs are distinct in their amino acid sequence, they often contain a proline residue in the A- or P-site, since prolines are the slowest nucleophilic substrates for peptide bond formation [199]. RAPs can lead to stalling at three different steps: during elongation at the peptidyl-transfer step (ErmCl-erythromycin, SecM, VemP and MifM), the translocation step (CGS1-S-adenosylmethionine (AdoMet)) and at the translation termination step (TnaC-tryptophan and arginine attenuator peptide (AAP-arginine)) [107]. In the following sections, the focus will be laid on the arresting peptides SecM, TnaC and VemP since those can be used as RAPs in prokaryotic systems.
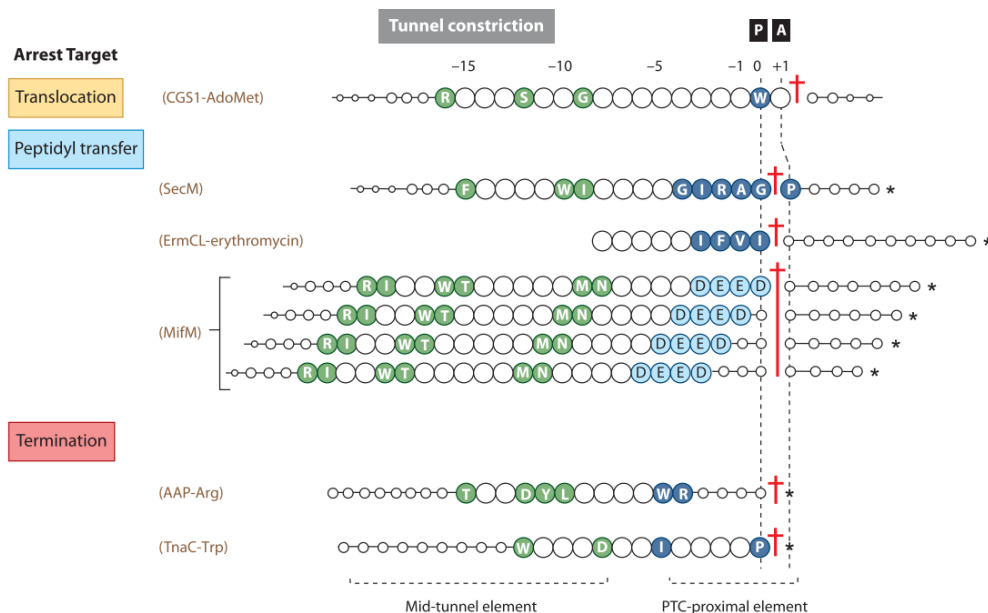
Figure 1.4: ***Overview of arrest peptides.*** *Sequence positions of arrest-essential amino acids in RAPs. The arrest point is shown with a red cross, defining the P-site as position 0 and the A-site as position +1. Residues important for stalling are highlighted green, if located in the mid-tunnel region and in blue, if located close to the PTC. Figure reprinted with permission from Ito and Chiba [107]. © 2013 Annual Reviews, Inc.*

1.3.2 SecM

The *E. coli* secretion monitor (SecM) is the most widely used RAP in co-translational folding studies [27, 28, 51, 56, 188] and is a 170 amino acids long protein encoded by an open reading frame (ORF) preceding *secA*, which encodes an ATPase driving protein export across the SecYEG translocon [209, 107]. The essential stalling sequence of SecM is FXXXXWIXXXXGIRAGP ($\mathbf{F}^{150}$STPV$\mathbf{WI}$S-QAQ$\mathbf{GIRAGP}^{166}$) [180]. Two stalling mechanisms of SecM have been identified, one with the peptidyl-glycyl-tRNA in the P-site and one with this tRNA in the P/E-site and the prolyl-tRNA in the A/P-site [280]. Due to the occupation of the A-site, the stalled RNCs are resistant to puromycin [176]. Essential residues of the SecM peptide are Arg163, Gly165 and Pro166 since mutation of these can abolish the stalling completely [180]. Arg163 is located at a confined space between A2451, C2452 and U2506 of the 23S rRNA [280]. While Trp155, Ile156 and Gly165 are important residues as mutation results in reduced stalling efficiency, Phe159, Gly161, Ile162 and Ala164 are only partially required [180]. Strong density was observable between the carbonyl oxygen of Ile162 and G2505, as well as the sidechain of Ile162 and C2610 [280]. Mutational studies on the ribosomal site revealed that residues A2053, A2062, A2503 within domain V of 23S rRNA are essential and insertion of one adenine residue within the consecutive residues (A749-A753) in hairpin 35 of domain II also abolished the arrest [259, 180]. Furthermore, the ribosomal proteins L22 and L4 are involved within the stalling mechanism. Mutations of Gly91 and Ala93 in L22 resulted in release of stalling [180] and thus problably are involved in SecM binding. In case of L4 Arg61 directly interacts with Trp155 of SecM [280]. Beside those, the tRNA also plays a role in efficient SecM stalling. The CCA-end of the P-tRNA forms canonical base pairs with G2251 and G2252 within the 23S rRNA and the peptidyl bond linkage between G165 contacts U2585, which is together with U2506 in an uniduced state that disfavors A-site accommodation [280]. When the ribosome stalls on the *secM-secA* mRNA, the SD sequence of *secA*, which is usually located within a stem-loop structure (Figure 1.5a) gets exposed (Figure 1.5b) [170].

Translation of *secA* can proceed and the translation arrest is released when the Sec machinery tries to export the NC of SecM [181, 25]. Functional SecA, an ATPase, then recruits the accessible N-terminal region of the SecM nascent peptide to the SecYEG complex, where it is translocated across the cytoplasm membrane, thereby the stalling is released and SecM is then proteolyzed in the periplasmic space [160, 179] (Figure 1.5c). Therefore, SecM is a cis-specific up-regulator in its native function and also a cis-chaperone by bringing the *secM-secA* mRNA to the membrane, where the localized biosynthesis of SecA, a multi-conformation protein, may favor its folding into an active conformation [182, 108].
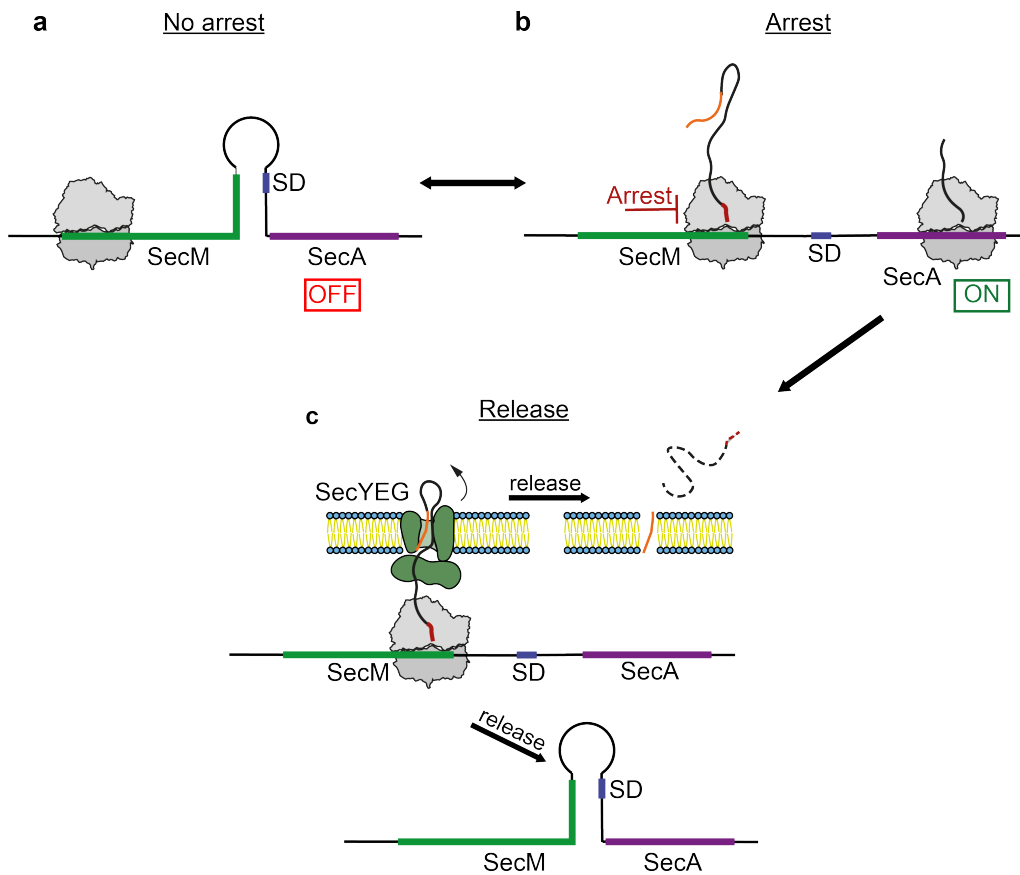
Figure 1.5: *Arrest regulation of SecM. **a)** The Shine-Dalgarno (SD) sequence of SecA is hidden within a stem-loop. **b)** Expression of the SecM mRNA leads to disruption of the stem-loop and exposes the SD sequence of SecA, due to interaction of the nascent peptide with the ribosomal tunnel (red thick line). **c)** The N-terminal signal sequence (orange) is recognized by the SecA-SecYEG-complex and translocated across the cytoplasm membrane. The stalling is released through the pulling force and the completed product of SecM is proteolyzed within the periplasmic space. Modified after Ito and Chiba [107], Ito et al. [108].*

1.3.3 The tryptophan-sensing TnaC

The tna operon of *E. coli* encodes the peptide TnaC of 24 amino acids with the essential residues (**W**FNI**D**NK**I**VDHR**P**). Tryptophanase (tnaA) and tryptophan permease (tnaB) are located downstream of tnaC and both are inducible by tryptophan [107]. The translational arrest occurs at the point of termination. In the absence of tryptophan, transcription terminates prematurely before *tnaA* at a Rho-dependent terminator, which is located in the spacer region between *tnaC* and *tnaA*. In the presence of tryptophan, however, the translation continues into *tnaAB* and peptide release by RF2 is inhibited [79] (Figure 1.6). The RNC is created by stalling of Pro24 within the P-site and the UGA stop codon in the A-site [107] (Figure 1.6).
Bischoff *et al.* [18] were able to identify two tryptophan molecules within the ribosomal tunnel. They are located 15-20 Å from the PTC within two hydrophobic
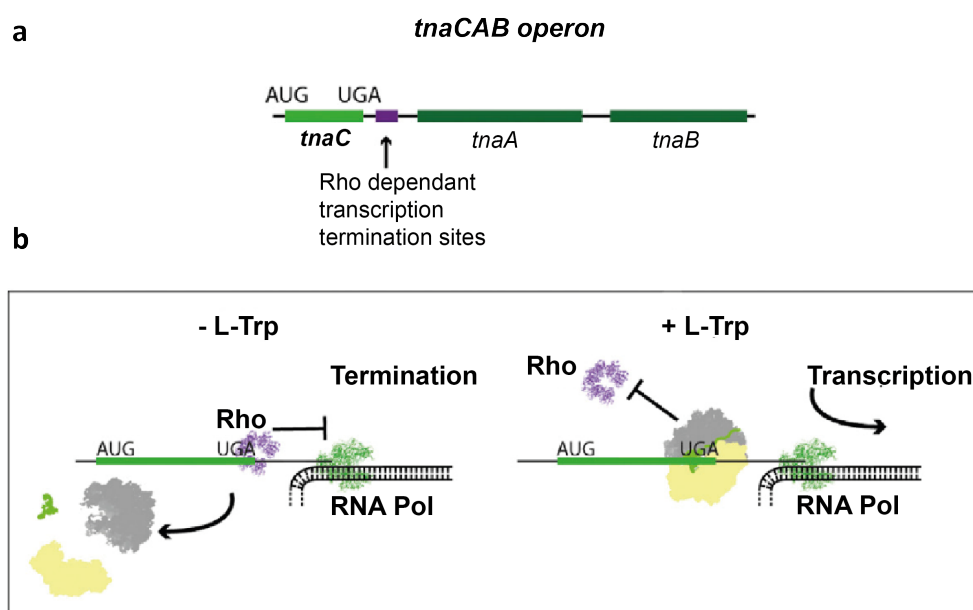
Figure 1.6: *Arrest regulation of TnaC. a) The tnaCAB operon. b) In the absence of tryptophan transcription is terminated and Rho can bind to the termination site. In the presence of tryptophan the ribosome stalls with the stop codon in the A-site and transcription of tnaA and tnaB is executed. Figure adapted with permission from Bischoff et al. [18] © 2014 Lukas Bischoff, Otto Berninghausen, Roland Beckmann. Published by Elsevier Inc.*

pockets formed by the TnaC peptide between the residues Asp21 and Asn17 and nucleotides of the 23S rRNA. The first tryptophan is positioned in a stacking interaction to the 23S rRNA base U2586, while the second tryptophan is located in the area of A2058/A2059 and the base pair of U2609/A752 [18]. Additionally, several contacts between the TnaC and the ribosomal tunnel have been identified using cryo-EM and mutation studies. These contacts include Thr9-L22 (Arg95), Trp12-L22 (Arg92), Phe13-23S rRNA (A751), Asn14-L4 (R61), Asp16-L22 (K90, R92), Lys18-23S rRNA (U2609/A752), Asp21-23S rRNA (A2062), His22-23S rRNA (U2505), Arg23-23S rRNA (U2505, C2452, U2506) and Pro24-23S rRNA (U2585) [228, 107, 18]. The conserved amino acids within the TnaC peptide appear in specific interactions with the ribosomal exit tunnel, thus two binding pockets for tryptophan can be formed, which turns the ribosome in a sensor for tryptophan [18]. Hence, binding of tryptophan results in allosteric silencing of the PTC [18].

1.3.4   *Vibrio* export monitoring peptide (VemP)

The export monitoring polypeptide, VemP, from *Vibrio alginolyticus* acts like SecM as a cis-up-regulator for secDF2, a membrane protein complex with 12 transmembrane segments and large periplasmic domains [191]. The translation of SecDF2, is up-regulated if the organism encounters a low Na^+ environment (63 mM or lower) and is therefore an essential mechanism for environmental adaption

in marine bacteria [105] (Figure 1.7). The VemP sequence can be fused C-terminal
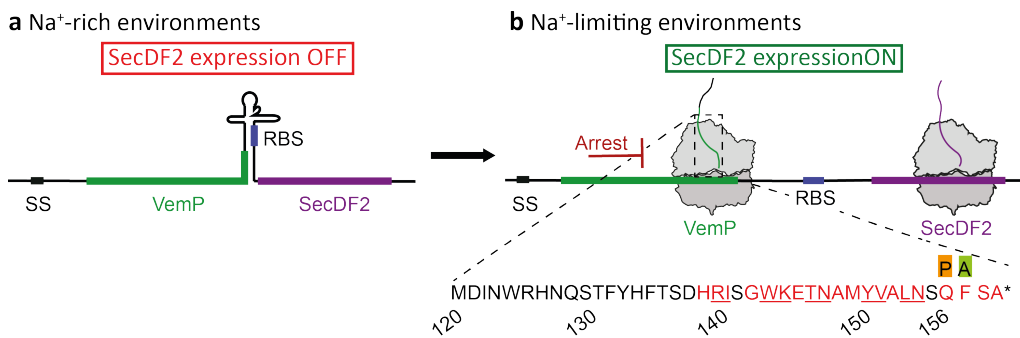
Figure 1.7: ***Arrest regulation of VemP. a)*** *Representation of the VemP-SecDF2 mRNA with N-terminal signal sequence (SS), VemP gene (green), stem-loop structure and SecDF2 gene (purple). The ribosomal binding site of SecDF2 is hidden in a stem-loop structure in Na^+-rich environments.* ***b)*** *In Na^+-limiting environments the ribosome is arrested with Gln156 in the P-site and the incoming aminoacyl-tRNA of Phe157 in the A-site. The disruption of the mRNA secondary structure leads to accessibility of the RBS and expression of SecDF2. Amino acid residues within the ribosomal tunnel are shown in one letter code with critical residues (red underlinied) and important residues (red) highlighted. The asterisk (*) indicates the stop codon. Figure modified after Su* et al. *[249].*
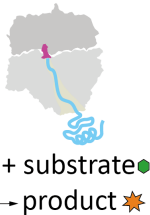
to a protein of interest to study RNCs and even be expressed in *E. coli.* Ishii *et al.* [105] argued that at least the 27 C-terminal residues of VemP are required for the elongation arrest, since an amber stop codon introduced before this sequence resulted in no translational arrest. Therefore, mutation studies were performed on those 27 long F^{131}YHFTSD**HRI**S**GWKETNAMYVALN**S**Q**FSA sequence and revealed the importance for several residues on the translation arrest [105]. Additionally, toeprinting assays revealed that Gln156 is located in the P-site and Phe157 within the A-site [105], which would make it resistant to puromycin. Su *et al.* [249] solved the structure of VemP using cryo-EM showing an extremely compact formation within the ribosomal tunnel of two α-helices connected by an α-turn and a loop. Due to this extensive compaction 37 amino acids of VemP are located within the upper two thirds (approx. 50-55 Å) of the ribosomal tunnel and 51 aa would be required to totally span the exit tunnel, in contrast to 34 aa of SecM [249]. The cryo-EM structure of the arrest peptide was solved with the full-length VemP sequence lacking only the first 25 amino acids, the signal sequence, since pulse-chase experiments showed low stalling efficiency for the residues 131-156 *in vivo*. Thereby electron density for the residues 120-156 was detectable within the tunnel.

1.4   Investigating co–translational folding using RNCs

To study co-ranslational protein folding, both *in vivo* and *in vitro* produced RNCs have been used. In the following sections, the studies will be separated by

the different techniques used. Due to immense number of publications on RNCs only a selection of studies can be represented here. Table 1.1 shows an overview of different methods to study co-translational protein folding. For further information the reviews of Javed *et al.* [115] and Komar [130] are recommended.

Table 1.1: ***Methods for studying co-translational folding.*** *Different methods for the investigation of co-translational protein folding, schematic illustration and comments on the requirements and detection for this method. Modified after Komar [130].*

Method	Schematic illustration	Requirements and detection
Limited proteolysis	+ protease	Preparation of stalled RNCs necessary. Analysis of proteolytic digested NC by SDS-PAGE. Investigation of NC inside and outside of ribosomal tunnel possible.
Enzymatic activity	+ substrate → product	Detection of chemiluminescence, fluorescence or proteolytic digest. Suitable for *in vivo* and *in vitro* measurements. Not suitable for investigations inside the ribosomal tunnel.
Ligand binding	+ ligand	Possible for *in vivo* and *in vitro* measurements. Analysis of ligand binding via liquid scintillation counting using labeled ligand. Not suitable for investigation inside the ribosomal tunnel.
Binding of antibodies	+ mAb	Preparation and isolation of stalled RNCs necessary. Suitable for *in vivo* and *in vitro* detection. Detection of antibody binding by ELISA, chromatography, etc.

Cysteine accessibil-
ity assay

+ PEG-MAL
(5 kDa)

Analysis of RNCs *in vitro* by
mass shift in SDS-PAGE. Suit-
able for investigation of folding
within the tunnel.

AP-based force-
measurements

folding drives
GFP translation

Suitable for *in vivo* and *in vitro*
studies. Analysis via detected
fluorescence of GFP, cloned to
the C-terminus of the AP. GFP
is expressed, if the NC folds and
applies a force, which releases
the stalling of the AP. Folding
in and outside of the tunnel can
be measured.

Detection of disul-
fide bonds
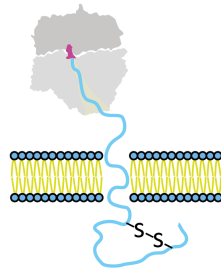
Suitable for *in vivo* detection of
intra- and intermolecular disul-
fide bond fornation. Isolation of
RNCs required, detection of the
NC on non-reducing and reduc-
ing SDS-PAGE. Not suitable for
investigation inside the riboso-
mal tunnel.

FRET

Suitable for the detection of NC
in vivo and *in vitro*. Detection
of FRET efficiency. Suitable for
detection of the NC inside and
outside of the ribosomal tunnel.

Cryo-EM

Purified RNCs, prepared *in vivo*
or *in vitro*, can be measured. Vi-
sualization of the electron den-
sity of the NC inside the tunnel
and outside, if the NC obtains a
rigid structure.

NMR spectroscopy	 isotopically labeled NC	Isotopically labeled, purified RNCs prepared *in vivo* and *in vitro* can be measured. Detection of the NC inside the ribosomal tunnel with solid-state NMR enhanced by DNP, outside of the tunnel also by solution-state NMR.

1.4.1 Studies using biochemical methods

Biochemical methods to study co-translational folding include, limited proteolysis using a proteinase, the detection of enzymatic activity, analysis of antibody binding or ligand binding as well as cysteine accessibility measurements using pegylation assays and *in vitro* arrest-peptide based force measurement assays.

1.4.1.1 Limited proteolysis

Limited proteolysis studies analyze the partial digestion of ribosome-arrested nascent chains using for example proteinase K, trypsin or chymotrypsin. Compact, folded protein structures are resistant to mild concentrations of protease, thus addition of protease after different time points to a translation system provides an indication when folding of the nascent protein can occur. For this method stalled RNCs are required and structure formation within the ribosomal tunnel as well as outside can be detected. Netzer and Hartl [184] investigated the folding of a fusion protein of human H-Ras and mouse dihydrofolate reductase (DHFR), both monomeric single-domain cytosolic proteins of approx. 20 kDa. The protease-resistant Ras domain occurred (in a eukaryotic *in vitro* translation system and in the cytosol of living cells) before the full-length fusion protein is synthesized, indicating co-translational folding of Ras [184, 130]. Frydman *et al.* [71] investigated the co-translational folding of the 62 kDa firefly luciferase in a eukaryotic translation system. The formation of the full-length luciferase was observed after 12 min and after 8 min the formation of a 22 kDa protease-resistant fragment was observed, indicating the independent co-translational folding of this N-terminal subdomain. The ribosome-arrested P22 tailspike protein, a homotrimer of 72 kDa subunits and γB crystallin were also treated with protease K, resulting in a protease-resistant fragment of 47 kDa as the intact TMS and a about 80-90 residues big fragment, corresponding the size of the NTD of γB crystallin [51, 22].

1.4.1.2 Enzymatic acitvity

The activity of ribosome-bound enzymes, can also provide an insight in co-translational folding, since the activity can be measured at different time-points of translation in a cell-free system or on selectively arrested NCs. Kolb *et al.* [128] investigated the co-translational folding of firefly luciferase in a cell-free system. They were able to monitor the activity of firefly luciferase continuously. While the ribosome-bound luciferase showed no enzymatic activity, the activity was recovered after addition of puromycin within seconds. The last 12 C-terminal amino acids are important for the activity and are covered by the ribosome, the immediate gain in activity is thus explainable with the fact, that the enzyme is already completely folded on the ribosome.

The auto-catalytic cleavage of the Semiliki forest virus (SFV) containing a chymotrypsin-like protease domain was investigated *in vivo* in prokaryotes and eukaryotes [185]. Nicola *et al.* [185] showed that the cleavage occurred 43 amino acids beyond the protease domain was translated, concluding that the protein cleaves itself almost immediately after leaving the ribosome. In addition, it was shown by Kelkar *et al.* [122] that sequestration of 10-15 amino acids within the ribosomal tunnel, prevents stable β-barrel formation within GFP. With 10 of 11 β-strands outside of the tunnel, a stable non-native conformation was obtained, which folded efficiently into the native structure after release from the ribosome.

1.4.1.3 Ligand binding and detection via conformational epitopes

The binding of ligands to ribosome-bound proteins as well as the detection of structural epitope formation via conformational antibodies provides an additional method for the investigation of co-translational folding. Komar *et al.* [132, 133] showed that the globin NC was able to bind heme, after 86 amino acids have been synthesized, while shorter chains lengths (34, 65 and 75) were not able to bind heme. Therefore, the 86 amino acid long ribosome-bound NC acquired a spatial structure allowing the interaction with heme, or heme attachment promoted the formation of the proper tertiary structure in the NC [133].

The study from Hamlin and Zabin [83] investigated the binding of anti-β-galactosidase to ribosome bound β-galactosidase. Due to detected binding events they argued that the NC has to fold before release of the ribosome. However, using a polyclonal immune serum, the specificity to a native conformation was not given. Beside using the above mentioned limited proteolytic digestion method on P22 tailspike protein, the conformation of the NC was also studied using monoclonal antibodies [51, 34]. Both studies detected that antibodies recognizing ribosome-bound NCs *in vivo* did not bind to early *in vitro* intermediates obtained within dilution of denaturant. However, the antibodies recognizing the *in vitro* intermediates, also recognized the *in vivo* ribosome-bound intermediates. Thus the folding pathway *in vitro* and *in vivo* seem to follow different routes.

1.4.1.4 *Cysteine accessibility and in vitro arrest–peptide based force–measurement assays*

The accessibility of cysteine residues within RNCs can be studied using polyethylene glycol maleimide (PEG-MAL). PEG-MAL binds to cysteine residues and increases the size of the peptide by for example 5 kDa, depending on the used PEG-MAL. This mass shift can be visualized on SDS-PAGE. Depending on the position of the cysteine residue within the NC and thus in the ribosomal tunnel, the accessibility of PEG-MAL varies. This method was used in various studies by Carol Deutsch [151, 152, 153, 135, 256]. Lu and Deutsch [152] showed using PEG-Mal and calmodulin maleimide (CaM-MAL) and an all-extended NC as tape measure, that the functional length of the eukaryotic ribosomal tunnel is 99-112 Å and that the S6 transmembrane segment in voltage-gated potassium (Kv) channels obtains a compact structure within the tunnel. In a following study, this tape measure was fused with alanine residues located at different positions inside the NC to follow the relative compaction within different zones inside the tunnel [152]. The zone nearest to the PTC and closer to the exit supported compaction of the NC, while the region between did not significantly contribute compaction [152]. Using a β-hairpin and an α-helical hairpin from the Kv channel Kosolapov and Deutsch [135] showed that minimalistic tertiary structures were able to form in the last 20 Å of the ribosomal tunnel, the vestibule region. Additionally, Tu and Deutsch [256] showed that all α-helical segments of Kv1.3 formed in the vestibule region.

Within Gunnar von Heijnes group an arrest-peptide based force measurement assay is used to study co-translational folding of proteins. In this approach an arrest peptide (AP), like SecM, is cloned to a protein of interest and the length of the linker between the AP and the proteins varies. If co-translational folding occurs, a pulling-force of the NC is exerted, which releases the NC from the ribosome. Using pulse-chase experiments or a GFP cloned to the C-terminus of the AP, the release of the protein can be visualized by fluorescent band on SDS-PAGE. With this method, a model for the membrane integration of a 19-residue leucine-alanine based segments, fused to the C-terminus of the leader peptidase, could be generated [106]. A linker length between 30 and 40 residues, showed an increase in released protein, indicating the formation of secondary structure. Placing helix-breaking proline residues in the middle part of the 19-residue segment, reduced the pulling force [106]. Additionally, using pulling-force profiles it was shown, that the second of the two zinc-finger domains in the yeast ADR1 protein folds co-translationally, which was also supported by cryo-EM demonstrating α-helical structures deep inside the vestibule [186]. Further, the folding of nine fast-folding domains with a molecular weight below 10 kDa was investigated with this AP-force measurement assays and showed that all nine domains initiated folding while located within the ribosomal exit tunnel, regardless of whether the domains contained α-helical, β-sheet structure or both [161]. The

principle of this *in vitro* arrest-peptide based force-measurement assays including
a force profile are shown in Figure 1.8a-d. The effect of TF, GroEL/ES on di-
hydrofolate reductase (DHFR) was also studied using pulling-force profiles [187].
DHFR folded only after emerging the ribosome and both the TF and GroEL
reduced the pulling-force, thus preventing folding. While the GroEL/ES com-
plex did not have an effect on DHFR and the pulling-force of ADR1 was also
unaffected by the presence of chaperones [187].

1.4.1.5 Disulfide bond formation

Possible disulfide bond formation can be detected within NC as well. Since
this chapter is even of greater relevance for this thesis, co-translational disulfide
bond formation is explained in section 1.6.4 after explanation of disulfide bond
formation in general.

1.4.2 Studies using Flourescence Resonance Energy Transfer (FRET)

Through the incorporation of non-natural amino acids via chemically modified
aminoacyl-tRNAs, a fluorescent dye can be attached to the NC. An advantage
of this technique is the high sensitivity, obtaining reliable measurements with a
sample containing only 1-10 nM fluorophore [117]. The fluorophore is also very
sensitive to its surrounding, changes in the probe emission intensity, lifetime,
energy (wavelength) and rotational freedom can be correlated to changes of the
NC environment [117]. Fluorescence resonance energy transfer (FRET) measure-
ments, a phenomenon which reflects the energy transfer of an excited donor dye
to a sufficiently close appropriated acceptor dye without the loss of a photon,
can be used to study co-translational protein folding. The FRET efficiency of
a NC harboring a donor and acceptor dye will increase if the dyes are close in
space, due to a folding event. An unstructured NC will thus exhibit a lower
FRET efficiency. Woolhead *et al.* [271] investigated the co-translational folding
of a 20 residue long transmembrane sequence (TMS) using εBOF as donor and
εBOP as acceptor dye, which was incorporated via an amber stop codon located
C-terminal of the donor lysine codon. The TMS formed a compact conforma-
tion near the PTC and also when integrated into the lipid bilayer, indicating
an α-helical structure. However, the TMS unfolded after release from the free
ribosome, indicating that the protein folding is induced and stabilized by the
ribosome. With the same approach, Woolhead *et al.* [270] were able to describe a
model for the SecM stalling mechanism. Using BOF-Met as the first residue and
BOP-Lys at the beginning of the third helix of the five-helix protein domain of
N5-glutamine methyltransferase HemK of different lengths in a reconstituted *in
vitro* translation system, the co-translational folding of this protein was monitored
in real-time [101]. Holtkamp *et al.* [101] identified a non-native, ribosome-bound
compact formation of the HemK70 construct, which unfolded after treatment
with puromycin. Further, the HemK70 NC was also resistant to thermolysin di-

gestion, since the N-terminal BOF residue was located inside the ribosome exit tunnel or in close proximity. Holtkamp *et al.* [101] also used photoinduced electron transfer (PET) between the N-terminal BOF and a Trp residue at different positions. If the Trp and BOF residue are in close contact, fluorescence quenching of BOF by Trp due to van der Waals contacts can be observed. These measurements also supported the existence of a compact transient state, which rearranged into the native state with increasing chain lengths [101]. A follow-up study on HemK using the same approach, provided a model for the co-translational folding showing the formation of compact structures within the ribosomal tunnel (Figure 1.8e-g) [171]. Mercier and Rodnina [171] explained that the formation of helices is the most parsimonious interpretation, but did not exclude the possibility of native-like intermediates resembling a molten globule or nonspecific collapse of the NC. Nevertheless, all these investigations by FRET supported the formation of compact structures stabilized by the ribosome.

1.4.3   Studies using cryo–Electron Microscopy (cryo–EM)

Using cryo-electon microscopy (cryo-EM) it is possible to visualize the electron density of the NC within the ribosomal tunnel. The stalling mechanism of the APs, like SecM [280], TnaC [229] and VemP [249] were all investigated by cryo-EM and are described in the sections 1.3.2-1.3.4 above. Primarily the group from Roland Beckmann and Daniel Wilson were involved in the visualization of α-helical structures within the ribosomal tunnel. Starting with a short peptide containing five Glu-Ala-Ala-Ala-Lys repeats forming an α-helix within the exit tunnel, the above mentioned zing-finger domain of ADR1 [186] and the three-helix bundle of spectrin domains, which obtained an α-helical structure within the vestibule of the ribosomal tunnel and was also investigated by AP based force measurements [188] are all structures obtained by these two groups. Recently, Javed *et al.* [114] showed that the fifth domain of ABP-120 filamin protein (FLN5), an immunoglobulin-like protein, revealed different states close to the ribosomal exit, assuming that the NC is highly dynamic and adopts a range of trajectories within the vestibule. The co-translational protein folding of FLN5 was also investigated by NMR spectroscopy [27].
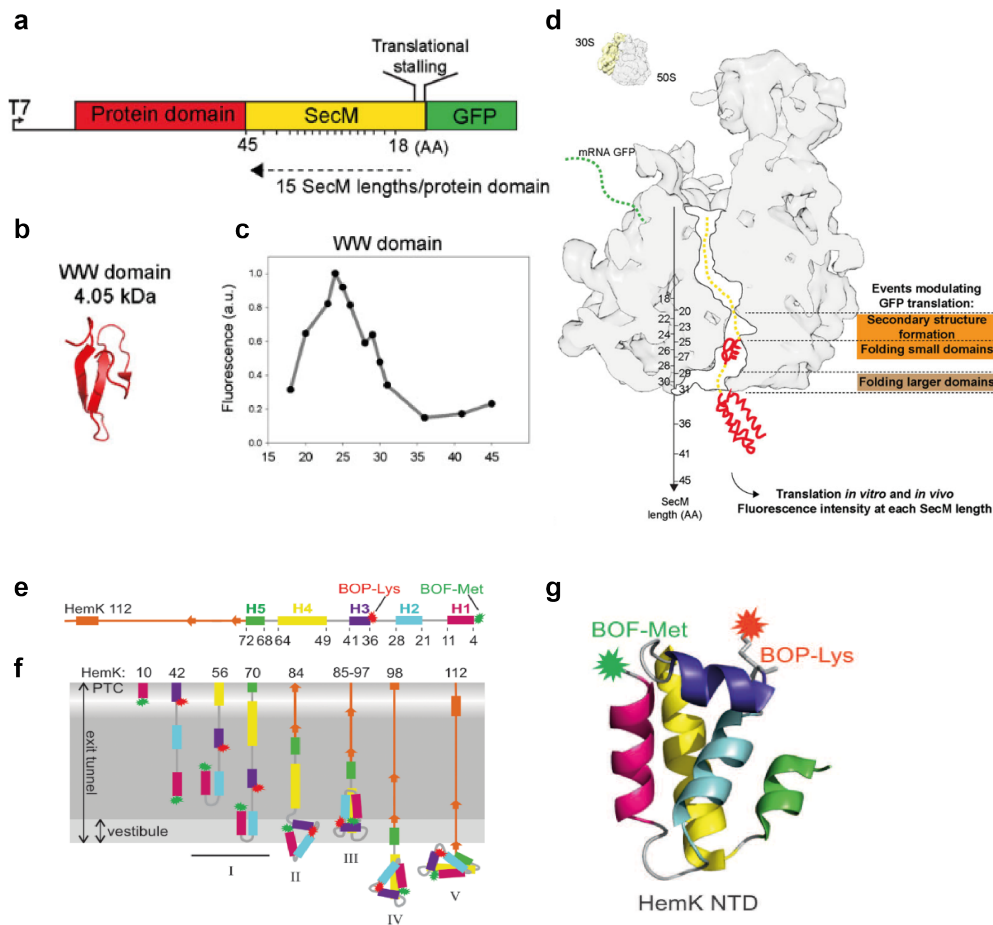
Figure 1.8: ***AP-based force measurement assay and structure of HemK with FRET dyes. a)*** *Schematic principle of the AP-based force measurement assay. GFP and SecM are cloned to the C-terminus of a protein domain. After folding, the translation arrest is released and GFP can be expressed.* ***b)*** *Representative structure of the WW domain and* ***c)*** *the measured GFP fluorescence of this domain with increasing linker length.* ***d)*** *Representation of the E. coli ribosomal exit tunnel showing the principle of the AP-based pulling force measurements and zones suitable for protein folding.* ***e)*** *Represenatation of HemK 112, with α-helices indicated as colored boxes.* ***f)*** *Model of co-translational folding for HemK with tertiary structure formation. FRET donor shown in green, FRET acceptor in red.* ***g)*** *X-ray structure of the NTD of HemK (PDB: 1T43). Figures adapted with permission from Marino* et al. *[161], Mercier and Rodnina [171], Holtkamp* et al. *[101]. © 2016 Federation of European Biochemical Societies. © 2018, American Chemical Society. © 2015, American Association for the Advancement of Science.*

.

1.4.4 Studies using Nuclear Magnetic Resonance Spectroscopy (NMR Spectroscopy)

Especially the group from John Christodoulou is pioneering in the detection of RNCs by solution-state NMR and a good overview of the challenges to detect the NC by NMR is provided by Waudby *et al.* [262]. The first ^{13}C, ^{15}N isotopi-

cally labeled NC studied by NMR spectroscopy was the tandem Ig domain repeat (domain 5 and 6) from the gelation factor ABP-120, also referred as Ig2-RNC [103]. The NC was expressed within a coupled transcription-translation *E. coli* cell-free system, supplied with a ^{13}C, ^{15}N-labeled amino acid mix. Hsu *et al.* [103] demonstrated using SOFAST-HMQC [223] that the Ig2 NTD obtained a native-like structure, since 45 cross-peaks were resolved in the ^{1}H-^{15}N correlation spectrum and showed high similarity to isolated Ig2 in the presence of 70S ribosomes. The CTD, however, was largely unfolded and flexible, when attached to the ribosome. The linewidths of the Ig2-RNC were 59 ± 29 Hz and are comparable to linewidths of the isolated Ig2 in presence of 70S ribosomes with a value of 41 ± 11 Hz.

In addition to the ^{1}H-^{15}N backbone correlation, the ^{1}H-^{13}C correlation of the methyl groups, reflecting also tertiary structure formation, is a powerful tool to study the folding of NC. While the residues located within the loop region showed severely line-broadening in the ^{1}H-^{15}N correlation spectrum of Ig2-RNC, the residues affected by line-broadening in the ^{1}H-^{13}C correlation spectrum were located within the interior of the domain structure [103, 102]. The line-broadening of Ile might be an effect of altered ring current effects from nearby aromatic side chains, since residues without an aromatic residue in proximity were not affected [102].

Using SecM and an expression system based on autoinduction, Cabrita *et al.* [28] were able to express ^{13}C, ^{15}N-labeled RNCs *in vivo*. This approach generated RNCs with yields of 50-200 pmol/L and significantly reduced the costs of isotopically labeled RNCs. This methods was further optimized and the procedure of the *in vivo* generation of ^{13}C, ^{15}N-labeled RNCs with expression, purification and useful solution-state NMR experiments is described in detail by Cassaignau *et al.* [30]. Using this *in vivo* generated RNCs, the domain 5 of the Ig2-RNC construct (ddFLN5) was arrested to the ribosome, the ^{1}H-^{15}N correlation spectrum revealed that the NC of ddFLN5 is highly dynamic and only a low population of partially folded species could be observed [28].

Simultaneously to Cabrita *et al.* [28], Rutkowska *et al.* [217] also used SecM for the generation of ^{13}C, ^{15}N uniformly and selectively labeled RNCs *in vivo*. Rutkowska *et al.* [217] investigated the co-translational folding of ribonuclease from *Bacillus amyloliquefaciens* (Barnase) using the 36 C-terminal residues of SecM and a 50-residue linker derived from RNA polymerase subunit β (RpoB). This BarnaseΔ50 construct was attached to the ribosome and 72 of 100 residues could be assigned.

The Src-homology (SH3) domain of α-spectrin, its folding-deficient point mutant m10 and two C-terminally truncated segments of SH3 were also studied by solution-state NMR [49]. Eichmann *et al.* [49] obtained negligible average chemical shift derivations between the ribosome arrested and the released SH3 domain and line-broadening was distributed along the entire amino acid sequence. Fur-

ther, with linewidths of the ribsome being > 1000 Hz, the observed increased line width of approx. 10 Hz by intermolecular interaction of the SH3 NC and the ribosome indicated that less than 1% of the SH3 domain interacted with the ribosome. The ribosome-arrested SH3 domain exhibits a compact, native-like β-sheet structure after the entire sequence emerges from the ribosome, the ribosome surface did not stabilize secondary or tertiary structures assuming that the folding of SH3 appears in a domainwise manner without significant influence of the ribosome [49].

In 2016 Cabrita *et al.* [27] investigated the co-translational folding of ddFLN5 using 12 RNCs in which the ddFLN6 linker varied between 21 and 110 residues (including SecM). The RNCs were both uniformly ^{15}N labeled and Ileδ1-^{13}CH$_3$ labeled. With a linker lengths above 47 residues, only the native-like like structure of ddFLN5 was detectable. Hence, the native structure was only observed after the entire domain emerged the ribosomal exit tunnel. Cabrita *et al.* [27] argued that the ribosome modulated the folding process, in such a way that the ribosome inhibits the acquisition of a stable structure, since the isolated, truncated versions of ddFLN5 were able to fold.

In addition to the solution-state NMR studies, two publications studying the interaction with the ribosome using solid-state NMR exist at time of writing this thesis. While Gelis *et al.* [75] did not use RNCs but studied the interaction of the ^{13}C-^{15}N His and Tyr-labeled initiation factor IF1 with the ribosome, this publication used for the first time centrifugal device suitable to pellet the ribosome into the MAS rotor, which is now used as a standard technique. The only solid-state NMR enhanced by DNP study using uniformly labeled ^{13}C, ^{15}N RNCs up to date, was performed in the group of Hartmut Oschkinat [144]. The disulfide oxidoreductase A (DsbA) signal peptide of 20 amino acids in length was attached to SecM. The chemical shifts of Ala, Ile, Pro, Val and Ser in a POST-C7 double quantum to single-quantum (DQ-SQ) spectrum [94] were compared to the average chemical shifts obtained in α-helical-, β-strand-, or coil-like conformations. Mutating the Ser residues within the SecM stalling motif to Ala, Lange *et al.* [144] could identify the two remaining Ser residues within DsbA. Together with the obtained average chemical shift of the above mentioned residues they argued that the DsbA signal peptide does not obtain an α-helical structure inside the ribosomal exit tunnel. However, the folding of the RNC could have been influenced by the usage of Ulp1, which might induce a force on the NC.

The RNC constructs studied by NMR spectroscopy are shown schematically in Table 1.2. To sum up, the advantage of NMR spectroscopy for the investigation of co-translational folding is the ability to study the structure and dynamic of the NC in atomic detail and huge progress was made in the preparation of isotopically labeled RNCs. A disadvantage is the low sensitivity requiring sample concentrations a factor 10^4 larger than in other biophysical methods as FRET or cryo-EM.

Table 1.2: *RNC constructs investigated by NMR spectroscopy.*

Study	RNC construct

Hsu *et al.* [103, 102]

ddFLN5	ddFLN6 linker	BstNI
107 aa	89 aa	20 aa

Cabrita *et al.* [28]

6x His	ddFLN5	ddFLN6 linker	SecM
	105 aa	90 aa	17 aa

Rutkowska *et al.* [217]

Strep	Strep	Strep	TEV	Barnase	RpoB	SecM
				110 aa	50 aa	36 aa

Eichmann *et al.* [49]

Strep	Strep	Strep	Smt3	SH3/SH3 m10	TEV	lin-ker	SecM
			97 aa	63 aa		10 aa	27 aa

Strep	Strep	Strep	Smt3	SH3-T2	SecM
			97 aa	58 aa	27 aa

Strep	Strep	Strep	Smt3	SH3-T1	SecM
			97 aa	41 aa	27 aa

Cabrita *et al.* [27]

6x His	ddFLN5	multiple ddFLN6 linker	SecM
	105 aa	0-90 aa	21 aa

Lange *et al.* [144]

Strep	Strep	Strep	SUMO	DsbA signal sequence	SecM
			143 aa	20 aa	17 aa

1.5   Model protein γB crystallin (GBC)

The vertebrate eye lens fiber cells consist of three major protein components: α-, β- and γ- crystallins. α-crystallins can act as both a molecular chaperone and as a structural protein like the β- and γ-crystallins [20]. Those lenticular proteins cannot be replaced; thus they have to remain in their native conformation for the whole lifetime. Therefore, crystallins need to be thermodynamically stable and efficiently captured and refolded by protein chaperons [20]. To ensure the transparency of the eye lens, the crystallins are soluble at high cellular concentrations. γB crystallins are present at concentrations up to 860 mg/ml within bovine eye lenses to ensure an refractive index of 1.49 [240]. Unfortunately, oxidation,

deamination or cleavage of the lens proteins can lead to protein aggregation or non-native protein-protein interactions which results in clouding of the lens and cataract. About 20 million people in 2010 are suffering from cataract, which makes this lens altering process responsible for 51% of world blindness (WHO 2019). γ-crystallins are the main component of the core region of the eye lens with γB crystallin (GBC), also known as γ-II crystallin from calf eyes being very well studied. In contrast to β-crystallins, which are structural similar to γ-crystalins, GBC is a monomeric protein (174 amino acids, 21 kDa), consisting of two homolog domains. Each domain consists of two intercalating Greek-key motifs forming a sandwich of four antiparallel β-sheets [268]. The domains are connected via a six-residue linker [268] and are further stabilized by hydrophobic interaction including the residues Met43, Phe56, Ile81, Val132, Leu145 and Val170 [240]. The C-terminal residues Phe173 and Tyr174 are crucial in preventing intermolecular dimerization [190]. Figure 1.9 shows the X-ray structure at 1.47 Å of bovine GBC, the hydrophobic core between the two domains and the C-terminal aromatic extension are also depicted.
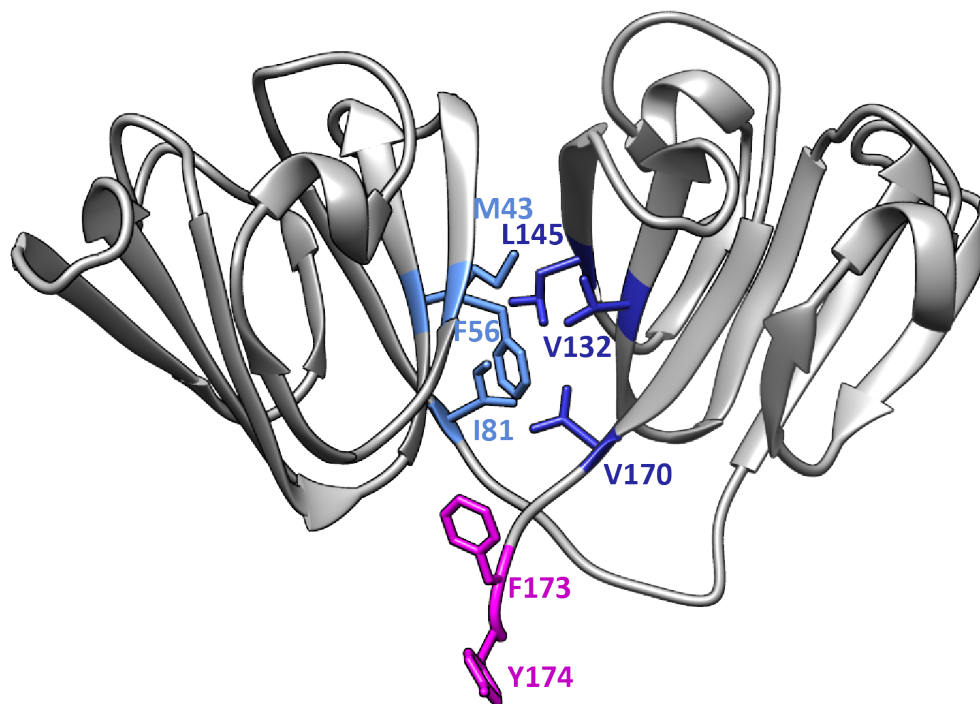
Figure 1.9: *Hydrophobic core of γB crystallin. 1.47 Å X-ray structure of bovine γB crystallin showing the hydrophobic core residues relevant for interdomain interaction. N-terminal residues are shown in light blue, C-terminal residues in darkblue. Aromatic C-terminal residues are highlighted in pink (PDB: 4GCR) (visualized using UCSF Chimera [204]).*

The NTD is not only more stable than the CTD, it also contains six out of seven cysteine residues. Usually a mammalian intracellular protein of 175 amino acids would contain around three cysteine residues [55]. The cysteines in GBC are

namely cysteine 15, 18, 22, 32, 41, 78 in the NTD and cysteine 109 in the CTD and are shown in Figure 1.10. Although six out of seven cysteine residues are located in the NTD, the overall sulfur contribution with five methionine in the CTD and only two in the NTD is nearly evenly distributed between the two domains [268]. All cysteines except of cysteine 32 are located in a β-strand or are part of a β-bridge.
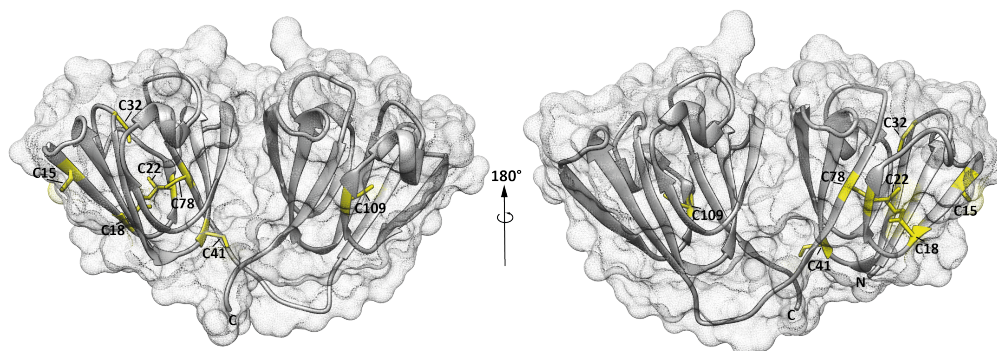
Figure 1.10: *1.47 Å X-ray structure of bovine γB crystallin. Cysteine residues are highlighted in yellow with Cys22 (C22) exhibiting two different conformations (visualized using UCSF Chimera [204].)*

Table 1.3 shows the solvent accessibility of the cysteine residues, calculated from crystallographic data [159]. Cys15 exhibits the highest solvent accessibility, while Cys22 and Cys41 are partly exposed and Cys32 and Cys78 are completely buried inside the protein [159].

Table 1.3: *Solvent accessibility of cysteine residues in GBC [159].*

Cysteine residue	Cys15	Cys18	Cys22	Cys32	Cys41	Cys78	Cys109
Solvent accessibility ($Å^2$)	65	2	19	0	12	0	5

The sulfur-containing residues in GBC are located close to aromatic residues and might undergo S-π interactions. Proteins with alternating aromatic and sulfur-containing amino acids have been identified in several globular proteins [175]. Those S-π interaction may not only contribute to the overall protein stability [175], but also protect the UV-sensitive aromatic residues from photo-oxidation [268]. Since GBC absorbs strongly in the UV region of the radiation spectrum, it has been considered that it could function as a filter for the retina [241]. Cys22 is unique to γB crystallins and was able to form a disulfide bridge to Cys18, visible in two X-ray structures [268, 178]. However, the occurrence of this disulfide bridge is still under debate. While Wistow *et al.* [268] argued that the disulfide bridge might be an effect of the relatively aged crystals, Najmudin *et al.* [178] interpreted that during storage or data collection the high energy disulfide bridge is formed due to rotation of the Cys22 side chain. McDermott and co-workers, were also

able to detect an intramolecular disulfide bridge of GBC using protein isolated from calf lens by reaction with 5,5'-dithiobis-2-nitrobenzoic acid (DTNB). This disulfide bridge is also resistant to reduction with 10 mM 2-mercaptoethanol at 20°C and pH 7.4 [169]. At pH 8.2 the protein was only completely reduced with 2 mM DTT after 4-5 h [169]. Additionally, Pande *et al.* [196] observed a disulfide bond using Raman spectroscopy, which disappeared after treatment with 2 mM DTT. They argued that the most likely candidates for this disulfide bond are the Cys18/Cys22 or Cys22/Cys78 pairs. The disulfide bridge between Cys18 and Cys22 has already been observed via X-ray crystallography [268, 178]. But with an torsion angle (χ3) of 94.5° between Cys22 and Cys78 a disulfide linkage, which usually adopts ±90° seems also likely [196]. Also mixed disulfide bridges with 2-mercaptoethanol and oxidized and reduced glutathione can be formed [239]. Slingsby and Miller [239] identified that one to three glutathione can be bound to GBC as a mixed disulfide. However, this study does not provide any information, which cysteine residues are involved and only Cys32 and Cys78 can be excluded due to their inaccessibility to bulky reagents [239]. Comparing these results to the solvent accessibility of the cysteines in Table 1.3, it is most likely that Cys15, Cys22 and Cys41 are involved in mixed disulfides. All in all, the behavior of the cysteine residues within GBC are still not fully understood, but oxidation of the cysteines is among other factors related to cataract [3]. However, since both GBC domains fold independently, are rich in disulfide bridges and β-sheets, GBC is an appropriate protein to study co-translational β-strand formation as well as co-translational oxidation.

1.6 Disulfide bond formation

A disulfide bridge or disulfide bond within a protein is a covalent linkage between the sulfur atoms of thiol groups (-SH) of two cysteine residues. This cysteine dimer is also named as cystine. The disulfide bond is formed via oxidation of the two thiol groups. As shown in equation 1.

$$R_1 - SH + R_2 - SH \leftrightarrow R_1 - S - S - R_2 + 2H^+ + 2e^- \tag{1}$$
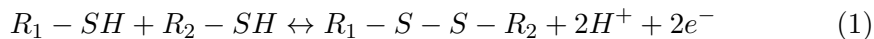
Around 40% of the human protein-encoding genes are predicted to contain either a signal sequence or transmembrane segment and may thus be modified by disulfide bonds [57]. Disulfide bridges are not correlated to a certain size or function of a protein, although it has been shown that small proteins (< 50 aa) contain a larger number of cystines, probably to compensate the low number of hydrophobic contacts [203]. The prevailing opinion of disulfide bond formation is that they have been introduced during evolution in order to increase thermal stability to function also in an extracellular environment [203]. An argument for this theory is that usually cytosolic proteins do not contain disulfide bridges,

however certain thermophilic organisms exhibit paired cytosolic cysteine residues [252, 12, 118]. The disulfide bonds in the bovine pancreatic trypsin inhibitor (BPTI) led to an increased stability of 2.5-5.1 kcal/mol for each cosslink [251]. Additionally, an introduced disulfide bridge in the T4 lysozyme enhanced the stability by increasing the T_m by up to 23.4°C [165]. The increased stability is associated with the loss of conformational space and thus decrease the entropy of the unfolded state [45]. The change in entropy upon introduction of a disulfide bond can be described by equation 2.

$$\Delta S = -2.1 - \frac{3}{2} \cdot R \cdot ln(n) \tag{2}$$

with R as the universal gas constant (R=8.314 $JK^{-1}mol^{-1}$) and n the number of residues between the crosslink. Using equation 2 and 3 the change of the Gibbs free energy ΔG upon addition of a disulfide bridge can be determined. If the effect of the disulfide bonds is only entropic, $\Delta\Delta$H would become zero and $\Delta\Delta$G could be correlated directly to T$\Delta\Delta$S. But, this assumption neglects the influence of hydrophobic effects, the effect of the enthalpy on disulfide bridges as well as the influence on the folding pathway [200]. Therefore, the effect of disulfide bonds on proteins is more complicated to predict, especially if proteins do not obtain a two-state model but unfold with different intermediate states [45].

$$\Delta\Delta G = \Delta(\Delta H - T\Delta S) \tag{3}$$

The number of possible disulfide bridges (DSBs) within a protein can be determined by equation 4.

$$DSB = \frac{n!}{\frac{n}{2} \cdot 2^{n/2}} \tag{4}$$

With n being the number of cysteine residues within the protein. This means that a protein with six cysteine residues, as in the NTD of GBC, could obtain 15 different disulfide bridges and lysozyme with eight cysteine residues would obtain already 105 different disulfide bonds. However, it seems unlikely that all different conformations are sampled by a protein and not all disulfide bridges contribute to the stability of a protein. Especially if the disulfide bond is located within a rigid region of the protein, it seems to lower the stability by decreasing the T_m [164]. Dani *et al.* [42], analyzed the stability of 47 disulfide bonds introduced by mutation within 24 proteins. 30 disulfides resulted in increased thermodynamic stability, 3 showed no change and 14 mutations led to a decrease in stability. Stabilizing disulfides showed the following features:

1. Location within regions of medium to high mobility

2. Location close to protein surface with depths of 3-5 Å

3. > 25 residues located between the crosslink

4. No steric overlap with surrounding atoms

Further, disulfide bonds exhibit a torsion angle χ3, which is defined by the rotation of the two Cβ atoms about the Sγ-Sγ bond, of -87 and +97 degree as shown by Craig and Dombkowski [36] analyzing 1505 native disulfide bonds in 331 non-homologous proteins. These values differ slightly from the often cited ±90°. The torsion angle χ1, which is defined by the N-Cα-Cβ-Sγ bond occupies values at ±60° and ±180°, while χ2 defined by the Cα-Cβ-Sγ- Sγ peaks at ±60° [45]. In total the five torsion angles (Figure 1.11a) make up 20 possible cystine configurations, which were identified in 6874 disulfide bonds of 2776 X-ray structures [226]. Those 20 configurations can be classified and grouped in hook, spiral and staple (Figure 1.11b). Additionally, the sign of the central angle χ3 describes the disulfide bond as left-handed (LH) or right-handed (RH). An overview of the classification of the 20 different configurations of disulfide bonds can be found in Schmidt *et al.* [226]. Three configurations occur predominantly in allosteric cystines, which are disulfide bonds that lead to conformational change when they are formed or broken and thereby control the function of a protein [226]. The most common configuration is the –RHstaple, followed by the –LHhook an -/+RHhook [93].
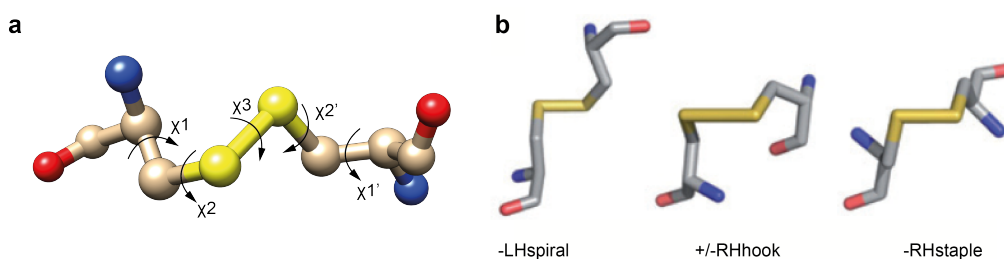
Figure 1.11: *Classification of disulfide bonds depending on their dihedral bond angle.* ***a)*** *Five torsion angles that describe the disulfide bond (χ1, χ2, χ3, χ1', χ2').* ***b)*** *The three basic disulfide groups: spiral, hook and staple as determined by the torsion angles. The sign of the central torsion angle χ3 is used to distinguish between right-handed (RH) and left-handed (LH) disulfide bond. PDB entry 1DL5 was used to generate Figure 1.11a, visualized using USCF Chimera [204]. Figure 1.11b reprinted with permission from Hogg [93]. © 2013, Springer Nature.*

Additional to the specific dihedral angles, disulfide bonds exhibit specific bond lengths between the two cysteines. Petersen *et al.* [203] identified that the most frequent Cα-Cα distance in left-handed disulfide bridges peaks at 5.8 Å, while right-handed disulfide bonds exhibit values of 5.4 Å and 4.0 Å. The Cβ-Cβ distances of 72 disulfide bonds within 22 proteins showed a range of 3.5-4.5 Å [247]. According to the study from Petersen *et al.* [203] free cysteine residues seem to prefer an α-helical structure, while disulfides seem to be located preferentially in extended conformation of β-sheets. Further, they identified that free cysteine residues are close to histidine and methionine residues both sequential and in space, while the charged residues lysine, glutamate and aspartate as well as

the polar residue asparagine show a decreased occurrence close to free cysteines. Cysteine residues which are part of a disulfide bond seem to be located close to tryptophan and arginine residues. The location close to aromatic amino acids can result in S-π interactions as mentioned before.

1.6.1   Prevention of disulfide bond formation and cysteine modifications in prokaryotes

In prokaryotes disulfide bond formation usually occurs within the periplasm of gram-negative bacteria. Although some proteins haboring disulfide bonds were identified in the *E. coli* cytoplasm, they are rare and several pathways exist to keep cytoplasmic proteins reduced. These pathways include glutathione (GSH), glutaredoxin (Grx) and thioredoxin (Trx).

1.6.1.1 Glutathione (GSH)

Glutathione (GSH) is a tripeptide of glutamic acid, cysteine and glycine with a γ-peptide bond between the carboxy side chain group of glutamic acid and cysteine (Figure 1.12). This γ-peptide bond protects the tripeptide from intracellular peptidases [243]. The molecular mass of GSH is 307 Da and at physicological pH the carboxy groups are negatively charged, while the amino group is positively charged. Since GSH is less prone to oxidation than cysteine, it is an ideal candidate to maintain the intracellular redox potential and the thiol group of GSH provides a strong electron-donating character [40]. Through oxidation two molecules of GSH form a disulfide bridge and the glutathione disulfide (GSSG) is formed (Figure 1.12). The redox potential of a reaction can be calculated using
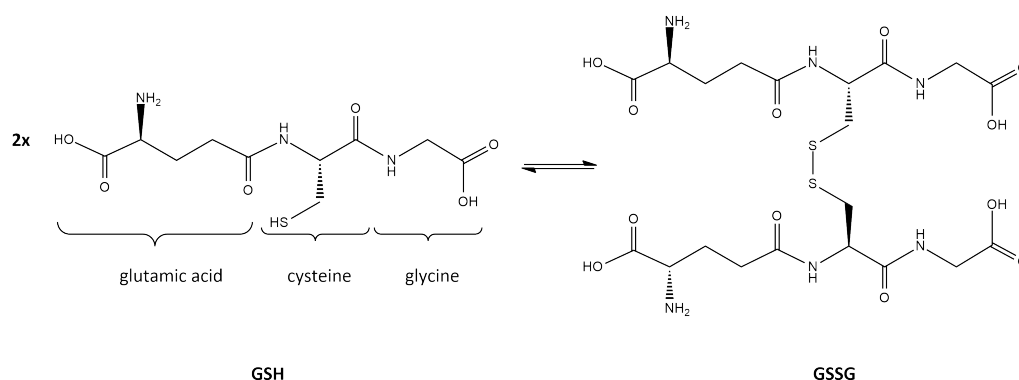
Figure 1.12: *Structure of glutathione.* *Structural formula of glutathione (GSH) (left) and glutathione disulfide (GSSG) (right).*

the Nernst equation (equation 5).

$$E = E^0 + \frac{RT}{z_e F} ln \frac{a_{Ox}}{a_{Red}} \tag{5}$$

With E^0 as the standard electron potential, R the universal gas constant (R=8.314 JK^{-1}mol^{-1}), T the temperature in Kelvin, F the Faraday constant (F=

96485.33 Cmol^{-1}), z_e the number of electron transferred and a the chemical activity. The lower the redox potential of an environment the more likely new synthesized compounds get reduced. At room temperature and pH 7.0 the Nernst equation for the GSH/GSSG redox couple can be written as followed (equation 6)[40]:

$$E = -240 - \frac{59.1}{2} log \frac{[GSH]^2}{[GSSG]} mV \tag{6}$$

GSH is the most abundant low-molecular–mass thiol with concentrations of 5 mM in the *E. coli* cytoplasm and the GSH/GSSG ratio can range from 50/1 to 200/1 [8, 41, 173]. This results in a redox potential of -240 mV in *E. coli* assuming a GSH/GSSG ratio of 200 and a GSH concentration of 5 mM. The GSH/GSSG ratio is usually closely regulated, but as can be seen from the Nernst equation, changes of the GSH concentration can influence the redox potential without changing the GSH/GSSG ratio [40]. A GSH concentration of 10 mM or 1 mM within an *E. coli* cells results in a difference of 29 mV in the redox potential without change of the GSH/GSSG ratio. GSH is found in prokaryotic and eukaryotic cells, however gram-positive bacteria usually do not contain GSH but generate other low-molecular weight-thiols such as mycothiol in *Actinabacteria* [54, 53, 243]. Glutathione is synthesized in the cell via two enzymes the γ-glutamyl-cysteine synthetase (γ-GCS) and GSH-synthetase (GS). The first one catalyzes bond formation between the γ-carboxy group of glutamic acid and cysteine, while the latter catalyzes bond formation between this dipeptide with glycine [243]. The cellular functions of GSH include the reduction of reactive oxygen species (ROS) by forming a thiyl radical (-S$^{\bullet}$), which is unstable under physiological pH and subsequently forms the GSSG dimer [41]. GSSG reacts easily with free thiol groups and is therefore highly toxic to cells [243]. To prevent high levels of GSSG in the cell, glutathione reductase (GOR) catalyzes the reduction of GSSG to GSH requiring NADPH. The three major forms of glutathione within a cell are the reduced form GSH, the disulfide GSSG and mixed disulfide with protein (GSS-protein or PSSGs), although 99.5% of glutathione exists in the reduced form and the mixed disulfide does not exceed 1% [54, 242, 243]. The mechanism which creates this reversible bond to protein cysteine residues is called S-glutathionylation and results in S-glutathionylated proteins (PSSGs) [41]. S-glutathionylation is considered to be a post-translational modification and usually occurs in response to oxidative stress, although this modification has also been observed under physiological conditions [32, 237, 41]. Protein thiols can also be oxidized to sulphenic (RSOH), sulphinic (RSO$_2$H) or sulphonic (RSO$_3$H) acid in the presence of ROS and an oxidative environment. Hence, in *E. coli* this formation does not usually occur in the cytosplasm but in the periplasm [41]. While sulphenic acid can be reduced by GSH, the oxidation to sulphenic and sulphonic acids remains irreversible [41]. Next to these modifications, protein thiols can also form intra- or inter-disulfide bridges. Figure 1.13 shows possible

mechanisms of protein S-glutathionylation and deglutathionylation within the cell. Glutathionylation occurs especially at cysteine residues with high accessibility and a low pK_a. Several parameter influencing the S-glutathionylation and S-nitrosylation site are shown in Figure 1.14a.
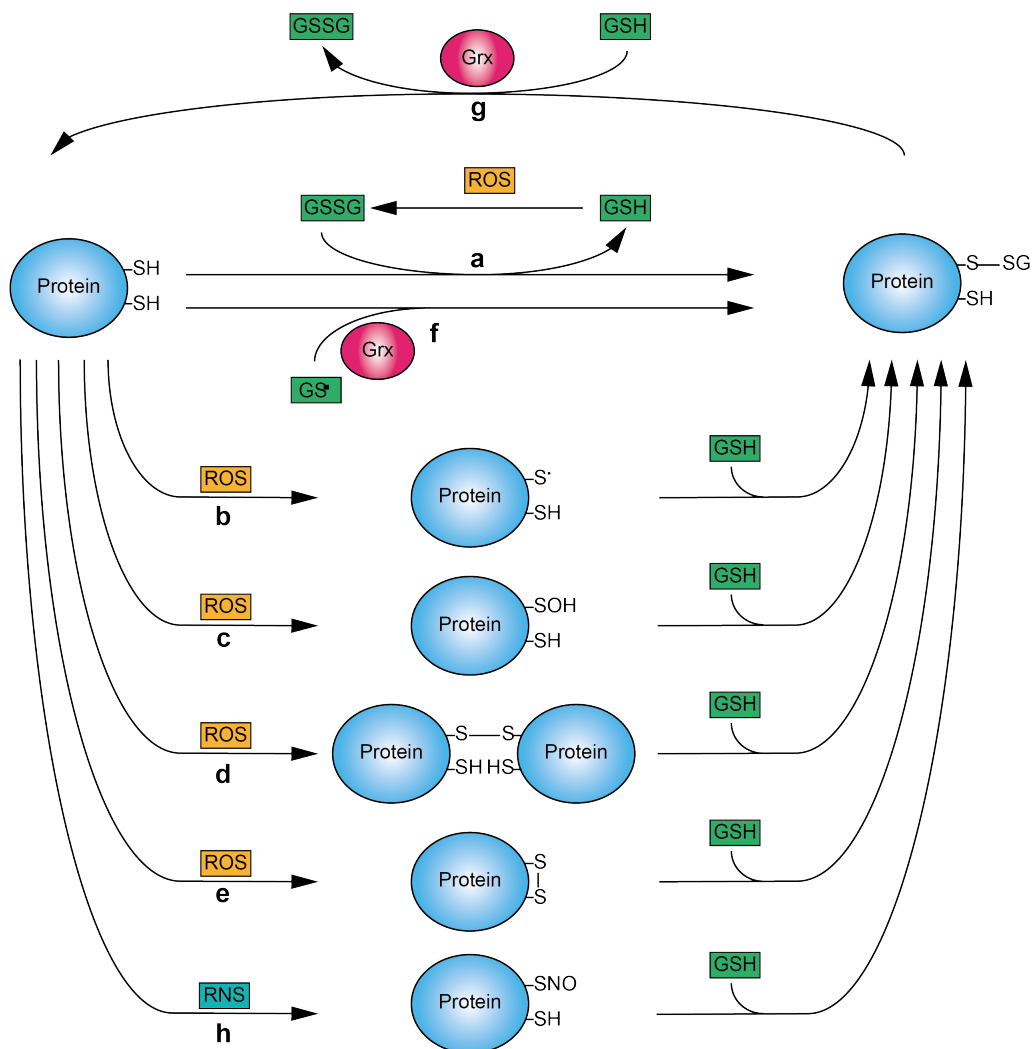
Figure 1.13: *S-glutathioylation,   S-nitrosylation   and   deglutathionylation mechanisms.a)* S-glutathionylation via GSSG increase in the GSS-G/GSH ratio.   Reactive oxygen species (ROS) catalyze generation of thiyl radicals (**b**), sulphenic acid (**c**), protein dimerization (**d**) or intra-disulfide bond formation (**e**), which then become glutathionylated by GSH in the absence of GSSG. The glutathionylation via GS• is catalyzed by glutaredoxin (GRX) (**f**), which usually acts as a reductant (**g**). Reactive nitrogen species (RNS) react with cystein residue to form a S-nitrosylated protein (**h**). Further reaction with GSH results in a S-glutathionylated protein. Figure modified after Dalle-Donne et al. [41].*

1.6.1.2 S-nitrosylation/S-nitrosation

Aside from S-glutathionylation as post-translational modification, S-nitrosation or S-nitrosylation of several proteins (protein-SNO) has also been recorded [163, 87, 82, 88, 125]. A good overview of the S-nitrosylation sites of different pro-

teins as well as the effect of this modifications is given in the review by Hess *et al.* [88]. S-nitrosylation might not only occur as cellular response to nitrosative stress, but might also be implicated in the regulation of numerous signaling pathways [41, 88].

The terms S-nitrosylation or S-nitrosation are often used synonymously, although the latter is the proper term for the formation of a nitroso-group (R-NO) by a one-eletron oxidation from a NO$^\bullet$ radical [244]. But also other reactive nitrogen species (RNS), like N_2O_3 or $^+$NO are involved in the formation of protein-SNO [2] (Figure 1.14b). For the reaction of a NO$^\bullet$ radical with a thiol an electron sink (for example oxygen or transition metals) is required [244]. NO$^\bullet$ itself can be produced by nitric oxidase synthases (NOS) and is involved in many physiological processes as a signaling molecule [2]. Under aerobic conditions NO$^\bullet$ can react to N_2O_3, which reacts with water to NO_2^- and with a thiolate to S-nitrosothiol [244]. Thiol radicals can directly react with the NO$^\bullet$ to form the protein-SNO. The measured rate constants for the reaction of NO$^\bullet$ with thiyl radicals of cysteine and glutathione are in the range of (2-3) x 10^9 $M^{-1}s^{-1}$, and hence only a factor of 2-3 times lower than the diffusion-controlled limit of water [157]. The reaction of NO$^\bullet$ with thiols to form nitrosating thiols is impossible in the absence of oxygen and relatively slow under physiological conditions, resulting in a very unlikely biosynthetic pathway [78]. The direct S-nitrosylation of GSH to form GSNO with N_2O_3 is two orders of magnitude slower and in the range of 6.6 x 10^7 $M^{-1}s^{-1}$ [244, 78]. Another mechanism for the formation of protein-SNO involves binding of NO$^\bullet$ to ferric heme. The Fe^{3+}-NO, exhibits a significant Fe^{2+}-NO$^+$ character, the loss of NO$^+$ results in the formation of nitrosothiol and heme reduction [244, 2]. Also possible is the transnitrosation of proteins via GSNO, which was monitored using HPLC by Hogg [92].

The reactivity of the cysteine residues depends on its pK_a value. Cysteines with a low pK_a and an exposed environment are more likely to become nitrosylated than cysteine residues within the inner core of a protein and a high pK_a value. If cysteine residues are adjacent to basic amino acids (His, Lys, Arg) or charged residues as (Ser, His), this could be a possible site for S-nitrosylation [41]. Parameter influencing the reactivity of Cys for S-glutathionylation (Figure 1.14a) are also valid for S-nitrosation.

While the protein-SNO can also react with glutathione forming S-glutathionylated protein (Figure 1.13), denitrosylation can be carried out by the oxidoreductase thioredoxin [88].

1.6.1.3 Thioredoxin (Trx)

Another mechanism to prevent unwanted disulfide bond formation within the cytoplasm is the thioredoxin/thioredoxin-reductase system. Thioredoxins (Trx) are small proteins of around 10-12 kDa and present in all kingdoms of life [16]. Two different Trx are present within *E. coli*, yeast and mammalian cells [98].
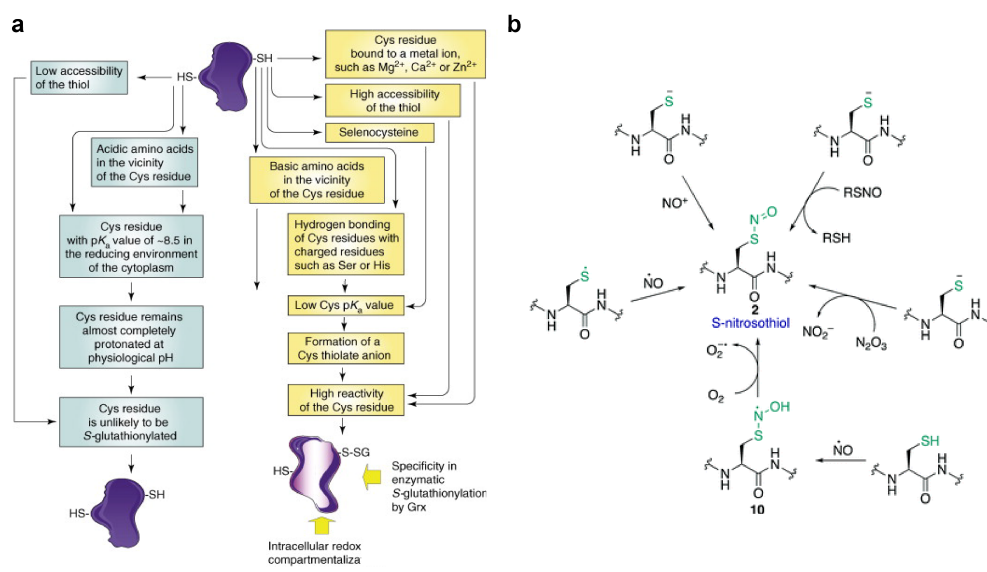
Figure 1.14: *Important parameters for S-glutathionylation and S-nitrosylation mechanism. **a)** Parameters influencing the reactivity of cysteine residues and hence the specificity of protein S-glutathionylation, but also S-nitrosylation. **b)** Mechanism of S-nitrosothiol formation in proteins. Figure 1.14a reprinted with permission from Dalle-Donne et al. [41]. © 2008 Elsevier Ltd. All rights reserved. Figure 1.14b reprinted with permission from Alcock et al. [2]. Published by The Royal Society of Chemistry.*

E. coli Trx1 (encoded by *trxA*) is 12 kDa in size and was first discovered as an electron donor for ribonucleotide reductase (RNR) in 1964 [146]. The mammalian Trx1 is localized in the nucleus/cytosol, Trx2 (encoded by *trxC*) is 15.5 kDa in size and localized within mitochondria [246, 16]. Intracellular concentrations of Trx vary between 1-10 µM in bovine tissue and 15 µM in bacteria [99, 222]. Compared to GSH the Trx concentration is therefore 100-1000 fold lower. Trx have a characteristic tertiary structure, the so-called Trx fold motif, which consists of four-stranded β-sheets surrounded by three α-helices with the basic βαβαββα-topology found in glutaredoxins [100, 97]. An additional feature of this Trx motif is a cis-proline located before the third β-sheet and the Cys-X-X-Cys active site, which is located on the loop between β-sheet one and α-helix one [100]. Depending on the amino acids between these two cysteine residues, the redox potential can differ dramatically. The strongest redox potential with $\Delta E'_0$=-270 mV in *E. coli* is obtained by the cytosolic Trx with an Cys-Gly-Pro-Cys active site [139]. Reduction of a disulfide bond is generated by a ping-pong mechanism [96], where the N-terminal cysteine (Cys32) in the active site possesses an unusual low pK_a value of 6.7 and the C-terminal cysteine (Cys35) a rather high pK_a value of 9.0 [120]. The low pK_a value of the N-terminal cysteine in the active site is related to the proximity of the carboxygroup of Asp26 and the ε-amino group of Lys57, which keep the cysteine residue deprotonated at physiological conditions [47]. The mechanism of disulfide reduction is shown schematically in Figure 1.15a.

After reduction of the protein, oxidized Trx is reduced by thioredoxin-reductase (TrxR) using NADPH and FAD. Beside Trx and GSH, glutaredoxins are also involved in the reduction of disulfide bridges.
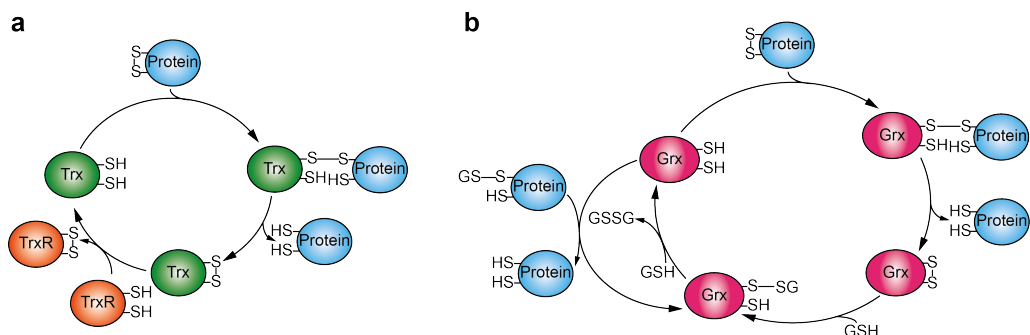
Figure 1.15: *Schematic mechanism of the redox cycle of thioredoxins (Trx) and glutaredoxins (Grx). **a)** Trx reduces a disulfide bond within a protein using both active site cysteine residues. A mixed disulfide with the target protein is formed as intermediate. The oxidized Trx is reduced by thioredoxin-reductase (TrxR). **b)** Grx reduces a disulfide bond within a protein forming an mixed disulfide. Oxidized Grx is reduced by GSH leading to a mixed disulfide, which is further reduced by another GSH. Grx can also reduce S-glutathionylated proteins. Figure modified after Berndt and Holmgren [15].*

1.6.1.4 Glutaredoxin (Grx)

Glutaredoxins (Grxs) exists in all organisms containing GSH, as GSH is used as electron donor. Hence Grxs are not found in gram-positive bacteria. Grxs also belong to the thioredoxin family and exhibit the basic Trx motif of βαβαββα-structural element [100, 97]. Two different Grxs are distinguished, the dithiol Grxs with Cys-Pro-Tyr-Cys and the monothiol Grxs with Cys-Gly-Phe-Ser in the active site [15]. Grxs are general thiol-disulfide oxidoreductases and reduce protein disulfide bridges but also protein-GSH mixed disulfides [15]. *E. coli* contains three dithiol Grxs 1-3 [95, 7, 15], while mammalian cells contain the two dithiol proteins (Grx1 and Grx2) [98], and one monothiol Grx (Grx5) [267]. While Grx1 is localized within the cytosol of mammalian cells, Grx2 is found in the nucleus and mitochondria [98]. Monothiol Grxs, reduce GSH-mixed disulfide proteins only with the N-terminal cysteine residues in the active site, whereas dithiol Grxs use both active site cysteine residues for the reduction of disulfide bonds [219]. The size of the Grxs varies between 9-13 kDa for Grx1, Grx3 and Grx4 and Grx2 is a larger protein with 24 kDa [15]. Grx2 exhibits an overall GSH-S-transferase structure and only the N-terminal part forms the classical Trx-fold glutaredoxin domain [273]. Grx4 is the only essential member of the Trx family proteins in *E. coli* [59, 15]. The standard state redox potential for Grx1 and Grx3 from *E. coli*, both containing Cys-Pro-Tyr-Cys in the active site were determined to be E°'= -233 mV and -198 mV, respectively [6]. The redox potential of human Grx were also determined to lie in this range, with -232 mV and -221 mV for hGrx1 and

hGrx2, with Cys-Pro-Tyr-Cys and Cys-Ser-Tyr-Cys in the active site [219]. A schematic mechanism for the reduction of glutathionylated-protein and a protein containing a disulfide bridge is shown in Figure 1.15b. Since GSH, Trx, and Grxs keep the cytosol a reductive environment due to their high redox potential, the periplasm is a more oxidizing environment favoring disulfide bridge formation.

1.6.2   Disulfide bond formation in prokaryotes

In prokaryotes, disulfide bond formation usually occurs within the periplasm of gram-negative bacteria, due to the reductive environment of the cytoplasm. Disulfide bond formation is catalyzed and corrected by proteins named Dsbs (Dsb for disulfide bond) (Figure 1.16). One of the probably best studied DSB-forming systems is the DsbAB system of *E. coli*. DsbA is a 21 kDa monomeric protein and contains the classical Trx-fold including the CXXC motif [195]. The cysteine residues within the CXXC motif are in a disulfide form. DsbA catalyzed the disulfide bond of a protein substrate by a thiol-exchange reaction [119] and is one of the strongest thiol oxidants with a redox potential of -119 mV arising from the unusual low pK_a of 3.5 for Cys30 [173, 183].

The reduced DsbA is then re-oxidized to its active disulfide state by the inner-membrane protein DsbB. DsbB, a 20 kDa protein, passes then the electrons it gained from DsbA to ubiquinone and subsequently to molecular oxygen via the electron transport chain under aerobic conditions. Under anaerobic conditions, the electrons are passed to menaquinone and then to fumarate or nitrate reductase [9, 250]. DsbB does not obtain a Trx-fold, but six cysteines of which four are essential for the reoxidation of DsbA and two are also located within a CXXC motif [143, 11, 112]. Since DsbA is a very powerful, but not specific oxidant and will catalyze even the formation of non-native disulfide bonds. Those proteins are isomerized back to their native state by the DsbCD system. DsbC, a 23 kDa homodimeric-protein, consists of two domains, a C-terminal thioredoxin with a CXXC motif and a N-terminal dimerization domain [168]. In contrast to DsbA, DsbC is kept in the reduced state within the periplasm, but possesses a similar redox potential of -130 mV [173]. The reduced form of DsbC interacts with the misfolded protein via a thiol-disulfide exchange. If the disulfide bond within the substrate protein is only reshuffled, DsbC remains in its active reduced state. In case, the disulfide bond within the substrate needs to be reduced, DsbC remains oxidized until its reduction by DsbD.

Additionally to DsbC a second protein disulfide isomerase, DsbG, is also present in the periplasm, although with a four-fold less abundance [173]. DsbG shares a 24% amino acid identity to DsbC and a similar redox potential between -127 and -129 mV [173]. DsbD, an inner-membrane protein, reduces both isomerases by direct thiol-disulfide exchange. The electrons for this reaction are received from cytoplasmic Trx.
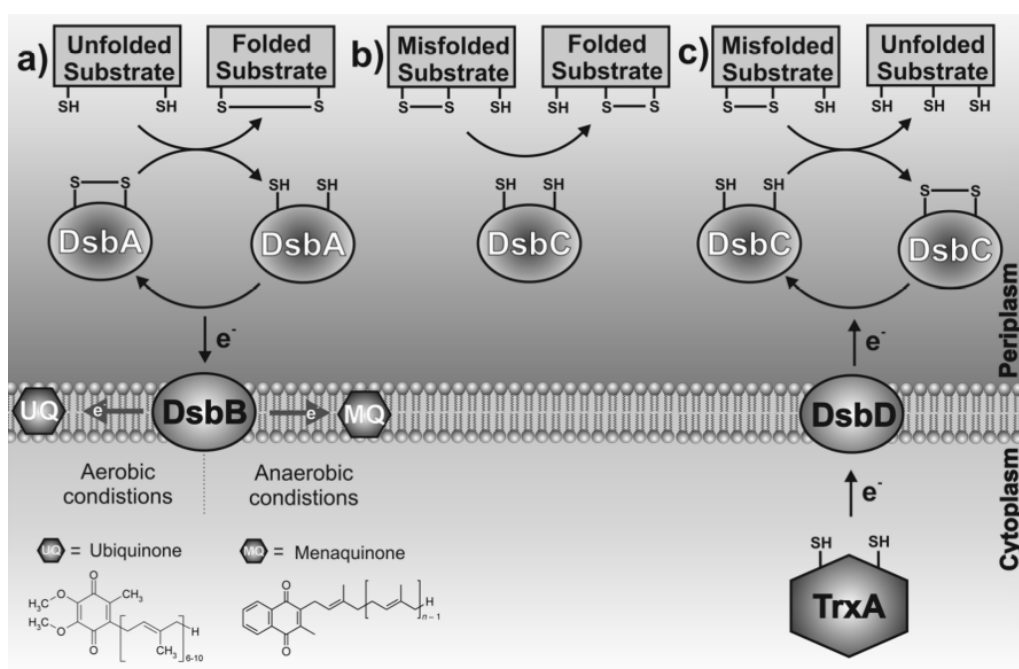
Figure 1.16: *Disulfide bond forming system in the E. coli periplasm. a) Oxidation of a target protein by DsbA via thiol-disulfide exchange. DsbB is reoxidized by DsbB that transfers the electrons to ubiquinone under aerobic and to manaquinone under anerobic conditions. b) Non-native disulfide bonds can be reshuffled by DsbC via thiol-disulfide exchange, remaining the DsbC in a reduced form. c) Reduction of non-native disulfide bonds can be reduced by DsbC. The oxidized DsbC is re-reduced by DsbD provided by electrons from thioredoxin (TrxA). Figure reprinted with permission from Silvers, R., Schlepckow, K., Wirmer-Bartoschek, J. & Schwalbe, H. [238]. © 2011, Springer Science Business Media, LLC.*

1.6.3   Disulfide bond formation in eukaryotes

For the sake of completeness, the disulfide bond formation within eukaryotes should be mentioned here, briefly. In eukaryotes the oxidative folding takes part in the lumen of the endoplasmic reticulum (ER) with a redox potential of -180 mV [272]. The oxidative folding pathway of *S. cerevisiae* requires the two proteins: Ero1p (ER oxidoreduction 1 protein) and PDI (protein disulfide isomerase). Ero1p is a 65 kDa, flavine adenine dinucleotide (FAD)-containing, glycosylated luminal ER protein, which is tightly associated with the ER membrane and functions as a oxidase similar to the bacterial DsbB protein [64, 205, 272, 233]. Ero1p transfers oxidizing equivalents directly to PDI, which then oxidizes substrate proteins [65, 254, 233]. Both reactions occur via direct thiol-disulfide exchange reactions [65]. Ero1p contains four essential cysteine residues, which form two active site cysteine pairs: Cys100-Cys105 and Cys352-Cys355 [66].

PDI, a 57 kDa protein is essential for *S. cerevisiae* viability [272] and contains two thioredoxin domains with two active sites consisting of a CXXC motif. The redox potential of those two sites are -188 mV and – 152 mV [158]. Beside the oxidase function, PDI is also involved in reshuffling of disulfide bonds of misfolded

proteins. A second pathway, including a small ER oxidase (Erv2) has been iden-
tified in yeast. Erv2, a 22 kDa membrane-associated ER protein binds FAD.
Using molecular oxygen as electron acceptor, Erv2 can catalyze the formation of
disulfide bonds as for example in PDI via the FAD cofactor [232, 233]. Figure
1.17 shows an overview of the redox potentials of the major cellular proteins,
low-molecular weight thiols (DTT, GSH), the cofactor ubiqinone and molecular
oxygen.
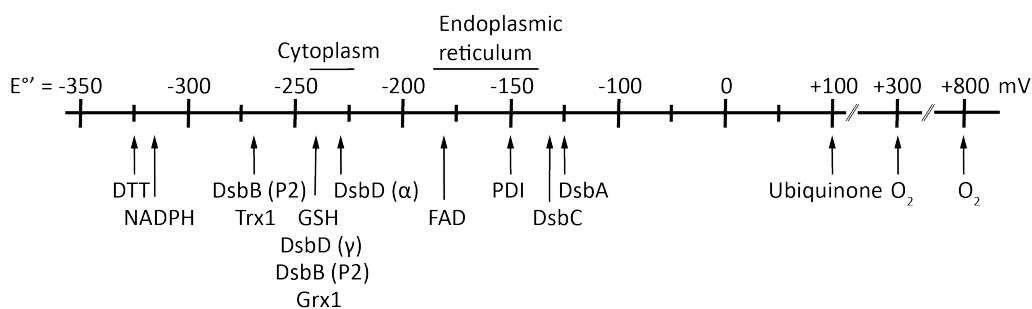
Figure 1.17: *Redox pontential of major cellular redox components. Redox pon-
tential of major cellular redox proteins, low-molecular weight thiols (DTT,
GSH), the cofactor ubiquinone and the O_2-H_2O_2 and O_2-H_2O redox pair
with +300 and +800 mV, respectively. Figure modified after Sevier and
Kaiser [233].*

1.6.4   Disulfide bond formation and protein folding

Two models describing the relationship between protein folding and disulfide
bond formation exist [265]:

(I) the folded precursors mechanism in which cysteine residues are brought in
proximity after a structure is formed

(II) the quasi-stochastic model in which cysteine residues pair in an unfolded
precursor and influence further protein folding

While Welker *et al.* [265] argued that the quasi-stochastic mechanism might be
generally favored over the folded-precursor mechanism, especially for folding-
coupled reshuffling steps as in RNase A, Qin *et al.* [206] showed with coarse-
grained molecular simulations that the bovine pancreatic trypsin inhibitor (BPTI)
follows the folded-precursor mechanism. BPTI is basically the model protein in
the field of oxidative folding. Experiments by Creighton [37] at a pH of 8.7,
thus around the pKa of the cysteine side chain, observed intermediates with non-
native disulfide species. At a almost neutral pH of 7.3, which is below the pKa
of cysteine residues only native disulfide bonds accumulated during the folding
process [264]. These experiments, however, were not able to solve the question if
disulfide bond guides protein folding or not.
Jansens *et al.* [113] showed with pulse-chase experiments that presumably non-
native inter-domain disulfide bonds are prominent in folding intermediates of the

low-density lipoprotein receptor (LDL-R), which are re-arranged into the native intra-domain disulfides. This observation is also in favor of the quasi-stochastic model. In addition, Land *et al.* [142] also observed early disulfide bond formation in glycoprotein gp120 and gp160, which are then isomerized into the native conformation and only after disulfide bond formation in gp120 the cleavage of the leader peptide occurred.

In contrast to LDL-R and the glycoproteins, the 27th Ig domain from human cardiac titin does also contain a non-native inter-domain disulfide at an early stage in the protein folding pathway, however this disulfide bond is not re-arranged to a native one and can trigger misfolding [136]. Further, PDI catalyzed the disulfide bond formation of Ig at a late stage of protein folding, since the mixed-disulfide bond to PDI remained until Ig had attained its near-native state [136]. These observations led to the assumption that protein folding drives the disulfide bond formation in Ig.

While the measurements on Ig were performed by atomic force microscopy on a full-length protein *in vitro*, Robinson *et al.* [212] used an eukaryotic *in vitro* translation system and stalled the β2-microglobulin (β2M) with different linker-lengths on the ribosome. β2M obtains a disulfide bonds only after the entire sequence of the protein domain was translocated into the ER. This result is in favor of the folded-precursor model. However, the cysteine residues involved in disulfide bond formation within β2M are Cys45 and Cys100 and thus 55 amino acids apart from each other. In a recent study Robinson *et al.* [211] investigated three structurally diverse proteins β2M, prolactin and the disintegrin domain of ADAM metallopeptidase domain 10 (ADAM10) using a similar approach. Their results indicated that β2M and prolactin follow the folded-precursor mechanism. In domains lacking secondary structure, however, disulfide bonds can be formed also through non-native disulfide bond formation before protein folding occurs, hence providing evidence for the quasi-stochastic model in the disintegrin domain. With a lower cysteine content, the quasi-stochastic model was also retained. In summary, depending on the protein, its secondary structure and the distance of the cysteine residues within the primary sequence, either one of the model is observed or even both models can show contributions at different stages of the protein folding process [156].

Beside the relationship between protein folding and disulfide bond formation, also the question arises at which state in protein synthesis, co-translationally or post-translationally, disulfide bond formation takes place. The studies by Robinson *et al.* [212, 211] with arrested NCs, showed that co-translational disulfide bond formation can occur. Further, Hoffmann *et al.* [90], who studied the influence of the TF on disulfide bond formation, were also able to detect disulfide bonds within translation arrested constructs of β-lactamase, barnase and m10, the folding-deficient point mutant of the small β-sheet Src-homology 3 (SH3) domain of α-spectrin. And also the studies by Bergman and Kuehl [13], Jansens

et al. [113] and Peters and Davidson [202] were able to detect disulfide bond formation within nascent protein chains. In addition, PDI was cross-linked to NC of prolactin [127] and using a PDI-deficient *in vitro* translation/translocation system showed that PDI is essential for co-translational disulfide bond formation in γ-gliadin [24]. These two experiments also favor co-translational disulfide bond formation.

1.7   Motivation

Buhr *et al.* [22] showed that codon usage had an influence on the co-translational folding of GBC. The unmodified version with the gene sequence from *B. taurus* was in the fully reduced state and translated with 1.8 aa/s, whereas the harmonized version, which was optimized to the *E. coli* codon usage, was partly oxidized and translated with 2.2 aa/s. Also the resistance to protease K differed, with 86% protease K resistance for H and 65% for the U variant.

Since, both proteins were treated the same way, the differences might occur during translation of the proteins inside the cell. Therefore, we wondered if possible disulfide bond formation could happen at an early state of translation or even inside the ribosomal exit tunnel (Figure 1.18). This early disulfide bond might be protected in the fast-translated variant H and would be reduced by the *E. coli* cytosol in the slow-translated variant U.
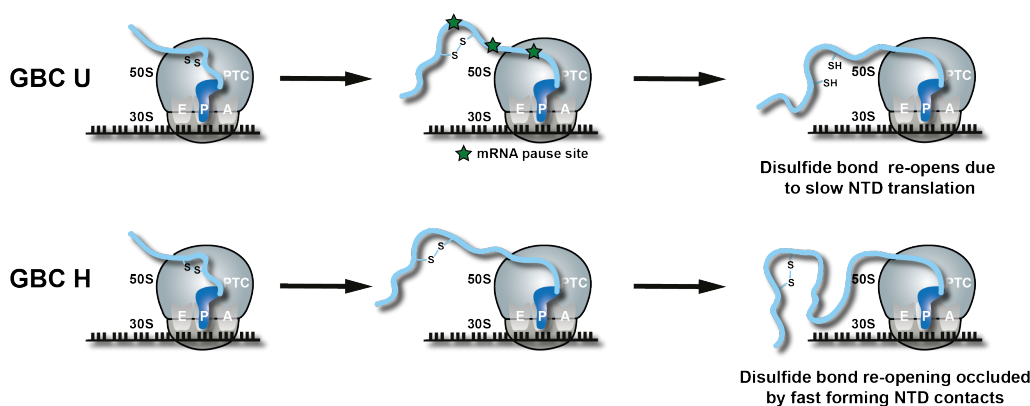
Figure 1.18: ***Model for co-translational GBC oxidation.*** *Hypothesis exlaining the conformational differences in gene variants U and H of GBC. A disulfide bond in the ribosomal exit tunnel or at an early state of translation is formed in both gene variants. The disulfide bond is reduced by the E. coli cytosol in the slow translated version U, whereas this bond is protected by fast-forming NTD contacts or the CTD in the gene variant H.*

To determine, if the NTD can fold indepedently and whether the CTD of GBC is required for the observed oxidation in the NTD, the single NTD was studied using solution-state NMR. To study the oxidation of GBC at an early state of translation or inside the ribosomal exit tunnel, translation-arrested GBC constructs were designed. These constructs were studied by mass spectrometry, solid-state NMR

enhanced by DNP and cryo-EM in respect to their oxidation state. The conformations of the nascent polypeptide chain were simulated using flexible-meccano, to invesitagte if the ribosomal tunnel provides enough space for disulfide bond formation. In addition, the full-length GBC was oxidized with $CuCl_2$ to investigate which disulfide bonds could potentially form. All three topics with addressed questions in this thesis are summarized in Figure 1.19.
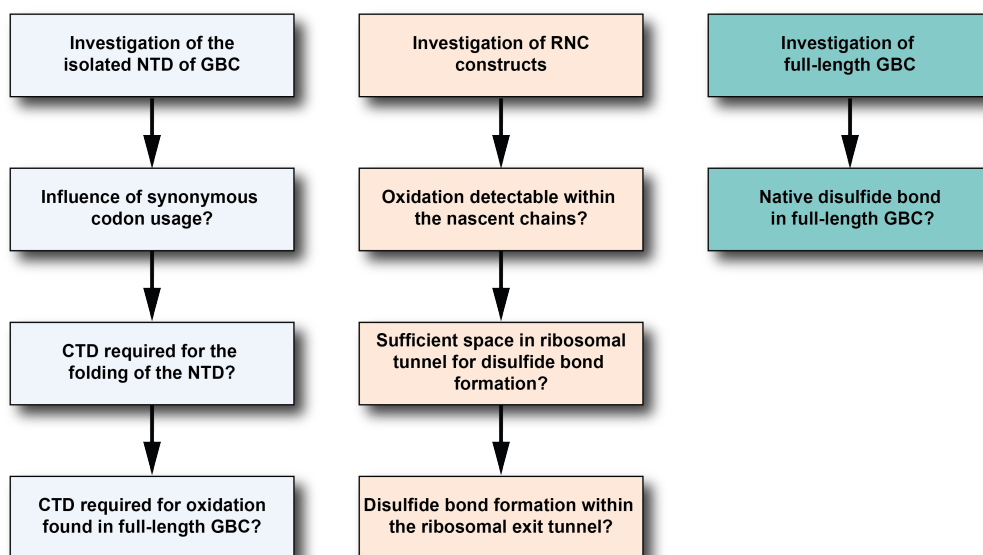
Figure 1.19: ***Different topics with questions addressed in this thesis.*** *The NTD of GBC (light blue) and the full-length GBC (green) were studied by liquid-state NMR. RNCs of GBC (dark yellow) were investigated by solid-state NMR, cryo-EM and mass spectrometry.*

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals, enzymes and standards

Unless further clarified, all chemicals were obtained from Carl ROTH (Karlsruhe, Germany), Sigma-Aldrich (Steinheim, Germany) or Merck KGaA (Darmstadt, Germany). Enzymes and DNA ladders were purchased from New England Biolabs GmbH (NEB) (Frankfurt, Germany). Low-molecular weight marker was obtained from Roche (Basel, Switzerland) and the spectra multicolor broad range protein ladder for western blots was bought from Thermo Fisher Scientific (Waltham, USA).

2.1.2 *E. coli* strains

Competent *E. coli* strains such as DH5α and BL21(DE3) were obtained from NEB (Frankfurt, Germany). BL21(DE3)Δtig::Kan cells were provided by Bernd Buckau and JE28 cells by Suparna Sanyal.

2.1.3 Other materials

Gene sequences and primer were obtained by Eurofins Genomics GmbH (Ebersberg, Germany) or GenScript (New Jersey, USA). DNA sequencing was performed by Eurofins Genomics GmbH (Ebersberg, Germany). Protein purification was carried out on a Äktapurifier, Äktabasic or Ätkapure system from GE Healthcare (Chicago, USA). Consumable material was provided by Sarstedt (Nürnbrecht, Germany) and neoLab® Migge GmbH (Heidelberg, Germany). DNA purification kits were obtained from Qiagen (Hilden, Germany).

2.2 Molecular Cloning

2.2.1 Cloning of the N–terminal domain of GBC

The DNA sequences of the full-length harmonized (H) and unmodified (U) GBC containing a C-terminal 6x histidine tag on the expression vector pET15b were provided by Dr. Florian Buhr. The sequences for both full-length constructs are listed in the appendix. The NTD constructs were designed with Pro83 as last amino acid performing two consecutive PCRs. Table 2.1 displays the DNA sequence and melting points of the forward (.fwd) and two reverse (.rev) primers used to design the NTD constructs. All primers were ordered at Eurofins (Ebersberg, Germany). The PCR was performed using the forward primers and the first set of reverse primers in Table 2.1 under the conditions shown in Table 2.2. The total volume of the reaction was 50 µl. The obtained PCR products were loaded

on a 2% (w/v) preparative agarose gel in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA). After agarose gel electrophoresis (section 2.2.7.) PCR products of the corresponding size were cut out of the gel and extracted using the Qiagen gel extraction kit (Hilden, Germany). DNA concentration after the extraction was around 20-30 ng/µl. These DNA fragments were used as a template for another PCR with the second set of reverse primers. After gel extraction the purified PCR products were ligated into the linearized and dephosporylated expression vector pET15b.

Table 2.1: *Primers used to design the NTD of GBC. Name, DNA sequence and melting temperature (T_M) of primers used for the production of N-terminal domain of GBC.*

Primer name	DNA Sequence 5'→3'	Melting point (T_M in °C)
WT.fwd	TATACCATGGGGAAGATCACT	57.5
H.fwd_corr	TATACCATGGGAAAAATCACTTT	55.5
WT83-1.rev	ATGATGCGGGATGAGGCGG	61.6
WT83-2.rev	GATCCTCGAGTCAATGATGA TGATGATGATGCGGGATGAGG	79.0
H83-1.rev	ATGATGTGGGATTAAGCGACAGC	62.9
H83-2.rev	GATCCTCGAGTCAATGATGATGAT GATGATGGTGGGATTAAGCGG	79.9

Table 2.2: *Thermocycling conditions for routine PCR with Phusion. Temperature (°C), time (s) and repetition of each PCR step.*

Step	Temperature (°C)	Time (s)	Repetitons
Initial Denaturation	98	30	1x
Denaturation	98	10	30x
Annealing	55	30	30x
Elongation	72	1min per kb	30x
Final Elongation	72	5min	1x
Pause	4	∞	

2.2.2 Restriction enzyme digestion and dephosphorylation

The pET15b expression vector and inserts obtained by PCR were digested using the restriction enzymes NcoI-HF and XhoI from NEB (Frankfurt, Germany). The recognition sequences of these enzymes are shown in Table 2.3. For the digestion 0.7-1.5 µg of the insert was incubated for 3-4 hrs at 37°C using 20,000 U/ml NcoI-HF and XhoI in Cut Smart buffer from NEB (Frankfurt, Germany). To

prevent ligation of the empty vector, dephosphorylation of the plasmid was performed using CIP purchased from NEB (Frankfurt, Germany). To 1 pmol of DNA, the corresponding volume of 10x CutSmart buffer and 1 unit of CIP were added and the volume was adjusted with H_2O. The reaction was incubated at 37°C for 1 h and afterwards purified using the QIAquick PCR purification kit by Qiagen (Hilden, Germany).

Table 2.3: *Recognition sequence of NcoI-HF and XhoI.*

Restriction enzyme	Recognition sequence
NcoI-HF	C-CATGG
XhoI	C-TCGAG

2.2.3   Ligation

Ligation was performed according to the NEB Ligation protocol with T4 DNA Ligase. 50 ng of vector were combined with a 3-fold molar excess of insert and adjusted to 10 µl with H_2O, when 2x Quick Ligation buffer was used, or to 18 µl, when 10x T4 DNA Ligase buffer was used. 10 µl of 2x Quick Liagtion buffer or 2 µl of 10x ligation buffer were added and the reaction was mixed. 1 µl of Quick T4 DNA Ligase or T4 DNA Ligase (400,000 U/ml) was added and the reaction was incubated at RT for 5 min or 10 min, respectively. If the T4 DNA Ligase was used, the reaction was stopped by heat inactivation at 65°C for 10 min. 5 µl of the liagation mix were used for transformation.

2.2.4   Site directed mutagenesis

To create different mutants, like exchange of cysteine to alanine, site-directed mutagenesis was carried out using the QuikChange II or QuikChange Lightning Site-Directed-Mutagenesis Kit from Agilent Technologies (Santa Clara, USA). Primers used for the mutagenesis reactions are shown in Table 2.4. The primers were designed using the web-based QuikChange Primer Design tool from Agilent Technologies (Santa Clara, USA). Volumes and concentrations used for the mutagenesis reactions using the lightning kit are shown in Table 2.5. PCR reaction was carried out using the heating program shown in Table 2.6. If the QuikChange II kit was used, the quicksolution was omitted and the condition shown in Table 2.6 in brackets were used.

2.2.5   Transformation and preparation of overnight cultures

50 µl of chemical competent cells, (DH5α, BL21De3 or BL21(De3)Δtig::Kan) were thawn on ice for 10 min. 2 µl (100-150 ng) of plasmid or 5 µl of ligation reaction were added and mixed by pippeting up and down. The cells were kept on ice for futher 30 min, and afterwards heat shocked for 30-45 s at 42°C in a water bath (GFL, Burgwedel, Germany). Cells were placed on ice for 2 min

Table 2.4: ***Primers for site-directed mutagenesis.*** *Name, DNA sequence and melting point (T_M) of primers used for site-directed mutagenesis.*

Primer name	DNA Sequence 5'→3'	Melting point (T_M in °C)
W32SecM C15A.fwd	GGG CTT CCA GGG CCA CGC TTA CGA GTG CAG CAG TG	82.5
W32SecM C15A.rev	CAC TGC TGC ACT CGT AAG CGT GGC CCT GGA AGC CC	82.5
W32SecM C33A.fwd	CCC TAT TTC AGC CGC GCT TTC AGC ACG CCC GTC	80.3
W32SecM C33A.rev	GAC GGG CGT GCT GAA AGC GCG GCT GAA ATA GGG	80.3
W32SecM C22A.fwd	GAG TGC AGC AGT GAC GCT CCC AAC CTG CAG CCC	81.5
W32SecM C22A.rev	GGG CTG CAG GTT GGG AGC GTC ACT GCT GCA CTC	81.5

Table 2.5: ***Reaction mixture for site-directed mutagenesis.*** *Compounds, volume and concentration of solutions used for site-directed mutagenesis PCR.*

Compound	Volume (µl)	final concentration or amount
Forward primer	1.25	0.625 µM
Reverse primer	1.25	0.625 µM
dNTPs	1	
Quicksolution	0.6	
10x buffer	2	1x
DNA Template	2	20 ng
H$_2$O	11.5	
Lightning enzyme	0.4	

and 450 µl of SOC medium from NEB (Frankfurt, Germany) were added. Cells were incubated at 37°C for 1 h under continous shaking 160 rpm (Infors HT, Bottmingen, Switzerland) or at 600 rpm using a thermostat (HLC, BioTech, Germany). If a ligation reaction or a site directed mutagenesis reaction was transformed, cells were centrifuged at 5,000 x *g* at 4°C for 3 mins. 400 µl of supernatant were discarded and the remaining 100 µl were used to resuspend the cells. The cells were then placed on agar plates containing 100 µg/ml ampicilin or 100 µg/ml ampicilin and 30 µg/ml for BL21(De3)Δtig::Kan cells. If plasmid was transformed directly, the centrifugation step was left out and 50-100 µl of the cells were plated directly on the corresponding agar plates. The cells were incubated at 37°C overnight in an incubator (Binder, Tuttlingen, Germany).

Table 2.6: *__Thermocycling conditions for site-directed mutagenesis.__ Thermocycling conditions using the QuikChange II or QuikChange Lightning Site-Directed-Mutagenesis Kit by Agilent. Conditions for the QuikChange kit are shown in brackets.*

Step	Temperature (°C)	Time (s)	Repetitons
Initial Denaturation	95	120 (30)	1x
Denaturation	95	20 (30)	18x
Annealing	60 (55)	10 (60)	18x
Elongation	68	60 per kb of plasmid length	18x
Final Elongation		n	1x
Pause	4	∞	

2.2.6   Analytical agarose gel electrophoresis

PCR products and plasmid DNA digested with restriction enzymes were loaded on a 1-2% (w/v) agarose gel in TAE buffer. DNA samples were mixed with DNA loading dye (6x) and loaded on the gel. Typically 500 ng DNA were loaded per lane and the Quick-Load 2-log DNA Ladder (New England Biolabs, USA) was used as size reference. Agarose gel electrophoreses was performed at 120 V for 30-45 min. Afterwards the gel was incubated in GelRed staining solution (1:10,000) (Biotium, USA) for 10 min at RT and visualized on the UV-transilluminator of a Gel iX20 Imager (Intas Science Imaging, Germany) or o the Molecular Imager® Geldoc™ XR+ system (Bio-Rad, Hercules, USA).

2.2.7   Preparative agarose gel electrophoresis

For the purification of DNA using preparative gel electrophoresis, a 1-2% (w/v) agarose gel was prepared in TAE buffer and 1:10,000 GelRed (Biotium, USA) was added to the gel solution. The DNA samples were mixed with DNA loading dye (6x) and up to 30 µl/well were loaded on the gel. Agarose gel electrophoreses was performed at 120 V for 30-45 min and bands were visualized with the UV-transilluminator of a Gel iX20 Imager (Intas Science Imaging, Germany). Bands of the corresponding size were cut out and purified using the QIAquick Gel Extraction Kit (Hilden, Germany).

2.2.8   Minipreparation of plasmid DNA

For the isolation of plasmid DNA, a minipreparation was performed using the Mini-Prep-Kit from Quiagen (Hilden, Germany). 5 ml LB medium with 100 µg/ml ampicilin or 100 µg/ml ampicilin and 30 µg/ml for BL21(DE3)Δtig::Kan cells were inoculated either with freshly plated cells or cryo-cultures and incubated at 37°C at 160 rpm (Infors HT, Bottmingen, Switzerland). The plasmids

were isolated and purified using the QIAquick Mini Prep Kit (Hilden, Germany). Plasmids were sequenced by Eurofins Genomics GmbH (Ebersberg, Germany).

2.2.9   Generation of chemical competent cells

Chemically competent cells were produced as described by the Griffitts lab (Brigham Young University). Briefly, a 5 ml overnight culture of BL21(DE3)Δtig ::Kan cells in LB was incubated at 37°C and 160 rpm, as described above. 100 ml of LB were inoculated with 1 ml of the saturated overnight culture and shaked at 37°C and 160 rpm until an $OD_{600\ nm}$ of 0.4 was reached. The cells were placed on an ice for 5 min and kept cool from this point on. The cells were divided in two tubes with approx. 40 ml and centrifuged at 7,600 x g for 10 min. The cell pellets were resupended with 15 ml of cold Mg^{2+}/Ca^{2+} solution (appendix) and incubated for 30 min on ice. The cell solution was centrifuged at 7,600 x g for 10 min and each cell pellet was resupended with 1.6 ml of cold 100 mM $CaCl_2$ solution and afterwards incubated for 20 min on ice. Both tubes were combined and 0.5 ml of cold 80% (v/v) glycerol were added to the cells and mixed. 100 µl aliquots were taken, flash-frozen in liquid nitrogen and stored at -80°C.

2.3   Protein expression and purification of GBC

2.3.1   Expression and purification of full-length GBC and oxidation of full-length GBC

Uniformly labeled GBC was expressed and purified as described previously [22]. After backbone assignment experiments, the protein was diluted to 10 µM and oxidized with 2 µM $CuCl_2$ under continuously stirring and air supply overnight. The protein was loaded on a gelfiltration column (HiLoad 26&60 Superdex 75 prep grade) (GE Healthcare, Chicago, USA) in NMR buffer (50 mM Tris, 200 mM NaCl, pH 8.0) and the monomeric protein was collected and concentrated to approx. 0.8 mM.

2.3.2   Expression of the unlabeled N–terminal domain of GBC

Freshly transformed *E. coli* cells (BL21(DE3)) with pET15b plasmid harboring either the U or H sequence of GBC were incubated overnight in 50 mL LB medium containing 100 µg/ml ampicillin at 37°C and 160 rpm (Infors HT, Bottmingen, Switzerland). The 2 L main culture, containing 100 µg/ml ampicillin, was inoculated to a start $OD_{600\ nm}$ of 0.1-0.2. Expression cultures were incubated at 37°C and 130-140 rpm (Infors HT, Bottmingen, Switzerland). At an $OD_{600\ nm}$ of 0.7-0.8 the protein expression was induced by addition of 1mM IPTG. The cells were harvested after 3 hrs of expression by centrifugation (5,000 x g, 15 min, 4°C). The pellets were flash frozen in liquid nitrogen and stored at -80°C.

2.3.3   Expression of ^{13}C, ^{15}N and ^{15}N labeled N–terminal domain of GBC

For the expression isotopically labeled GBC, a 50 mL overnight culture in LB medium containing 100 µg/L ampicillin was pelleted by centrifugation (4,000 x g, 15 min, 4°C) and resuspended in a small volume of the final expression medium to achieve a start $OD_{600\ nm}$ of 0.1-0.2. For uniformly ^{15}N-labeling the preculture was either resuspended in 250 mL of a rich medium Celltone Complete (98%, ^{15}N, Cambridge Isotope Labs (Tewksbury, USA)) or in 2 L M9 minimal medium containing 1 g/L ^{15}NH$_4$Cl. For additional ^{13}C labeling, 2 g/L ^{13}C-glucose was added. The main culture in rich media was incubated at 37°C at 160 rpm and the 2 L M9 minimal cultures were grown at 37°C at 130-140 rpm. At OD_{600nm} of around 0.8, expression was induced by addition of 1 mM IPTG. After 3 hrs of expression (8 hrs in rich medium) the cells were harvested by centrifugation (5,000 x g, 15 min, 4°C). The pellets were flash frozen in liquid nitrogen and stored at -80°C.

2.3.4   Cell lysis

The cell pellets were resuspended in GBC lysis buffer (50 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 8.0). One complete EDTA-free protease inhibitor tablet (Roche, Basel, Switzerland) was added per 100 mL cell lysate. The cells were disrupted using an M-110P microfluidizer (Microfluidics, Westwood, USA) at 1000 Bar in approx. 6 cycles. Cell lysate was centrifuged twice at 20,000 x g, 4°C for 20 min to remove insoluble cell components. The supernatant was further purified by Ni-NTA affinity chromatography.

2.3.5   Purification of the N–terminal domain of GBC

All purification steps were carried out at 25°C. Cell lysate were loaded on a 5 mL HisTrap HP Ni-NTA column (GE Healthcare, Chicago, USA) equilibrated with GBC lysis buffer (50 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 8.0) with a flow-rate of 2 ml/min. The column was washed with GBC lysate buffer until a baseline at 280 nm was reached. 6x His-tagged GBC was eluted with GBC elution buffer (50 mM sodium phosphate, 500 mM NaCl, 250 mM imidazole, pH 8.0) at a flow-rate of 3 ml/min. Protein fractions were analyzed by SDS-PAGE, the purest fractions were pooled and concentrated to a volume of 10 mL using a Vivaspin 20 centrifugation concentrator (Sartorius, Göttingen, Germany) with a MWCO of 3 kDa. The concentrated protein fractions were loaded on a 320 ml gelfiltration column (HiLoad 26&60 Superdex 75 prep grade) (GE Healthcare, Chicago, USA) equilibrated with GBC NMR buffer (50 mM Tris, 200 mM NaCl, pH 8.0). Protein fractions were analyzed by SDS-PAGE and the purest fractions were pooled. The final sample was concentrated using a Vivaspin 20 and Vivaspin 2 centrifugation concentrator with a molecular

weight cut-off (MWCO) of 3 kDa to a final concentration between 0.5-1 mM.

2.4   Protein expression and purification of ribosomes and RNCs

2.4.1   Expression and purification of ^{15}N uniformly labeled 70S ribosomes

E. coli (JE28) cells [48] were grown in M9 media containing ^{15}N NH$_4$Cl (1 g/L) and 30 µg/ml Kanamycin at 37°C . At an OD$_{600 \text{ nm}}$ of around 0.8 the cells were cooled on ice for 15 min and harvested by centrifugation (5,000 x *g*, 30 min, 4°C). Cells were lyzed and purified by Ni-NTA affinity chromatography as previously described [48]. Ribosome fractions were pooled and loaded on a 10%-40% (w/v) sucrose gradient. The buffer was exchanged to JE28 buffer (20 mM Tris-HCl, 10 mM MgCl$_2$, 150 mM KCl, 30 mM NH$_4$Cl, pH 7.6). The final sample was concentrated to approx. 20µM and diluted to 11µM for the DNP-enhanced solid-state NMR experiments. ^{15}N labeled ribosomes were used to optimize the glycerol cushions required for the solid-state NMR experiments.

2.4.2   Expression of unlabeled RNCs

E. coli BL21(DE3)Δtig::Kan cells were transformed with the protein encoding plasmids (10x-GBC-SecM derivatives on pET15b). Cells were grown in LB medium at 30°C at 130-140 rpm. At an OD$_{600 \text{ nm}}$ of 1, expression was induced by addition of 1 mM IPTG. After 1h of expression the cells were cooled on ice for about 15 min and harvested by centrifugation (5,000 x *g*, 4°C, 20min). Cell pellets were flash frozen in liquid nitrogen and stored at -80°C.

2.4.3   Expression of ^{13}C, ^{15}N uniformly labeled RNCs

The cells were transformed with the protein encoding plasmids (10x-GBC-SecM derivatives on pET15b). RNCs were expressed in *E. coli* BL21(DE3)Δtig:: Kan for selectively labeled RNCs or in BL21(DE3) cells for uniformly labeled RNCs. For uniformly labeling the cells were grown in MDG medium (appendix) at 30°C and 130-140 rpm according to Cabrita *et al.* [28] until OD$_{600 \text{ nm}}$ of 5. The cells were harvested by centrifugation (2,200 x *g*, 4°C, 20 min) and resuspended in M9 media without nitrogen and carbon source. The OD$_{600 \text{ nm}}$ was monitored and after 60 min ^{15}NH$_4$Cl (1 g/L), ^{13}C-glucose (2 g/L), 150 mg/L rifampicin and 1mM IPTG were added to the medium. After 1 h of expression the cells were cooled on ice for about 15 min and harvested by centrifugation (5,000 x *g*, 4°C, 20 min). Cell pellets were flash frozen in liquid nitrogen and stored at -80°C.

2.4.4   Expression of ^{13}C, ^{15}N cysteine labeled RNCs

For selectively labeled RNCs, the cells were grown in LB medium at 30°C according to Rutkowska *et al.* [217] to an OD$_{600 \text{ nm}}$ of 1. Cells were cooled

on ice and harvested by centrifugation (2,700 x g, 4°C, 20 min). Afterwards the cells were resuspended in M9 salts and centrifuged again. This washing step was repeated once more. The cell pellet was resuspended in 1/4 of the volume in M9 media supplemented with all amino acids and ^{13}C, ^{15}N cysteine (20 mg/L) (appendix). The expression was induced with 1mM IPTG and 150 mg/L rifampicin was added 8 min after induction with IPTG. After 1 h the cells were cooled on ice and harvested by centrifugation (5000 x g, 30 min, 4°C). The cell pellets were frozen in liquid nitrogen and stored at -80°C.

2.4.5   Purification of RNCs

RNCs were purified as described by Cassaignau *et al.* [30] with the exception of the sucrose cushion step. The cells were resupended in RNC lysis buffer (50 mM HEPES, 500 mM KOAc, 12 mM MgOAc) supplemented with Dnase I, lysozyme and protease inhibitor tablet (Roche, Basel, Switzerland). The cells were disrupted using a M-100P microfluidizer (1000 bar, Microfluidics Westwood, USA). The lysate was centrifuged twice (38,400 x g, 4°C, 30 min) and the supernatant was loaded on a 5 ml HisTrap HP Ni-NTA column (Histrap, GE Healtcare, Chicago, USA) equilibrated with RNC lysis buffer. The column was washed with 4% elution buffer (20mM Imidazole) and RNC were eluted in a linear gradient to 100% elution buffer (50 mM HEPES, 500 mM KOAc, 12 mM MgOAc, 500 mM Imidazole). RNCs containing fractions were pooled and concentrated using a 100-kDa-cutoff concentrator (Vivaspin, Sartorius, Göttingen, Germany) at 3,500 x g at 4°C. Approximately 1.6 nmol was loaded on a 10%-40% (w/v) sucrose gradient (35 ml, 25 × 89 mm tubes, SW28 Beckmann Coulter, Brea, USA). The concentration was determined by measuring the absorbance at 260 nm, with one optical unit corresponds to 24 pmol/ml 70S ribosomes [19]. The gradients were spun at 23,000 rpm (Beckman Coulter, SW 32 Ti) at 4°C for 16 hrs. Afterwards the gradients were fractionated in 1.13 ml fractions using a Piston Gradient Fractionator (Biocomp Instruments, Fredericton, USA). The fractions were loaded on a 15% SDS-PAGE gel or NuPAGE Bis-Tris 4%-12% gel with a neutral pH from ThermoFischer Scientific (Waltham, USA), the purest fractions were pooled and the buffer exchanged to Tico (10 mM HEPES, 30 mM NH$_4$Cl, 10 mM MgCl$_2$, 1 mM EDTA, pH 7.5). The final sample was concentrated to 10-20 μM. Around 7 mg RNCs per L expression culture were obtained.

2.4.6   Purification of ^{13}C, ^{15}N cysteine labeled flow-through (FT) ribosomes

The flow-through during the Ni-NTA affinity chromatography of a U24SecM sample was collected and pelleted for 4 hrs at 225,000 x g and 4°C. The pellet was resuspended in Tico buffer (10 mM HEPES, 30 mM NH$_4$Cl, 10 mM MgCl$_2$, 1 mM EDTA, pH 7.5) and 1.9 nmol were loaded on a 10%-40% (w/v) sucrose gradient. 70S containing ribosome fractions were pooled and the buffer exchanged to Tico

buffer (via dialysis as for U32SecM cryo-EM, for all other samples the buffer was exchanged using protein concentrators). 200 µl with a concentration of 24 µM were loaded on a 500 µl glycerol cushion with the 15 mM AMUPol and 30% (v/v) glycerol.

2.5 Biochemical techniques

2.5.1 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE)

Protein purity was analyzed using SDS-PAGE. Samples were either mixed with reducing 6x SDS loading buffer (375 mM Tris, pH 6.8, 10% (w/v) SDS, 50% glycerol, 10% 2-mercaptoethanol, 0.03% bromphenol blue) or 2x SDS reducing loading buffer (250 mM Tris, pH 6.8, 4% SDS, 20% glycerol, 0.25% bromophenol blue, 10% 2-mercaptoethanol) or non-reducing 6x SDS or 2x SDS loading buffer without 2-mercaptoethanol. Typically 10 µl were loaded per gel and as size reference the low-molecular weight marker from GE Healthcare (Freiburg, Germany) or the Spectra Multicolor Broad Range Protein Ladder from ThermoFisher Scientific (Waltham, USA) was applied. For the RNCs, NuPAGE Bis-Tris 4%-12% gels with a neutral pH from ThermoFischer Scientific (Waltham, USA) and for isolated GBC gels composing of a 5% (v/v) stacking (pH 6.8) and 15% (v/v) resolving phase (pH 8.8) were used. The bands were separated using constant voltages of 180-200 V for 40-50 min, the gel was stained with Coomassie Brilliant Blue staining solution for 20 min and destained with water. Gels were digitalized using a Gel iX20 Imager (Intas Science Imaging, Germany) or Molecular Imager® Geldoc™ XR+ system (Bio-Rad, Hercules, USA). Buffer recipes required for SDS-PAGE are listed in the appendix.

2.5.2 Western blotting

After SDS-PAGE the gels were loaded on a western-blot sandwich for specific protein detection. The setup is shown in Figure 2.1. One filter paper (Thermo Fisher Scientific, Waltham, USA) was equilibrated in anode buffer (75 mM Tris, 20% (v/v) methanol, pH 7.4) and loaded on the anode plate. The Immobilon-P PVDF membrane (0.45 µm poresize, Merck, Darmstadt, Germany) was activated in 100% methanol for 1 min and afterwards equilibrated in cathode buffer (40 mM aminocaproic acid, 25 mM Tris, 20% (v/v) methanol, pH 9.0). The membrane was put on the filter paper, equilibrated in cathode buffer. The SDS-PAGE gel was loaded on the membrane and another filter paper was equilibrated in cathode buffer and laied on top. The western blotting was performed at 1 mA per cm^2 membrane for 1 h at RT. The membrane was blocked with PBS+M (5% (w/v) milk solution in 1x PBS (10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, 137 mM NaCl, pH 7.4)) for 30 min at RT. Afterwards the membrane was incubated with

the first antibody (Poly-His, Sigma-Aldrich, Merck, Darmstadt, Germany) in PBS+M (1:3000) at 4°C over night. The membrane was washed three times for 10 min with PBST (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, pH 7.4, 0.05% (v/v) Tween-20) at RT and then incubated with the second antibody (goat-anti mouse-hrp, Jackson ImmunoResearch Laboratories, West Grove, USA) in PBS+M (1:5000) for 2 hrs at 4°C. The membrane was washed again for three times, each for 10 min at RT with PBST and developed using ECL Western Blotting Substrate (Pierce[TM], Thermo Fisher Scientific, Waltham, USA) and the chemoluminescence was visualized on an Intras Advanced flourescence and ECL imager (Intas Science Imaging, Göttingen Germany).
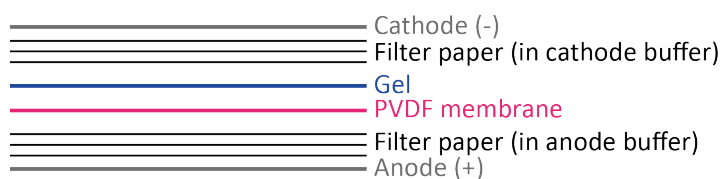
Figure 2.1: *Assembly of a semi-dry western blot.*

2.5.3   Solid phase peptide synthesis (SPPS)

The U32 peptides were synthesized by solid phase peptide synthesis using standard Fmoc (Fluorenylmethyloxycarbonyl) chemistry according to Merrifield [172]. The C-terminal residue attached to the 2-chlorotrityl resin was obtained from Merck (Darmstadt, Germany). Synthesis was performed in a 0.1 mM scale using Fmoc-amino acids purchased from Merck (Darmstad, Germany) and ^{13}C, ^{15}N Fmoc-cysteine from Cambridge Isotope Labs (Tewksbury, USA).

2.5.4   Analytical and preparative RP-HPLC

Analytical RP-HPLC runs for the full-length and one-domain GBC protein were performed on a Jasco HPLC system using a PerfectSil 300 C4-RP column (250 × 4.6 mm) as described in Buhr *et al.* [22]. Preparative RP-HPLC runs for the purification of the U32 peptide were performed on a Shimadzu HPLC system with diode-array detection using a Orbit 100 C18 column (250 x 20 mm). Gradient: 20-60% B in 40 min, 29% B isocratic (10 min), flow rate 8 ml/min. Buffer A: H_2O, 0.1% TFA. Buffer B: Acetonitrile, 0.1% TFA. U32 peptides were characterized using mass spectrometry (MALDI) and analytical HPLC.

2.6   Biophysical techniques

2.6.1   Liquid-state NMR Spectroscopy

Liquid state NMR experiments were performed on Bruker spectrometers (AV600, AV800, AV900 and AV950) (Rheinstetten, Germany). Backbone assignment of GBC was done by performing ^{1}H-^{15}N BEST-TROSY and the BEST-TROSY ver-

sions of the standard triple resonance experiments (HNCACB, HN(CO)CACB, HNCO, CC(CO)NH) on a Bruker spectrometer (AV600 and AV800) equipped with a TCI cryo probe. All spectra were recorded at 298 K. The NMR spectra were acquired and processed using the software Topspin 3.5 (Bruker BioSpin, Rheinstetten, Germany). The 3D spectra were recorded with an exponential weighting of sampling scheme and processed with Topspin 3.5. Data was analyzed with Sparky 3.114 [77], Cara 1.8.4.2 [123] or Topspin 3.5-4.08. Sample concentrations varied between approx. 300 µM for the peptide from solid phase peptide synthesis and proteins expressed in ^{15}N rich media to 1.2 mM for protein expressed in minimal media. The peptides spectra were recorded at 277 K and the protein samples at 298 K. All samples contained 10% D_2O and 0.1% DSS. All spectra were referred to the DSS signal.

2.6.2   DNP–enhanced solid-state NMR Spectroscopy

A 200 µl sample was loaded on a 500 µl cushion of Tico buffer in D8-^{13}C depleted-glycerol:D_2O:H_2O (42:45.3:14.5) (v/v) with a total AMUPOL concentration of 15 mM (final volume 700 µl). The sample was sedimented at 28,000 rpm (Beckman Coulter, SW 28) and 4°C for 15 hrs using a centrifugation device as previously described [75]. All DNP-enhanced solid-state NMR experiments were carried out on a Bruker Avance spectrometer operating at 9.4 T (400 MHz as ^{1}H Larmor frequency). Due large molecular weight of the RNCs, only a few (2-3.4 nmol) can be packed into rotor, which is two orders in magnitude lower than required for conventional solid-state NMR. Therefore, DNP enhancement is indispensable to increase signal to noise for ribosome samples that cannot be concentrated above 30 µM. The spectrometer was equipped with a customized Bruker low-temperature HCN triple resonance probehead. The high power microwave was generated from a CPI gyrotron (Palo Alto, USA) operating at 9.4 T and was directed to the sample position in the solid-state NMR probe head through corrugated waveguides. The effective microwave power emitted into the MAS stator was about 12 W. In all experiments MAS was stabilized at 8 kHz, sample temperature was kept at about 110 K and an interscan delay of 3 s was used. ^{1}H-^{13}C CP 1D and ^{13}C-^{13}C DQ-SQ experiments were acquired with the similar parameters as previously reported [116]. Briefly, for the CP step, a ramped (80-100%) ^{1}H RF field at 52 kHz as maximum power and a ^{13}C RF pulse at 48.5 kHz were applied during 0.8 ms contact time. In the ^{13}C-^{13}C DQ-SQ experiments, the DQ coherence was excited and reconverted back to the SQ coherence by a train of POST-C7 recoupling pulses of 0.5 ms total length with field strength at 56 kHz. The DQ evolution period was rotor "synchronized" permitting a large spectral window of the indirect dimension. During the ^{13}C detection period, SPINAL64 ^{1}H decoupling pulses with 100 kHz field strength were applied. The ^{13}C-^{13}C DQ-SQ 2D spectra were acquired with 60 and 2432 points on F1 (DQ, 556 ppm) and F2 (SQ, 296 ppm) dimension respectively and were processed with a 1024*8192

(F1*F2) matrix. Qsine (ssb = 2) and Gaussian (lb = -50 Hz, GB = 0.04) window functions were used for processing the indirect and direct ^{13}C dimensions, respectively. The ^{1}H-^{15}N CP experiment used for optimizing DNP sample preparations was acquired with 512 points and a spectral window of 600 ppm. During the 400 µs CP contact time, a ramped (80-100%) ^{1}H RF field at 62 kHz as maximum power and a ^{15}N RF pulse at 47 kHz were applied. In the ^{15}N-^{13}C TEDOR 2D experiment, the recoupling time was set to 0.5 ms (4 rotor periods). In the ^{15}N-^{13}C DCP 2D experiment, the ^{15}N-^{13}C magnetization transfer was realized by applying a ramped (90-100%) ^{13}C RF pulse with 28 kHz field strength and a ^{15}N RF pulse with 40 kHz field strength during the 5.6 ms contact time. The ^{15}N-^{13}C TEDOR spectrum was acquired with 192 and 1024 points on F1 (197 ppm) and F2 (296 ppm) dimension respectively and were processed with a 1024*4096 (F1*F2) matrix. The ^{15}N-^{13}C DCP spectrum was acquired with 192 and 2048 points on F1 (197 ppm) and F2 (296 ppm) dimensions respectively, and were processed with a 1024*4096 (F1*F2) matrix. The Qsine (ssb = 2) and Gaussian (lb = -20 Hz, GB = 0.05) window functions were used for processing the indirect (^{15}N) and direct (^{13}C) dimensions. respectively. All ^{13}C chemical shifts were referenced to DSS in water indirectly using an alanine powder sample. All ^{15}N chemical shifts were referenced indirectly to liquid ammonia. All DNP-enhanced MAS solid-state NMR experiments were performed by Dr. Jiafei Mao (Prof. Clemens Glaubitz group, (Goethe University of Frankfurt am Main)). This part of methods was written by Dr. Jiafei Mao.

2.6.3   Spectral Analysis of DNP–enhanced solid-state NMR

All chemical shift assignments on ^{13}C-^{13}C DQ-SQ spectra were carried out in Bruker Topspin 3.5 (Bruker, Rheinstetten) and Sparky [77]. Because the Cβ chemical shift is the mostly indicative parameter for the Cys chemical states and Cα chemical shift is sensitive to Cys backbone conformation, we have focused primarily on assigning the signals in the Cα-Cβ region in our spectra. We first assigned those resolved Cβ signals in the Cα/Cβ(DQ)-Cβ(SQ) region and the resolved Cα signals in the Cα/Cβ(DQ)-Cα(SQ) region. Based on these assignments, the correlated Cα and Cβ signals were then extracted from the spectra. The solid-state NMR data analysis were performed together with Dr. Mao. Integration of the Cβ region for the determination of the relative integration ratio was carried out with Topspin 3.5 (Bruker, Rheinstetten) using the a+- mode. For the calculation of the integration error, the chemical shift region was varied slightly +2 ppm for the oxidized and -2 ppm for the reduced region in the SQ chemical shift dimension. The integration region of the oxidized signals, was between 49.68-35 ppm (SQ), for the overlapping region between 35-33 ppm (SQ) and for the reduced region signals between 32-23.12 ppm (SQ). The DQ chemical shift dimension for the integraton regions was between 121-69 ppm. Five different integration areas were used and the mean and standard error was calculated

using R [208].

2.6.4 Chemical shift determination of S-nitroso glutathione, glutathione and gluathione disulfide

Commercially available S-nitroso-glutathione (Sigma-Aldrich, St. Louis, USA), glutathione (Carl Roth, Karlsruhe, Germany) and glutathione disulfide (Carl Roth, Karlsruhe, Germany) was resolved in GBC NMR buffer (50 mM Tris, 200 mM NaCl, 10% D2O, pH 8.0) with 1 mM Trimethylsilylpropanoic acid (TSP) to a concentration of approximately 50 mM. ^{13}C 1D NMR spectrum was acquired with 2048 scans and 65536 points at 298 K on a Bruker spectrometer (AV500HD). The spectrum was processed with Topspin 4.08 and the chemical shifts were referenced to TSP. Cα and Cβ chemical shifts of the cysteine residue within the different glutathione variants were determined.

2.7 Secondary structure analysis of cysteine residues within RNCs

The Cα and Cβ chemical shifts of the cysteine residues within the ^{13}C-^{13}C DQ-SQ of the RNCs were extracted and saved as a .csv file. The files were loaded into R [208]. Chemical shift data of cysteine residues within different structural elements derived from the Biological Magnetic Resonance Bank (BMRB) [257] were taken from [261]. The Cα chemical shifts were plotted as a function of the Cβ chemical shits. An ellipse containing 90% of the data was drawn around each structural element. The individual Cα and Cβ of the RNCs were plotted within this data set. Cβ chemical shifts above 35 ppm correspond to oxidized cysteine residues, Cβ chemical shifts between 35-32 ppm are in the overlapping region, and can be either oxidized or reduced and Cβ chemical shifts below 32 ppm correspond to reduced cysteine residues. The R script for generating this Cα-Cβ chemical shift plots is shown in the appendix.

2.8 Cryo–EM

2.8.1 Sample preparation, collection of micrographs and image processing

A 3.5 µl aliquot of an approximately 50 nM U32SecM sample in Tico buffer was applied to a glow-discharged (20 s) Quantifoil R1/2 holey carbon grid (Cu 300 mesh, Quantifoil Mirco Tools GmbH, Jena, Germany) coated with an approx. 5 nm thick carbon film [80]. Grids were plunge-frozen in liquid ethane by using a Vitrobot Mark IV (FEI, Hillsboro, USA) with a blotting time of 8-9 s, a blotting force of -1 to -5, at 4°C and 100% humidity. Dose-fractionated movie stacks (36 frames, 0.2 s each) were collected at a nominal magnification of 130,000 x (1.05 Å pixel size) with a 300 kV Titan Krios (FEI) microscope in nanoprobe EFTEM

mode, equipped with a K2 Summit detector (Gatan, Pleasanton, USA) and a GIF Quantum s.e. post-column energy filter in zero loss peak mode. The defocus values varied from -0.7 μm to -3.5 μm. Data were recorded using SerialEM [162] in a semi-automated way. Image processing was done in Relion 3 [224, 283] using Shawn Zheng's MotionCor2 [281] for beam induced motion correction and Niko Grigorieff's CTFFIND4.1 [214] to estimate the contrast transfer function parameters. 436,999 particles were automatically picked using 2D class references from manually picked particles. Bad particles were removed using reference-free 2D classification and the remaining 326,344 particles were further 3D classified in 6 classes with an angular sampling rate of 7.5°. The reference map for the 3D classification was generated from the 70S Ribosome-SecM (PDB 3JBV) model in USCF Chimera [204] and low-pass filtered to 60 Å in Relion. Classes with low density for the 30S subunit were removed and 266,000 particles were used for further 3D refinement using an initial angular sampling of 7.5° and local angular searches of 1.8°. The refined structure was further post-processed using a soft mask generated by extending the binarized 15 Å lowpass filtered map and addition of a cosine-shaped edge yielding a structure of 3.52 Å. B-Factor for sharpening of the map was automatically determined by Relion (-102.4). We used the CTF refinement functionality of Relion 3.1 to estimate per-particle defocus values with a search range of 200 nm and without beamtilt estimation. CTF refinement was followed by Bayesian polishing using all frames of the movie stacks and parameters that were determined during a training run using 10,000 particles. CTF refinement and Bayesian polishing improved the resolution of the reconstruction to 2.98 Å. In order to resolve the structural heterogeneity of the reconstruction we 3D-classified the particles into 4 classes using a mask comprising the large subunit and all tRNA states, plus a regularization parameter of 4, and without performing image alignment. Besides one very low-populated class with potential junk particles, we identified two classes showing clear tRNA density in the P-Site (196,000 particles and 13,000 particles) and one class that showed clear tRNA density in the P-site but slightly shifted compared to the other classes (55,000 particles). We refined the highly populated class with tRNA density in the P-site and the class with shifted tRNA density in the P-site independently. Refinement of the high populated class converged at 2.83 Å resolution, showing the tRNA located in the P-site, whereas the other class converged at 3.19 Å resolution, showing the tRNA density in P/E and in the A/P* site. To further improve the resolution of the 2.83 Å reconstruction we estimated the magnification anisotropy, refined the defocus per-particle and the astigmatism per-micrograph followed by estimation of beamtilt and trefoil with the CTF refinement of Relion 3.1. The final 3D refinement and reconstruction that corrected these effects converged at 2.58 Å resolution. An overview of the processing pipeline is given in Figure 2.2. This part of methods was written by Linda Schulte and Julian Reitz.
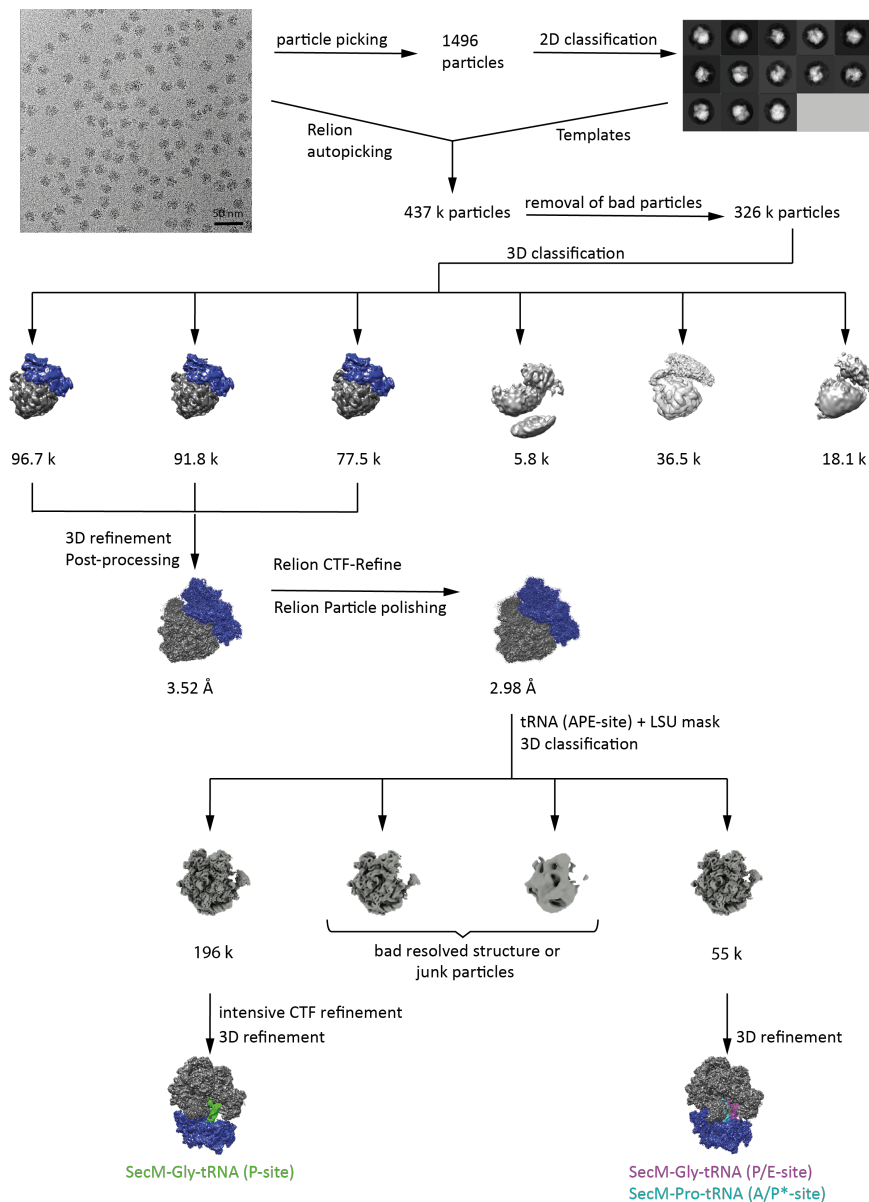
Figure 2.2: *Overview of the processing-pipeline for cryo-EM structure of U32SecM. Arrows indicate the direction of processing. The micrograph shows the expected size of the ribosome particles under cryo conditions. 1,496 particles were picked manually and were further used to generate 2D classes as template for the auto-picking in Relion. 437,000 particles were picked and after several rounds of 2D classification, 326,000 good particles remained. After 3D classification three classes with both 50S subunit (grey) and 30S subunit (blue) were used for 3D refinement and post-processing. The obtained 3.52 Å structure was CTF-refined and polished to a 2.98 Å structure using Relion. All tRNA sites and the 50S subunit were used as a mask for masked 3D classification. Four distinct classes were obtained, with one class showing clear P-site tRNA, one class showing P/E-site and two classes with low resolved structure or junk. The class with tRNA density in the P-site was intensively CTF refined and both classes with clear tRNA density were 3D refined and post-processed using Relion. The final resolution was 2.58 Å for the ribosome with P-site tRNA (green) and 3.19 Å for the ribosome with P/E-site tRNA (pink) and A/P*-site tRNA (cyan).*

2.8.2   Model building

As an initial reference model for the SecM-stalled RNC, the cryo-EM structure of the 70S *E. coli* ribosome (PDB 3JBU) was chosen. The model was fitted into the cryo-EM electron density map of the *E. coli* BL21(DE3) ribosome using the Fit in Map tool of UCSF Chimera [204]. To focus in detail on the structural analysis of the 50S ribosomal subunit, the 30S subunit could be removed from the model. Therefore, structural dynamics between both subunits were not analyzed. For an initial real-space refinement of all 50S subunit RNA, tRNA and protein components the Rigid Body Fit Zone tool of COOT [50] was used. Since the density of the nascent chain (NC) was rather faint in the B-factor corrected sharpened map, we mutated the ten N-terminal residues of the NC sequence to the GBC sequence using COOT. Afterwards, the model was refined using the Real-space refine (cryo-EM) tool of PHENIX [1]. Model validation and calculation of FSC curves were carried out using the PHENIX Comprehensive validation (cryo-EM) tool. Furthermore, the validated model was manually analyzed using COOT. Ramachandran outliers, unusual rotamers and potential clashes were inspected and adjusted if necessary. In addition, protein and RNA sequence deviations of *E. coli* K12 and *E. coli* BL21(DE3) ribosomal 50S subunit were corrected manually. Subsequently, several cycles of refinement and validation were carried out using COOT and PHENIX. The cryo-EM electron density maps of GBC-SecM-RNC (50S subunit) have been deposited in the Electron Microscopy Data Bank under accession code 4531. Their associated atomic models have been deposited in the Protein Data Bank under accession code 6QDW. To build the NC, we then used the unsharpened map, which showed the NC density much more clear. We rebuilt the previously refined nascent chain (PDB: 6QDW) and added GBC residues until the density became too faint for proper model building. Model validation and calculation of FSC curves were carried out using the PHENIX Comprehensive validation (cryo-EM) tool. Furthermore, the validated model was manually analyzed using COOT. Ramachandran outliers, unusual rotamers and potential clashes were inspected and adjusted if necessary. Subsequently, several cycles of refinement and validation were carried out using COOT and PHENIX. The cryo-EM electron density maps of GBC-SecM-RNC (50S subunit) have been deposited in the Electron Microscopy Data Bank under accession code 10891. Their associated atomic models have been deposited in the Protein Data Bank under accession code 6YS3. Statistical parameters of the two structures are given in Table 2.7.

Table 2.7: *Experimental conditions of data collection, refinement and model statistics of 6QDW and 6YSE.*

Data Collection	6QDW	6YSE
Particles	196,254	196,254
Pixel size (Å)	1.05	1.05
Defocus range (µM)	0.7-3.5	0.7-3.5
Voltage (kV)	300	300
Electron dose (e⁻/Å²)	60	60
Model Composition		
Protein residues	3127	3136
RNA bases	3091	3091
Refinement		
Resolution (Å, 0.143 FSC)	2.83	2.58
Map sharpening B-factor	-44.45	-34.53
Validation		
RMSZ, bonds	1.30	1.24
RMSZ, angles	1.15	1.11
Rotamer outliers (%)	1.0	0.8
Ramachandran outliers (%)	0.1	0.0
Ramachandran favoured (%)	93	93
Correct sugar pucker (%)	100	100
Good backbone conformations (%)	73	73
Scores		
MolProbity	1.88	1.81
Clash score, all atoms	6.27	6.44

2.9   LC-MS/MS measurements of RNCs

Purified protein-samples were initially diluted in 50 mM ammonium bicarbonate buffer and treated with 10 µg RNAse A (Qiagen) for 10 min, followed by the alkylation of free cysteine residues with iodoacetamide (Thermo Fisher, Waltham, USA) for 20 min in the dark. Proteins were then digested in solution by the addition of 1 µg trypsin (Serva, Heidelberg, Germany) and incubation for 3.5 hrs at 37°C. Proteolytic digests were finally desalted using C18 ZipTips (Merck, Darmstadt, Germany) according to the manufacturer's instructions, dried in a vacuum concentrator (Eppendorf, Hamburg, Germany) and reconstituted in water/acetonitrile 95/5/0.1 (v/v) with 0.1% formic acid. Subsequent LC-MS/MS analyses were carried out either on a nanoElute coupled to an Impact II mass spectrometer (Bruker Daltonics, Massachusetts, USA) or an Ultimate3000 RSLCnano coupled

to a Fusion Lumos mass spectrometer (Thermo Fisher) using the parameters
given in Table 2.8. Mass spectra were acquired over the mass range 150-2200
m/z (Impact II), 350-1600 m/z (Fusion Lumos) and sequence information was
acquired by a computer-controlled, dynamic method with a fixed cycle time of 3 s
and intensity dependent acquisition speed for MS/MS-spectra of the most abun-
dant candidate ions. The resulting files were exported to the open MZML format
(Impact II) and recalibrated, and database-, PTM- and crosslink searches were
performed using the Metamorpheus search engine against a combined database
containing the *E. coli* proteome (obtained from Uniprot, 11/2019), the respective
SecM constructs and common contaminants.

The search parameters used were as follows: <20ppm mass tolerance for precur-
sor and fragment ions, fully tryptic cleavages with up to two missed cleavages
and heavy labeled cysteine, carbamidomethylation (C), glutathione (C), oxida-
tion (M) as variable modifications. Further modifications were analysed with
GPTMD (common artifacts and biological modifications) as well as disulfide
crosslinks searched for light and heavy labeled cysteines. This part of methods
was written by Jakob Meier-Credo.

Table 2.8: *System, columns and experimental conditions for mass spectrom-*
etry performance on RNCs.

System	nanoElute / Impact II	Ultimate3000 / Fusion Lumos
Columns	Triart C18, 5*0.3 mm, 1.9 µm (YMC) Reprosil C18, 16*0.0075 mm, 1.9 µm (Pepsep)	Pepmap C18, 20*0.075 mm, 3 µm (Thermo Fisher) Pepmap C18, 500*0.075 mm, >2 µm (Thermo Fisher)
Mobile phase	A: Water + 0.1% FA B: Acetonitrile + 0.1% FA	A: Water + 0.1% FA B: 80% Acetonitrile + 0.1% FA
Gradient	Linear 2-35% B, 90 min	Linear 4-48% B, 178 min
Flow rate	300 ng/min	
Temperature	55°C	

2.10   Simulation of conformational sampling of newly synthesized polypeptide in the ribosomal exit tunnel

The degrees of freedom available to the nascent polypeptide chain were sim-
ulated using an adapted version of the statistical coil sampling engine flexible-
meccano [14, 194]. Backbone degrees of freedom were defined by amino-acid-
specific potentials derived from loop-regions extracted from a database of high-
resolution crystal structures [150], steric clashes were taken into consideration
using a residue-specific hard sphere model [147]. Conformational sampling of

U32SecM was initiated at position S151 of SecM and the polypeptide was built in the direction from C to N, using a self-avoiding algorithm that randomly samples the backbone potentials. Only self-avoiding conformations and conformers that avoid steric clash with all heavy atoms from the coordinates of the tunnel (determined from cryo-EM (PDB: 6YSE and 4UG0)) are retained for further analysis. The calculation was repeated for 2,0000 conformations.

The potential of cysteine amino acids (Cys15, Cys18, Cys22 and Cys32) to form disulfide bridges was assessed by measuring the distances between cysteine Cβ atoms within each conformation, with a threshold of 4.5 Å. The same calculations were repeated in the presence and absence of the tunnel to assess the relative importance of internal degrees of freedom and steric interaction with amino acids forming the tunnel wall. This part of methods was written by Prof. Dr. Martin Blackledge.

3   Results and Discussion

3.1   Investigation of the NTD of GBC

Two synonymous variants of the NTD of GBC were studied to determine if the NTD itself is responsible for conformational differences as detected in the full-length protein [22]. The translation frequencies of the two NTD variants are shown in Figure 3.1. U is translated with a significantly lower frequency in *E.coli* than in *B. taurus*. Especially in the first 28 amino acids, where three of the six cysteine residues in the NTD are located, are the differences in the translation speed obvious. The H variant is adapted to the codon usage of the *E. coli* host and does not show this slow translation frequency in the first 28 amino acids.
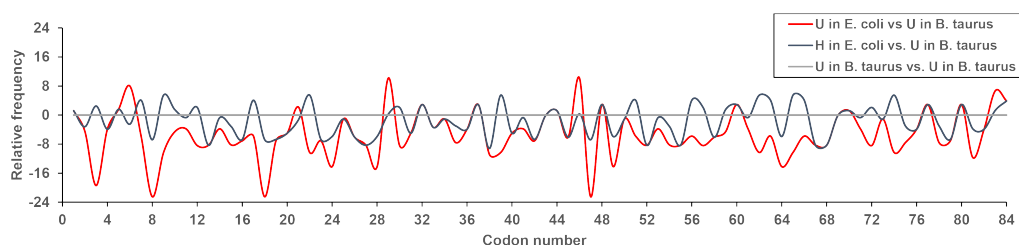
Figure 3.1: *Codon usage of the NTD of GBC. Relative differences in the codon usage of the variants U (red) and H (blue) expressed in E. coli to the codon usage of U in B. taurus (grey). Residues 1-83 are shown inluding the N-terminal methionine.*

Purification of the NTD resulted in monomeric and to a lesser extent also in dimeric protein, especially after concentrating the sample to NMR conditions (approx. 1 mM) (Figure 3.2a). The dimer formation could not be prevented by lowering the pH of the purification buffers from 8.0 to 6.5 and increase the salt concentration from 200 mM to 500 mM NaCl (Figure 3.2b). After concentrating the sample and recording NMR spectra at 298K, the protein started to precipitate. Addition of DTT to the SDS-loading buffer did not show any effect on the final NMR sample as detected by SDS-PAGE (Figure 3.3a). Treatment with 100 mM DTT and 3 h incubation at 90°C resulted in only one visible protein band on the SDS-PAGE gel (Figure 3.3b). Hence, not only disulfide bonds might be involved in the dimerization of the protein. Both proteins were expressed several times in LB, M9 minimal medium and in rich medium and the purified proteins were measured in ^{1}H-^{15}N correlation experiments. In most cases both protein ^{1}H-^{15}N BEST-TROSY spectra looked completely identical. Addition of DTT to the purified protein did not show any changes of the backbone signals, so we assumed that the protein is in a reduced conformation.

Differences between U and H were recorded twice out of five different protein samples. Once for protein expressed M9 minimal medium and once expressed in rich medium. The individual purification steps for the proteins expressed in rich
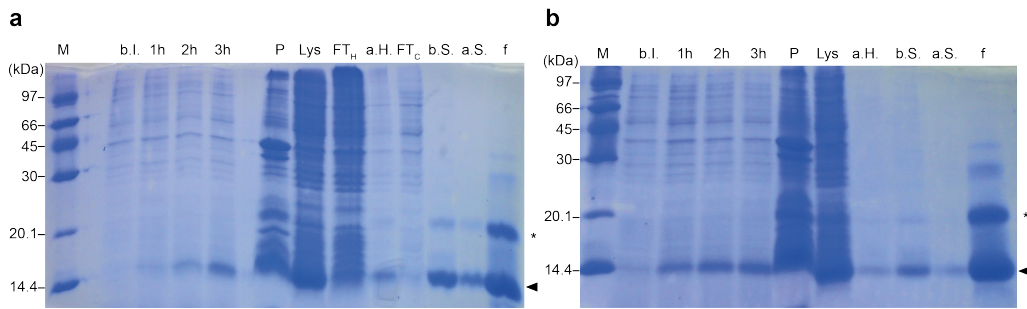
Figure 3.2: *SDS-PAGE gel of NTD GBC (U) after individual purification steps under different pH and salt conditions. a) 15% SDS-PAGE gel of unlabeled NTD of GBC (U) purification steps with purifaction buffer at pH 8.0 and 200 mM NaCl in final NMR buffer. b) 15% SDS-PAGE gel of unlabeled NTD of GBC (U) with purification buffer pH 6.5 and 500 mM NaCl in final NMR buffer. Samples from left to right: Protein marker (M), before Induction (b.I.), samples were taken after 1h, 2h and 3h after induction. Cell lysate (Lys), pellet (P), flow-through during nickel-affinity chromatography (FT_H), after Ni-NTA affinity chromatography (a.H.), flow-through during concentrating, before SEC (b. S.), after SEC (a. S.), final NMR sample (f). Monomeric protein is labeled with a triangle (◄) and dimer with an asterik (*). For the SDS-PAGE the LMW-SDS Marker from GE Healthcare (Freiburg, Germany) was used.*
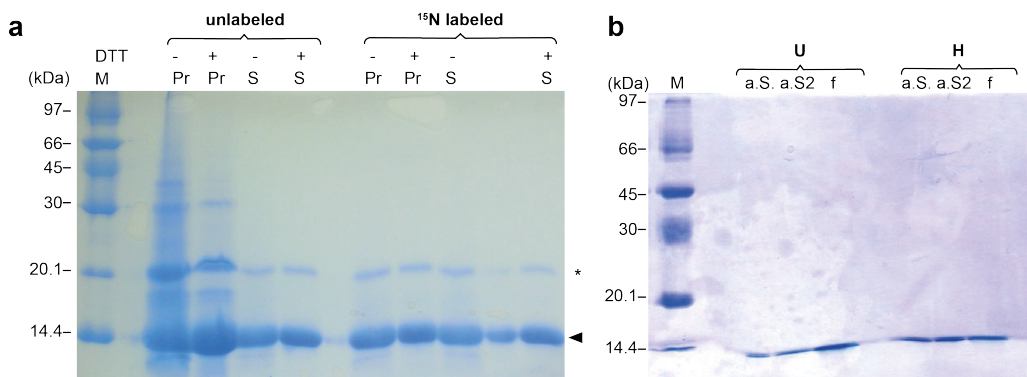
Figure 3.3: *SDS-PAGE gel of NTD GBC (U and H) after treatment with DTT. a) 15% SDS-PAGE gel of unlabeled and ^{15}N labeled NTD of GBC (U) after addition of 100 mM DTT. Samples from left to right: Protein marker (M), precipitate (Pr) of final NMR sample without DTT, precipitate (Pr) of final sample with DTT, supernatant of final sample (S) without DTT and supernatant of final sample with DTT. Same order for unlabeled and ^{15}N labeled protein. b) 15% SDS-PAGE gel of ^{15}N labeled NTD of GBC (U and H) after treatment with 100 mM DTT and incubation at 90°C. Samples from left to right: Protein marker (M), after SEC (a. S.), after SEC protein with slightly higher MW (a. S2), final NMR sample (f). Monomeric protein is labeled with a triangle (◄) and dimer with an asterik (*). For the SDS-PAGE the LMW-SDS Marker from GE Healthcare (Freiburg, Germany) was used.*

medium are shown on the SDS-PAGE gel (Figure 3.4) and the 1D and 2D NMR spectra of both are displayed in Figure 3.5. The spectra of the two proteins differ both in the 1D ^{1}H spectrum in the methyl range and in the amide protein region (Figure 3.5a) and the latter region is also shown in the ^{1}H-^{15}N BEST-TROSY spectrum (Figure 3.5b). The obtained differences could be removed by addition of DTT (Figure 3.6) and are therefore related to cysteine oxidation. In contrast to the study by Buhr *et al.* [22] addition of DTT did not have an effect on H, but on the spectrum of U. The effect of DTT is both visible in the 1D ^{1}H (Figure 3.6a) and in the 2D ^{1}H-^{15}N BEST-TROSY spectrum (Figure 3.6b). Since the signals in methyl region in the 1D ^{1}H experiment are broader in the U protein without DTT and the linewidth decreases by addition of DTT, we assumed that U existed to a higher extent as a dimer than the H protein (Figure 3.4 and that addition of DTT converted the dimer into monomers.
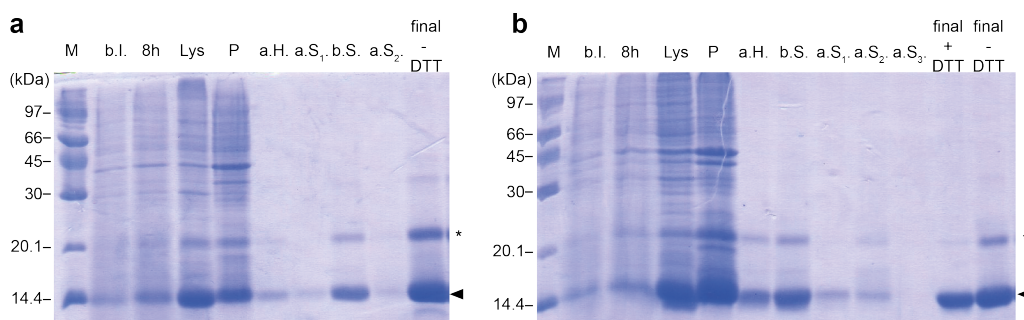
Figure 3.4: ***SDS-PAGE gel of NTD GBC (U and H) NMR samples. a)*** *15% SDS-PAGE gel of ^{15}N labeled NTD of GBC U after individual purification steps. Samples from left to right: Protein marker (M), before induction (b. I.), 8h after inducton (8h), cell lysate (Lys), cell pellet (P), after Ni-NTA chromatography (a. H.), after size exclusion chromatography (a. S$_1$), before size exclusion chromatography (b. S.), after size exclusion chromatography with slightly higher molecular weight (a. S$_2$), final NMR sample (final).* ***b)*** *15% SDS-PAGE gel of $_{15}$N labeled NTD of GBC H after individual purification steps. Samples from left to right: Protein marker (M), before induction (b. I.), 8h after inducton (8h), cell lysate (Lys), cell pellet (P), after Ni-NTA chromatography (a. H.), before size exclusion chromatography (b. S.), after size exclusion chromatography (a. S$_1$), after size exclusion chromatography with slightly higher molecular weight (a. S$_2$), final NMR sample + DTT (final) and final NMR sample - DTT. Monomeric protein is labeled with a triangle (◄) and dimer with an asterik (*). For the SDS-PAGE the LMW-SDS Marker from GE Healthcare (Freiburg, Germany) was used.*

The well dispersed signals in the ^{1}H-^{15}N BEST-TROSY spectra (Figure 3.5) indicate that both NTDs of U and H are folded. Hence, the CTD is not required for the protein folding of the NTD as published previously [216, 234]. In a previous study no dimerization of the NTD alone was observed [167]. However, the NTD was expressed including the peptide-linker (amino acids 83-88) and the protein was studied by gel filtration and analytical ultracentrifugation. These experiments were performed with a significant lower protein concentration (0.4 mg/ml) than during NMR spectroscopy (10 mg/ml). Mayr *et al.* [166] showed,
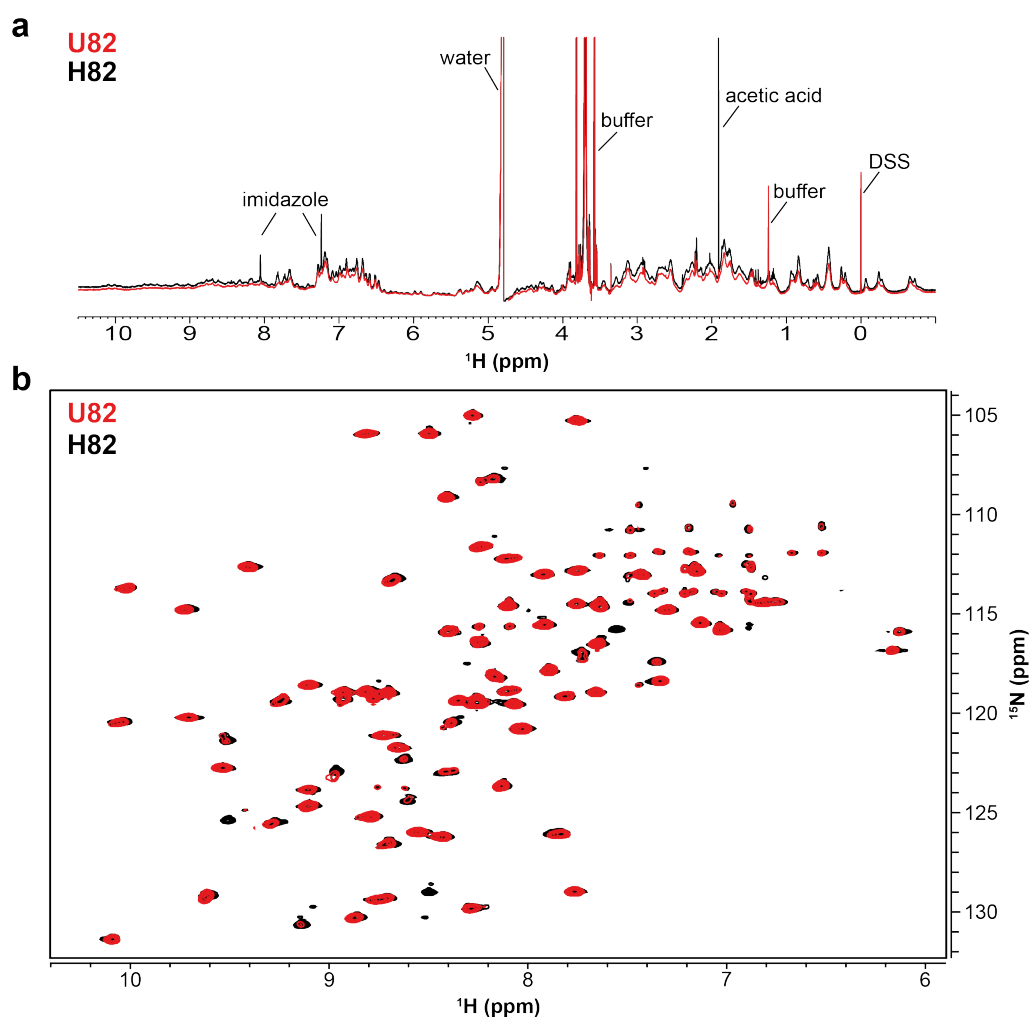
Figure 3.5: *Obtained differences by NMR spectroscopy of the GBC NTD variants U and H. a)* 1D ^1H NMR spectrum of U82 (red) and H82 (black). Signals were referenced to DSS. Parameters for acquisition: NS=32, TD=32768, T=298K. **b)** 2D ^1H-^{15}N BEST-TROSY spectrum of U82 (red) and H82 (black). Parameters for acquisition: NS=32, TD= 1176, 512, T=298K. Concentration of the protein U was 1 mM and of the protein H 0.9 mM. Residual imidazole and acetic acid signals as well as water, buffer and DSS signals are labeled accordingly.
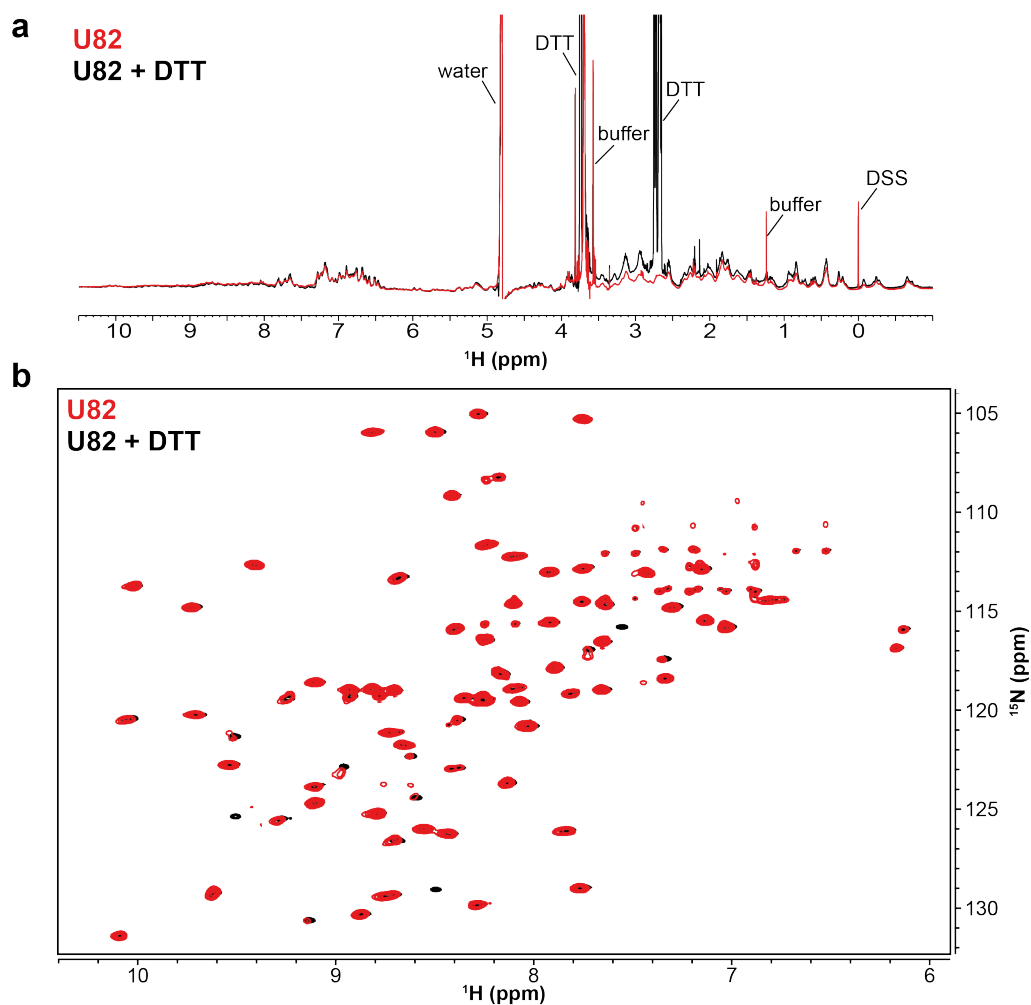
Figure 3.6: *1D ^1H and 2D ^1H-^{15}N BEST-TROSY spectrum of U82 and U82 after addition of DTT. a)* *1D ^1H NMR spectrum of U82 (red) and U82 after addition of DTT (black). Signals were referenced to DSS. Parameters for acquisition: NS=32, TD=32768, T=298K. **b)** 2D ^1H-^{15}N BEST-TROSY spectrum of U82(red) and U82 after addition of DTT (black). Parameters for acquisition: NS=32, TD= 1176, 512, T=298K. Concentration of the protein U was 1 mM. Final DTT concentration was approx. 38 mM.*

that mutation of the linker peptide did not change the stability of the full-length GBC and argued that domain-interface rather than the connecting peptide determines the domain association. But, probably the linker stabilizes the NTD to form a homodimer. To determine which residues were influenced by addition of DTT to the protein, triple resonance experiments were performed and the protein ^{1}H-^{15}N BEST-TROSY spectrum could be assigned (Figure 3.7). The assignment was performed on H protein treated with DTT to obtain only one reduced conformation. The chemical shifts of the backbone assignment are listed in the appendix.
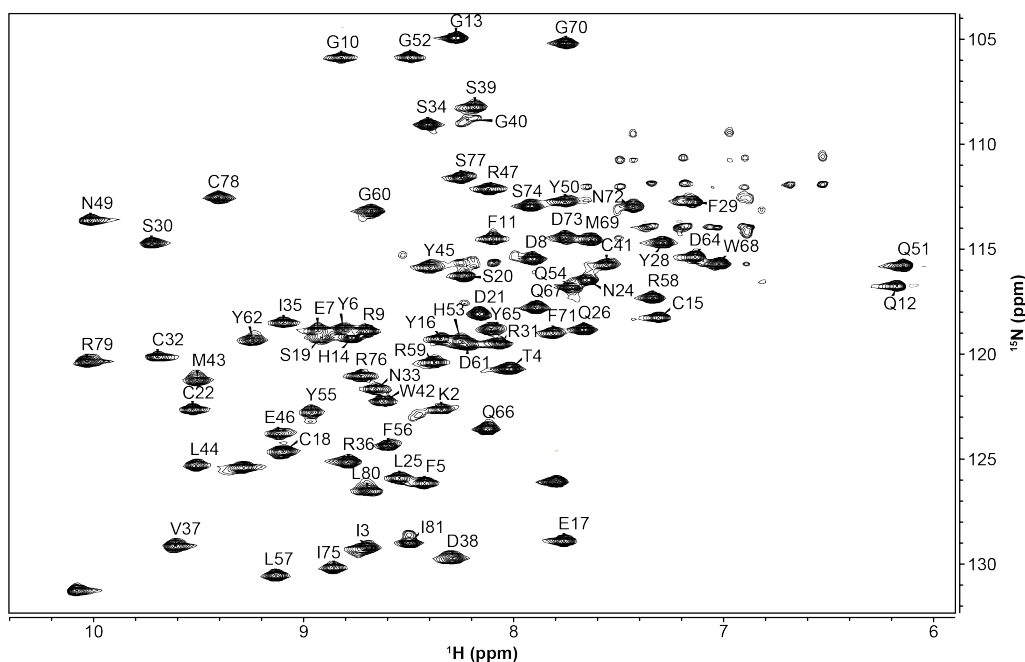
Figure 3.7: *Assigned backbone signals in the 2D ^{1}H-^{15}N BEST-TROSY of reduced H82. Paramter for acquisition: NS=8, TD= 1024, 200. T=298K. Protein concentraion of H was 1.1 mM.*

After assignment the missing signals in the U spectrum could be assigned to the amino acids Cys41, Met43, Leu44, Gln 54, Tyr 55, Phe 56, Leu57, Arg 58 and Ile81 (Figure 3.7). These residues are located at the inter-domain interface between the NTD and CTD (Figure 3.8a) and Met43, Phe56 and Ile81 are also involved in hydrophobic interaction [241] (Figure 3.8b). After addition of unlabeled CTD to the ^{15}N labeled NTD, these signals disappeared and addition of the unlabeled NTD to ^{15}N labeled CTD showed also clear interaction between the two domains (appendix Figure 5.1), although no heterodimer formation was observed in a previous study adding recombinant CTD to the proteolytically prepared NTD [234]. In this study, the connecting peptide linker was attached to the NTD, hence the six-residue peptide linker prevents homodimerization of the NTD, but is not required for inter-domain interaction.

In summary, the NTD is able to fold independently and H is completely reducued. U exists mostly in a reduced form, since addition of DTT did not show

an effect of the backbone signals in the ^{1}H-^{15}N BEST-TROSY spectrum in three out of five experiments. Twice addition of DTT decreased the dimer formation in U with Cys41 involved. However, an effect of DTT addition on Cys18 or Cys22, which were shown to be involved in a disulfide bridge in the X-ray structure [177], was not detected. The CTD is translated significantly slower than the NTD [216] and most of the translational pauses occur during the translation of the CTD [131]. The recognized conformational difference between U and H in the full-length spectrum by Buhr *et al.* [22] might thus occur while the CTD is translated (Figure 3.9. Probably the oxidation could also occur at an earlier state during translation while part of the NTD is still inside the ribosomal exit tunnel.
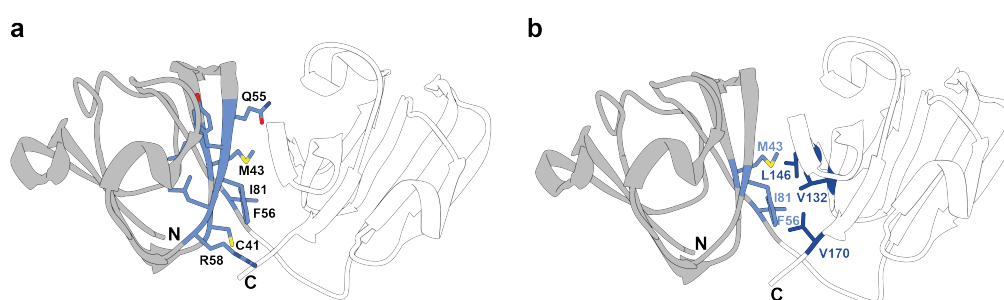
Figure 3.8: ***DTT sensitive residues in the NTD of GBC. a)*** *Residues appearing after the addtion of DTT to the protein U (light blue) shown on GBC structure (PDB: 4QWA). Influenced residues were: Cys41, Met43, Leu44, Gln54, Tyr55, Phe56, Leu57, Arg58 and Ile81.* ***b)*** *GBC structure with hydrophobic core residues relevant for inter-domain interaction. NTD residues involved in inter-domain interactions are Met43, Phe56 and Ile81 (light blue) and CTD residues are Val132, Leu146 and Val170 (dark blue). For better visualization amino acids are labeled in the one letter code. NTD is shown in grey and CTD in white (visualized using UCSF Chimera [204]).*
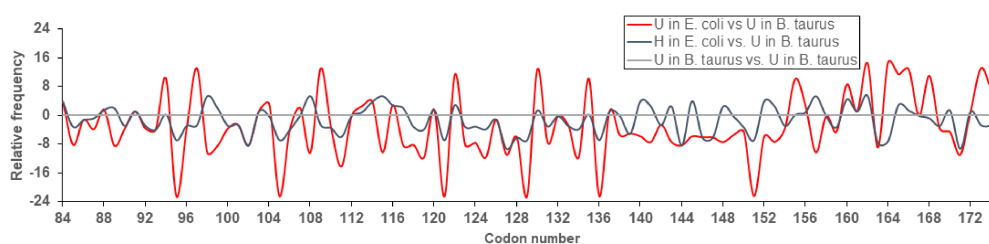
Figure 3.9: ***Codon usage of the CTD of GBC.*** *Relative differences in the codon usage of the variants U (red) and H (blue) expressed in E. coli to the codon usage of U in B. taurus (grey). Residues 84-175 are shown inluding the N-terminal methionine.*

3.2   Disulfide bond formation within the ribosomal exit tunnel

3.2.1   RNC constructs for cryo-EM and solid-state NMR experiments

To study how the oxidation state of cysteine residues change during translation, different lengths of the NTD of GBC were fused with the 17-aa long SecM stalling motif [180] (Figure 3.10). The focus was set on the NTD, as six out of seven cysteine residues are located within this domain and only these cysteine residues were prone to oxidation [22]. Six different constructs were designed and studied by solid-state NMR enhanced by DNP, cryo-EM, flexible-meccano simulations and mass spectrometry. Labeling of the RNC constructs excluded the N-terminal methionine residue to be consistent with other published studies. The constructs contained four to six cysteine residues and for the shortest construct U32SecM cysteine to alanine mutations were carried out using site-directed mutagenesis (section 2.2.4). DNA sequence and amino acid sequence of all constructs are listed in the appendix. The mutated RNCs name only the cysteine residues left, while all other cysteine residues are mutated to alanine, hence U32SecM C18 contains only Cys18. The construct lengths are visualized on the structure of isolated GBC (PDB: 4W9B), reflecting the obtained secondary structure of the non-ribosome bound GBC (Figure 3.11). The integrity of the RNC was confirmed by western blot analysis (Figure 5.2).
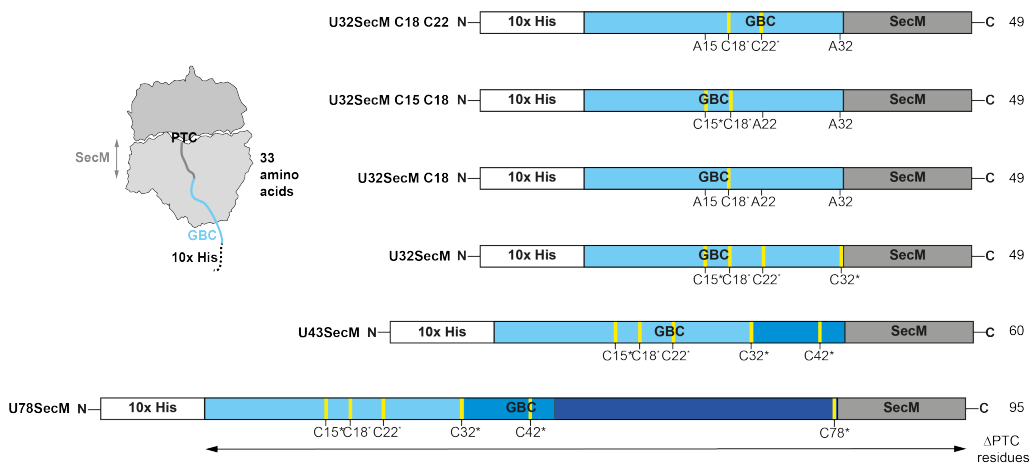
Figure 3.10: ***RNC GBC constructs with variable polypeptide chain lengths.*** *Different lengths of the GBC gene were fused with an N-terminal 10x histidine purification tag (white) and the 17-amino-acid SecM stalling sequence (dark grey). Numbering excludes the N-terminal methionine of GBC to be consistent with other published studies. Residues 1-32 (light blue), 33-43 (blue), 44-78 (dark blue). ^{13}C labeled cysteine residues are marked with an asterisk (*) and a yellow line.*
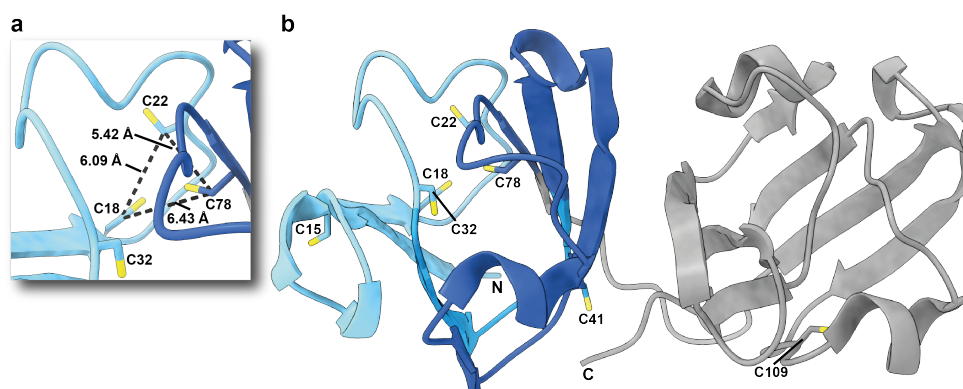
Figure 3.11: *RNC GBC constructs displayed on GBC structure. **a)** Detailed view on Cys18, Cys22 and Cys78 and distances between these Cys residues. **b)** Complete structure of full-length GBC with coloring corresponding to RNC chain length. Numbering excludes the N-terminal methionine of GBC to be consistent with other published studies. Residues 1-32 (light blue), 33-43 (blue), 44-78 (dark blue) and 79-174 (grey) (PDB: 4W9B). Residues 79-174 were not used as RNC constructs.*

3.2.2   Control of leaky expression

To verify if leak expression of the NC occurs before addition of IPTG and whether RNCs can be separated from free ribosomes using Ni-NTA affinity chromatography, the cell lysate of induced and uninduced cells were loaded onto a Ni-NTA column. Therefore, BL21(DE3)Δtig cells were grown in uninduced and induced medium with 1 mM ITPG. Two constructs U32SecM C15A C32A and U79SecM were used for this control experiment. Only one RNC construct is representated as both RNC constructs showed similar results. Figure 3.12 displays the Ni-NTA affinity chromatogram of the RNC U32SecM C15A C32A with and without induction of IPTG.

This experiment clearly proves that only in the induced sample ribosomes elute with an absorbance ratio of $A_{260\ nm}/A_{280\ nm}$ of ~ 2.0. This absorbance ratio is characteristic to ribosomes or samples with high nucleotide content such as RNA and DNA. The undesired DNA was digested by addition of DNase to the cell lysate and 60% of the ribosomal mass is made out of rRNA, hence RNCs elute. Thus, leaky expression was successfully suppressed, unspecific binding of free ribosomes to the Ni-NTA column could be ruled out and the purification strategy is suitable as a first step RNCs isolation.

a   uninduced (without IPTG)
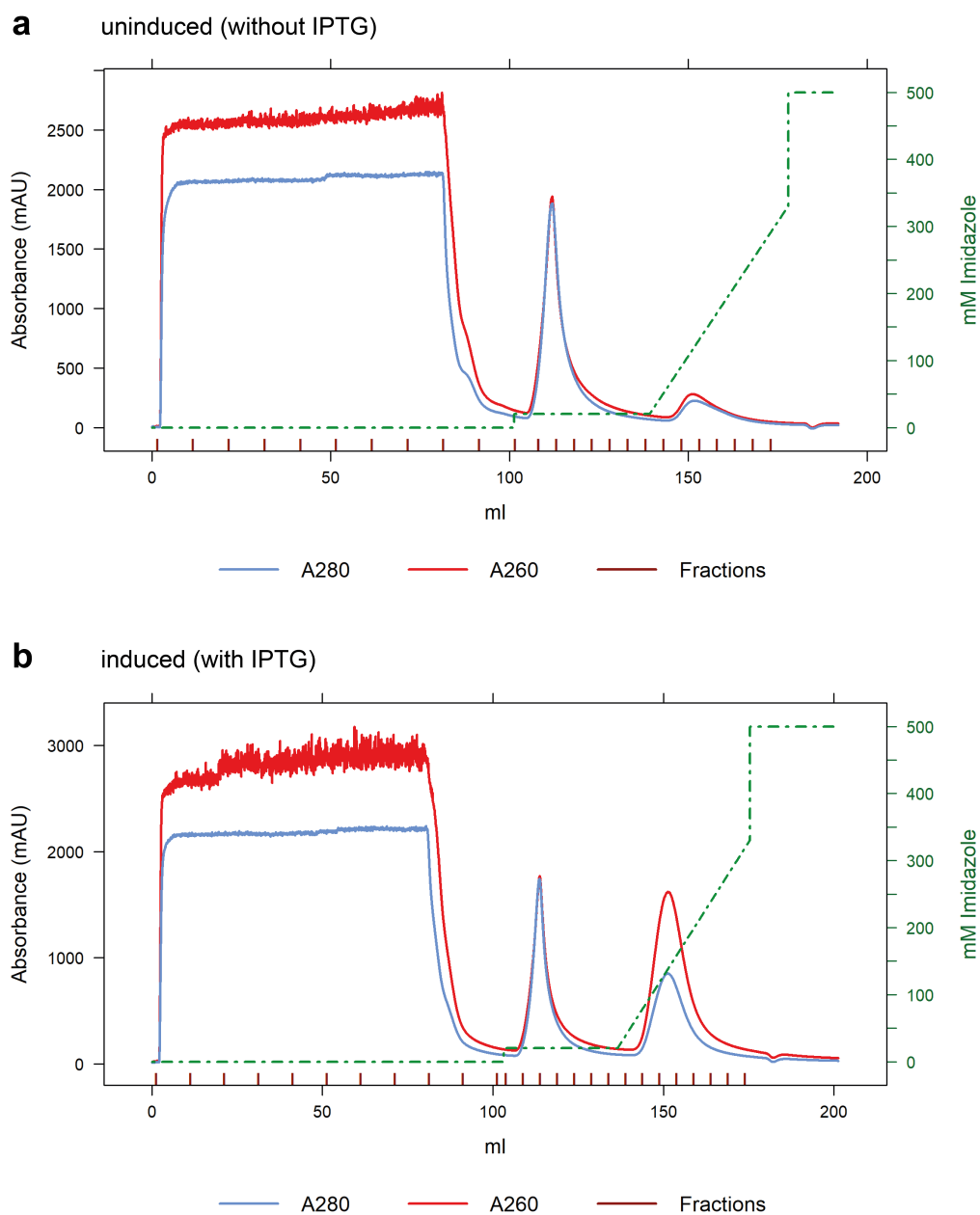
b   induced (with IPTG)

Figure 3.12: ***Control of leaky expression using Ni-NTA affinity chromatography.*** *Chromatogram of Ni-NTA affinity chromatography of uninduced (without IPTG) (**a**) and induced (with IPTG) (**b**) U32SecM C15A C32A cells after cell lysis. Absorbance was recorded at 260 nm (red) and 280 nm (blue). Bound protein or RNCs were eluted by application of a linear gradient up to 500 mM imidazol.*

3.2.3   Exemplary purification on U78SecM

All RNC constructs were purified by Ni-NTA affinity chromatography and a subsequent sucrose gradient ultracentifugation step as schematically illustrated (Figure 3.13). The results of this purification procedure are shown exemplary with U78SecM in this section. Figure 3.14 displays the Ni-NTA affinity chromatogram of the ^{13}C, ^{15}N cysteine labeled U78SecM RNC. This longest construct was purified using 10 mM imidazole within the lysis buffer, which was later excluded from all other RNC purifications. A strong signal corresponding to the size of the NC bound to tRNA, is visible within the flow-through (FT) fraction on the western blot (Figure 3.15b). In order to increase the binding of the RNC to the Ni-NTA resin, imidazole in the lysis buffer was omitted and the loading speed was reduced from 2.5 ml/min to 1.5 ml/min. Thereby no RNC signal was detected in the flow-through fractions of the other purifications. Another approach used by Cassaignau *et al.* [30] is the batch procedure.
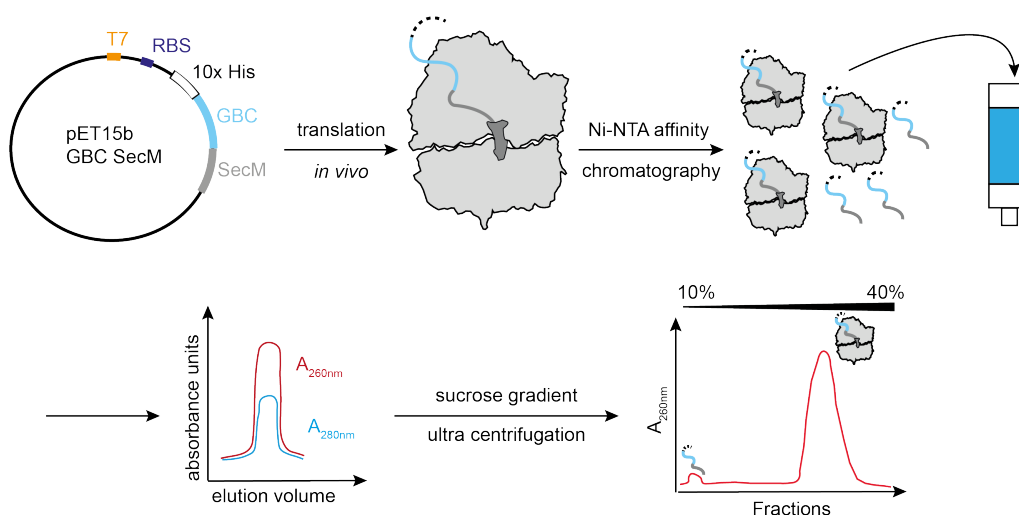
Figure 3.13: *Schematical representation of RNCs production. RNC constructs contain a N-terminal 10x histidine purification tag (white and dashed line), the GBC segment (light blue) and the C-terminal 17-amino-acid SecM stalling sequence (dark grey) and are expressed in vivo. RNCs are purified by Ni-NTA affinity chromatography revealing an absorption ratio of $A_{260\ nm}/A_{280\ nm}$ of 2, and a subsequent sucrose gradient ultra centrifugation step.*

Within the pellet fractions (P1 and P2) and inclusion body fraction (Inc) is free nascent chain detectable (Figure 3.15b). This is either due to incomplete cell lysis or insolubility of the released NC in this lysis buffer. The U78SecM NC exhibits an isoelectric point (pI) of 7.19, the selectively labeled medium has a pH of 7.4 and all purification steps are performed on a pH of 7.5 to keep the ribosome intact. Therefore, free NC could be localized to inclusion bodies within the *E. coli* cell or precipitate after cell disruption. RNC containing fractions elute in fractions 36-54 (Figure 3.14). The highest absorbance increase is detectable in fraction
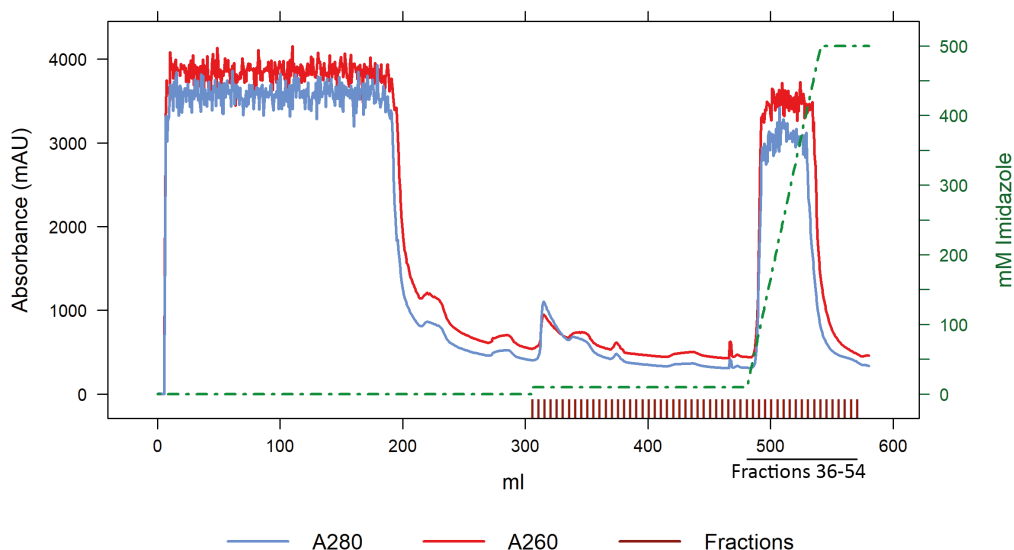
Figure 3.14: ***Ni-NTA affinity chromatogram of ^{13}C, ^{15}N cysteine labeled U78SecM.*** *The absorbance at 260 nm (red) and 280 nm (blue) was recorded. Bound protein or RNCs were eluted by application of a linear gradient (20-500 mM imidazol) Fractions 36-54 were analyzed using SDS-PAGE.*

37 and 38 with an imidazole concentration between 83 and 88 mM. Fractions 48-50 show an intense signal of free NC in the western blot, while absorbance in this fractions is rather low. Hence, the NC elutes after the RNCs. The free NC is removed from the RNCs in a subsequent purification subsequent step using sucrose gradient ultracentrifugation. The pooled fractions are concentrated using a centrifugation concentrator with a MWCO of 100 kDa, large enough that the free NC (MW 12.80 kDa) can pass the membrane removing also free NCs from RNCs.
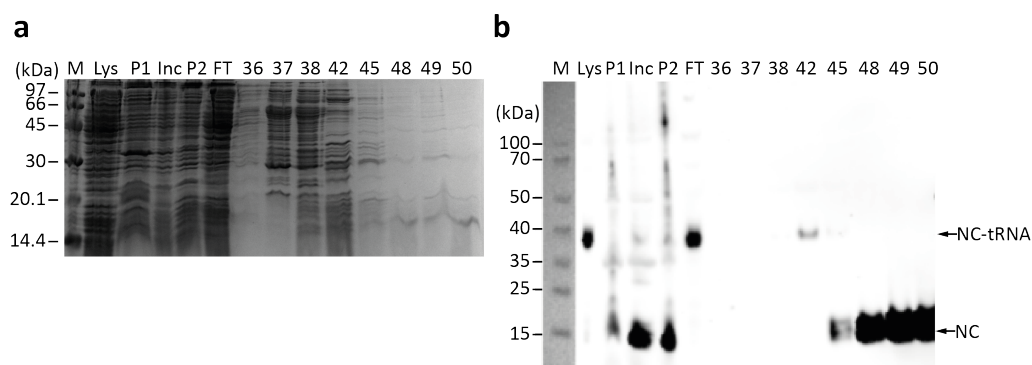
Figure 3.15: *SDS-PAGE gel and western blot of ^{13}C, ^{15}N cysteine labeled U78SecM after Ni-NTA affinity chromatography.* *a)* *15% SDS gel of ^{13}C, ^{15}N cysteine labeled U78SecM after Ni-NTA chromatography stained with coomassie blue.* *b)* *Western blot of ^{13}C, ^{15}N cysteine labeled U78SecM after Ni-NTA chromatography. Samples from left to right: protein marker (M), cell lysate (Lys), pellet 1 (P1), inclusion body (Inc), pellet 2 (P2), flow-through (FT), Fractions 36-38, 42, 45, 48-50. Fractions 37-48 were pooled and used for further purification. For the SDS-PAGE the LMW-SDS Marker from GE Healthcare (Freiburg, Germany) and for western blotting the Spectra^{TM} Multicolor Broad Range Protein Ladder from ThermoFisher Scientific (Waltham, USA) were used. Size of the NC and the NC bound to tRNA (NC-tRNA) are labeled accordingly.*
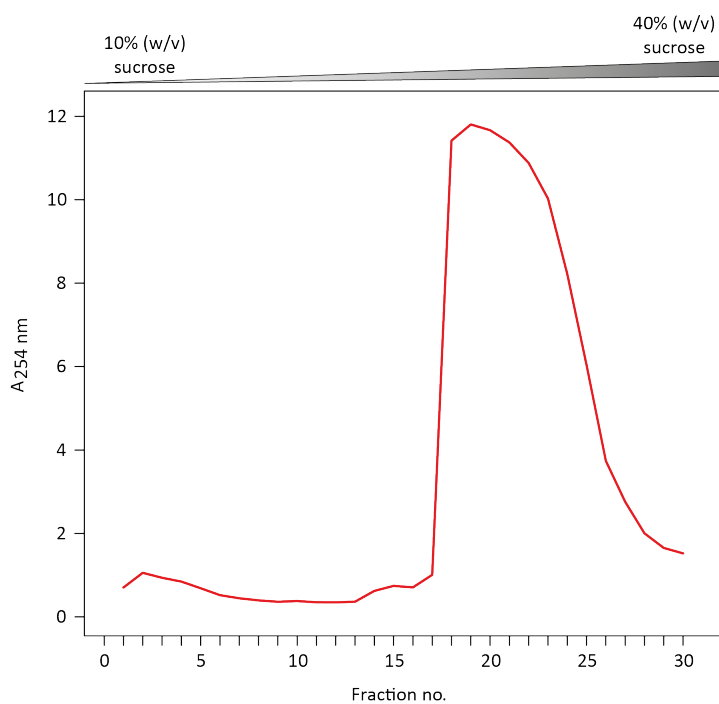
Figure 3.16: *Absorbance profile of ^{13}C, ^{15}N cysteine labeled U78SecM after 10%-40% (w/v) sucrose gradient ultracentrifugation.* *Absorbance was detected at 254 nm. Fractions were collected automatically using a Piston Gradient Fractionator (Biocomp Instruments, Fredericton, USA).*

The sucrose gradient ultracentrifugation is performed as described in section 2.4.5. Figure 3.16 illustrates the absorbance at 254 nm of one of six centrifugation tubes used per run. The highest absorbance is detectable in fractions 18-26. On the SDS-PAGE gel, these fractions show the expected characteristic band pattern of ribosomes and on the western blot a signal at the size of the NC with bound tRNA (Figure 3.17). In addition, a faint band at the size of the free NC is also detectable but probably due to the incubation at 70°C in SDS loading buffer. In fraction 2 at the top of the gradient, free NC is detectable. Hence, NC and RNCs could be easily separated by sucrose gradient ultracentrifugation. In summary, the established purification procedure for RNCs consists of two steps i) using Ni-NTA affinity chromatography ribosomes from RNCs can be separated ii) using sucrose gradient ultracentrifugation free NC and RNCs can be separated. The sucrose cushion step to pellet the RNCs as used by Cassaignau *et al.* [30] could be omitted, reducing the purification time and preventing the harsh conditions to dissolve the ribosomes again.
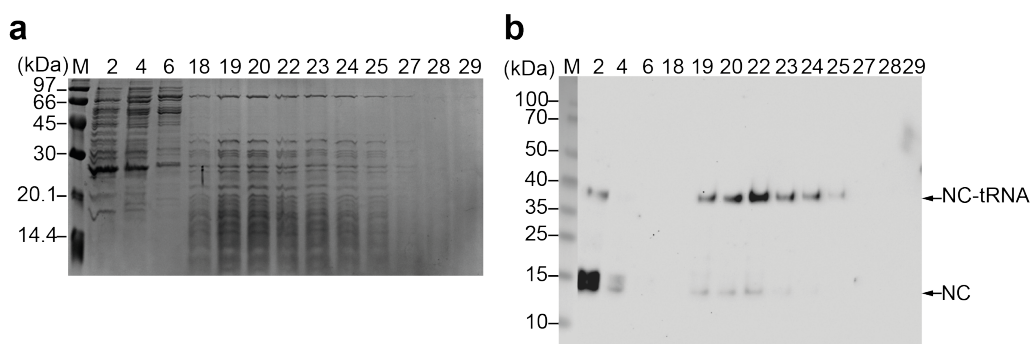
Figure 3.17: ***SDS-PAGE gel and western blot of ^{13}C, ^{15}N cysteine labeled U78SecM after sucrose gradient ultracentrifugation.*** *a) 15% SDS-PAGE gel of ^{13}C, ^{15}N cysteine labeled U78SecM after sucrose gradient ultracentrifugation stained with coomassie blue.* ***b)*** *Western blot of ^{13}C, ^{15}N cysteine labeled U78SecM after sucrose gradient ultracentrifugation. Samples from left to right: protein marker (M), Fractions 2, 4, 6, 18-20, 22-25, 27-28. Fractions 19-24 were pooled and used for solid-state NMR enhanced by DNP. For the SDS-PAGE the LMW-SDS Marker from GE Healthcare (Freiburg, Germany) and for western blotting the Spectra^{TM} Multicolor Broad Range Protein Ladder from ThermoFisher Scientific (Waltham, USA) were used. Size of the NC and the NC bound to tRNA (NC-tRNA) are labeled accordingly.*

3.2.4 Optimization of the glycerol cushion for solid–state NMR experiments enhanced by DNP

3.2.4.1 Purification of ^{15}N labeled ribosomes from JE28 cells

To get a sufficient amount of RNCs to perform solid-state NMR experiments and a good signal to noise ratio, different glycerol cushion concentrations were tested. JE28 cells were grown in minimal media supplemented with ^{15}N-NH$_4$Cl purified according to section 2.4.1 and then used as a test sample to calculate

the ^{1}H enhancement factor ε and signal intensity. Figure 3.18 shows the Ni-NTA affinity chromatogram of ribosomes containing a 6x His-tag on the ribosomal protein L12 [48] and Figure 3.19 the corresponding fractions on a 12% SDS-PAGE gel.
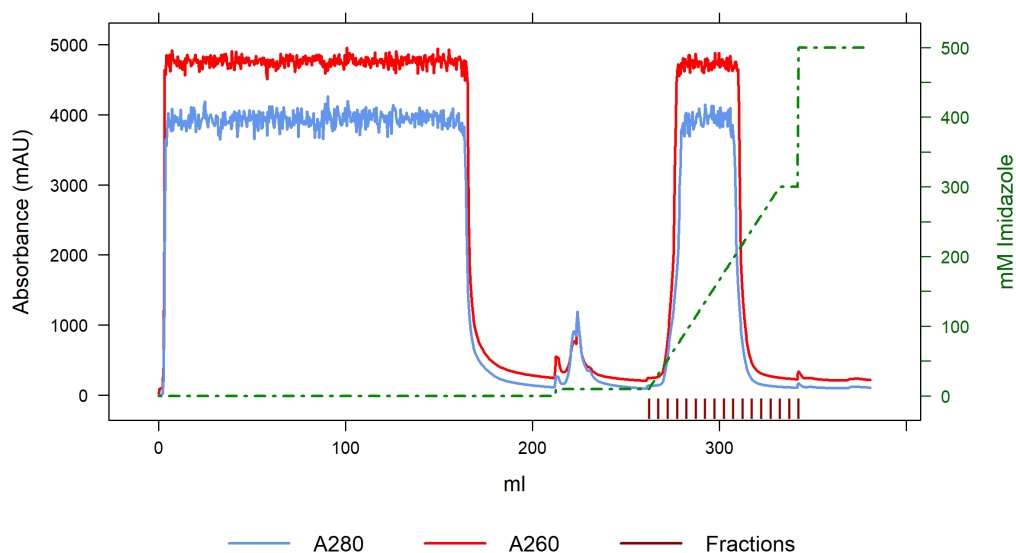
Figure 3.18: *Ni-NTA affinity chromatogram of ^{15}N labeled ribosomes from JE28 cells. The absorbance was recorded at 260 nm (red) and 280 nm (blue). Bound protein or RNCs was eluted by application of a linear gradient up to 500 mM imidazol.*

After affinity purification, the ribosome containing fractions 4-11 were pooled and approx. 2 nmol per tube were loaded onto a 10%-40% (w/v) sucrose gradient (Figure 3.20a). As these experiments were performed with a SW41 rotor (Beckman Coulter, Brea, USA) and fractionized manually, the absorbance profile differs from the once performed with a SW32 rotor (Beckman Coulter, Brea, USA) on the sucrose gradient station. The absorbance was recorded at 254 nm and normalized to the maximum absorbance value.

The fractions of this exemplary sucrose gradient were loaded on a 12% SDS-PAGE gel (Figure 3.20b). Fractions 12 and 13 were taken from the bottom of the centrifugation tube, which could not be reached by the stamp of the gradient fractionizer. Fractions 8-11 were pooled, sucrose removed and concentrated to approx. 12 µM for the glycerol cushion experiment. As can be seen from the 12% SDS silver stained gel, only minor contaminations (visible as thin protein bands between 40 and 60 kDa) can be detected.
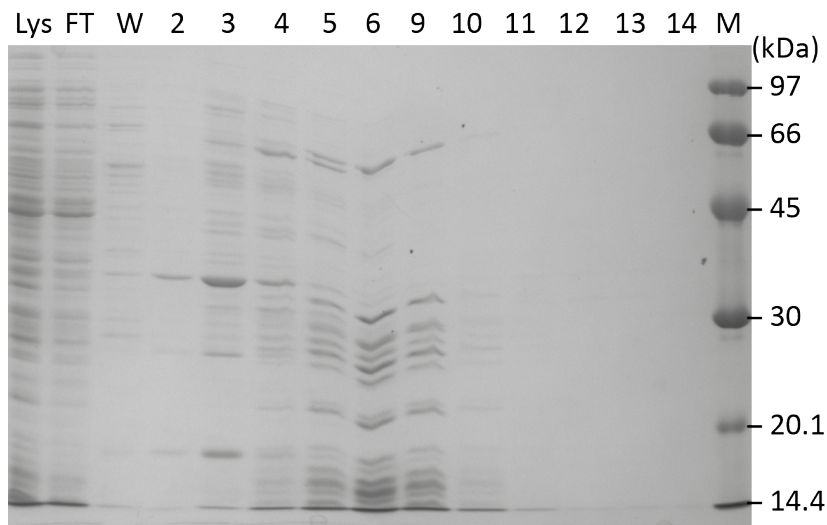
Figure 3.19: *12% SDS-PAGS gel of ^{15}N labeled ribosomes after Ni-NTA affinity chromatography.* *Samples from left to right: Cell lysate (Lys), flow-through (FT), wash fraction with 5 mM Imidazol (W), Fractions 2, 3, 4, 5, 6, 9, 10-14, protein marker (M). The fractions 4-11 were pooled and loaded onto a sucrose gradient. The LMW-SDS Marker from GE Healthcare (Freiburg, Germany) was used. The gel was stained with coomassie blue.*
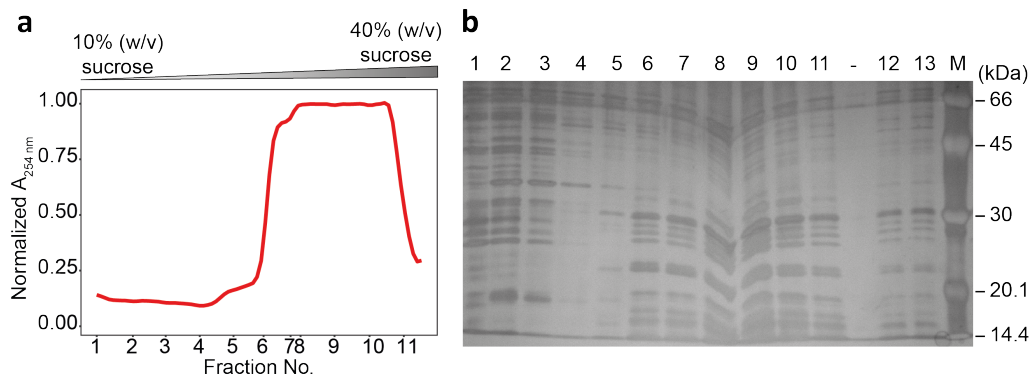
Figure 3.20: *Sucrose gradient ultracentrifugation of ^{15}N labeled ribosomes.* *a)* *Absorbance profile of ^{15}N labeled ribosomes after 10%-40% (w/v) sucrose gradient ultracentrifugation. Absorbance was recorded at 254 nm and normalized to the maximum value. Fractions were collected manually.* *b)* *Silver stained 12% SDS-PAGE gel of ^{15}N labeled ribosomes after after 10%-40% (w/v) sucrose gradient ultracentrifugation. Fractions 1-11 fractionized using the fractionizer. Fractions 12 and 13 were taken from the bottom of the centrifugation tube, protein marker (M). The fractions 8-11 were pooled, sucrose removed and the ribosomes used for further experiments. The LMW-SDS Marker from GE Healthcare (Freiburg, Germany) was used.*

3.2.4.2 Determination of the ^1H enhancement factor (ϵ) and signal intensity using different glycerol cushions

The purified ^{15}N uniformly labeled ribosomes were analyzed by different glycerol cushions varying in AMUPoL and glycerol concentrations (Table 3.1). The relative ^{15}N signal intensity (protein and nucleic acid) with and without microwave (MW) irradiation was normalized to the highest value in each category (MW off or MW on) and the ^1H enhancement factor ε was calculated (done by Dr. Jiafei Mao). Table 3.1 presents the corresponding values and Figure 3.21 shows a plot of the DNP enhancement factor ε versus signal intensity without DNP enhancement. It is obvious that, even though the highest ε value (169) is observed using a glycerol cushion of 40% (v/v), the signal intensity without DNP enhancement is the lowest within all samples. Indeed, even with such a high DNP enhancement, the signal intensity of this sample under DNP condition remains the second lowest in our screening. Therefore, using such a high glycerol cushion results in a lower concentration of ribosomes within the rotor. On the other hand, the signal intensity of the sample lacking glycerol shows the highest value, if the MW is off due to the high concentration of ribosomes within the rotor, but the DNP enhancement is rather low.

Table 3.1: *Optimization of glycerol cushion.* *Glycerol concentration in % (v/v), AMUPoL concentration (mM), ^1H enhancement factor (ϵ), relative signal intensity without and with microwave (MW) normalized to the highest signal intensity in each category.*

Sample number	Glycerol concentration in % (v/v)	AMUPol concentration (mM)	^1H enhancement (ϵ)	Relative signal intensity (MW off)	Relative signal intensity (MW on)
I	40	7.5	169	0.20	0.65
II	40	15	130	0.23	0.59
III	30	15	131	0.39	1.00
IV	20	15	41	0.97	0.78
V	10	15	67	0.65	0.85
VI	0	15	35	1.00	0.68

Ravera *et al.* [210] observed an ε of approx. 40 sedimenting the polarizing agent TOTAPOL with the 480 kDa iron-storage protein complex apoferritin (ApoF) without any glass-forming agent. Sedimentation resulted in a layer of "glassy-like" protein, since the frozen but not sedimented sample obtained an enhancement factor of only 2. The cryo-protected and sedimented sample with glycerol showed an ε of 70 though [210]. In the ribosomes samples, a more than four-fold reduction in ε ranging between 0% (v/v) and 40% (v/v) glycerol concentration is detectable. Lange *et al.* [145] investigated the effect of different TOTAPOL concentrations in uniformly labeled ^{13}C, ^{15}N proline by comparing relaxation times, enhancement
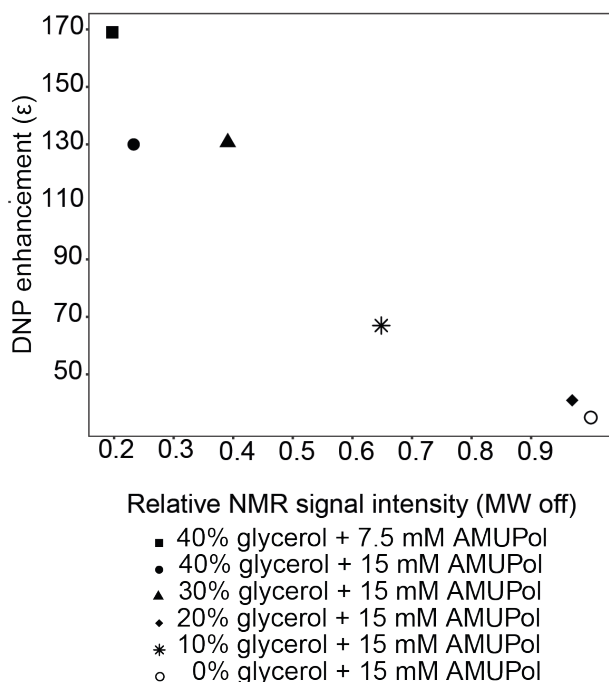
Figure 3.21: *Analysis of glycerol cushions. DNP enhancement factor and relative signal intensity of ^{15}N labeled ribosomes without microwave depending on different glycerol cushion conditions.*

factor (ε) and NMR signal intensity. Signal-to-noise ratio (S/N) was either measured with a constant number of scans or total experiment time. In the former the S/N ratio decreased with increasing concentration of TOTAPOL and in the latter, the maximum S/N ratio was observed at a concentration of 26 mM TOTAPOL and decreased with increasing radical concentration. The presence of a radical also leads to a shortening of the bulk spin-lattice relaxation time T_1 and the spin-spin relaxation T_{2*} [145]. With a short T_1 a faster repetition of rate in proton-carbon CP experiments is possible, which would increase the S/N per unit time [145]. Shorter T_{2*} values, however highlight line broadening [145]. Lange *et al.* [145] observed no change in the linewidth in a TOTAPOL concentration range of 0 to 26 mM, which is the commonly used for biological samples. The authors also estimated that nuclei signals closer than 10 Å to a radical center are bleached due to the paramagnetic effects. These effects would be more severe in samples with a high concentration of biradical and/or without glass-forming agent. These samples would thereby show a reduced number of detectable nuclei in the enhanced spectrum and a reduced apparent DNP enhancement factor. The obtained ε of 130-169 using AMUPol here is approx. 10-times higher compared to a different RNC sample using a 60% (v/v) glycerol and 30 mM TOTAPOL [144].

The best DNP enhanced signals for ribosomes samples are obtained using the 30% (v/v) glycerol cushion. A lower AMUPol concentration of 7.5 mM as in sample I (Table 3.1) could potentially improved ε and reduce sample costs, but

the effect was not studied for 30% (v/v) glycerol and therefore, all subsequent RNC samples were prepared with 15 mM AMUPol.

3.2.5   Experiments on uniformly ^{13}C, ^{15}N U78SecM

Within the ^{13}C-^{13}C DQ-SQ 2D spectrum of the uniformly ^{13}C, ^{15}N labeled U78SecM (Figure 3.22) the chemical shifts of different types of RNA bases, ribose as well as some characteristic amino acids could be assigned. ^{13}C signals correlating directly bonded carbons (therefore short ^{13}C-^{13}C distance) appear as a pair at the F1 (DQ) chemical shift as the sum of their SQ chemical shift. For example, the C1' and C2' of ribose have a chemical shift around 94 and 77 ppm. Since they are connected by a bond, both signals appear at a DQ chemical shift of $94 + 77 = 171$ ppm. The high signal intensity of ribose is due to the scrambling of ^{13}C-glucose, even at rather low level, into the highly abundant ribose. The prokaryotic ribosome contains three different rRNAs, the 5S, 16S and 23S rRNA consisting of 120, 1500 and 2900 nucleotides, respectively. Hence, 4520 ribose are compared with 107 amino acids of the NC. Beside the ribose signals also the carbon atoms of the RNA bases are detectable and are labeled with a lowercase letter, while the amino acids are labeled with an uppercase letter in Figure 3.22. The side chain signals of aromatic amino acids are overlapping with the rRNA bases.

The only residues well resolved from other amino acids are alanine, serine and threonine. However, since U78SecM consists of two alanines, nine serine and two threonine residues, those signals cannot be unambiguously assigned. The ^{13}C uniformly labeled 20-amino-acids' long signal sequence of disulfide oxidoreductase A (DsbA) was used by Lange et al. [144] to study the secondary structure within the ribosomal tunnel. DsbA was fused to SecM to keep the NC arrested to the ribosome. In that study two different mutations of DsbA were used containing each time two serine residues while the other two were mutated to alanine. With this approach two serine residues could be identified per experiment and the chemical shift of those serines but also of non-overlapping residues types of alanine, isoleucine, valine and proline were compared to averaged chemical shifts from the Protein Chemical shift Database [279]. Although this signal sequence of DsbA was predicted to obtain an α-helical structure, isoleucine and valine residues were not involved in an α-helix and only Ser16 and Ser18 showed a minor helical content. However, the potential structural pertubations on the NC by introducing Ser/Ala mutations remain unclear. Lange et al. [144] used a different expression method for the production of RNCs. Starting with IPTG addition to the LB medium, changing then the medium to minimal medium without ^{13}C-glucose and ^{15}N-NH$_4$Cl. After 5 min rifampicin and after additional 10 min incubation, ^{13}C-glucose and ^{15}N NH$_4$Cl were added. In theory, this procedure also expresses unlabeled RNCs. Scrambling of ^{13}C-glucose into ribose was also detectable using this expression procedure, however the spectra region of the rRNA bases were not
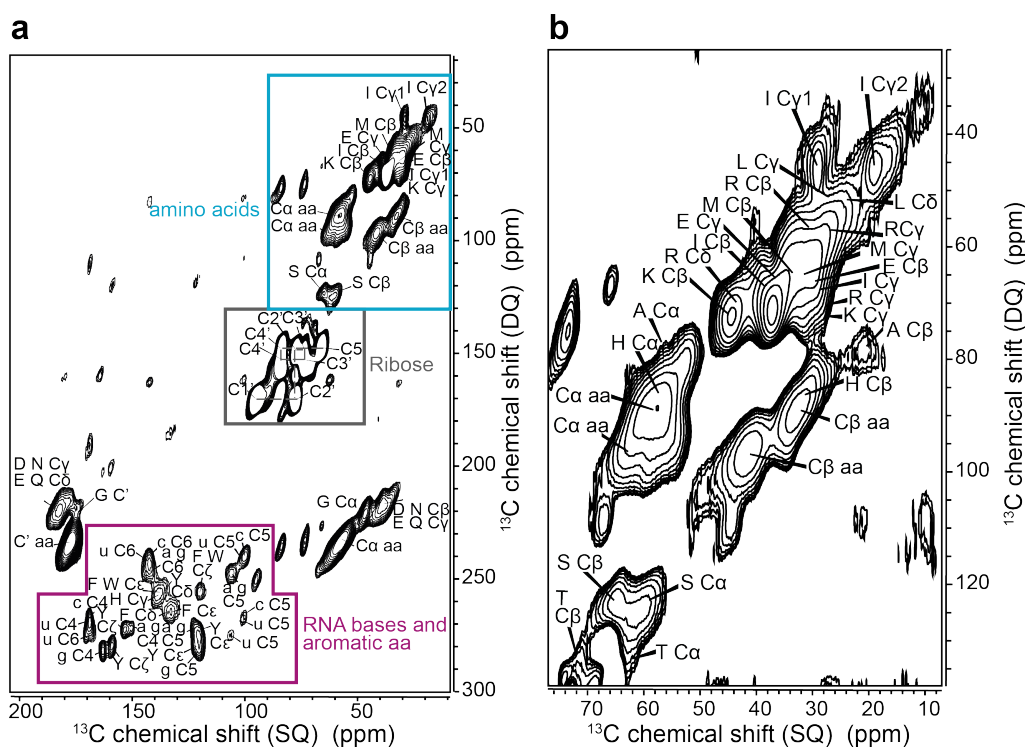
Figure 3.22: ***POST-C7 ^{13}C-^{13}C DQ-SQ 2D spectrum of the uniformly ^{13}C, ^{15}N labeled U78SecM. a)*** *Overall spectrum, showing the cross peaks of RNA bases (small letter code) and aromatic amino acids (capital letter code), the ribose cross peaks and the cross peaks of the amino acid region. Aromatic amino acids and RNA bases are shown in a pink box. The ribose signals can be unambiguously assigned and are shown in a grey box. Ca-Cγ of some characteristic amino acids are displayed in a cyan box.* ***b)*** *Amino acid region of the ^{13}C-^{13}C DQ-SQ spectrum. Exemplary amino acid chemical shifts are labeled with the one letter code. Conditions: 10 mM HEPES, 30 mM NH4Cl, 12 mM MgCl$_2$, 1 mM EDTA, pH 7.5 in D8-glycerol:D20:H20 (42.5:45.3:14.5). 9.4 T (400 MHz for protons), T = 100 K, MAS frequency = 8 kHz, ns = 256, 4 nmol U78SecM.*

shown in this publication. Additionally, Lange *et al.* [144] recorded impurities of DNA and peptidoglycan signals, probably due to binding of TOTAPOL to peptidoglycans. Those impurities have not been detected in our samples, indicating a higher sample quality.

Comparing the alanine, serine and threonine residues to the average chemical shifts (Table 3.2) provides a hint of the secondary structure. The localization of the serine and threonine residues in the NTD of GBC is shown in Figure 3.23a (PDB: 4W9B). The alanine residues are only located within the SecM sequence. In the structure of the NTD no serine nor threonine residues are arranged in an α-helical conformation. The averaged chemical shifts from Table 3.2 for secondary structure elements are shown in the ^{13}C-^{13}C DQ-SQ spectrum (Figure 3.23b). Since the DQ dimension is less resolved than the SQ dimension the averaged chemical shifts were placed only regarding the SQ chemical shift dimension. The Cβ chemical shifts of alanine indicates a β-strand, while the Cα chemical shift

of alanine cannot be assigned. Serine signals could not distinguish different conformations, since the signals are strongly overlapping. Seven serine residues are located within U78 and two are in the SecM sequence. Although the Cβ chemical shift of serine overlaps almost exactly with the averaged Cβ chemical shift of serine in a β-strand, an unambiguous identification of the secondary structure is impossible. For the threonine residues an α-helical conformation can be ruled out, since no signal is obtained at the position of the averaged chemical shift. One threonine is located inside a β-strand of GBC and the other threonine is located in SecM. Since both β-strand and coil conformation for threonine could be obtained and the averaged chemical shifts for this two conformations are strongly overlapping it is not clear whether this sequence of GBC is folded on the ribosome or not. Therefore, the certainty of this kind of analysis as performed by Lange *et al.* [144] is debatable.

Table 3.2: *Chemical shifts of alanine, serine and threonine.* *Averaged $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts values of alanine, serine and threonine (in ppm) categorized to secondary structure conformation derived from the Protein chemical shift database [279].*

Amino acid	^{13}C chemical shift of	Coil (ppm)	α-helix (ppm)	β-strand (ppm)
Alanine	Cα	52.65 ± 1.64	54.83 ± 1.05	51.53 ± 1.91
	Cβ	19.06 ± 1.26	18.26 ± 0.88	21.14 ± 2.05
Serine	Cα	53.38 ± 1.69	60.88 ± 1.61	57.54 ± 1.40
	Cβ	64.03 ± 1.27	63.08 ± 1.12	65.16 ± 1.51
Threonine	Cα	61.64 ± 2.07	65.61 ± 2.39	61.06 ± 1.59
	Cβ	70.12 ± 1.33	68.66 ± 1.17	70.75 ± 1.51

Summing up, some characteristic amino acids within U78SecM could be identified. The secondary structure of U78SecM is most likely within a β-strand or random-coil conformation, however an α-helical conformation cannot be ruled out completely. RNA base signals were detected indicating a certain level of ^{13}C scrambling into these chemical moieties. However, this scrambling were only detectable within the uniform labeled U78SecM sample but not in selectively cysteine labeled samples as used in our main work. Although, uniformly labeling of the NC is in principle possible, the structural analysis using solid-state NMR enhanced by DNP is difficult due to the severe overlap and weak dispersion of signals within the ^{13}C-^{13}C DQ-SQ spectrum.

Therefore, we decided to use ^{13}C, ^{15}N cysteine labeled RNCs and to assign one specific cysteine residue the neighboring C-terminal amino acid should be labeled with a ^{15}N isotope. The carbonyl carbon chemical shifts of this particular cysteine residue can be assigned using a N-C TEDOR experiment [89] (Figure 3.24a). Additionally, the Cα chemical shifts of all cysteine residues can also be detected
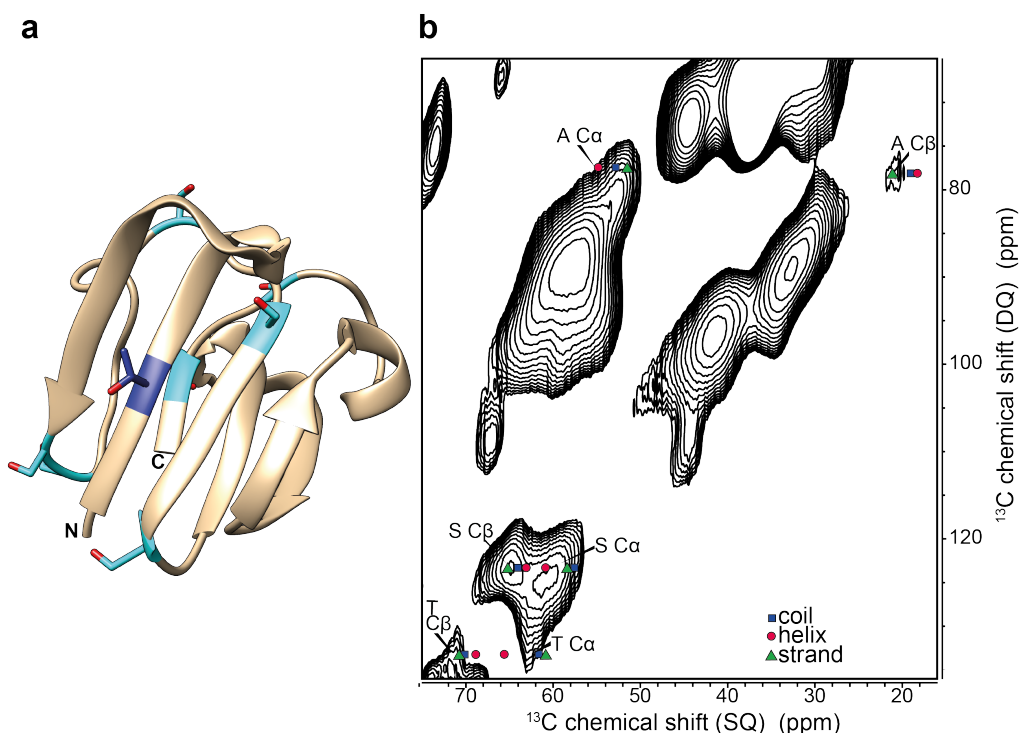
Figure 3.23: *Analysis of secondary structure elements in uniformly labeled U78SecM. a) Structure of GBC residue 1-78 (PDB: 4W9B). Serine residues in cyan, threonine residue in blue. b) POST-C7 ^{13}C-^{13}C 2D DQ-SQ spectrum of alanine, serine and threonine region of uniformly ^{13}C, ^{15}N labeled nascent chain U78SecM. Averaged chemical shifts [279] of these amino acids for coil (blue square), α-helical (red circle) and β-strand (green triangle) secondary structure are shown. Amino acids are labeled in the one letter code.*

using this experiment. For unambiguous identification of the carbon chemical shifts of the cysteine residue with a ^{15}N labeled neighbor a NCOCX experiment [198] is the way of choice (Figure 3.24b).

Within the NCOCX experiments, the magnetization is transferred from amid proton to nitrogen, further to the carbonyl carbon of the preceding amino acids and then via spin diffusion to carbon nuclei in the side chains (Figure 3.24b). Both experiments were carried out as a first test on the uniformly labeled U78SecM sample (Figure 3.24c and 3.24d). N-C', N-Cα correlations of amino acids could be identified, but also correlations within RNA bases. The intensities of RNA base signals are comparable to those of the N-C' or N-Cα signals within the TEDOR experiment, hence ^{15}N and ^{13}C labeling of neighboring atoms within the RNA occurs. In the NCOCX experiments also correlation between the ribose and N9 of purine or N1 of pyrimide are visible. This is due to the fact, that the N-C' coupling with around 15Hz [26, 35] is very similar to the N9-C1' and N1-C1' coupling with 11 and 12 Hz, respectively [72].

In this uniform labeled U78SecM sample, it was not possible to assign single amino acids, due to the severe overlap of signals. However, we also tested the

a

NC-TEDOR

b

NCOCX

c

d
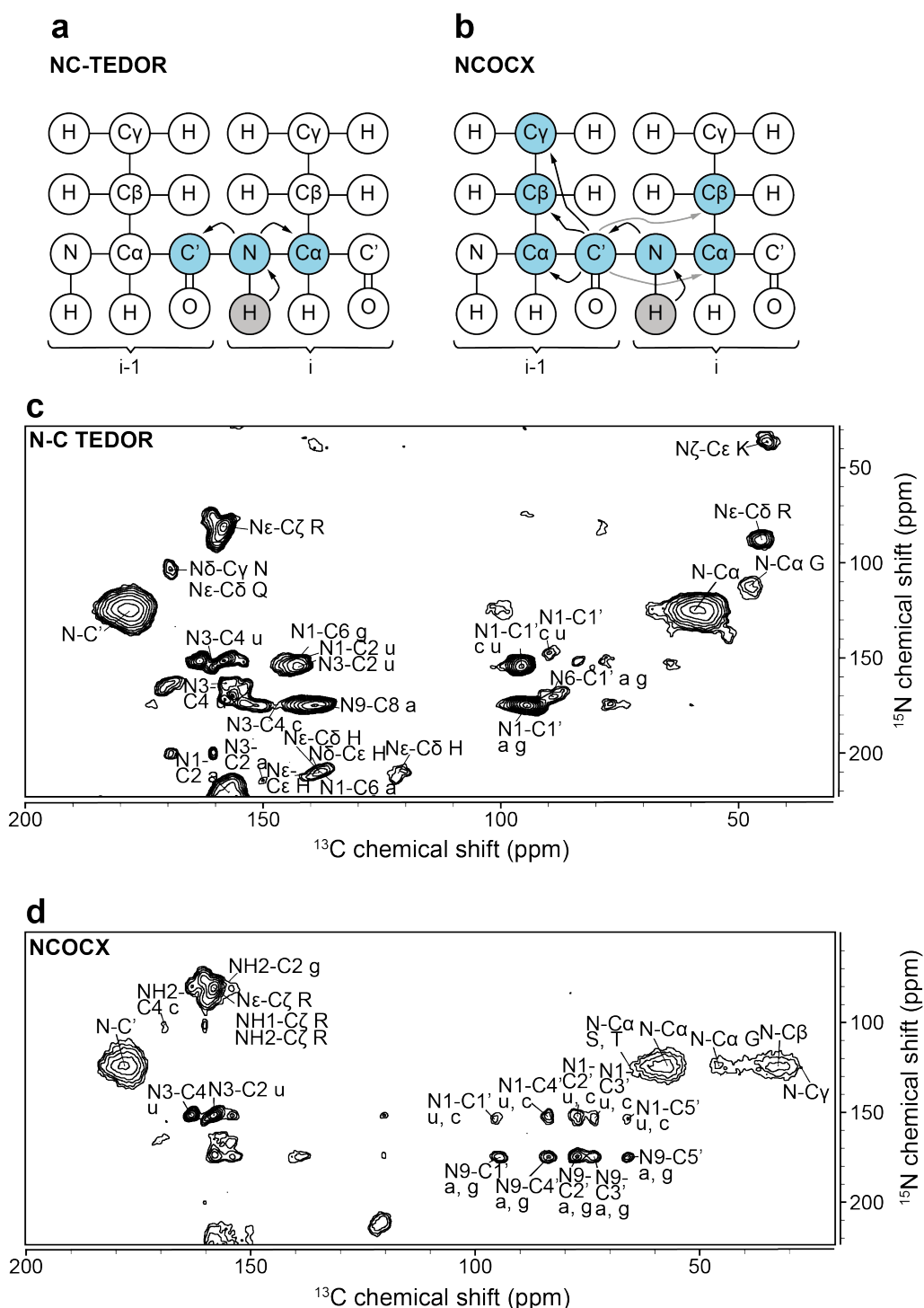
Figure 3.24: ***NC-TEDOR and NCOCX correlation experiment on*** ^{15}N, ^{13}C ***labeled U78SecM. a)*** *Schematic magnetization transfer in NC-TEDOR experiment.* ***b)*** *Schematic magnetization transfer in NCOCX experiment. Depending on the mixing time either the correlation between N_i to CO_{i-1} and carbon atoms of residue i are detected (black arrows) or with longer mixing times correlation between nearby carbons are detected (grey arrows). Blue atoms are detected, grey atoms are used for sensitivity enhancement.* ***c)*** *2D N-C TEDOR and* ***d)*** *2D NCOCX experiment on ^{15}N-^{13}C uniformly labeled U78SecM.Nucleotides are labeled wit lowercase and amino acids with uppercase letters. Parameters for TEDOR experiment: NS=128; TD= 1024, 192; MAS frequency = 8 kHz; 100 K. Parameters for NCOCX experiment: NS=128, TD= 2048, 192; MAS frequency =8 kHz; 100K.*

NCOCX experiment as a 1D version on a selectively ^{13}C, ^{15}N cysteine and ^{15}N proline U32SecM sample. Using this labeling method the carbon chemical shifts of only Cys22 should be detectable. The spectrum of this experiment is shown in Figure 3.25. Two different mixing times (50 ms and 100 ms) were used and one signal corresponding to the C' chemical shift of cysteine can be identified. However, the other carbon chemical shifts of cysteine cannot be assigned, since the signals hardly differ from noise. The longer mixing time increases the signal intensity a bit, but still it is not possible to unambiguously assign a signal to the Cα and Cβ chemical shift of cysteine, since more than the expected number of signals is observed. There might be two oxidized Cβ chemical shift at 43.37 ppm and 34.79 ppm and reduced Cβ chemical shifts at 32.14 ppm, 30.15 ppm and 28.63 ppm. This would mean, that the NC does not obtain a homogenous conformation inside the ribosomal exit tunnel. Also the two oxidized Cβ chemical shifts would suggest that either different disulfide bonds are formed or that the cysteine residue is in any other way chemically modified.
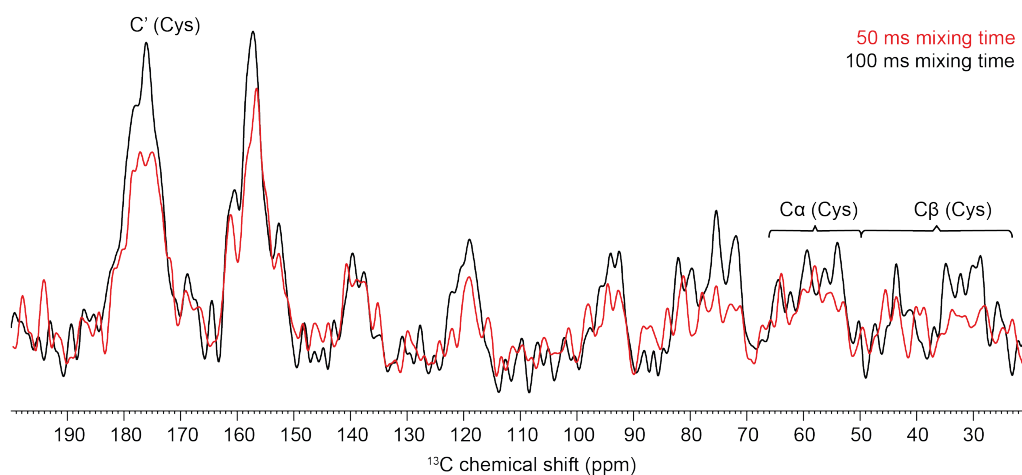
Figure 3.25: *1D NCOCX of ^{13}C, ^{15}N cysteine and ^{15}N proline labeled U32SecM. 50 ms mixing time (red), 100 ms mixing time (black). Parameters for acquisition: NS=8192, TD=2048, MAS frequency =8 kHz, 100K.*

The 3-4 nmol sample in the rotor are not sufficient enough for a NCOCX experiment even enhanced by DNP. Since the DQ-SQ experiments were sufficient for identifying the Cα and Cβ chemical shifts of cysteine residues, this experiment was used on different RNC constructs and secondary structure elements and oxidation ratios were compared. First, however, we tested if the detected signals are indeed NC signals and not isotopically labeled ribosomal proteins. In addition, different cysteine concentration in the expression medium used for the selective cysteine labeling were tested in respect to the growth rate of *E. coli.*

3.2.6   Optimization of the cysteine concentration in selectively labeled media

In order to selectively detect only NC cysteine residues, the expression of the RNCs was performed in a kind of full medium containing all amino acids, nucleosides or bases (appendix). Since the BL21(De3)Δtig cells are grown in LB medium to an $OD_{600\ nm}$=1 and the expression is carried out in selectively labeled medium for only one hour as described by Rutkowska *et al.* [217], the cysteine concentration used in our initial protocol could be reduced to save costs. To investigate suitable cysteine concentrations within the medium, BL21(De3)Δtig were grown in selectively labeled media containing different concentrations of cysteine and the growth of the cells was monitored. The concentration ranged from 0 mg/L to 50 mg/L, the latter being the concentration used in the initial protocol. Figure 3.26 displays the cell growth of BL21(De3)Δtig cells transformed with U78SecM on pET15b in selective media with different concentrations of cysteine. As can be seen from this figure, there is no influence of the cell growth of any cysteine concentration. All concentrations of cysteine seem suitable for the cell viability, since cysteine can also be synthesized from serine if not enough cysteine is in the medium [263]. Therefore, the cells are only transferred to the selectively labeled medium directly before induction, which is similar to previous published methods for selectively labeled FGFR3 kinase domain [258, 221], which used an excess of unlabeled amino acids and the addition of labeled amino acid only prior to induction. Although the cells grow in medium lacking one amino acids, the pathway for synthesis of this amino acid should be already turned on and therefore reduce the labeling efficiency. Both studies, however, showed incorporation of $> 90\%$ of the ^{15}N, ^{13}C selectively labeled amino acid. Neither of all mentioned protocols above used deficient cells to prevent formation of the amino acids by another pathway. Due to the short expression time of only one hour and the good viability of the cells in all concentrations, the cysteine concentration was lowered slightly to 30 mg/L, which still leads to a reduction of 40% of the cost for this selective labeled amino acids. Of course, this simple cell growth experiment is not comparable to a detailed investigation of the isotope incorporation by NMR spectroscopy. This however, is beyond the scope of this thesis.
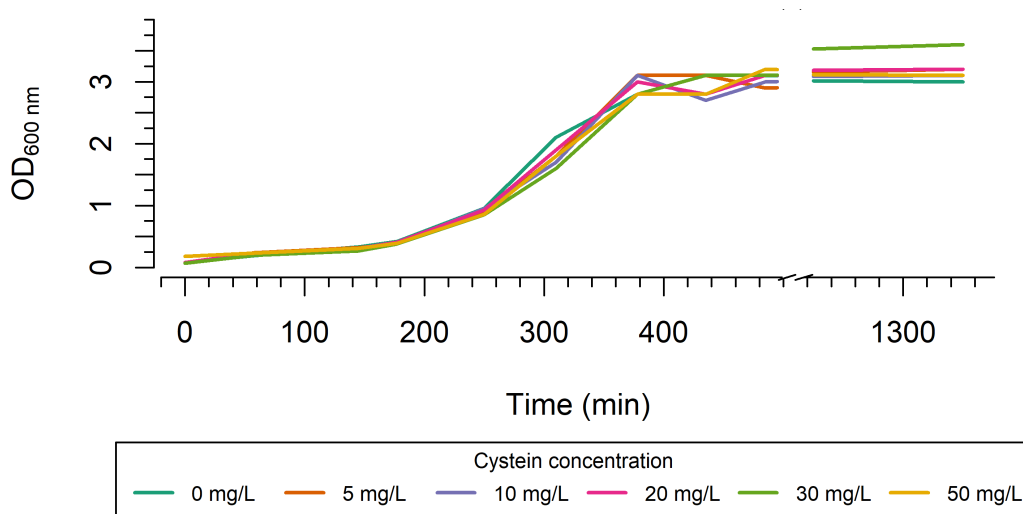
Figure 3.26: *Optimization of cysteine concentration. Growth curve of E. coli BL21(De3)Δ tig cells transformed with U78SecM on pET15b in selective medium containing different cysteine concentrations. The OD_{600nm} was measured after different time points. Cells were grown at 37°C shaking with 160 rpm.*

3.2.7 Control of ^{13}C scrambling into the ribosome

To confirm that only NC cysteines and not ribosomal proteins are isotopically labeled, ribosomes from the flow-through fraction of the Ni-NTA affinity chromatography were used as control. These flow-through (FT) ribosomes were derived from selectively ^{13}C cysteine labeled medium as the selectively labeled RNCs. After Ni-NTA affinity chromatography, the FT ribosomes were pelleted by ultracentrifugation and further purified using sucrose gradient ultracentrifugation as described in section 2.4.6. The first ultracentrifugation step was performed to concentrate the ribosomes and to remove all soluble non-ribosomal proteins from the sample. In the sucrose gradient ultracentrifugation step the integrity of the ribosomes was confirmed. No signal for the purified FT ribosomes was detected, hence no RNCs were present in this sample.

Figure 3.27 shows the ^{13}C-^{13}C DQ-SQ 2D spectrum of FT ribosomes overlaid with the U32SecM sample. The intensity of both spectra were calibrated according to the ribose signal intensity. No Cα and Cβ chemical shifts of cysteine residues can be detected as seen in the 1D ^{13}C slice, concluding that no detectable scrambling occurs to ribosomal proteins.
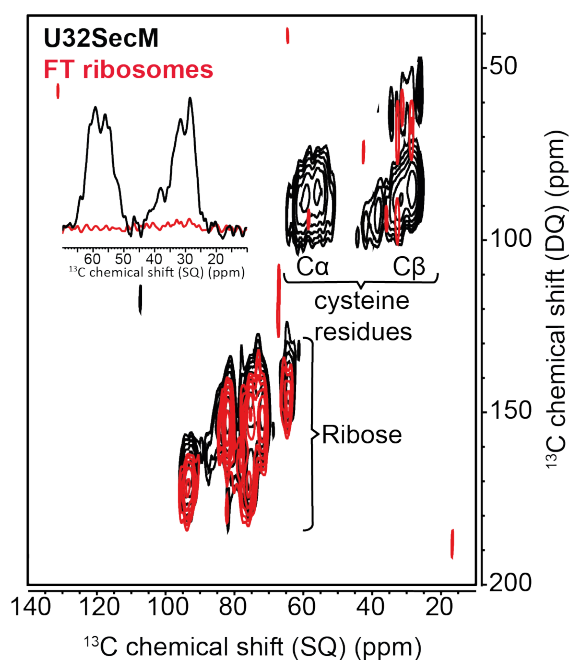
Figure 3.27: *Analysis of scrambling into ribosomal proteins. Overlay of POST-C7 ^{13}C-^{13}C DQ-SQ 2D spectra of the ^{13}C, ^{15}N cysteine labeled nascent chain U32SecM (black) and flow-through ribosomes grown in selectively labeled medium (red). 1D ^{13} slice of Cα-Cβ cross peak region of cysteine residues of flow-through ribosomes (red) and U32SecM (black).*

3.2.8   Flexible–meccano simulations on U32SecM

To test whether the conformational space of the ribosomal exit tunnel is sufficiently large for disulfide bond formation and multiple conformations of the NC, and thereby confirming the initial hypothesis, the conformational space of the NC inside the tunnel was simulated using flexible-meccano [14]. This method relies on amino-acid-specific (φ/ψ) sampling in coil regions of protein structure. The flexible-meccano simulations were performed by Prof. Dr. Martin Blackledge. The dimensions of the ribosomal exit tunnel were taken from our obtained cryo-EM structure (PDB: 6YS3), and the NC was sequentially built from Ser151 of SecM by adding 26 amino acids of the U32SecM construct to the N-terminal end. Hence, the simulation of the NC include the amino acids Asp8-Cys32 of GBC and Phe150-Ser151 or Phe150-Trp155 of SecM. Indeed, the ribosomal exit tunnel provides sufficient space for different disulfide bridges to form including a disulfide bond between Cys22 and Cys32 inside the ribosomal exit tunnel close to the constriction site (Figure 3.28). Figure 3.28a presents the ensemble of different U32SecM NCs and Figure 3.28b-f the different disulfide bridges in U32SecM that can be obtained. Cβ-Cβ distances below 4.5 Å were considered as potential disulfide bonds. Further, we were able to observe that the ribosomal exit tunnel modulates the sampling of the NC, with Cys32 and Phe150 showing a higher propensity to sample an α-helical conformation inside the ribosomal exit tunnel of *E. coli* than in solution (Figure 3.29). While simulation of the U32SecM NC

inside the ribosomal exit tunnel of *H. sapiens* (Figure 3.29) revealed differences in the sampling of Arg31, Cys32 and Phe150 compared to the *E. coli* ribosome. Although all disulfide bridges of U32SecM could also be formed in the eukaryotic ribosome (Figure 3.30). Hence, disulfide formation within eukaryotic ribosomes would require further investigation, since different structures might be obtained inside the ribosomal exit tunnel. Since disulfide bond formation might be possible in the ribosomal exit tunnel and the cysteine chemical shift of the NC alone in the ribosome could be identified in ^{13}C-^{13}C DQ-SQ experiments, we tested whether the oxidation level changes with elongating chain lengths.
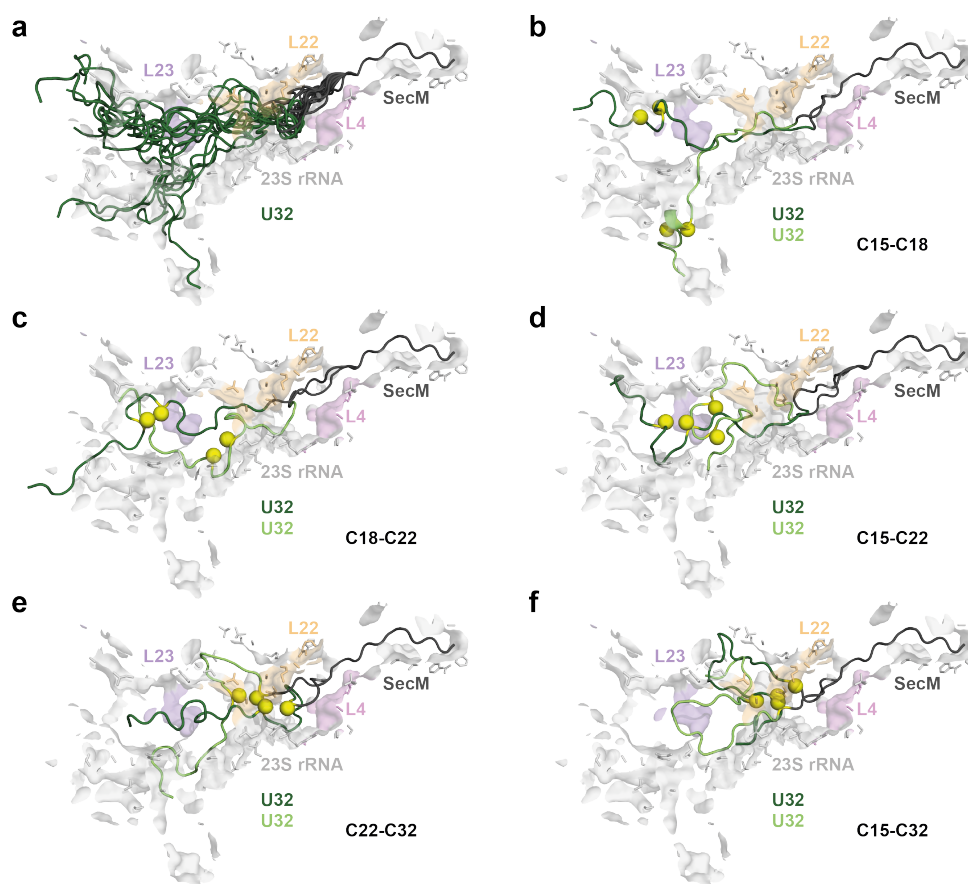
Figure 3.28: *Simulation of U32SecM within the ribosomal exit tunnel of E. coli using flexible-meccano. a) Ensemble of structures within the tunnel. U32SecM chains with Cys15-Cy18 (b), Cys18-Cys22 (c), Cys15-Cys22 (d), Cys22-Cys32 (e) and Cys15-Cys32 (f) in close proximity for possible disulfide bridge. Cysteine residues are highlighted with a yellow ball. SecM in dark grey, U32 sequence in green. Ribosomal proteins L4 (light pink), L22 (light orange) and L23 (light violet) and 23S rRNA (light grey) shape the ribosomal tunnel. The simulation included residues from Asp8 in GBC to Trp155 in SecM. The tunnel dimensions of PDB 6YSE were used.*
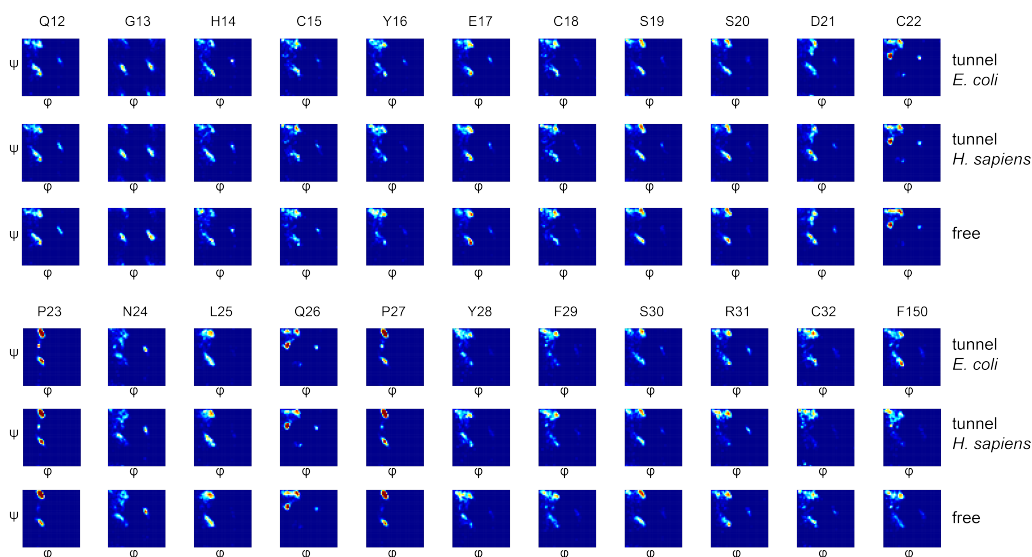
Figure 3.29: *Ramachandran plots of the simulated U32SecM NC in the ribosomal exit tunnel and free in solution. The tunnel dimensions of E. coli (PDB: 6YS3) and H. sapiens (PDB: 4UG0) were used for the simulation of the NC using flexible-meccano [14]. High populated states of the NC in red and low populated in blue.*

3.2.9 Solid–state NMR on ^{13}C, ^{15}N cysteine labeled GBC RNCs with different chain lengths

As demonstrated in the previous section, we were able to detect selectively the ^{13}C NMR signals of NC cysteine residues in RNCs. The ribosomal exit tunnel seem suitable to accommodate different conformations of the short NC and different disulfide bonds. We then monitored the oxidation ratio of the further elongating NC and determined the secondary structure of the NC cysteine residues.

The different RNC constructs and U32SecM mutants are shown in Figure 3.10. All samples were treated identically and the ^{13}C-^{13}C DQ-SQ 2D spectra of the these samples are shown in Figure 3.31. The Cβ chemical shift of the individual signals were assigned directly on the resolved signals and the corresponding Cα chemical shift was calculated indirectly from the DQ (Cα + Cβ) and SQ (Cβ) chemical shift. Vice versa, the DQ and SQ chemical shift of the Cα signal was used to calculate the Cβ chemical shift as in U32SecM C18 C22 for Cα7 and Cα8. However, since the Cα chemical shifts are less resolved, signals calculated using the Cα chemical shift contain a higher error. Cβ chemical shifts above 35 ppm are attributed to oxidized cysteine residues. Reduced cysteine residues show Cβ chemical shifts less than 32 ppm and in the overlapping region (35-32 ppm) the oxidation state of cysteine residues are ambiguous. First, in all samples oxidized species were detected. For the U32SecM C18 C22 mutant (Fig 3.31a) and for the U32SecM C15 C18 mutant (Figure 3.31b), this might suggest that a Cys18-Cys22 or Cys15-Cys18 disulfide bond was formed. A disulfide bond linking Cys15
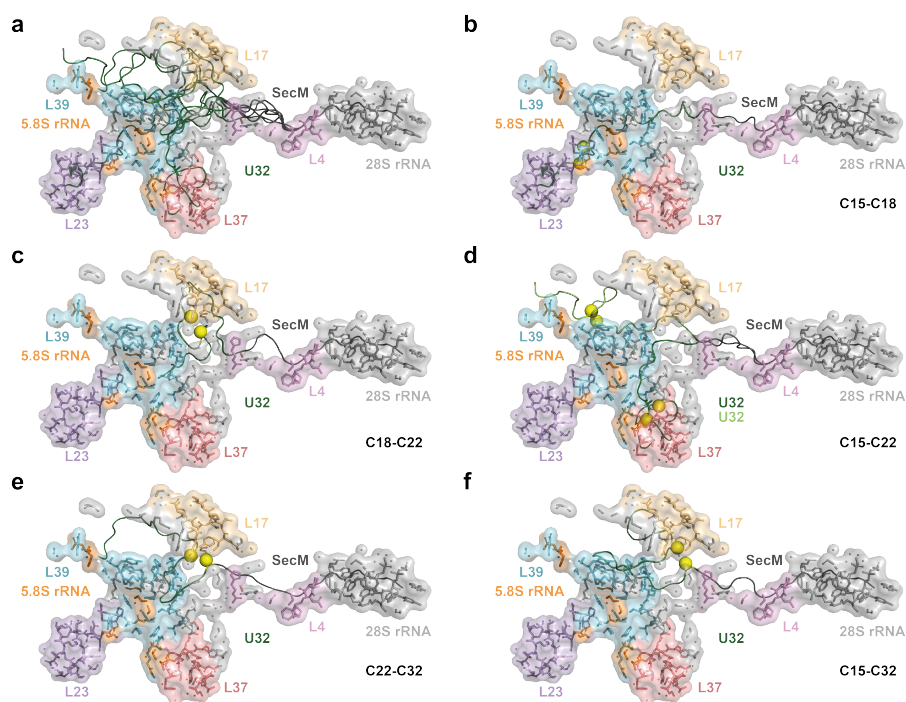
Figure 3.30: *Simulation of U32SecM within the ribosomal exit tunnel of H. sapiens using flexible-meccano. a) Ensemble of structures within the tunnel. U32SecM chains with Cys15-Cy18 (**b**), Cys18-Cys22 (**c**), Cys15-Cys22 (**d**), Cys22-Cys32 (**e**) and Cys15-Cys32 (**f**) in close proximity for possible disulfide bridge. Cysteine residues are highlighted with a yellow ball. U32 in green and lightgreen. Ribosomal proteins L4 (light pink), L17 (light orange) and L23 (light violet), L37 (salmon), L39 (cyan), 5.8S rRNA (orange) and 28S rRNA (light grey) shape the ribosomal tunnel. The ribosomal protein L17 of H. sapiens is the structural equivalent to L22 in E. coli ribosomes. The simulation included residues from Asp8 in GBC to Trp155 in SecM. The tunnel dimensions of PDB 4UG0 were used.*

and Cys18 is not compatible with the conformation of the full-length protein, in which both side chains face different directions and the distance between the Cβ carbons is above 10 Å. Although these cysteines would be part of the common CXXC motif, as found in oxidoreductases such as thioredoxin. A disulfide bond between Cys18 and Cys22 was identified in a previous X-ray structure and both Cβ carbons were only 4.8 Å (PDB: 4GCR) apart[177]. However, also the single cysteine mutant U32SecM C18 shows oxidized signals. Since no intramolecular disulfide bond can form in this mutated NC, the oxidized cysteine is involved solely in intermolecular disulfide bond formation.

These chemical modifications of cysteine include S-glutathionylation and S-nitrosylation. S-glutathionylation, a post-translational modification, includes the addition of the tripeptide glutathione via an intermolecular disulfide bond. Since the concentration of reduced glutathione is around 5 mM in the *E. coli* cytosol and the ratio between reduced glutathione to glutathione disulfide (GSH/GSSG) is in a range between 50/1 and 200/1 [173], the addition of glutathione to Cys18

seems possible in cells. Besides, S-glutathionylation of full-length GBC was indeed detected previously [148, 277]. Liang and Pelletier [148] detected one glutathione bound to one GBC molecule and argued that it is most likely bound to Cys15, due to the high static accessibility of Cys15 [159] (Table 1.3). Cys18 is usually not accessible to glutathione in the full-length GBC (Table 1.3). In the ribosomal exit tunnel, the short NC is not likely to fold in a similar way as the full-length protein, therefore Cys18 could be accessible to glutathione addition in this specific environment. Here, S-glutathionylation of U32SecM C18 indicates that native-like β-strands do not form inside the ribosomal tunnel, which would otherwise bury the side chain of Cys18. While Liang and Pelletier [148] used oxidized glutathione (GSSG) for the detection of mixed disulfides in GBC, it was also shown that reduced glutathione can bind to GBC [277, 239] when added in 26 molar excess and incubated for 1-2 hrs at room temperature. Yu *et al.* [277] demonstrated that two cysteine residues might be involved in the formation of mixed-disulfides with reduced glutathione. In addition, the authors suggested that the cluster of Cys18, Cys22 and Cys78 might be involved in redox reactions with glutathione and the mixed-disulfides as intermediates.

S-nitrosylation can be mediated by S-nitrosoglutathione (GSNO) and is also considered as a post-translational modification which might be an established response to oxidative/nitrosative stress [76]. Since the media and buffers did not contain any nitric oxide, nitrogen dioxide, GSNO or any glutathione, chemical modifications of cysteine residues must take place within the cell. In general, these modifications are reduced after release of the ribosome within the highly reductive environment of the *E. coli* cytosol. Thus heterologously expressed proteins can typically be produced only in reduced form in *E. coli*, although exceptions have been reported [8, 230, 111, 10]. In addition, S-glutathionylated proteins are also reduced by Grx[15]. GSH might be small enough to enter the ribosomal tunnel and thus bind to the NC. Grx with a molecular weight between 9-13 kDa, however cannot enter the ribosomal exit tunnel and reduce the mixed disulfide with GSH.

The ^{13}C-^{13}C DQ-SQ spectrum of the different RNCs showed not only oxidized cysteine signals within each constructs, but also the number of cysteine signals exceed the number of cysteine residues in the NC. This could be attributed plausibly to the distinct intermolecular single-site modifications, the complex intramolecular disulfide bonds, and multiple sidechain/backbone conformations within the ribsosomal exit tunnel as modeled with flexible-meccano (Figure 3.28) and also demonstrated for domain 5 and 6 constructs of ABP-120 filamin protein [114]. The U32SecM and U43SecM constructs exhibit multiple Cβ signals (Figure 3.31d and 3.31e), while U78SecM (Figure 3.31f) shows the same number of signals as cysteine residues in this construct. U78SecM may adopt a more defined structure and/or carry less chemical modifications, since some of the cysteine residues might be already buried within the protein interior.

Plotting the Cα chemical shift versus the Cβ chemical shift gives a hint about the obtained secondary structure of the cysteine residues [235, 261] (Figure 3.32). The cysteine residues within the full-length GBC are mostly located within β-sheets, only Cys22 and Cys32 are located within a random-coil structure (Figure 3.11). For all RNC constructs, except U32SecM, all signals obtain a Cα-Cβ chemical shift in the β-sheet or random-coil region. U32SecM has also one signal within the region of reduced α-helix (Figure 3.32d). However, the Cβ chemical shift of this particular signal was calculated using the DQ and SQ chemical shift of the Cα signal. Hence, this signal has a higher error. Again, while U32SecM and U43SecM show a lot of signals, U78SecM shows only six signals as the number cysteine residues in this construct. These signals are also located within β-sheet and coil structure as the full-length GBC. However, four residues are oxidized, whereas the full-length GBC is fully reduced [22].

The Cβ chemical shifts of S-gluathionylated or S-nitrosylated cysteine residues cannot be distinguished from cysteine residues involved in disulfide bonds (Figure 3.33a). This complexity could indeed partially explain the strikingly high heterogeneity observed in our solid state NMR spectra. Comparing these chemical shift dristribution plots of the RNCs to the plots of GSH, GSSG and GSNO shows, that the signals of the cysteine are also located within the β-sheet or random-coil region (Figure 3.33b). Hence, it is not possible to distinguish between inter- and intramolecular disulfide bonds using this solid-state NMR method.

In principle, the disulfide-bond linked cysteine residues can be identified by interresidue ^{13}Cβ-^{13}Cβ contact or through ^{77}Se-^{77}Se correlation of ^{77}Se (I=1/2) substituted cysteine residues. The first approach was not successfull, possibly due to the relatively long Cβ-Cβ distance (4.4 Å) in our case and/or the similar Cβ chemical shifts. The diselenide bond within a 37-residue spider protein (ϰ-ACTX-Hv1c) was observed through one bond ^{77}Se-^{77}Se scalar coupling in a two-dimensional ^{77}Se-^{77}Se COSY experiment in liquid-state NMR [174]. However, in our case, the amount of sample might not be sufficient for this experiment. This strategy also needs an auxotrophic *E. coli* strain for protein expression, which often compromise the protein yield. An alternative approach to detect the connected cysteine residues is mass spectrometry. Here, the RNCs were also subjected to LC-MS/MS analysis (collaboration with Jakob Meier-Credo and Julian Langer).

In addition, using solid-state NMR, we also studied if the oxidation ratio between the different constructs changes with elongating chain length (Figure 3.34). The oxidation ratio was determined by integrating the signal intensity over different spectral regions characteristic to the oxidized, reduced and ambiguous states and then deriving the ratio by dividing these values by the total integral over all cysteine signals. The error bars reflect slightly different integration regions, as described in the method section. The error of the different integration areas is negligible with a maximum error of 1.2% for U32SecM C18 (Table 3.3). The

U32SecM mutants with one and two cysteine residues show a very similar oxidation ratio of 16-19%. The oxidation level increases to 25% for U32SecM with four cysteine residues and further to 38% for U43SecM carrying five cysteine residues. U78SecM with six cysteine residues shows an oxidation level of 34% similar to that of U43SecM. While 25% oxidation ratio in U32SecM corresponds to in average one oxidized cysteine per RNC, two cysteine residues would be oxidized in the U43SecM and U78SecM, which could possibly be a disulfide bond. However, also the cysteine modifications with glutathione and a nitroso group are count into the oxidation ratio. To conclude a more general trend with the elongating chain length further DNP solid-state NMR studies on more constructs with varying lengths and cysteine numbers are required, which is not compatible with the current experimental thoroughput. Between U43SecM and U78SecM the oxidation ratio decreases slightly, further experiments are needed to verify, if the oxidation ratio would further decrease with longer chain lengths.

Table 3.3: *Relative integration ratio (Rel. int. ratio) and standard error of integration area variation of different RNCs.*

Construct name	Integration region	Rel. int. ratio (mean)	Standard error
U32SecM C18	oxidized	0.190	0.012
U32SecM C18	reduced	0.599	0.011
U32SecM C18	overlapping	0.211	0.002
U32SecM C18 C22	oxidized	0.163	0.009
U32SecM C18 C22	reduced	0.636	0.005
U32SecM C18 C22	overlapping	0.201	0.004
U32SecM C15 C18	oxidized	0.186	0.005
U32SecM C15 C18	reduced	0.606	0.007
U32SecM C15 C18	overlapping	0.208	0.002
U32SecM	oxidized	0.253	0.002
U32SecM	reduced	0.554	0.001
U32SecM	overlapping	0.193	0.002
U43SecM	oxidized	0.382	0.003
U43SecM	reduced	0.434	0.002
U43SecM	overlapping	0.184	0.001
U78SecM	oxidized	0.343	0.004
U78SecM	reduced	0.486	0.003
U78SecM	overlapping	0.172	0.002

To study if disulfide bonds can form inside the ribosomal exit tunnel as suggested by the flexible-meccano simulations (Figure 3.28), a U32SecM construct, carrying four cysteine residues, was further studied by solid-state NMR and cryo-EM. An

oxidation ratio of 63% (Figure 3.35 and Table 3.4) was observed, corresponding
to at least two oxidized cysteine residues. This is in line with the formation of a
disulfide bond. However, the single-site chemical modifications of cysteines, such
as S-glutathionylation and S-nitrosylation, could also contribute to this observed
oxidation ratio. The ^{13}C-^{13}C DQ-SQ 2D spectrum is represented in Figure 3.35a
and the corresponding Cα-Cβ chemical shift plot in Figure 3.35b. Due to signal
overlapping, the Cβ chemical shifts of signals 7-10 were indirectly determined from
the DQ and SQ chemical shift of Cα and thereby show a lower precision. Some
signals are located in the Cα-Cβ region of oxidized α-helix conformation (Figure
3.35). This would implicate that non-native structures are obtained within the
ribosomal exit tunnel, since all cysteine residues are located within β-sheets and
random-coil in the folded full-length GBC. Non-native compact structures were
also found in the N^5-glutamine metyltransferase HemK as studied by FRET and
PET [171]. As mentioned above, to further map the disulfide bond formation or
other modifications in the different constructs, the RNCs were analyzed by LC-
MS/MS. In addition, we also studied the binding of glutathione to U32 peptides
using liquid-state NMR.

Table 3.4: *Relative integration ratio (Rel. int. ratio) and standard error of integration area variation of U32SecM cryo-EM sample.*

Construct name	Integration region	Rel. int. ratio (mean)	Standard error
U32SecM for cryo-EM	oxidized	0.629	0.005
U32SecM for cryo-EM	reduced	0.190	0.009
U32SecM for cryo-EM	overlapping	0.181	0.004
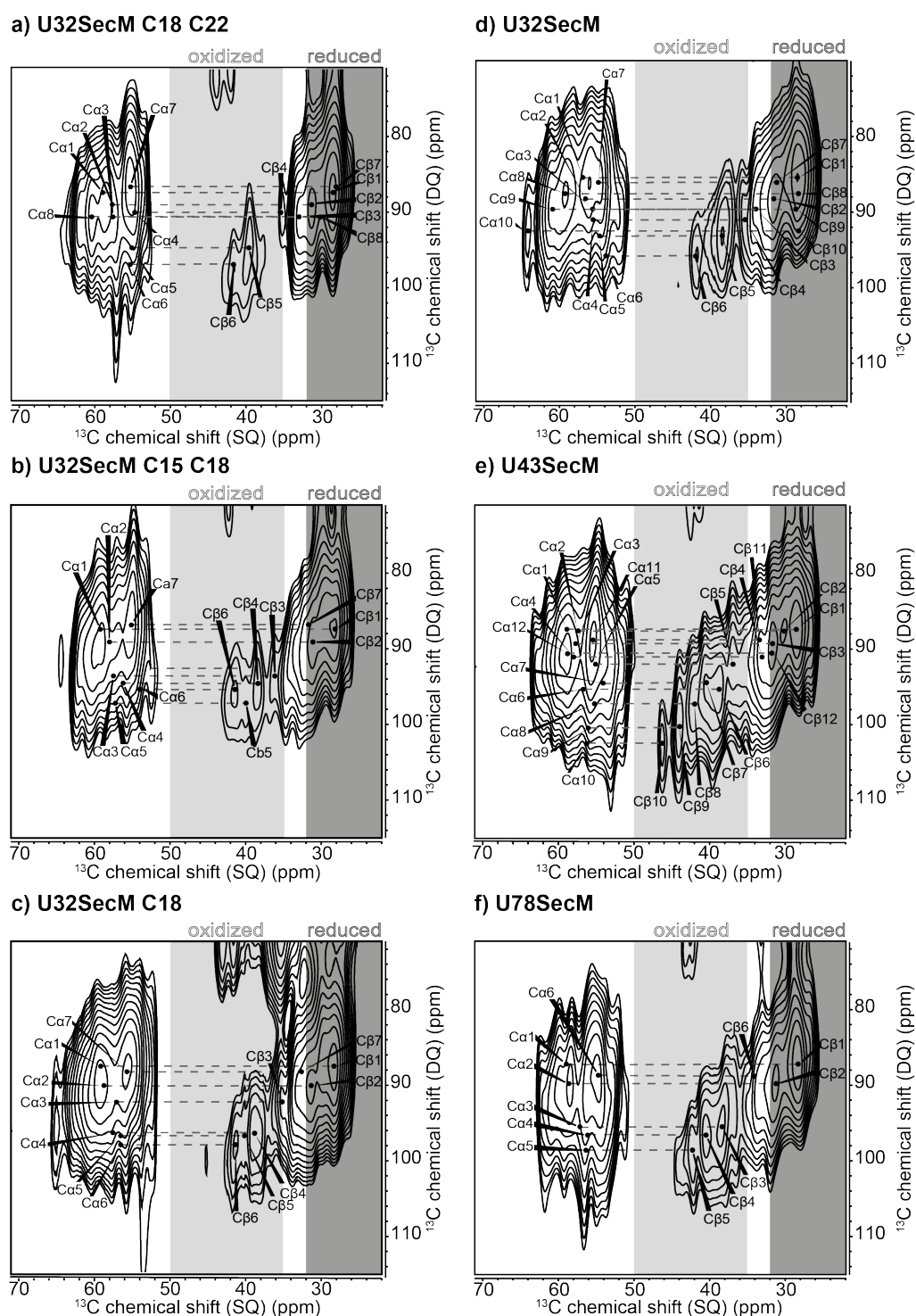
Figure 3.31: *^{13}C-^{13}C 2D DQ-SQ spectrum of all RNC constructs. Cysteine Cα-Cβ cross-peak regions of the ^{13}C-^{13}C DQ-SQ spectrum of ^{13}C, ^{15}N cysteine labeled RNCs.* **a)** *U32SecM C18 C22,* **b)** *U32SecM C15 C18,* **c)** *U32SecM C18,* **d)** *U32SecM,* **e)** *U43SecM and* **f)** *U78SecM (numbering scheme as given in Figure 3.10). The cross-peak region for oxidized cysteine residues is highlighted in light grey, and for reduced cysteine residues in dark grey.*
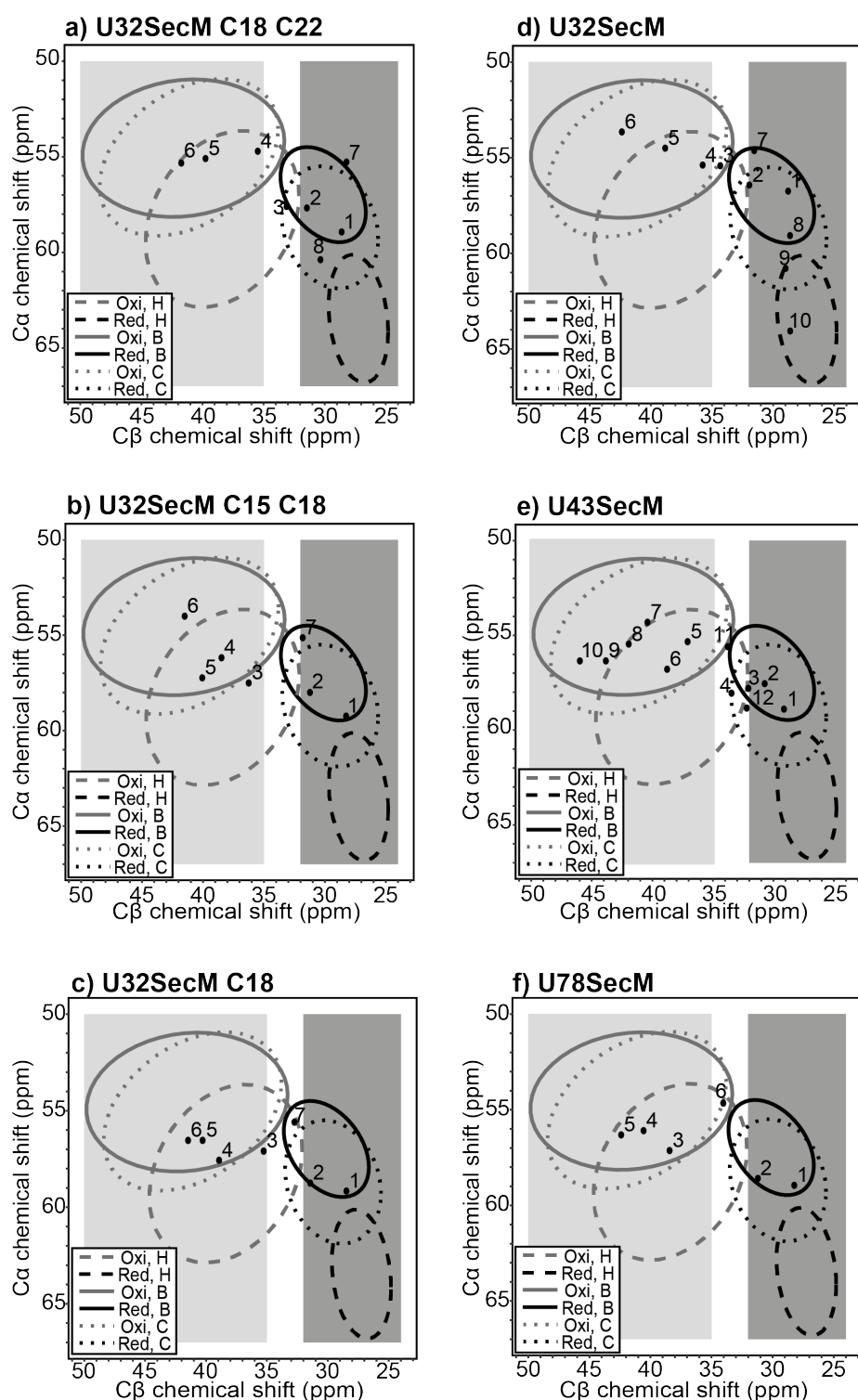
Figure 3.32: *Structural analysis of selectively cysteine labeled RNCs.* Cα and Cβ chemical shifts from RNC constructs compared to values derived from a chemical shift database [261]. **a)** U32SecM C18 C22, **b)** U32SecM C15 C18, **c)** U32SecM C18, **d)** U32SecM, **e)** U43SecM and **f)** U78SecM (numbering scheme as given in Figure 3.10). Ellipses show 90% of the corresponding chemical shifts. Secondary structures: H = helix, B = β-strand and C =coil. The cross-peak region for oxidized cysteine residues is highlighted in light grey, and for reduced cysteine in dark grey.
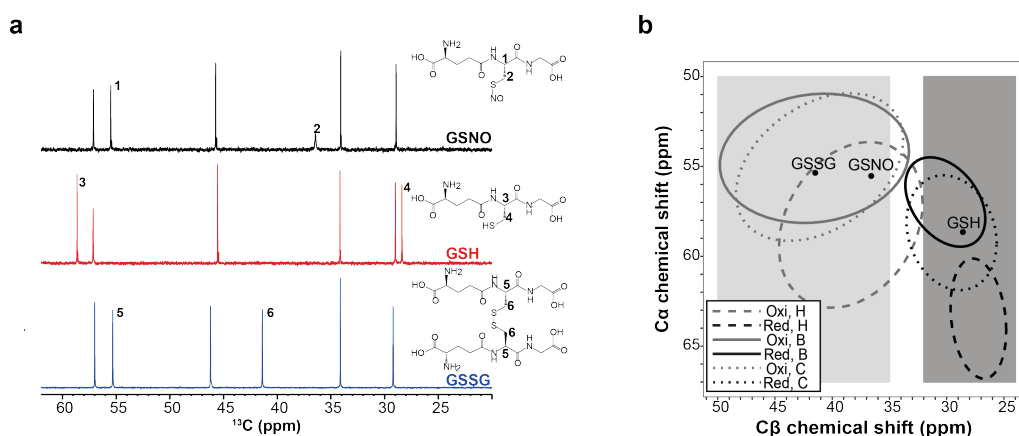
Figure 3.33: *Chemical shifts of cysteine modifications.* *a)* 1D ^{13}C NMR spectrum of S-nitroso glutathione (black), glutathione (GSH) (red) and glutathione disulfide (GSSG) (blue). Cα and Cβ chemical shifts of cysteine in odd numbers and even numbers, respectively. (*1*=55.489 ppm, *2*= 36.512 ppm; *3*=58.618 ppm, *4*=28.492 ppm; *5*=55.323 ppm, *6*=41.390 ppm). *b)* Cα and Cβ chemical shifts of GSSG, GSNO and GSH compared to values derived from a chemical shift database [261]. Ellipses show 90% of the corresponding chemical shifts. Secondary structures: H = helix, B = β-strand and C=coil. The cross-peak region for oxidized cysteine residues is highlighted in light grey and for reduced cysteine in dark grey.
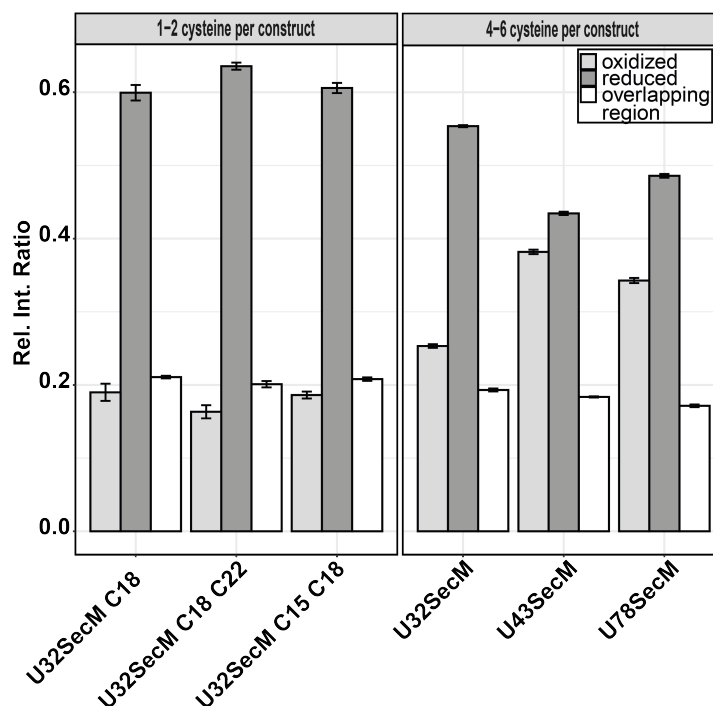
Figure 3.34: *Oxidation ratios of GBC RNCs.* Relative integration ratio (Rel. int. ratio) of oxidized, reduced or overlapping regions, the last of which does not allow for differentiation, of all RNC spectra. Constructs contain 1-2 cysteine residues (left), 4-6 cysteine residues (right). Signal intensity of oxidized cysteine residues (>35 ppm), reduced cysteine residues (<32 ppm) and the overlapping region (35–32 ppm) was divided by the total signal intensity of Cβ chemical shifts. Error bars were calculated using different integration regions (2.6.3).
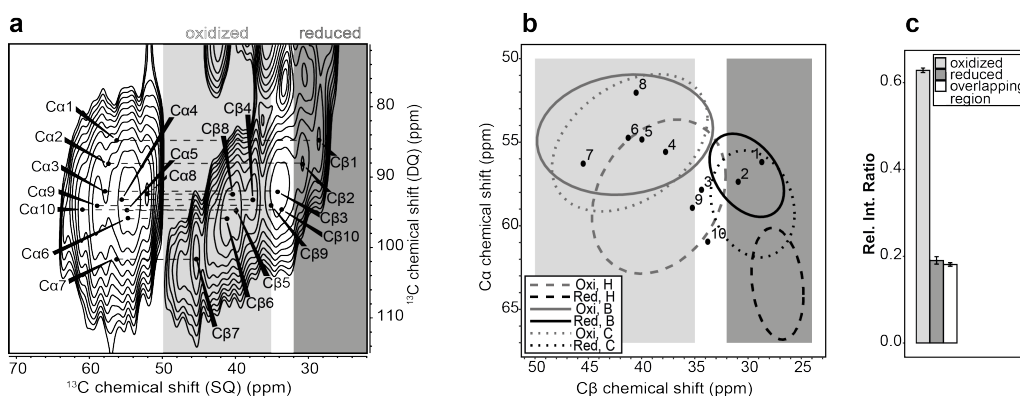
Figure 3.35: *Analysis of ^{13}C-^{13}C DQ-SQ solid-state NMR spectrum of ^{13}C, ^{15}N cysteine labeled U32SecM cryo-EM sample. **a)** Cysteine Cα-Cβ cross peak regions of the ^{13}C-^{13}C DQ-SQ 2D spectrum of ^{13}C, ^{15}N cysteine labeled U32SecM. **b)** Cα and Cβ chemical shifts from U32SecM compared to values derived from a chemical shift database [261]. Ellipses contain 90% of the corresponding chemical shifts within each group [261] (H = helix, B = β-strand, C = coil). The spectral region for oxidized Cβ cysteines in light grey, and for reduced cysteine in dark grey. **c)** Relative integration ratio (Rel. int. ratio) of oxidized (light grey) or reduced residues (dark grey) or the overlapping region (white), which does not allow for differentiation, of U32SecM cryo-EM sample. Signal intensity of oxidized cysteine residues (>35 ppm), reduced cysteine residues (<32 ppm) and the overlapping region (35-32 ppm) was divided by the total signal intensity of Cβ chemical shifts. Error bars were calculated using different integration regions (2.6.3).*

3.2.10   Investigation of glutathione binding to U32 peptides

To verify whether glutathione can bind to GBC fragments and to assign the chemical shifts of reduced and oxidized GBC fragments, U32 peptides containing two cysteine residues per construct were synthesized using solid state peptide synthesis (SPPS). One cysteine per construct was ^{13}C, ^{15}N labeled, while the other one was unlabeled (Figure 3.36). Another construct containing cysteine Cys15 and Cys18 with Cys15 ^{13}C, ^{15}N labeled was also synthesized, but could not be purified as monomer.
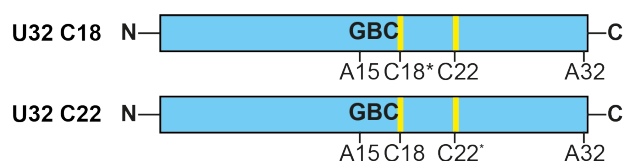
Figure 3.36: *Labeling scheme of U32 peptides synthesized by solid phase peptide synthesis. Each construct contains two cysteine residues (yellow line) with one cysteine ^{13}C, ^{15}N labeled marked with an asterisk (*).*

The chemical shifts of U32 C18 and U32 C22 peptides were detected in a 2D ^{1}H-^{13}C HSQC experiment (Figure 3.37a and 3.37b (black spectrum)) and the Cα chemical shift was plotted as a function of the Cβ chemical shift (Figure 3.37c). The signals of Cys18 and Cys22 are located within the reduced β-sheet or random-coil region and the 1D ^{1}H NMR spectrum of these peptides showed low chemical shift dispersion and line broadening of the amide protons, both indications of an unfolded conformation (Figure 5.3 in appendix). Comparing the Cα-Cβ plot of these peptides to the U32SecM RNC constructs, shows that a signal at the location of C18 can be detected in all RNC spectrums (signal 1 in Figure 3.32a-c, e-f and signal 8 in Figure 3.32d). The signal of Cys22 in the U32SecM C18 C22 constructs (Figure 3.32a) cannot be identified, and in the U32SecM construct it likely corresponds to signal 1. However, since the NC is located inside the ribosome exit tunnel and the U32 peptide is free in solution, comparing the chemical shifts is problematic.

Oxidation of the cysteine residues was achieved by addition of GSSG/GHS to the peptides. The chemical shifts of the oxidized cysteines were determined (Figure 3.37a and 3.37b (red spectrum)) and the Cα chemical shift was plotted as a function of the Cβ chemical shift (Figure 3.37c). The signals of both cysteines are also located within the β-strand or random-coil region. Oxidized Cys18 signal differs from the oxidized signals found in the U32SecM RNC constructs (Figure 3.32a-d). In the U43SecM construct a signal corresponding to the Cα-Cβ chemical shift of Cys18 in U32 can be found (signal 9). A signal at the location of the oxidized Cys22 is also not visible in the U32SecM RNC constructs. It is likely that signal 5 in U32SecM C18 C22 (Figure 3.32a) and U32SecM (Figure 3.32d) could correspond to Cys22, however, since the peptides have different environments the assignment of the RNC signals using the U32 peptides is not possible.
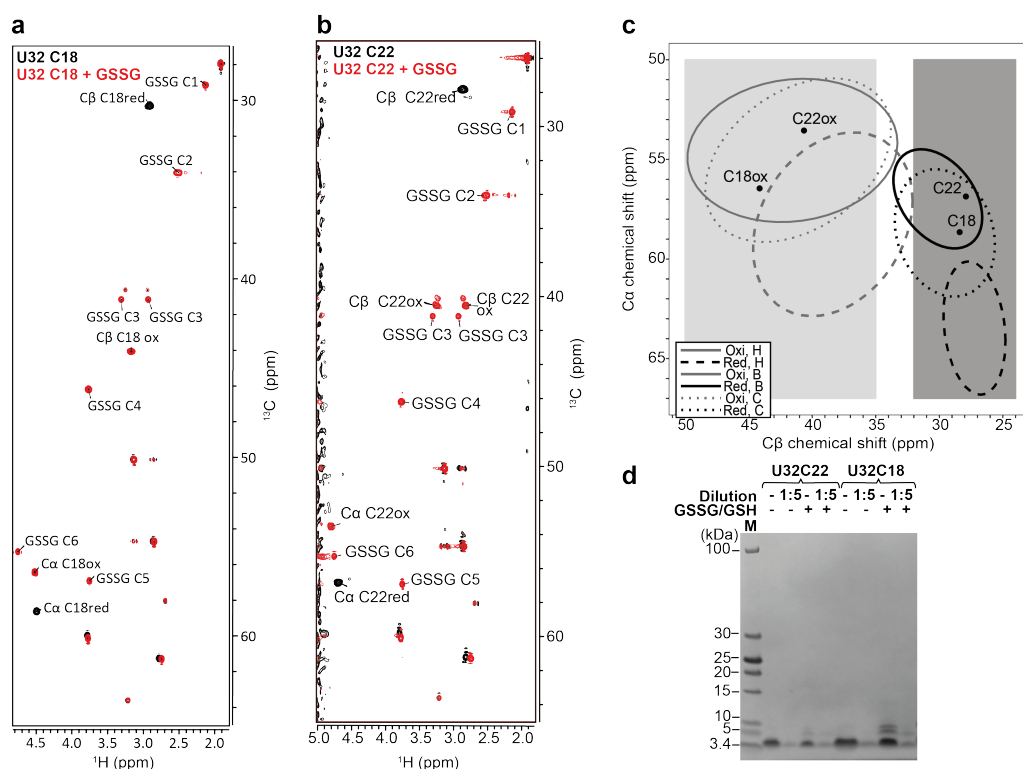
Figure 3.37: *Oxidation of U32 peptides after addition of GSSG/GSH. a) 2D ^1H-^{13}C HSQC spectrum of ^{13}C, ^{15}N C18 labeled U32 peptide (black) and after addition of GSSH/GSH (red). b) 2D ^1H-^{13}C HSQC spectrum of ^{13}C, ^{15}N C22 labeled U32 peptide (black) and after addition of GSSH/GSH (red). Protein concentration: approx. 100 µM U32, in Tico buffer (10 mM HEPES, 30 mM NH_4Cl, 30 mM $MgCl_2$, 1 mM EDTA, 10% D_2O, 50 µM DSS, pH 8.1. Acquisition parameters: NS=4, TD=1024, 64, T=277K. Final GSSG/GSH concentration: 5.77 mM GSSG and 0.57 mM GSH. Assignment of GSSG is shown in Figure 5.4. c) Cα and Cβ chemical shifts from U32 peptides C18 and C22 reduced and after oxidation with GSSG/GSH compared to values derived from a chemical shift database [261]. d) 4-12% NuPAGE gel of U32 C18 and C22 peptide samples with and without glutathione. Samples from left to right: Protein marker (M), U32 C22 peptide without GSSH/GSH, U32 C22 peptide in 1:5 dilution with water, U32 C22 peptide after addition of GSSG/GSH, U32 C22 peptide after GSSG/GSH addition in 1:5 dilution with water, U32 C18 peptide without GSSH/GSH, U32 C18 peptide in 1:5 dilution with water, U32 C18 peptide after addition of GSSG/GSH, U32 C22 peptide after GSSG/GSH addition in 1:5 diltion with water. The Protein marker PageRuler Unstained Low Range Protein Ladder from ThermoFisher Scientific (Waltham, USA) was used as reference.*

The chemical shifts of the free peptides are also located within the same region of the RNC constructs. The free peptides obtain an unfolded conformation, whereas the structure of the NC is unknown. Inducing the cysteine oxidation with GSSG/GSH leads not only to an intra-molecular disulfide bond, but also to dimer formation and to binding of glutathione (Figure 3.37d and Figure 5.5 in appendix) making it difficult to assign the chemical shift to the real oxidation state. Using these peptides, we could show that addition of oxidized glutathione

leads to oxidation of the cysteine residues and that glutathione can indeed bind to GBC fragments. To determine if intra-molecular or inter-molecular disulfides with glutathione are also formed in the RNC constructs, the RNC samples were investigated by LC-MS/MS.

3.2.11   Investigation of cysteine oxidation by mass spectrometry

To differentiate the oxidation states of cysteines within NCs, we investigated the NC after RNase digestion using liquid chromatography-coupled tandem mass spectrometry (LC-MS/MS). All measurements were carried out by Jakob Meier-Credo in the lab of Dr. Julian Langer.

In the absence of reducing agents during sample preparation, peptides with glutathione and nitrosylation adducts of the NC polypeptides were detected together with intramolecular disulfide bonds (Figure 3.38 and Table 3.5). Glutathione and nitric oxide bound to the side chain of Cys18 in the GFQGHAYECSSDAPN-LQPYFSR polypeptide of the U32SecM C18 construct (Figure 3.38a, b) were detected. It was previously demonstrated that rifampicin decreases the glutathione concentration within *E. coli* cells [149]. This could then result in an increased oxidation of proteins inside the cells. However, S-glutathionylation of the NC was detected in both, cells expressed in LB medium without addition of rifampicin and in selectively labeled medium with rifampicin. Also the addition of cysteine to the medium, as in the selectively labeled medium, increases the glutathione concentration, counteracting the effect of rifampicin [149]. The observed cysteine modifications are thus not necessarily linked to the cell medium. The double mutant U32SecM C18 C22 did not contain any disulfide bonds or cysteine modifications, except the alkylation with carbamidomethyl from the iodactamid treatment. The U32SecM C15 C18 double mutant did not show any cysteine modifications at all (Table 3.5). All cysteine residues in the U32SecM variant with an oxidation level of 25% can be modified by carbamidomethyl, hence some residues are in a reduced state. However, also two disulfide bonds were detected between Cys15 and Cys32 and between Cys22 and Cys32. With a trypsin cleavage site between both peptides, the differentiation between inter- and intramolecular disulfide bond is impossible.

An intra-molecular disulfide bond, without trypsin cleavage site between the two peptides, was detected between Cys32 and Cys41 in the polypeptide CNSIRVDS-GCWMFSTPVWISQAQGIR of the U43SecM (Figure 3.38c) and in the U78SecM construct (Table 3.5). Except for S-nitrosylation, none of these modifications was detected in control samples with reducing conditions. All detected disulfide bonds are visualized in Figure 3.38d. A non-native disulfide bond between Cys32 and Cys41 was also observed in the W42Q mutant of γD crystallin after interaction with oxidized γD crystallin [231]. This human γD shares 77% sequence identity to bovine γB crystallin and showed an oxidoreductase function using disulfide exchange to initiate aggregation of mutant crystallins [207, 231].
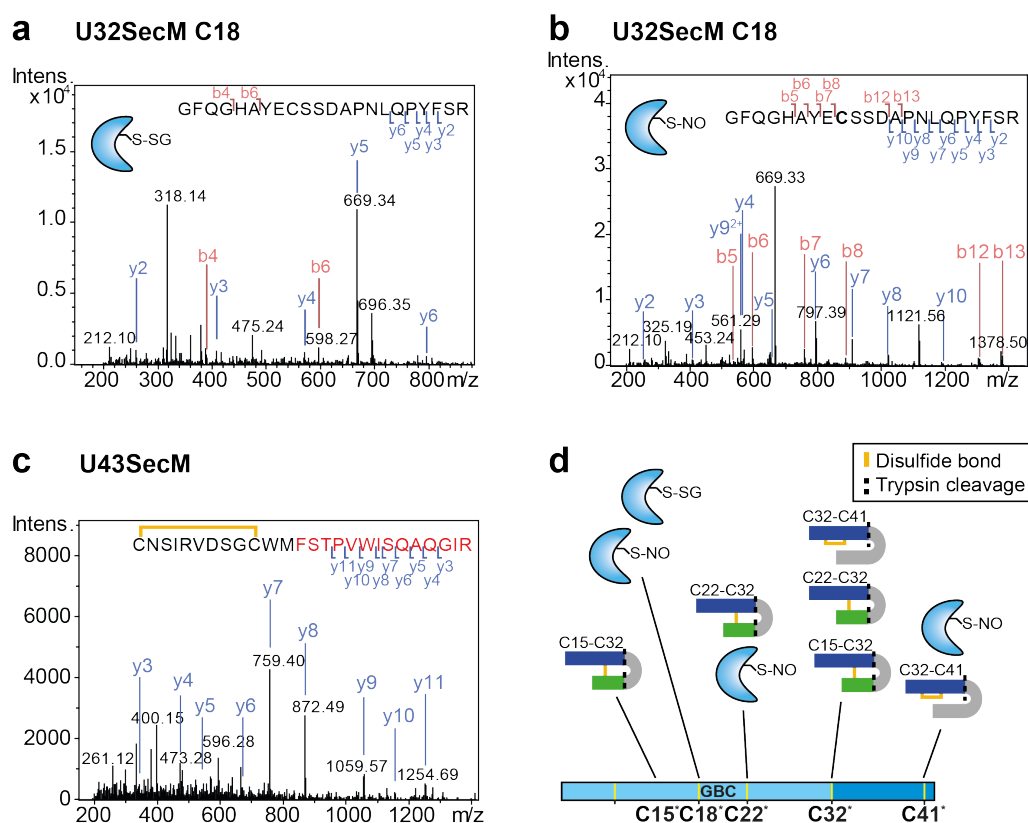
Figure 3.38: *LC-MS/MS investigation of RNC fragments. **a)** Fragment spectrum of the tryptic peptide GFQGHAYECSSDAPNLQPYFSR (precursor m/z 695.54 (4+)) of U32SecM C18 with a glutathione modification on Cys18. **b)** Fragment spectrum of the tryptic peptide GFQGHAYEC-SSDAPNLQPYFSR (precursor m/z 835.03 (3+)) of U32SecM C18 with a nitroso modification on Cys18. Matched fragment ions are indicated as dashes in the sequence and respective labels on peaks. **c)** Fragment spectrum of the tryptic peptide CNSIRVDSGCWMFSTPVWISQAQGIR (precursor m/z 980.46 (3+)) of U43SecM with a loop-linked disulfide bond (Cys32-Cys41). Matched fragment ions are indicated by dashes in the sequence and respective labels on peaks. **d)** Overview of cysteine oxidation states (S-nitrosylation (S-NO), S-glutathionylation (S-SG), disulfide bonding) of the NCs in RNCs as observed by LC/MS-MS. Cys78 in U78SecM was not involved in any disulfide bond formation. Trypsin cleavage site is shown as black dashed line, disulfide bond as yellow solid line.*

Interestingly, Cys32 is involved in several disulfide bonds, indicating a high reactivity of this particular cysteine residue. In the full-length GBC, Cys32 is completely buried inside the protein and has a solvent accessibility of $0\,\text{Å}^2$ (Table 1.3). The neighboring cysteine residues Cys22 and Cys41, which were also involved in disulfide bond formation, obtain a rather low solvent accessibility of 19 and $12\,\text{Å}^2$, respectively. Important for disulfide bond formation and cysteine modifications are also the proximity to basic (His, Lys, Arg) or charged amino acids (Ser, His) [41] (Figure 1.14). Two serine residues are located between Cys18 and Cys22 and one serine and arginine residue precedes Cys32. Cys78 is also in sequential proximity to one arginine and two serine residues.

Table 3.5: *Identified loop-linked disulfides and cysteine modifications with glutathione or nitric oxide in NCs by LC-MS/MS.* *Free, reduced cysteines were alkylated with iodacetamid (2.9).*

Construct name	Identified disulfide bond	Trypsin cleavage site between peptides	Cysteine modification
U32SecM C18	-	-	S-nitrosylation (Cys18), S-gluathionylation (Cys18), Carbamidomethyl (Cys18)
U32SecM C15 C18	-	-	-
U32SecM C18 C22	-	-	Carbamidomethyl (Cys18, Cys22)
U32SecM	Cys15-Cys32, Cys22-Cys32	yes	Carbamidomethyl (Cys15, Cys18, Cys22, Cys32)
U43SecM	Cys32-Cys41	no	S-nitrosylation (Cys22, Cys41), Carbamidomethyl (Cys15, Cys22, Cys32, Cys41)
U78SecM	Cys32-Cys41	no	S-nitrosylation (Cys41), Carbamidomethyl (Cys15, Cys22, Cys32, Cys41, Cys78)

The disulfide bond prediction tool DISULFIND [31] predicted a possible disulfide bond between Cys18 and Cys32 within the GBC sequence of U32 revealing also the high reactivity of Cys32. This software tool [31] uses global features like amino acid composition, number of cysteines and average cysteine conservations and peptide chain length but does not predict a structure. The other cysteine residues in this U32 sequence were predicted to be reduced, but with a very low confidence value. For U32SecM no disulfide bond was predicted, but the prediction was also of low confidence. For U43SecM and U78SecM no disulfide bond was predicted and each cysteine within prediction revealed a high confidence value. Another disulfide prediction tool DiANNA [60, 61, 62] predicted a bunch of different disulfide bonds in U32SecM, U43SecM and U78SecM. A disulfide bond between Cys18 and Cys32 was predicted in most of the sequences, demonstrating again the high reactivity of Cys32. These results and also the disulfide bonds detected by LC-MS/MS reveal that all cysteines within the N-terminal domain

of GBC are highly sensitive to oxidation in the NC.

Since the location of the cysteine residues inside the ribosomal tunnel are not
known, the proximity to ribosomal amino acids is unknown. However, struc-
tural information of the NC inside the ribosomal exit tunnel can be obtained by
cryo-EM.

3.2.12 Cryo–EM on U32SecM

To study if possible disulfide bond formation can occur inside the ribosomal
exit tunnel, the shortest construct U32SecM with in total four cysteine residues
was investigated by cryo-EM. A fully extended peptide, occupying 3.0-3.3 Å per
residue needs only 33 amino acid residues to span the 100 Å long ribosomal tun-
nel, while an α–helix which occupies 1.5 Å per amino acid needs around 67 amino
acids to transverse the tunnel [152, 84]. In the U32SecM construct only 34 amino
acids include all cysteine residues and the SecM stalling sequence. Hence, at least
three of four cysteine residues should be visible inside the ribosomal tunnel after
analyzing the cryo-EM studies.

Actually, 196,000 (77%) out of 266,000 particles, showed the Gly-tRNA within
the P-site of the ribosome, while 55,000 (21%) particles showed the ribosome in
the rotated state with electron density of two tRNAs, the Gly-tRNA in the P/E-
site and the Pro-tRNA in the A/P*-site, as described previously [280] (Figure
2.2). The resolution of the maximum obtained EM density was calculated at
2.53 Å with one bound tRNA and at 3.19 Å with two tRNAs bound. Zhang *et al.*
[280] observed a slightly different ratio between the two states, with 60,354 par-
ticles with the Gly-tRNA in the P-site and 41,501 particles with the Gly-tRNA
in the P/E-site and the Pro-tRNA in the A/P*-site. The decreased occupancy of
Gly-tRNA in the P-site might be the result of using an *in vitro* translation assay,
which may lead to the formation of polysomes, that are not observed when the
RNCs are produced *in vivo*. To remove the polysomes, Zhang *et al.* [280] added
20 U of RNase A to the last purification step. If SecM is used as stalling sequence
to investigate RNCs, *in vivo* produced RNCs seemed to be the preferred choice.
Since 21% of the particles are occupied with Pro-tRNA in the A/P*-site and ad-
ditionally A2602 of 23S rRNA and Arg163 of SecM block the entry of an incoming
aminoacyl-tRNA into the A-site [280] the SecM stalling sequence is resistant to
puromycin [176, 253]. The preferred way to produce RNCs, which can be released
by puromycin, is therefore the use of *in vitro* translation assays, either by using a
commercial system without release factors or with linearized *in vitro* translation
template missing a stop codon.

A cross section of the 50S subunit with detail view of the ribosomal exit tunnel
is shown in Figure 3.39a. Overall, 35 amino acids were modeled into the cryo-
EM density including all four cysteine residues of U32SecM. The distance between
Gly165 of SecM and the last modeled amino acid His14 of GBC is 65.59 Å, result-
ing in 1.87 Å per residue. This value is close to the 1.50 Å per residue, obtained

in an α- helical conformation and indeed a α-helix is visible in Figure 3.39b.
23S rRNA domains I to V form the majority of the tunnel surface[189]. Further-
more, the ribosomal proteins L4, L22 and L23 also contribute to the ribosomal
tunnel surface, with L22 and its long β–hairpin loop being the highest protein
contribution [189]. The ribosomal protein loops of L4 and L22 compose the con-
striction site at around 34.1 ± 4.8 Å from the PTC as averaged of 20 different
structures obtained from all three kingdoms of life and at around 32 Å in *E. coli*
[46]. L23 is located at the lower part of the tunnel and covers about 19.0 ± 2.8 Å
of the ribosomal tunnel in bacteria [46]. L23 is also present in eukaryotes but
the segment covering the tunnel is replaced by L39, which covers 31.6 ± 2.3 Å
[46]. Hence, co-translational protein folding within the ribosomal tunnel might
differ between eukaryotes and prokaryotes, especially with a second constriction
site found in eukaryotes [46]. While the vestibule region of prokaryotes is large
enough to provide a substantial degree of structure, this region might be too nar-
row in eukaryotes [46]. However, the slightly tighter tunnel radius in eukaryotes
might also favor α-helical conformations. Co-translational folding of GBC within
eukaryotes would therefore require further investigation.
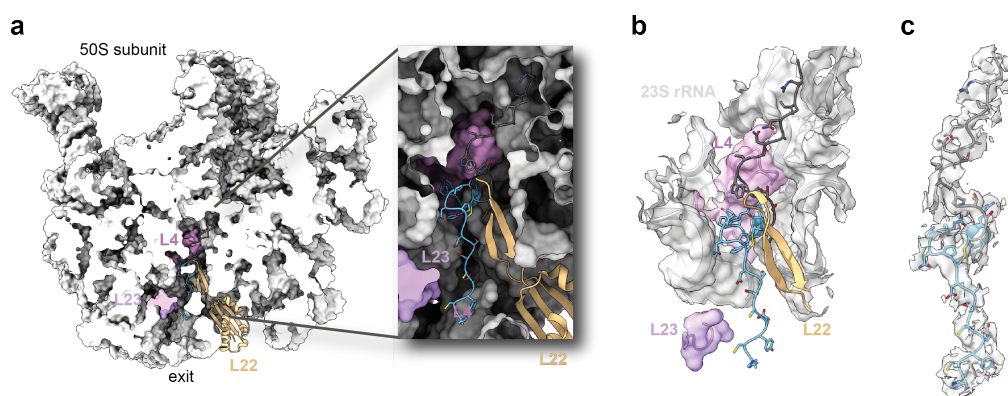
Figure 3.39: *Cryo-EM structure of U32SecM. a) Side view of ribosomal 50S sub-
unit with focus on the NC. The surfaces of nucleotides 753, 1323-1325
and 1616 were hidden to gain a full view on the NC. b) U32SecM in the
ribosomal exit tunnel with the surrounding area (within 20 Å) shown in
surface and cartoon representation. c) cryo-EM density of 2.7 Å around
the NC with model of U32SecM.*

The fusion zone of SecM to GBC in U32SecM is nearby the constriction site.
The amino acids Phe150 of SecM to Tyr28 form an α-helix and Cys32 and Cys22
are generating a disulfide bridge visible in Figure 3.39b and Figure 3.39c. Fur-
thermore, this α-helix includes cysteine residue Cys32, which is in line with the
observation made by solid-state NMR, showing signals in an α-helical and oxidized
conformation (Figure 3.35) and the ramachandran plot of the NC simulation us-
ing flexible-meccano [14] (Figure 3.29).
It is unclear if this helix is only formed in this GBC-SecM fusion protein and might
be only obtained, since SecM stalls the NC at this position. Within U43SecM
one signal is located at the border between reduced coil or β-sheet formation and

oxidized α-helix (Figure 3.32). Probably, the formation of the α-helix obtained in U32SecM is restricted to this construct with its location close to the constriction site and the interactions to the ribosome at this point. Interestingly, however, secondary structure prediction using PredictProtein [274] for U32 without SecM also predicts a α-helix between Cys32 and Lys25, which is not formed in the native full-length GBC. In a secondary structure prediction using the protein sequence of U43 without SecM the α-helix is then only predicted between Lys25 and Gln26 and Tyr28 and Phe29 and the α-helix disappears in the prediction of U78 completely. Hence, the α-helix in U32 could also be formed without the stalling sequence of SecM and elongation of NC leads to unwinding of this α-helical conformation.

Interactions of U32SecM within the ribosomal exit tunnel are shown in Figure 3.40. The stalling mechanism of SecM was established before [280, 179, 17]. In agreement with previous data [114], we detect a compact form of SecM with a distance of 22.41 Å between Arg163 and Trp155 (Figure 3.40a), which is below the expected 30 Å for an extended SecM chain. Arg163 is located close to U2506 of 23S rRNA, the π-stacking interaction between Trp155 and A753 of 23S rRNA (Figure 3.40b) is also found [280]. The backbone of Trp155 interacts with Arg61 of L4 (Figure 3.40b), resulting in a different orientation of this Trp side chain as described by Zhang *et al.* [280], but similar to Javed *et al.* [114].

Interactions between GBC and the ribosome involve hydrogen bonds between Arg31, Ser30 and nucleotides C1259 and U1260 of 23S rRNA (Figure 3.40c) and between the Asn24 side chain and the phosphate backbone of C461 and C462 of 23S rRNA (Figure 3.40d). Interactions in the lower region of the tunnel include hydrogen bonds of the carbonyl groups of Ser19 and Ser20 with the side chains of Arg84 and Arg95 of L22, respectively (Figure 3.40e), and of Tyr16 with the phosphate backbone of A508 of 23S rRNA (Figure 3.40f). Arg84 is also involved in interaction with RNC constructs containing the fifth and sixth domains from the gelation factor ABP-120 [114] and the Myc-tag [280]. In addition, also the interaction of the nucleotide A508 with the Myc sequence was described previously [280]. The interactions between the NC and the ribosomal exit tunnel are listed in Table 3.6 and shown schematically in Figure 3.41a. Most likely, all of these interactions help to stabilize the α-helical conformation of amino acids Phe150 to Tyr28 and the loop formed between Cys32 and Cys22, and bring these residues in close proximity to form a disulfide bridge (Figures 3.39 and 3.41).

Therefore, the ribosomal exit tunnel allows a non-native disulfide bond formation, since the Cβ-Cβ distance between Cys22 and Cys32 is 12.62 Å within the native full-length GBC (PDB: 4W9A), which is too far away to form a disulfide bond. Usually the Cβ-Cβ distance between cysteine residues involved in disulfide bond formation is in the range between 3.5-4.5 Å [247]. The Cβ-Cβ distance between Cys18 and Cys22 is 6.09 Å in the recombinat expressed full-length GBC (PDB: 4W9A) and 4.81 Å (PDB: 4GCR) in GBC isolated from bovine eye lenses. In

the lattter a disulfide bond was observed between Cys18 and Cys22 [177]. The distance between Cys15 and Cys18, which form a typical CXXC motif, common in thioredoxin and oxidoreductases, was measured at 10.20 Å and therefore disulfide bond formation is impossible in the full-length GBC between both cysteine side chains that are facing into different directions. A more likely disulfide bridge could be formed between Cys22 and Cys78 in the full-length GBC, since these residues are only 5.42 Å (PDB: 4W9A) apart from each other and the angle between these residues is 94.5° approaching the ±90° which is commonly obtained in disulfide bonds [196, 36]. In close proximity to Cys22 and Cys32 are Phe150 from SecM and Arg95 of L22 (Figure 3.40g). Basic amino acids, such as arginine, in the vicinity of cysteine residues can lead to the formation of Cys thiolate and the high reactivity of this thiolate to further disulfide bond formation [41]. Aromatic amino acids close to Cys residues could undergo S-π interactions and thereby contribute to the stability of the native GBC [175], which could increase the reactivity of the thiol. The formation of the observed disulfide bond between Cys32 and Cys22 might be enhanced by the tight space inside the ribosome and the surrounding amino acids.

Although the α-helix within U32SecM is not found in the native full-length GBC, the loop formed between the residues Cys32 and Cys22 is also partly observable in the native structure of GBC (Figure 3.41b) with an RMSD of 4.33 Å between both structures. So, probably a part of the native structure is already formed inside the ribosomal tunnel. Furthermore, the cysteine residues of a flexible-meccano simulation are located close to the Cys22 and Cys32 of the cryo-EM structure (Figure 3.41c). The Cys22 residues are located 3.72 Å and Cys32 residues 2.98 Å apart from each other. The RMSD between the simulation and the structure between the His15 and Trp155 is 5.45 Å. In summary, by using cryo-EM a non-native disulfide bond between Cys22 and Cys32 and a non-native α-helix including the residues Phe150 to Tyr28 could be identified, which might be an intermediate in the co-translational protein folding of GBC. The coupling between protein folding and disulfide bond formation is explained by two models: In the folded precursor model, cysteine residues are brought into close proximity after the structure is formed. In the quasi-stochastic model, cysteines pair in an unfolded conformation and influence subsequent protein folding [265]. A coarse-gained molecular simulation assumed protein folding guides disulfide bond formation in BPTI [206]. The oxidative folding of BPTI is pH dependent, while at pH 8.7 close to the pK$_a$ of the cysteine side chain non-native disulfide bonds were observed [37], only native disulfide bonds were detected at a neutral pH of 7.3 [264]. The non-native disulfide bond detected in the NC of GBC favors the quasi-stochastic model. The RNC constructs were measured near physiological conditions at pH 7.5, also below the pK$_a$ of the cysteine side chain. Thus, the detected non-native disulfide bond is rather influenced by interactions with the ribosomal tunnel, in particular with Arg95 of L22 then by the pH. Co-translational disulfide bond
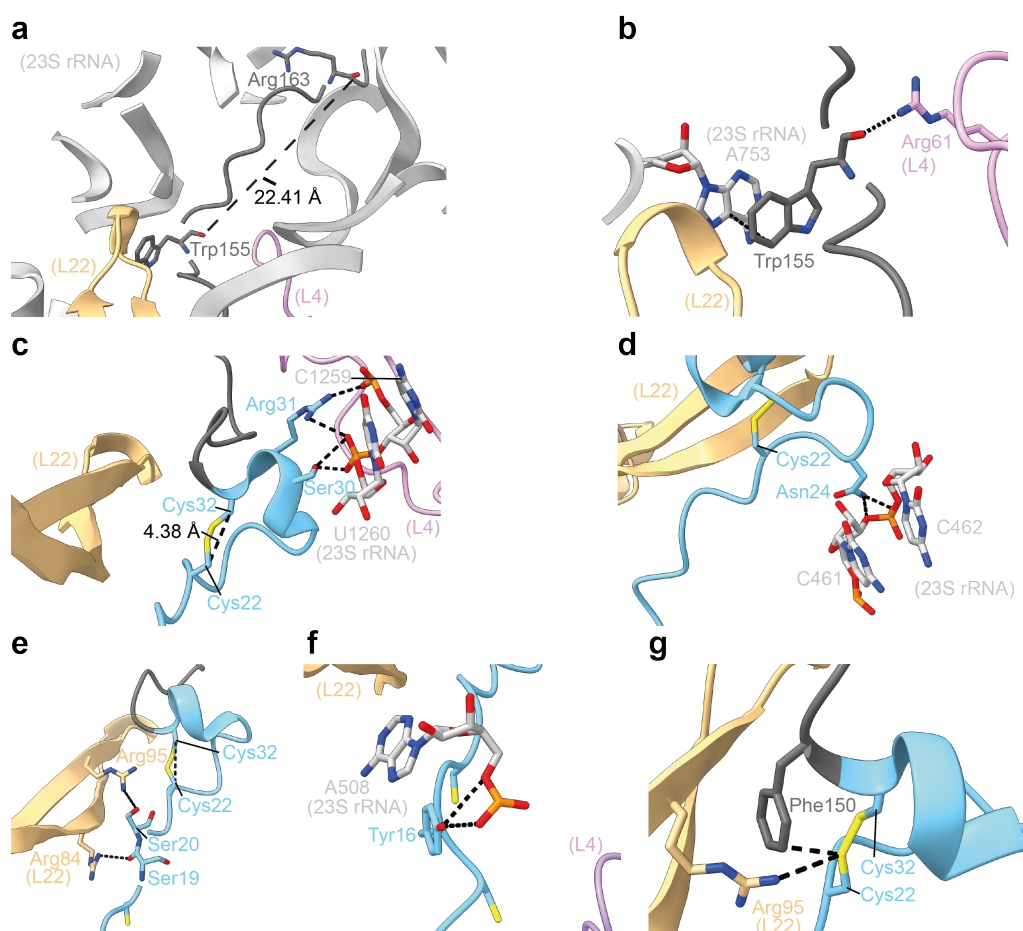
Figure 3.40: *Interactions between U32SecM and the ribosomal exit tunnel.*
a) Compaction of SecM between residues Arg163 and Trp155 (dark grey).
b) π-π stacking between Trp155 of SecM (dark grey) with A753 of 23S
rRNA (light grey). *c) Interaction of GBC residues Arg31 (light blue)*
with C1259 and U1260 of 23S rNA (light grey) and Ser30 with U1260. *d)*
Hydrogen bonding between the Asn24 side chain of GBC (light blue) with
the phosphate backbone of C461 and C462 (light grey). *e) Interaction*
between Ser20 and Ser19 of GBC (light blue) with Arg95 and Arg84 of
L22 (light orange), respectively. *f) Hydrogen bonding between Tyr16 of*
GBC (light blue) and the phosphate backbone of A508 of the 23S rRNA
(light grey). *g) Close proximity of Cys22 to Phe150 of SecM (dark grey)*
and Arg95 of L22 (light orange). The distance of Cys22 to Arg95 is 4.03
Å and to the side chain of Phe150 3.39 Å. L4 in light pink, L22 in light
orange, L23 in light purple, 23S rRNA in light grey, SecM in dark grey
and GBC in light blue. Nucleotides and amino acids are shown in one
letter and three letter code, respectively.

Table 3.6: *Interactions between U32SecM and the ribosome.*

NC residues	Interacting residue or ribonucletide		
Atom, residue	Atom, residue (three letter code) or ribonucleotide (one letter code)	Protein or RNA	Type of interaction
O(CO), Trp155	H(NH$_2$), Arg61	L4	Hydrogen bonding
Trp155	A753	23S rRNA	π-π-stacking
H (NH), Arg31	OP1 (PO$_4$), C1259	23S rRNA	Hydrogen bonding
H (NH$_2$), Arg31	OP2 (PO$_4$), U1260	23S rRNA	Hydrogen bonding
H (OH), Ser30	OP1 (PO$_4$), U1260; OP2 (PO$_4$), U1260	23S rRNA	Hydrogen bonding
H2 (NH$_2$), Asn24	OP1 (PO$_4$), C461; O3', C462	23S rRNA	Hydrogen bonding
S, Cys22	S, Cys32	NC	Disulfide bond
O (CO), Ser19	H (N$_2$), Arg84	L22	Hydrogen bonding
O (OH), Ser20	H (NH$_2$), Arg95	L22	Hydrogen bonding
H (OH), Tyr16	O (PO$_4$), A508	23S rRNA	Hydrogen bonding
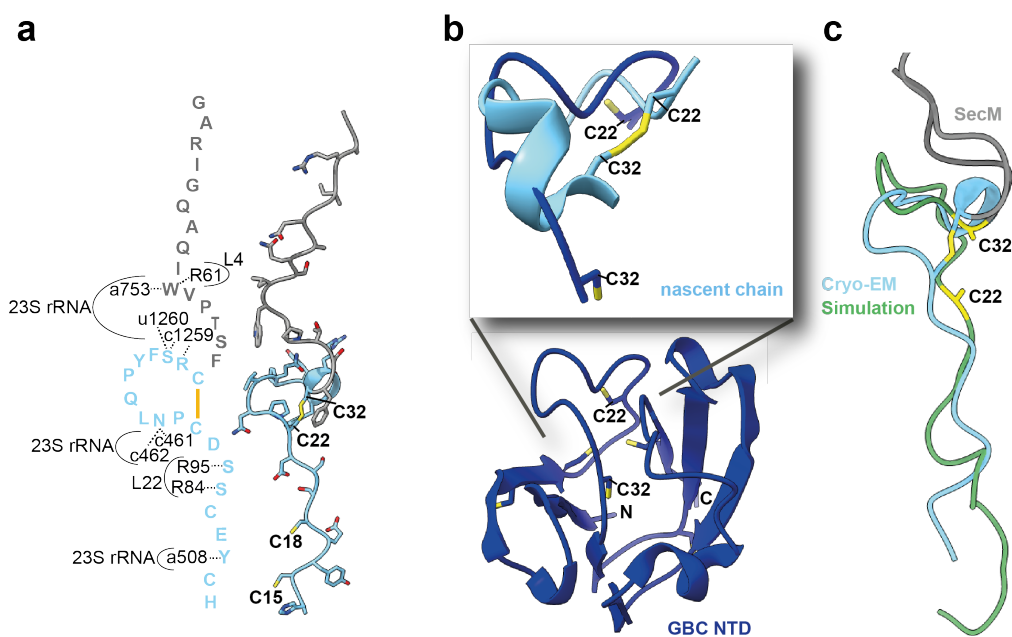
Figure 3.41: *Cryo-EM structure and structural analysis of U32SecM. a)*
Model of NC with interactions to ribosomal proteins and 23S rRNA rep-
resented. Nucleotides and amino acids are shown in lowercase and upper-
case one-letter code, respectively. SecM in grey and GBC in light blue. b)
Overlay of loop between Cys22 and Cys32 of the full-length GBC protein
(PDB: 4W9A) (dark blue) and the nascent chain U32SecM (PDB: 6YS3)
(light blue). RMSD=4.33 Å. c) Overlay of NC model obtained by cryo-EM
(light blue) and flexible-meccano simulation (green), with possible disulfide
bond between Cys22 and Cys32 (yellow). RMSD=5.45 Å.

formation in β2-microglobulin in eukaryotes occurred only after the entire pro-
tein domain entered the ER, favoring the precursor model [212]. Support for the
quasi-stochastic model comes from two recent reports: in a cell-free expression
system, arrested translation intermediates of the disintegrin domain of human
ADAM10 formed non-native disulfide bonds [211]. Further, Robinson *et al.* [211]
showed secondary structure rather than cysteine density influences the mecha-
nism of cysteine coupling. Non-native disulfide bridges between distant cysteine
residues have also been observed in the low-density lipoprotein receptor (LDL-R)
in HeLa cells [113]. The influence of the ribosomal exit tunnel on disulfide bond
formation, however, was not yet studied.

3.3 Investigation of disulfide bonding in the full-length GBC

One of the crystal structures (PDB: 4GCR) of GBC isolated from bovine
eye-lenses contains Cys18 and Cys22 in two different oxidation states: one with
the cysteine residues fully reduced and one with a disulfide bond between Cys18
and Cys22 [177]. In this form the side chain of Cys22 is turned to bring the

sulfur atoms in close proximity. To investigate the cysteine oxidation states in solution, liquid-state NMR analysis of full-length GBC expressed and purified from *E.coli* was conducted. The full-length protein U was prepared fully reduced and oxidized with $CuCl_2$ under continuous stirring and air supply overnight. Buhr *et al.* [22] showed that this oxidation of U resulted in a ^{15}N backbone spectrum similar to H-P1, which was partly oxidized and could be reduced by addition of DTT. Figure 3.42a shows the size-exclusion chromatogram and Figure 3.42b the SDS-PAGE gel with the corresponding fractions of the purified GBC after oxidation with Cu(II). Only one protein band is visible with a size of the full-length GBC (MW=23.17 kDa). The fractions were pooled, concentrated to a final concentration of approx. 800 µM and a ^1H-^{15}N BEST-TROSY spectrum recorded. Without 2-Mercaptoethanol in the SDS-loading buffer, the protein appeared at a slightly smaller molecular weight, indicating that the protein is either oxidized or that 2-Mercaptoethanol binds to the full-length GBC as shown previously [277]. Since this behavior was also observed on the reduced protein, indeed 2-Mercaptoethanol seems to bind to the protein.
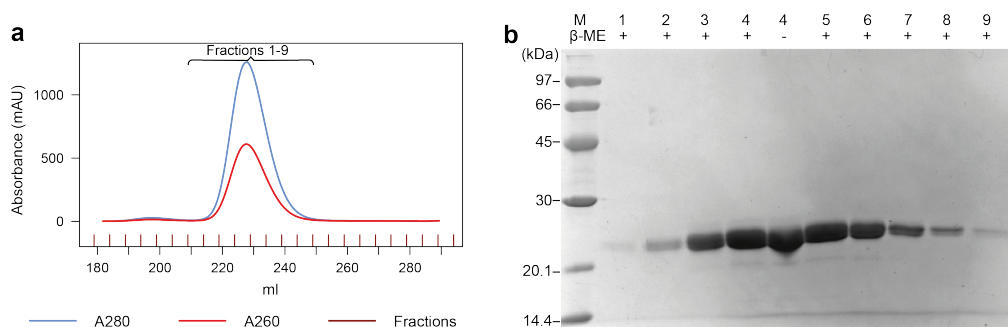
Figure 3.42: *Size-exclusion chromatogram and SDS-PAGE gel showing the corresponding fractions of GBC after oxidation. a) Size-exclusion chromatogram of full-length GBC after oxidation with $CuCl_2$. Gelfiltration column HiLoad 26&60 Superdex 75 prep grade was used (GE Healthcare, Chicago, USA). Absorbance at wavelength 280 nm (blue) and 260 nm (red). Fractions are shown as dark red line. b) 15% SDS-PAGE gel of size exclusion chromatography fractions of oxidized GBC. Samples from left to right: Protein marker (M), fractions 1-4 with β-Mercaptoethanol (β-ME) in SDS-loading buffer, Fraction 4 withouth β-ME in SDS-loading buffer and Fractions 5-9 with β-ME in SDS-loading buffer. For the SDS-PAGE the LMW-SDS Marker from GE Healthcare (Freiburg, Germany) was used.*

The 1D ^1H NMR spectra for the amide proton range are shown in Figure 3.43a. Small differences between the amid proton signals are visible, reflecting the same differences found between the U and H variant by Buhr *et al.* [22]. In the 2D ^1H-^{15}N BEST-TROSY spectrum (Figure 3.43b, c) the differences are more pronounced and again the spectrum of the oxidized protein U shows the same signal distribution as the H variant. Four cysteine residues (Cys15, Cys22, Cys41 and Cys78) showed peak doubling in the ^1H-^{15}N correlation spectrum. Except for Cys78 all these cysteines are solvent accessible. Cys15 exhibits the highest sol-

vent accessibility, which is three- to four-times the solvent accessibility of Cys22 and Cy41 (Table 1.3). Buhr *et al.* [22] also identified peak doubling for Cys32 in the ^{1}H-^{15}N BEST-TROSY spectrum using selectively cysteine labeled GBC. Due to signal overlap with Val132, the oxidized peak of Cys32 could not be unambiguously identified. Cys109 in the CTD did not show any chemical shift changes.
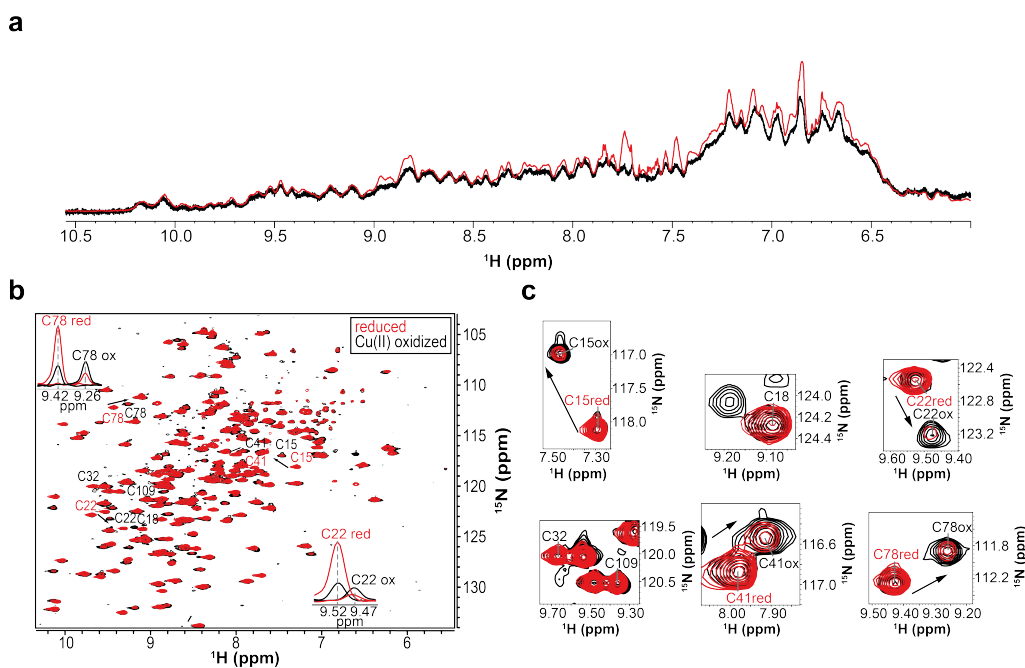
Figure 3.43: *1D ^{1}H and 2D backbone ^{1}H-^{15}N BEST-TROSY spectrum of full-length GBC cysteine residues reduced and after oxidation with Cu(II). a) 1D ^{1}H spectrum of amide proton region of full-length GBC reduced (red) and after oxidation with Cu(II)(black). Parameters for acquisition: NS=8, TD= 8192 for reduced and 32768 for oxidized GBC. T=298K. b) 2D backbone ^{1}H-^{15}N-BEST-TROSY spectrum of cysteine residues of full-length GBC reduced (red) and after oxidation with Cu(II) (black) and 1D slice of Cys22 and Cys78 peaks in panel. Positions of the cysteine residues are labeled in red and black accordingly. Parameters for acquisition: NS=16, TD=1024, 256, T=298K. c) Detail view of 2D backbone ^{1}H-^{15}N BEST-TROSY spectrum of reduced (red) and oxidized (black) GBC. Cysteine residues not changing position after oxidation (Cys18, Cys109) or that could not be unambiguously identified due to overlapping signals (Cys32) were labeled only once. Cu-(II) oxidation led to reduction in the protein yield and a decrease in the signal/noise ratio for the oxidized protein.*

To identify which cysteine residues might be involved in a disulfide bond, the Cα and Cβ chemical shifts were determined in triple-resonance experiments. The strip plots of the HNCACB spectra of the cysteine residues in GBC are displayed in Figure 3.44. The oxidized protein started to precipitate during data acquisition, resulting in a lower S/N ratio. Cβ chemical shifts below 32 ppm correspond to reduced cysteine residues, Cβ chemical shifts greater than 35 ppm are oxidized and between 32 and 35 ppm cysteine residues can be both, oxidized or reduced.

The Cα and Cβ chemical shift of Cy22 and Cys78 changed after oxidation with Cu(II).

Plotting the Cα against the Cβ chemical shifts indicates secondary structure and the oxidation state of the cysteine residues (Figure 3.45). Cys15 appears in the overlapping region, which could be both reduced and oxidized and also appears in the β-sheet or coil area. This cysteine residue has the highest solvent accessibility (Table 1.3) and is located in a β-strand with its side chain exposed to the solvent. Hence, the location of this signal in the Cα-Cβ plot seems reasonable. Cys18, Cys32, Cys41 and Cys109 are in a reduced β-sheet or coil conformation in the reduced protein. This seems also reasonable since all cysteine residues except for Cys32, which is localized in a coil, are involved in a β-strand within the crystal structure of GBC (PDB: 4GCR). With the exception for Cys32, for which Cα and Cβ chemical shift could not be identified due to peak overlap with Val132, the Cα and Cβ chemical shifts of these four cysteine residues do not change after oxidation with CuCl$_2$. The Cα and Cβ chemical shift of Cys22 changes dramatically from Cys22 positioned in a reduced β-sheet or coil region to an oxidized α-helical conformation. Cys78 shows also changes in the Cα and Cβ chemical shift and shifts from a reduced β-strand or oxidized α-helix to a position, where both conformations also overlap with a reduced coil conformation.

Cys78 is positioned indeed in a β-strand in the two crystal structures (PDB: 4GCR and 4W9A). The reactivity of cysteine residues is enhanced if the cysteine residues exhibit a high solvent accessibility or are located in the vicinity of basic amino acids like His, Lys and Arg, or charged residues as Ser [41] (Figure 1.14). The Cys22 in GBC is partly exposed and Cys18 and Cys78 have almost no or no solvent accessibility at all. These three cysteine residues are located within a Cβ-Cβ range below 6.43 Å to each other (Figure 3.11) and Ser20 is placed between these cysteines, hence disulfide shuffling between these three cysteine residues could in principle occur. The closest Cβ-Cβ distance is between Cys22 and Cys78 with 5.42 Å, but the clostest Sγ-Sγ distance is between Cys18 and Cys78. However, since Cys18, which was involved in a disulfide bridge with Cys22 in a crystal structure [177] (PDB:4GCR), exhibits a Cβ chemical shift below 32 ppm and did not show any change in the Cα and Cβ chemical shift after oxidation, the involvement of this residue in a disulfide bond seems unlikely. Since Cys22 and Cys78 are in close proximity in the crystal structure (PDB: 4W9A) with a Cβ-Cβ distance of 5.42 Å and exhibit an torsion angle (χ3) of 94.5°, which is the range of -87 and +97 degree determined by Craig and Dombkowski [36] analyzing 1505 native disulfides, a disulfide bond between these cysteines might occur. A possible disulfide bond between Cys22 and Cys78 was also proposed by Pande *et al.* [196], observing a disulfide bond in GBC using Raman spectroscopy.

Summing up, the disulfide bond which might be formed in the H protein is probably between Cys22 and Cy78. The most pronounced differences in the translation speed between U and H were in the first 70 amino acids, with U being

slower translated than H [22]. Probably, Cys22 is more exposed during translation in U and becomes reduced in the *E. coli* cytosol, while H Cys22 in H might be more reactive and/or protected. However, since synonymous codon usage did not show an effect on the oxidation states in the NTD alone (section 3.2), also the linker and CTD might be important for the observed difference between the U and H protein. Using cryo-EM and mass spectrometry, we showed that a non-native disulfide bond between Cys22-Cys32 and Cys32-Cys41 can form in arrested NCs. Although, some of these disulfide bonds might get reduced by the *E. coli* cytosol when exposed to the solvent, Cys22 could also rearrange to form a disulfide bond with Cys78 in the full-length GBC. Hence, future experiments should also focus on the involvement of the CTD and the linker between the two domains, both on isolated proteins and RNCs.
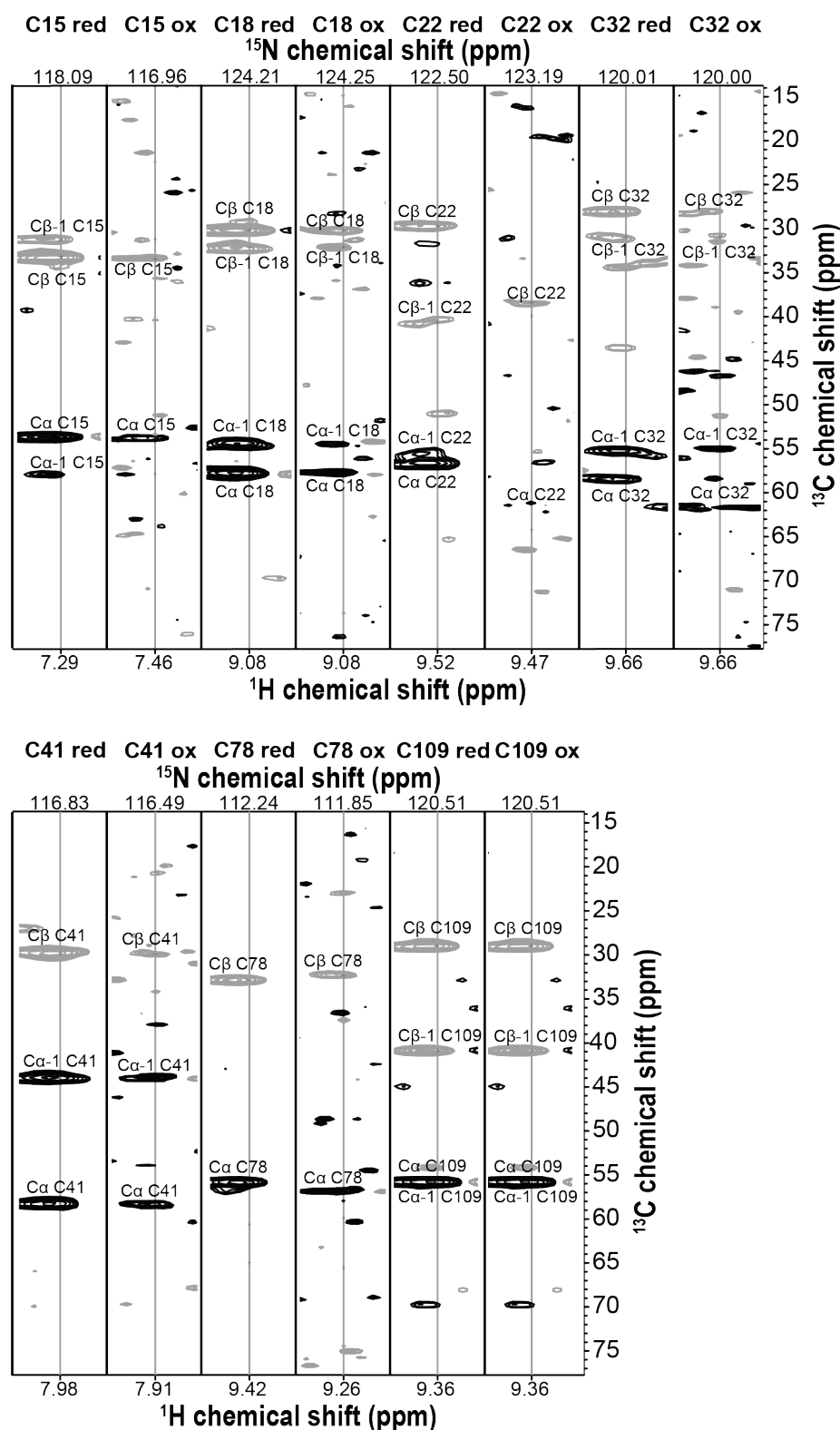
Figure 3.44: *^{1}H-^{13}C strip plots from the BEST-TROSY HNCACB spectrum of GBC full-length reduced (red) and after oxidation with Cu(II) (ox).* The $C\alpha$ resonances are phased to yield positive (black) cross-peaks, whereas the $C\alpha$ cross-peaks are negative (grey) The corresponding ^{15}N frequencies are at the top. After oxidation with Cu(II) and additional purification by size exclusion chromatography, approx. 30% of the protein yield was lost, leading to a decreased signal/noise ratio.
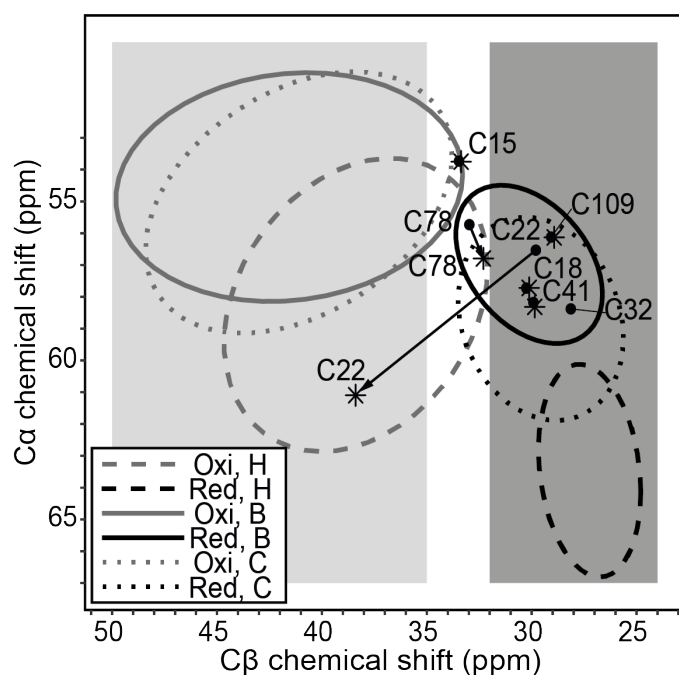
Figure 3.45: *Structural analysis of full-length GBC reduced and oxidized. Cα and Cα chemical shifts of the six cysteines of full-length GBC reduced (•) and after oxidation with Cu(II) (*) compared to values derived from a chemical shift database [261]. Ellipses contain 90% of the corresponding chemical shifts of each group (H = helix, B = β-strand, C = coil). The spectral region for oxidized Cβ cysteines is highlighted in light grey, and for reduced cysteine in dark grey. The overlapping region of both states is shown in white.*

4 Conclusion and outlook

The work presented here derives from a number of different samples, experiments and techniques, including liquid-state NMR on isolated proteins and peptides, solid-state NMR and cryo-EM on RNCs and mass spectrometry on tryptic digested GBC fragments. All techniques required different sample conditions, including selective isotopic labeling on the NC without the labeling of the ribosome, itself. This means 54 r-proteins with a total molecular weight of approx. 900 kDa [141] needed to be unlabeled, while the 7.2-12.7 kDa NC of GBC should obtain high yields of isotopic incorporation. Furthermore, solid-state NMR needed large amounts of RNCs (3 $\tilde{nmol}$), while for the preparation of the cryo-EM grids only a thousandth of this was required. For both techniques various sample conditions were studied, which together with the long sample preparation and measurement time for each construct delayed the finalization of this thesis. The discrepancies of the sample conditions and the different outcomes generated by various techniques had to be compared carefully. Variations in the protein yield of the different RNCs also affect the mass spectrometry results. With higher concentrations of RNCs more cysteine modifications were observed. Hence, also in RNCs with low sample concentrations, where no cysteine modification or disulfide bond was detected, both could still be present.

Nevertheless, the questions addressed in the beginning (Figure 1.19) concerning the effect of synonymous codon usage on the NTD and cysteine oxidation and disulfide bond formation inside the ribosomal exit tunnel and could be answered for the greater part within this thesis (Figure 4.1). An influence of synonymous codon usage was not observed in the NTD of GBC alone. The NTD was structured and able to fold independently, hence, the CTD is not required to assist the folding as expected. However, the CTD or the linker between the two domains might be important for the differences observed in the full-length GBC by Buhr *et al.* [22], although most differences in the translation frequency were observed in the first 70 amino acids of the NTD. Future experiments should address the influence of the linker and parts of the CTD with respect to the oxidation state of the NTD. In doing so, the region contributing to the observed differences in the cysteine oxidation states may be localized. Further, the influence of the translation kinetics on the folding of GBC or other proteins could be studied by varying the mRNA sequence. In this aspect, also the effect of adjacent codons (bicodons) should be investigated, as emerging evidence indicated that the elongation rate might be encoded by these bicodons [81, 73, 44]. These effects could be studied both on isolated proteins or attached to the ribosome with a linker spanning the ribosomal exit tunnel.

Since the majority of translational pausing occurred in the NTD [22], the possibility of cysteine oxidation during translation and even inside the ribosomal exit
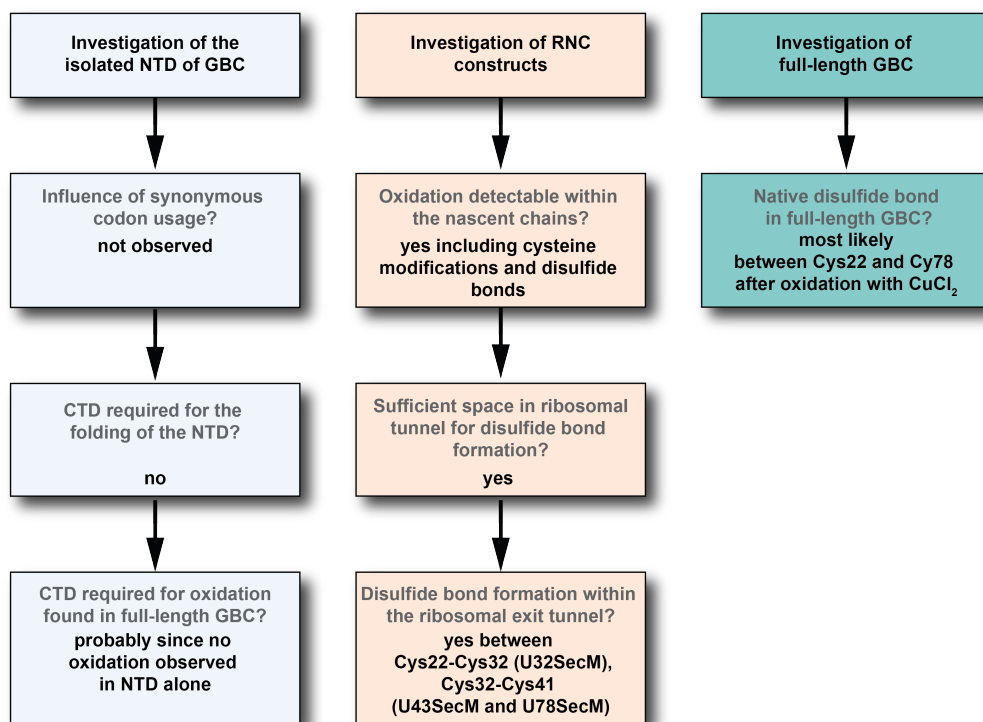
Figure 4.1: ***Topics and observed results discussed in this thesis.*** *The NTD of GBC (light blue) and the full-length GBC (green) were studied by liquid-state NMR. RNCs of GBC (dark yellow) were investigated by solid-state NMR, cryo-EM and mass spectrometry.*

tunnel was studied using RNCs. For the first time, selective ^{13}C, ^{15}N labeled RNCs were studied by solid-state NMR enhanced by DNP. Isotopic scrambling into rRNA occured in RNCs grown on ^{13}C-glucose and ^{15}N-NH$_4$Cl, although rifampicin was added after induction. Using the selective labeling approach described by Rutkowska *et al.* [217], but not further applied by the authors, we did not observe any labeling of ribosomal proteins or the rRNA. Although the NCOCX experiment with ^{13}C, ^{15}N labeled cysteine and ^{15}N labeled proline was not successful, additional optimization could make this experiment an elegant method to assign residues within the NC. The stalling and therefore the protein yield could be increased using the SecMstrong sequence [124]. Different labeled amino acids could be used to study the folding of NCs inside or after emerging the ribosomal tunnel. The small protein domains studied by Marino *et al.* [161] with arrest-peptide based force-measurement assays could be studied in the ribosomal tunnel and after emerging to determine the point at which these proteins start to fold and if these proteins obtain a native structure. Further, the co-translational folding of proteins could be also investigated by real-time FRET as described previously [101, 171]. With this approach compact structures formed during translation result in a high FRET efficiency.

All RNC constructs, even one mutant with only one cysteine residue showed 19% of cysteine oxidation. Hence, inter-disulfide bond formation within this

NC was detected. This was further confirmed by LC-MS/MS, detecting S-nitrosylated and S-glutathionylated cysteine residues in the NC. S-nitrosylation and S-glutathionylation are considered as post-translational cysteine modifications and can be mediated by S-nitroso-glutathione (GSNO) or glutathione (GSH) [76, 2]. In general, these modifications are reduced after release of the ribosome within the highly reductive environment of the *E. coli* cytosol.

Since the purification buffers did not contain glutathione or any nitric oxide-containing compounds, the observed modifications must have been formed inside the cell any maybe even during translation. Formation of mixed disulfides of GSH with distinct proteins was not observed *in vivo* [243]. However, glutathione adducts on GBC have in fact been previously reported [148, 277, 248] and in eukaryotic cells mixed disulfides with GSH are accumulated in certain situations [243]. Cysteine modifications within the NC could then promote the formation of an intra-molecular disulfide bond (Figure 4.2).

The ribosomal exit tunnel provides sufficient space for several disulfide bridges within U32SecM as calculated by flexible-meccano simulation [194]. LC-MS/MS also revealed different disulfide bonds within the NCs. In U43SecM and U78SecM one intra-molecular disulfide bond was observed between Cys32-Cys41.

In U32SecM a disulfide between Cys15-Cy32 and Cys22-Cys32 was detected, but a trypsin cleavage site was located between the peptides. Cys32 is involved in several disulfide bonds in the different RNC constructs and also prediction tools [31, 60, 61, 62] revealed disulfide bond formation including Cys32. Using cryo-EM a disulfide bond between Cys22 and Cys32 was indeed observed. This disulfide bridge and the compact formation, including an α-helical structure of the NC are stabilized by interactions between the NC and the ribosome. The disulfide bond between Cys22 and Cys32 might be shuffeled to Cys32 and Cys41 with further elongation of the NC.

A recent study by Duc *et al.* [46] revealed a second constriction site and a slightly smaller average radius in the lower part of eukaryotic ribosomes. Hence, disulfide formation within eukaryotic ribosomes would require further investigation. We showed, however that the ribosomal tunnel confines the polypeptide chain and allows disulfide bond formation as well as cysteine modifications (Figure 4.2). Some proteins or their co-translational folding intermediates could use such early formed disulfides as a means of escaping the reducing environment of the cytoplasm. This finding has implications for the oxidative folding pathway of the polypeptide chain inside the cell. While part of the chains may already form native disulfide bonds in the tunnel, non-native modifications will have to be reduced first and refold in the cytosol to attain the final native structure. The reduction of disulfide bonds in the *E. coli* cytosol is realized by the proteins Trx and Grx and the tripeptide GSH. The disulfide bond obtained in the ribosomal exit tunnel might be reduced after release to the *E. coli* cytosol with its high reductive environment. But as both proteins cannot enter the ribosomal exit

tunnel this limited access might favor disulfide bond formation at an early state of translation.

Also, the TF can bind to NCs, thereby delay co-translational folding, unfold already formed protein domains and postpone disulfide bond formation [91, 138]. Thus, the non-native disulfide bond could also be disrupted with futher elongation by the TF within the cell. Since the TF binds to NCs after 100 amino acids have been synthesized [192], only the folding of the longest construct U78SecM might be influenced by TF binding. Within a future study, involving longer constructs the influence of the TF should be considered and could be studied. Further, co-translational disulfide bond formation of different proteins could be investigated. For example β-lactamase, the small β-sheet Src-homology 3 (SH3) domain of α-spectrin and barnase designed with two disulfide bonds. All these proteins were already investigated using an oxidized *in vitro* translation system [90]. The authors detected increasing disulfide bond formation with increasing chain lengths. However, a linker was used to span the ribosomal exit tunnel and all constructs showed disulfide bond formation.
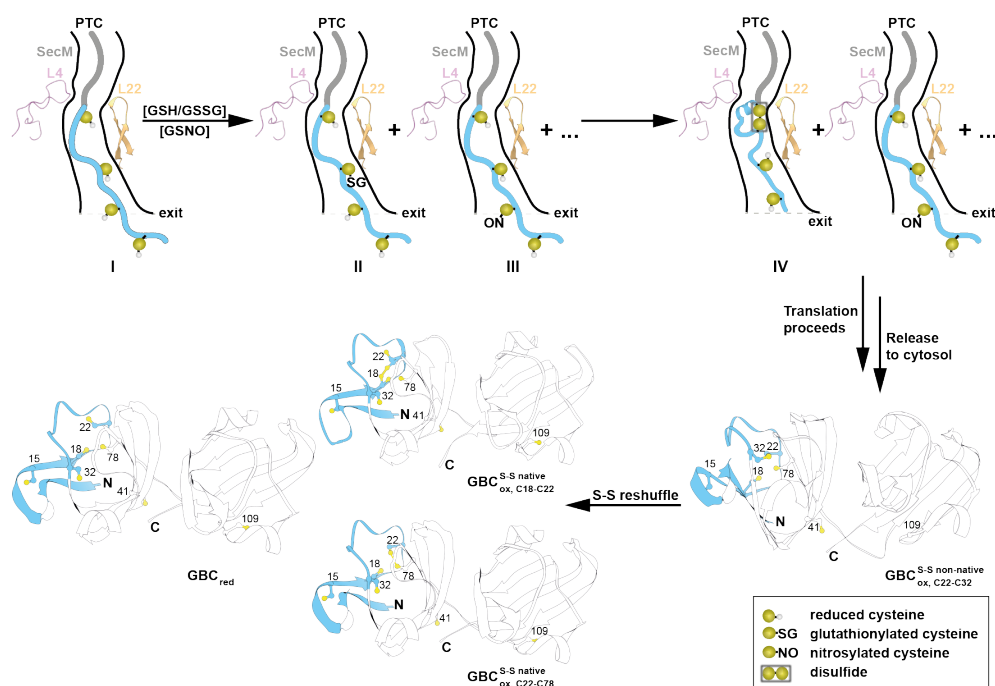
Figure 4.2: ***Disulfide bond formation within NCs and after release.*** *Free thiols of GBC NC **I** are modified by glutathione (GSH) **II**, or S-nitrosoglutathione (GSNO) **III**. Such modified cysteine residues with enhanced electrophilicity can be attacked by thiol groups of neighboring cysteine residues to form disulfide bonds **IV**. In the released NC, non-native disulfide are either reshuffeled to form native disulfide bonds or reduced depending on translation rate (Cys18-Cys22).*

Our investigations unexpectedly revealed that the detected disulfides in the exit tunnel are different, at least in part, to the disulfides in the native structure of

GBC where Cys78 is involved. Thus, the disulfide bond formed during translation has to be reduced and reoxidized and the early formation of a non-native disulfide will catalyze disulfide bond isomerization once the polypeptide chain is released (Figure 4.2). Such catalysis can increase the rate of disulfide bond formation and explain how differences in the rate of translation of proteins expressed from synonymous codons can be preserved in the native state of proteins.

Folding of the NC was studied inside and outside of the ribosomal exit tunnel. Since, the stalling sequence SecM is resistant to puromycin, the NC can only be released by destruction of the ribosome. Investigation of the folding before and after release needs either RNCs produced *in vitro* or a stalling sequence with a free ribosomal A-site, as the TnaC sequence. This stalling sequence was not used for the generation of isotopic labeled RNCs before, and would hence require further optimization. Another method would be the incorporation of the photocleavable 2-Nitrophenylalanine [201] into the SecM or SecMstrong [124] stalling sequence. Upon light irradiation at 365 nm the peptide bound will be cleaved and the NC released. This would be a very elegant and interesting method to observe protein folding before and after the release of the ribosome.

References

[1] Adams, P. D., P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart, 2010: PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, **66 (2)**, 213–221, doi:10.1107/S0907444909052925.

[2] Alcock, L. J., M. V. Perkins, and J. M. Chalker, 2018: Chemical methods for mapping cysteine oxidation. *Chemical Society Reviews*, **47 (1)**, 231–268, doi:10.1039/c7cs00607a.

[3] Anderson, E. I., D. D. Wright, and A. Spector, 1979: The state of sulfhydryl groups in normal and cataractous human lens proteins. II. Cortical and nuclear regions. *Experimental Eye Research*, **29 (3)**, 233–243, doi:10.1016/0014-4835(79)90004-6.

[4] Anfinsen, C. B., 1973: Principles that govern the folding of protein chains. *Science (New York, N.Y.)*, **181 (4096)**, 223–230, doi:10.1126/science.181.4096.223.

[5] Angov, E., C. J. Hillier, R. L. Kincaid, and J. a. Lyon, 2008: Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. *PLoS ONE*, **3 (5)**, 1–10, doi:10.1371/journal.pone.0002189.

[6] Aslund, F., K. D. Berndt, and A. Holmgren, 1997: Redox potentials of glutaredoxins and other thiol-disulfide oxidoreductases of the thioredoxin superfamily determined by direct protein- protein redox equilibria. *Journal of Biological Chemistry*, **272 (49)**, 30 780–30 786, doi:10.1074/jbc.272.49.30780.

[7] Åslund, F., B. Ehn, A. Miranda-Vizuete, C. Pueyo, and A. Holmgren, 1994: Two additional glutaredoxins exist in Escherichia coli: Glutaredoxin 3 is a hydrogen donor for ribonucleotide reductase in a thioredoxin/glutaredoxin 1 double mutant. *Proceedings of the National Academy of Sciences of the United States of America*, **91 (21)**, 9813–9817, doi:10.1073/pnas.91.21.9813.

[8] Åslund, F., M. Zheng, J. Beckwith, and G. Storz, 1999: Regulation of the OxyR transcription factor by hydrogen peroxide and the cellular thiol-disulfide status. *Proceedings of the National Academy of Sciences of the United States of America*, **96 (11)**, 6161–6165, doi:10.1073/pnas.96.11.6161.

[9] Bader, M., W. Muse, D. P. Ballou, C. Gassner, and J. C. Bardwell, 1999: Oxidative protein folding is driven by the electron transport system. *Cell*, **98 (2)**, 217–227, doi:10.1016/S0092-8674(00)81016-8.

[10] Banci, L., L. Barbieri, E. Luchinat, and E. Secci, 2013: Visualization of redox-controlled protein fold in living cells. *Chemistry and Biology*, **20 (6)**, 747–752, doi:10.1016/j.chembiol.2013.05.007.

[11] Bardwell, J. C., J. O. Lee, G. Jander, N. Martin, D. Belin, and J. Beckwith, 1993: A pathway for disulfide bond formation in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, **90 (3)**, 1038–1042, doi:10.1073/pnas.90.3.1038.

[12] Beeby, M., B. D. O'Connor, C. Ryttersgaard, D. R. Boutz, L. J. Perry, and T. O. Yeates, 2005: The genomics of disulfide bonding and protein stabilization in thermophiles. *PLoS Biology*, **3 (9)**, 1549–1558, doi:10.1371/journal.pbio.0030309.

[13] Bergman, L. W. and W. M. Kuehl, 1979: Formation of an intrachain disulfide bond on nascent immunoglobulin light chains. *The Journal of biological chemistry*, **254 (18)**, 8869–76.

[14] Bernado, P., L. Blanchard, P. Timmins, D. Marion, R. W. H. Ruigrok, and M. Blackledge, 2005: A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *Proceedings of the National Academy of Sciences*, **102 (47)**, 17 002–17 007, doi:10.1073/pnas.0506202102.

[15] Berndt, C. and A. Holmgren, 2009: Chapter 1. Oxidative Folding of Proteins in vivo . Chapter 1.1. Thioredoxins and the Regulation of Redox Conditions in Prokaryotes. *RCS Biomolecular Science*, Royal Society of Chemistry, 1–18, doi:10.1039/9781847559265-00001.

[16] Berndt, C., C. H. Lillig, and A. Holmgren, 2008: Thioredoxins and glutaredoxins as facilitators of protein folding. *Biochimica et Biophysica Acta - Molecular Cell Research*, **1783 (4)**, 641–650, doi:10.1016/j.bbamcr.2008.02.003.

[17] Bhushan, S., T. Hoffmann, B. Seidelt, J. Frauenfeld, T. Mielke, O. Berninghausen, D. N. Wilson, and R. Beckmann, 2011: SecM-Stalled Ribosomes Adopt an Altered Geometry at the Peptidyl Transferase Center. *PLoS Biology*, **9 (1)**, e1000 581, doi:10.1371/journal.pbio.1000581.

[18] Bischoff, L., O. Berninghausen, and R. Beckmann, 2014: Molecular Basis for the Ribosome Functioning as an L- Tryptophan Sensor. *Cell Reports*, 1–7, doi:10.1016/j.celrep.2014.09.011.

[19] Blaha, G., U. Stelzl, C. M. T. Spahn, R. K. Agrawal, J. Frank, and K. H. Nierhaus, 2000: [19] Preparation of Functional Ribosomal Complexes and Effect of Buffer Conditions on tRNA Positions Observed by Cryoelectron Microscopy. *Methods in Enzymology*, **317 (1998)**, 292–309, doi:10.1016/S0076-6879(00)17021-1.

[20] Bloemendal, H., W. de Jong, R. Jaenicke, N. H. Lubsen, C. Slingsby, and A. Tardieu, 2004: Ageing and vision: structure, stability and function of lens crystallins. *Progress in biophysics and molecular biology*, **86 (3)**, 407–85, doi:10.1016/j.pbiomolbio.2003.11.012.

[21] Bremer, H. and P. P. Dennis, 2008: Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. *EcoSal Plus*, **3 (1)**, 1–49, doi:10.1128/ecosal.5.2.3.

[22] Buhr, F., S. Jha, M. Thommen, J. Mittelstaet, F. Kutz, H. Schwalbe, M. V. Rodnina, and A. A. Komar, 2016: Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations. *Molecular Cell*, **61 (3)**, 341–351, doi:10.1016/j.molcel.2016.01.008.

[23] Bukau, B., E. Deuerling, C. Pfund, and E. A. Craig, 2000: Getting newly synthesized proteins into shape. *Cell*, **101 (2)**, 119–122, doi:10.1016/S0092-8674(00)80806-5.

[24] Bulleid, N. J. and R. B. Freedman, 1988: Defective co-translational formation of disulphide bonds in protein disulphide-isomerase-deficient microsomes. *Nature*, **335 (6191)**, 649–651, doi:10.1038/335649a0.

[25] Butkus, M. E., L. B. Prundeanu, and D. B. Oliver, 2003: Translocon "Pulling" of Nascent SecM Controls the Duration of Its Translational Pause and Secretion-Responsive secA Regulation. *Journal of Bacteriology*, **185 (22)**, 6719–6722, doi:10.1128/JB.185.22.6719-6722.2003.

[26] Bystrov, V. F., 1976: Spin-spin coupling and the conformational states of peptide systems. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **10 (2)**, 41–82, doi:10.1016/0079-6565(76)80001-5.

[27] Cabrita, L. D., A. M. E. Cassaignau, H. M. M. Launay, C. A. Waudby, T. Wlodarski, C. Camilloni, M.-E. Karyadi, A. L. Robertson, X. Wang, A. S. Wentink, L. S. Goodsell, C. A. Woolhead, M. Vendruscolo, C. M. Dobson, and J. Christodoulou, 2016: A structural ensemble of a ribosome–nascent chain complex during cotranslational protein folding. *Nature Structural & Molecular Biology*, **23 (4)**, 278–285, doi:10.1038/nsmb.3182.

[28] Cabrita, L. D., S.-T. D. Hsu, H. Launay, C. M. Dobson, and J. Christodoulou, 2009: Probing ribosome-nascent chain complexes produced in vivo by NMR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, **106 (52)**, 22 239–44, doi:10.1073/pnas.0903750106.

[29] Calloni, G., T. Chen, S. M. Schermann, H. C. Chang, P. Genevaux, F. Agostini, G. G. Tartaglia, M. Hayer-Hartl, and F. U. Hartl, 2012: DnaK Functions as a Central Hub in the E. coli Chaperone Network. *Cell Reports*, **1 (3)**, 251–264, doi:10.1016/j.celrep.2011.12.007.

[30] Cassaignau, A. M. E., H. M. M. Launay, M.-E. Karyadi, X. Wang, C. A. Waudby, A. Deckert, A. L. Robertson, J. Christodoulou, and L. D. Cabrita, 2016: A strategy for co-translational folding studies of ribosome-bound nascent chain complexes using NMR spectroscopy. *Nature Protocols*, **11 (8)**, 1492–1507, doi:10.1038/nprot.2016.101.

[31] Ceroni, A., A. Passerini, A. Vullo, and P. Frasconi, 2006: Disulfind: A disulfide bonding state and cysteine connectivity prediction server. *Nucleic Acids Research*, **34 (WEB. SERV. ISS.)**, 177–181, doi:10.1093/nar/gkl266.

[32] Chen, Y. R., C. L. Chen, D. R. Pfeiffer, and J. L. Zweier, 2007: Mitochondrial complex II in the post-ischemic heart: Oxidative injury and the role of protein S-glutathionylation. *Journal of Biological Chemistry*, **282 (45)**, 32 640–32 654, doi:10.1074/jbc.M702294200.

[33] Clark, P. L., 2004: Protein folding in the cell: Reshaping the folding funnel. *Trends in Biochemical Sciences*, **29 (10)**, 527–534, doi:10.1016/j.tibs.2004. 08.008.

[34] Clark, P. L. and J. King, 2001: A Newly Synthesized, Ribosome-bound Polypeptide Chain Adopts Conformations Dissimilar from Early in Vitro Refolding Intermediates. *Journal of Biological Chemistry*, **276 (27)**, 25 411–25 420, doi:10.1074/jbc.M008490200.

[35] Clubb, R. T. and G. Wagner, 1992: A triple-resonance pulse scheme for selectively correlating amide1HN and15N nuclei with the1Hα proton of the preceding residue. *Journal of Biomolecular NMR*, **2 (4)**, 389–394, doi:10. 1007/BF01874816.

[36] Craig, D. B. and A. A. Dombkowski, 2013: Disulfide by Design 2.0: A web-based tool for disulfide engineering in proteins. *BMC Bioinformatics*, **14 (1)**, 0–6, doi:10.1186/1471-2105-14-346.

[37] Creighton, T. E., 1990: Protein folding. *Biochemical Journal*, **270 (1)**, 1–16, doi:10.1042/bj2700001.

[38] Crombie, T., J. P. Boyle, J. R. Coggins, and A. J. Brown, 1994: The Folding of the Bifunctional TRP3 Protein in Yeast is Influenced by a Translational Pause which Lies in a Region of Structural Divergence with Escherichia coli Indoleglycerol-Phosphate Synthase. *European Journal of Biochemistry*, **226 (2)**, 657–664, doi:10.1111/j.1432-1033.1994.tb20093.x.

[39] Daggett, V. and A. R. Fersht, 2003: Is there a unifying mechanism for protein folding? *Trends in Biochemical Sciences*, **28 (1)**, 18–25, doi:10. 1016/S0968-0004(02)00012-9.

[40] Dalle-Donne, I., A. Milzani, N. Gagliano, R. Colombo, D. Giustarini, and R. Rossi, 2008: Molecular mechanisms and potential clinical significance of S-glutathionylation. *Antioxidants and Redox Signaling*, **10 (3)**, 445–473, doi:10.1089/ars.2007.1716.

[41] Dalle-Donne, I., R. Rossi, G. Colombo, D. Giustarini, and A. Milzani, 2009: Protein S-glutathionylation: a regulatory device from bacteria to humans. *Trends in Biochemical Sciences*, **34 (2)**, 85–96, doi:10.1016/j.tibs.2008.11. 002.

[42] Dani, V. S., C. Ramakrishnan, and R. Varadarajan, 2003: MODIP revisited: Re-evaluation and refinement of an automated procedure for modeling of disulfide bonds in proteins. *Protein Engineering*, **16 (3)**, 187–193, doi: 10.1093/proeng/gzg024.

[43] Deuerling, E., A. Schulze-Specking, T. Tomoyasu, A. Mogk, and B. Bukau, 1999: Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature*, **400 (6745)**, 693–696, doi:10.1038/23301.

[44] Diambra, L. A., 2017: Differential bicodon usage in lowly and highly abundant proteins. *PeerJ*, **5 (2014)**, e3081, doi:10.7717/peerj.3081.

[45] Dombkowski, A. A., K. Z. Sultana, and D. B. Craig, 2014: Protein disulfide engineering. *FEBS Letters*, **588 (2)**, 206–212, doi:10.1016/j.febslet.2013.11. 024.

[46] Duc, K. D., S. S. Batra, N. Bhattacharya, J. H. D. Cate, and Y. S. Song, 2019: Differences in the path to exit the ribosome across the three domains of life. *Nuclei*, 1–13, doi:10.1093/nar/gkz106.

[47] Dyson, H. J., M. F. Jeng, L. L. Tennant, I. Slaby, M. Lindell, D. S. Cui, S. Kuprin, and A. Holmgren, 1997: Effects of buried charged groups on cysteine thiol ionization and reactivity in Escherichia coli thioredoxin: Structural and functional characterization of mutants of Asp 26 and Lys 57. *Biochemistry*, **36 (9)**, 2622–2636, doi:10.1021/bi961801a.

[48] Ederth, J., C. S. Mandava, S. Dasgupta, and S. Sanyal, 2009: A single-step method for purification of active His-tagged ribosomes from a genetically engineered Escherichia coli. *Nucleic Acids Research*, **37 (2)**, doi:10.1093/nar/gkn992.

[49] Eichmann, C., S. Preissler, R. Riek, and E. Deuerling, 2010: Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, **107 (20)**, 9111–6, doi:10.1073/pnas.0914300107.

[50] Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan, 2010: Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, **66 (4)**, 486–501, doi:10.1107/S0907444910007493, `0402594v3`.

[51] Evans, M. S., I. M. Sander, and P. L. Clark, 2008: Cotranslational Folding Promotes β-Helix Formation and Avoids Aggregation In Vivo. *Journal of Molecular Biology*, **383 (3)**, 683–692, doi:10.1016/j.jmb.2008.07.035.

[52] Ewalt, K. L., J. P. Hendrick, W. A. Houry, and F. U. Hartl, 1997: In vivo observation of polypeptide flux through the bacterial chaperonin system. *Cell*, **90 (3)**, 491–500, doi:10.1016/S0092-8674(00)80509-7.

[53] Fahey, R. C., 2013: Glutathione analogs in prokaryotes. *Biochimica et Biophysica Acta - General Subjects*, **1830 (5)**, 3182–3198, doi:10.1016/j.bbagen.2012.10.006.

[54] Fahey, R. C., W. C. Brown, W. B. Adams, and M. B. Worsham, 1978: Occurrence of glutathione in bacteria. *Journal of Bacteriology*, **133 (3)**, 1126–1129.

[55] Fahey, R. C., J. S. Hunt, and G. C. Windham, 1977: On the cysteine and cystine content of proteins - Differences between Intracellular and Extracellular Proteins. *Journal of Molecular Evolution*, **10 (2)**, 155–160, doi:10.1007/BF01751808.

[56] Farías-Rico, J. A., F. Ruud Selin, I. Myronidi, M. Frühauf, and G. von Heijne, 2018: Effects of protein size, thermodynamic stability, and net charge on cotranslational folding on the ribosome. *Proceedings of the National Academy of Sciences*, **115 (40)**, E9280–E9287, doi:10.1073/pnas.1812756115.

[57] Fass, D. and C. Thorpe, 2017: Chemistry and Enzymology of Disulfide Cross-Linking in Proteins. *Chemical Reviews*, acs.chemrev.7b00123, doi:10.1021/acs.chemrev.7b00123.

[58] Fedorov, A. N. and T. O. Baldwin, 1999: Process of biosynthetic protein folding determines the rapid formation of native structure. *Journal of Molecular Biology*, **294 (2)**, 579–586, doi:10.1006/jmbi.1999.3281.

[59] Fernandes, A. P., M. Fladvad, C. Berndt, C. Andrésen, C. H. Lillig, P. Neubauer, M. Sunnerhagen, A. Holmgren, and A. Vlamis-Gardikas, 2005: A novel monothiol glutaredoxin (Grx4) from Escherichia coli can serve as a substrate for thioredoxin reductase. *Journal of Biological Chemistry*, **280 (26)**, 24 544–24 552, doi:10.1074/jbc.M500678200.

[60] Ferrè, F. and P. Clote, 2005: DiANNA: A web server for disulfide connectivity prediction. *Nucleic Acids Research*, **33 (SUPPL. 2)**, 230–232, doi:10.1093/nar/gki412.

[61] Ferrè, F. and P. Clote, 2005: Disulfide connectivity prediction using secondary structure information and diresidue frequencies. *Bioinformatics*, **21 (10)**, 2336–2346, doi:10.1093/bioinformatics/bti328.

[62] Ferrè, F. and P. Clote, 2006: DiANNA 1.1: An extension of the DiANNA web server for ternary cysteine classification. *Nucleic Acids Research*, **34 (WEB. SERV. ISS.)**, 182–185, doi:10.1093/nar/gkl189.

[63] Fersht, A. R., 2000: Transition-state structure as a unifying basis in protein-folding mechanisms: Contact order, chain topology, stability, and the extended nucleus mechanism. *Proceedings of the National Academy of Sciences*, **97 (4)**, 1525–1529, doi:10.1073/pnas.97.4.1525.

[64] Frand, A. R. and C. A. Kaiser, 1998: The ERO1 gene of yeast is required for oxidation of protein dithiols in the endoplasmic reticulum. *Molecular Cell*, **1 (2)**, 161–170, doi:10.1016/S1097-2765(00)80017-9.

[65] Frand, A. R. and C. A. Kaiser, 1999: Ero1p oxidizes protein disulfide isomerase in a pathway for disulfide bond formation in the endoplasmic reticulum. *Molecular Cell*, **4 (4)**, 469–477, doi:10.1016/S1097-2765(00)80198-7.

[66] Frand, A. R. and C. A. Kaiser, 2000: Two pairs of conserved cysteines are required for the oxidative activity of Ero1p in protein disulfide bond formation in the endoplasmic reticulum. *Molecular Biology of the Cell*, **11 (9)**, 2833–2843, doi:10.1091/mbc.11.9.2833.

[67] Frank, J. and R. K. Agrawal, 2000: A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature*, **406 (6793)**, 318–322, doi:10.1038/35018597.

[68] Frank, J., J. Zhu, P. Penczek, Y. Li, S. Srivastava, A. Verschoor, M. Radermacher, R. Grassucci, R. K. Lata, and R. K. Agrawal, 1995: A model of protein synthesis based on cryo-electron microscopy of the E. coli ribosome. *Nature*, **376 (6539)**, 441–444, doi:10.1038/376441a0.

[69] Frolova, L. Y. U., R. Y. U. Tsivkovskii, G. F. Sivolobova, N. Y. U. Oparina, O. I. Serpinsky, V. M. Blinov, S. I. Tatkov, and L. E. V. L. Kisselev, 1999: Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. *RNA*, **5**, 1014–1020, doi:10.1017/s135583829999043x.

[70] Frydman, J., 2001: Folding of Newly Translated Proteins In Vivo: The Role of Molecular Chaperones. *Annual Review of Biochemistry*, **70 (1)**, 603–647, doi:10.1146/annurev.biochem.70.1.603.

[71] Frydman, J., H. Erdjument-Bromage, P. Tempst, and F. U. Hartl, 1999: Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase. *Nat. Struct. Mol. Biol.*, **6 (7)**, 697–705, doi:10.1038/10754.

[72] Fürtig, B., C. Richter, J. Wöhnert, and H. Schwalbe, 2003: NMR spectroscopy of RNA. *ChemBioChem*, **4 (10)**, 936–962, doi:10.1002/cbic.200300700.

[73] Gamble, C. E., C. E. Brule, K. M. Dean, S. Fields, and E. J. Grayhack, 2016: Adjacent Codons Act in Concert to Modulate Translation Efficiency in Yeast. *Cell*, **166 (3)**, 679–690, doi:10.1016/j.cell.2016.05.070.

[74] Gao, H., Z. Zhou, U. Rawat, C. Huang, L. Bouakaz, C. Wang, J. Frank, and H. Song, 2007: RF3 Induces Ribosomal Conformational Changes Responsible for Dissociation of Class I Release Factors. *Cell*, **129**, 929–941, doi:10.1016/j.cell.2007.03.050.

[75] Gelis, I., V. Vitzthum, N. Dhimole, M. a. Caporini, A. Schedlbauer, D. Carnevale, S. R. Connell, P. Fucini, and G. Bodenhausen, 2013: Solid-state NMR enhanced by dynamic nuclear polarization as a novel tool for ribosome structural biology. *Journal of Biomolecular NMR*, **56 (2)**, 85–93, doi:10.1007/s10858-013-9721-2.

[76] Giustarini, D., A. Milzani, G. Aldini, M. Carini, R. Rossi, and I. Dalle-Donne, 2005: S -Nitrosation versus S -Glutathionylation of Protein Sulfhydryl Groups by S -Nitrosoglutathione. *Antioxidants & Redox Signaling*, **7 (7-8)**, 930–939, doi:10.1089/ars.2005.7.930.

[77] Goddard, T. D., Kneller, D. G., 2008: SPARKY 3. University of California, San Francisco.

[78] Goldstein, S. and G. Czapski, 1996: Mechanism of the nitrosation of thiols and amines by oxygenated $\bullet$NO solutions: The nature of the nitrosating intermediates. *Journal of the American Chemical Society*, **118 (14)**, 3419–3425, doi:10.1021/ja965410v.

[79] Gong, F. and C. Yanofsky, 2002: Analysis of tryptophanase operon expression in vitro. Accumulation of TnaC-peptidyl-tRNA in a release factor 2-depleted S-30 extract prevents Rho factor action, simulating induction. *Journal of Biological Chemistry*, **277 (19)**, 17 095–17 100, doi:10.1074/jbc.M201213200.

[80] Grassucci, R. A., D. J. Taylor, and J. Frank, 2007: Preparation of macromolecular complexes for cryo-electron microscopy. *Nature Protocols*, **2 (12)**, 3239–3246, doi:10.1038/nprot.2007.452.

[81] Guo, F. B., Y. N. Ye, H. L. Zhao, D. Lin, and W. Wei, 2012: Universal pattern and diverse strengths of successive synonymous codon bias in three domains of life, particularly among prokaryotic genomes. *DNA Research*, **19 (6)**, 477–485, doi:10.1093/dnares/dss027.

[82] Haendeler, J., 2006: Thioredoxin-1 and Posttranslational Modifications. *Antioxidants & Redox Signaling*, **8 (9-10)**, 1723–1728, doi:10.1089/ars. 2006.8.1723.

[83] Hamlin, J. and I. Zabin, 1972: -Galactosidase: immunological activity of ribosome-bound, growing polypeptide chains. *Proceedings of the National Academy of Sciences of the United States of America*, **69 (2)**, 412–416, doi:10.1073/pnas.69.2.412.

[84] Hardesty, B. and G. Kramer, 2000: Folding of a nascent peptide on the ribosome. *Progress in Nucleic Acid Research and Molecular Biology*, **66**, 41–66, doi:10.1016/s0079-6603(00)66026-9.

[85] Harms, J., F. Schluenzen, R. Zarivach, A. Bashan, S. Gat, I. Agmon, H. Bartels, F. Franceschi, and A. Yonath, 2001: High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell*, **107 (5)**, 679–688, doi:10.1016/S0092-8674(01)00546-3.

[86] Hayer-Hartl, M., A. Bracher, and F. U. Hartl, 2016: The GroEL-GroES Chaperonin Machine: A Nano-Cage for Protein Folding. *Trends in Biochemical Sciences*, **41 (1)**, 62–76, doi:10.1016/j.tibs.2015.07.009.

[87] Helmann, J. D., 2002: OxyR: a molecular code for redox sensing? *Science's STKE : signal transduction knowledge environment*, **2002 (157)**, 1–5, doi: 10.1126/stke.2002.157.pe46.

[88] Hess, D. T., A. Matsumoto, S. O. Kim, H. E. Marshall, and J. S. Stamler, 2005: Protein S-nitrosylation: Purview and parameters. *Nature Reviews Molecular Cell Biology*, **6 (2)**, 150–166, doi:10.1038/nrm1569.

[89] Hing, A. W., S. Vega, and J. Schaefer, 1992: Transferred-echo double-resonance NMR. *Journal of Magnetic Resonance (1969)*, **96 (1)**, 205–209, doi:10.1016/0022-2364(92)90305-Q.

[90] Hoffmann, A., A. H. Becker, B. Zachmann-Brand, E. Deuerling, B. Bukau, and G. Kramer, 2012: Concerted Action of the Ribosome and the Associated Chaperone Trigger Factor Confines Nascent Polypeptide Folding. *Molecular Cell*, **48 (1)**, 63–74, doi:10.1016/j.molcel.2012.07.018.

[91] Hoffmann, A., A. H. Becker, B. Zachmann-Brand, E. Deuerling, B. Bukau, and G. Kramer, 2012: Concerted Action of the Ribosome and the Associated Chaperone Trigger Factor Confines Nascent Polypeptide Folding. *Molecular Cell*, **48 (1)**, 63–74, doi:10.1016/j.molcel.2012.07.018.

[92] Hogg, N., 1999: The Kinetics of S-Transnitrosation—A Reversible Second-Order Reaction. *Analytical Biochemistry*, **272 (2)**, 257–262, doi:10.1006/ abio.1999.4199.

[93] Hogg, P. J., 2013: Targeting allosteric disulphide bonds in cancer. *Nature Reviews Cancer*, **13 (6)**, 425–431, doi:10.1038/nrc3519.

[94] Hohwy, M., H. J. Jakobsen, M. Edén, M. H. Levitt, and N. C. Nielsen, 1998: Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: A compensated C7 pulse sequence. *Journal of Chemical Physics*, **108 (7)**, 2686–2694, doi:10.1063/1.475661.

[95] Holmgren, A., 1979: Glutathione-dependent synthesis of deoxyribonucleotides. *J. Biol. Chem*, **254 (9)**, 3664–3671.

[96] Holmgren, A., 1979: Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. *The Journal of biological chemistry*, **254 (19)**, 9627–32, doi:10.10.79.

[97] Holmgren, A., 1995: Thioredoxin structure and mechanism: conformational changes on oxidation of the active-site sulfhydryls to a disulfide. *Structure*, **3 (3)**, 239–243, doi:10.1016/S0969-2126(01)00153-8.

[98] Holmgren, A., C. Johansson, C. Berndt, M. E. Lönn, C. Hudemann, and C. H. Lillig, 2005: Thiol redox control via thioredoxin and glutaredoxin systems. *Biochemical Society Transactions*, **33 (6)**, 1375–1377, doi:10.1042/BST20051375.

[99] Holmgren, A. and M. Luthman, 1978: Tissue Distribution and Subcellular Localization of Bovine Thioredoxin Determined by Radioimmunoassay†. *Biochemistry*, **17 (19)**, 4071–4077, doi:10.1021/bi00612a031.

[100] Holmgren, A., B. O. Soderberg, H. Eklund, and C. I. Branden, 1975: Three dimensional structure of Escherichia coli thioredoxin S2 to 2.8 Å resolution. *Proceedings of the National Academy of Sciences of the United States of America*, **72 (6)**, 2305–2309, doi:10.1073/pnas.72.6.2305.

[101] Holtkamp, W., G. Kokic, M. Jäger, J. Mittelstaet, A. A. Komar, and M. V. Rodnina, 2015: Cotranslational protein folding on the ribosome monitored in real time. *Science*, **350 (6264)**, 1104–1107, doi:10.1126/science.aad0344, `arXiv:1011.1669v3`.

[102] Hsu, S. T. D., L. D. Cabrita, P. Fucini, J. Christodoulou, and C. M. Dobson, 2009: Probing side-chain dynamics of a ribosome-bound nascent chain using methyl NMR spectroscopy. *Journal of the American Chemical Society*, **131 (24)**, 8366–8367, doi:10.1021/ja902778n.

[103] Hsu, S.-T. D., P. Fucini, L. D. Cabrita, H. Launay, C. M. Dobson, and J. Christodoulou, 2007: Structure and dynamics of a ribosome-bound nascent chain by NMR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, **104 (42)**, 16 516–21, doi:10.1073/pnas.0704664104.

[104] Ikemura, T., 1985: Codon usage and tRNA content in unicellular and multicellular organisms. *Molecular Biology and Evolution*, **2**, 13–34, doi:10.1093/oxfordjournals.molbev.a040335.

[105] Ishii, E., S. Chiba, N. Hashimoto, S. Kojima, M. Homma, K. Ito, Y. Akiyama, and H. Mori, 2015: Nascent chain-monitored remodeling of the Sec machinery for salinity adaptation of marine bacteria. *Proceedings of the National Academy of Sciences*, **112 (51)**, 201513 001, doi:10.1073/pnas.1513001112.

[106] Ismail, N., R. Hedman, N. Schiller, and G. Von Heijne, 2012: A biphasic pulling force acts on transmembrane helices during translocon-mediated membrane integration. *Nature Structural and Molecular Biology*, **19 (10)**, 1018–1023, doi:10.1038/nsmb.2376.

[107] Ito, K. and S. Chiba, 2013: Arrest Peptides: Cis -Acting Modulators of Translation. *Annual Review of Biochemistry*, **82 (1)**, 171–202, doi:10.1146/annurev-biochem-080211-105026.

[108] Ito, K., S. Chiba, and K. Pogliano, 2010: Divergent stalling sequences sense and control cellular physiology. *Biochemical and Biophysical Research Communications*, **393 (1)**, 1–5, doi:10.1016/j.bbrc.2010.01.073.

[109] Ito, K., M. Uno, and Y. Nakamura, 2000: A tripeptide 'anticodon' deciphers stop codons in messenger RNA. *Nature*, **403 (6770)**, 680–684, doi:10.1038/35001115.

[110] Jacobson, G. N. and P. L. Clark, 2016: Quality over quantity: Optimizing co-translational protein folding with non-'optimal' synonymous codons. *Current Opinion in Structural Biology*, **38**, 102–110, doi:10.1016/j.sbi.2016.06.002.

[111] Jakob, U., W. Muse, M. Eser, and J. C. Bardwell, 1999: Chaperone activity with a redox switch. *Cell*, **96 (3)**, 341–352, doi:10.1016/S0092-8674(00)80547-4.

[112] Jander, G., N. Martin, and J. Beckwith, 1994: Two cysteines in each periplasmic domain of the membrane protein DsbB are required for its function in protein disulfide bond formation. *The EMBO Journal*, **13 (21)**, 5121–5127, doi:10.1002/j.1460-2075.1994.tb06841.x.

[113] Jansens, A., E. Van Duijn, and I. Braakman, 2002: Coordinated non-vectorial folding in a newly synthesized multidomain protein. *Science*, **298 (5602)**, 2401–2403, doi:10.1126/science.1078376.

[114] Javed, A., L. D. Cabrita, A. M. Cassaignau, T. Wlodarski, J. Christodoulou, and E. V. Orlova, 2019: Visualising nascent chain dynamics at the ribosome exit tunnel by cryo-electron microscopy. *bioRxiv*, 722611, doi:10.1101/722611.

[115] Javed, A., J. Christodoulou, L. D. Cabrita, and E. V. Orlova, 2017: The ribosome and its role in protein folding: Looking through a magnifying glass. *Acta Crystallographica Section D: Structural Biology*, **73 (6)**, 509–521, doi:10.1107/S2059798317007446.

[116] Joedicke, L., J. Mao, G. Kuenze, C. Reinhart, T. Kalavacherla, H. R. Jonker, C. Richter, H. Schwalbe, J. Meiler, J. Preu, C. Glaubitz, and H. Michel, 2018: The molecular basis of subtype selectivity of human kinin G-protein-coupled receptors. *Nature Chemical Biology*, **14 (3)**, 284–290, doi:10.1038/nchembio.2551.

[117] Johnson, A. E., 2005: The co-translational folding and interactions of nascent protein chains: A new approach using fluorescence resonance energy transfer. *FEBS Letters*, **579 (4 SPEC. ISS.)**, 916–920, doi:10.1016/j.febslet.2004.11.046.

[118] Jorda, J. and T. O. Yeates, 2011: Widespread disulfide bonding in proteins from thermophilic archaea. *Archaea*, **2011**, doi:10.1155/2011/409156.

[119] Kadokura, H., H. Tian, T. Zander, J. C. Bardwell, and J. Beckwith, 2004: Snapshots of DsbA in Action: Detection of Proteins in the Process of Oxidative Folding. *Science*, **303 (5657)**, 534–537, doi:10.1126/science.1091724.

[120] Kallis, G. B. and A. Holmgren, 1980: Differential reactivity of the functional sulfhydryl groups of cysteine-32 and cysteine-32 present in the reduced form of thioredoxin from Escherichia coli. *Journal of Biological Chemistry*, **255 (21)**, 10 261–10 265.

[121] Keiler, K. C., 2015: Mechanisms of ribosome rescue in bacteria. *Nature Reviews Microbiology*, **13 (5)**, 285–297, doi:10.1038/nrmicro3438.

[122] Kelkar, D. A., A. Khushoo, Z. Yang, and W. R. Skach, 2012: Kinetic analysis of ribosome-bound fluorescent proteins reveals an early, stable, cotranslational folding intermediate. *Journal of Biological Chemistry*, **287 (4)**, 2568–2578, doi:10.1074/jbc.M111.318766.

[123] Keller, R., 2004: *The computer aided resonance assignment tutorial.* Cantina Verlag, 1–81 pp.

[124] Kempf, N., C. Remes, R. Ledesch, T. Züchner, H. Höfig, I. Ritter, A. Katranidis, and J. Fitter, 2017: A novel method to evaluate ribosomal performance in cell-free protein synthesis systems. *Scientific Reports*, **7 (April)**, 1–9, doi:10.1038/srep46753.

[125] Kim, S. O., K. Merchant, R. Nudelman, W. F. Beyer, T. Keng, J. DeAngelo, A. Hausladen, and J. S. Stamler, 2002: OxyR. *Cell*, **109 (3)**, 383–396, doi:10.1016/s0092-8674(02)00723-7.

[126] Kimchi-Sarfaty, C., J. M. Oh, I.-W. Kim, Z. E. Sauna, A. M. Calcagno, S. V. Ambudkar, and M. M. Gottesman, 2007: A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science (New York, N.Y.)*, **315 (5811)**, 525–528, doi:10.1126/science.1135308.

[127] Klappa, P., R. B. Freedman, and R. Zimmermann, 1995: Protein Disulphide Isomerase and a Lumenal Cyclophilin-Type Peptidyl Prolyl Cis-Trans Isomerase are in Transient Contact with Secretory Proteins During Late Stages of Translocation. *European Journal of Biochemistry*, **232 (3)**, 755–764, doi:10.1111/j.1432-1033.1995.0755a.x.

[128] Kolb, V., E. Makeyev, and A. Spirin, 1994: Folding of firefly luciferase during translation in a cell-free system. *The EMBO Journal*, **13 (15)**, 3631–3637, doi:10.1002/j.1460-2075.1994.tb06670.x.

[129] Komar, A. a., 2009: A pause for thought along the co-translational folding pathway. *Trends in biochemical sciences*, **34 (1)**, 16–24, doi:10.1016/j.tibs.2008.10.002.

[130] Komar, A. A., 2018: Unraveling co-translational protein folding: Concepts and methods. *Methods*, **137**, 71–81, doi:10.1016/j.ymeth.2017.11.007.

[131] Komar, A. A. and R. Jaenicke, 1995: Kinetics of translation of γB crystallin and its circularly permutated variant in an in vitro cell-free system: possible relations to codon distribution and protein folding. *FEBS letters*, **376**, 195–198, doi:10.1016/0014-5793(95)01275-0.

[132] Komar, A. A., A. Kommer, I. A. Krasheninnikov, and A. S. Spirin, 1993: Cotranslational heme binding to nascent globin chains. *FEBS Letters*, **326 (1-3)**, 261–263, doi:10.1016/0014-5793(93)81803-8.

[133] Komar, A. A., A. Kommer, I. A. Krasheninnikov, and A. S. Spirin, 1997: Cotranslational folding of globin. *Journal of Biological Chemistry*, **272 (16)**, 10 646–10 651, doi:10.1074/jbc.272.16.10646.

[134] Komar, A. A., T. Lesnik, and C. Reiss, 1999: Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. *FEBS Letters*, **462 (3)**, 387–391, doi:10.1016/S0014-5793(99)01566-5.

[135] Kosolapov, A. and C. Deutsch, 2009: Tertiary interactions within the ribosomal exit tunnel. *Nature structural & molecular biology*, **16 (4)**, 405–11, doi:10.1038/nsmb.1571.

[136] Kosuri, P., J. Alegre-Cebollada, J. Feng, A. Kaplan, A. Inglés-Prieto, C. L. Badilla, B. R. Stockwell, J. M. Sanchez-Ruiz, A. Holmgren, and J. M. Fernández, 2012: Protein folding drives disulfide formation. *Cell*, **151 (4)**, 794–806, doi:10.1016/j.cell.2012.09.036.

[137] Kramer, G., T. Rauch, W. Rist, S. Vorderwülbecke, H. Patzelt, A. Schulze-Specking, N. Ban, E. Deuerling, and B. Bukau, 2002: L23 protein functions as a chaperone docking site on the ribosome. *Nature*, **419 (6903)**, 171–174, doi:10.1038/nature01047.

[138] Kramer, G., A. Shiber, and B. Bukau, 2019: Mechanisms of Cotranslational Maturation of Newly Synthesized Proteins. *Annual Review of Biochemistry*, **88 (1)**, 337–364, doi:10.1146/annurev-biochem-013118-111717.

[139] Krause, G., J. Lundstrom, J. L. Barea, C. P. De la Cuesta, and A. Holmgren, 1991: Mimicking the active site of protein disulfide-isomerase by substitution of proline 34 in Escherichia coli thioredoxin. *Journal of Biological Chemistry*, **266 (15)**, 9494–9500.

[140] Kudva, R., H. Sandhu, M. Carroni, H. D. Bernstein, S. Biology, B. Branch, K. Diseases, and N. Institutes, 2018: The Shape of the Ribosome Exit Tunnel Affects Cotranslational Protein Folding. *eLife*, 1–15, doi:10.7554/eLife.36326.

[141] Kurland, C. G., 1960: Molecular characterization of ribonucleic acid from Escherichia coli ribosomes: I. Isolation and molecular weights. *Journal of Molecular Biology*, **2 (2)**, 83–91, doi:10.1016/S0022-2836(60)80029-0.

[142] Land, A., D. Zonneveld, and I. Braakman, 2003: Folding of HIV-1 Envelope glycoprotein involves extensive isomerization of disulfide bonds and conformation-dependent leader peptide cleavage. *The FASEB Journal*, **17 (9)**, 1058–1067, doi:10.1096/fj.02-0811com.

[143] Landeta, C., D. Boyd, and J. Beckwith, 2018: Disulfide bond formation in prokaryotes. *Nature Microbiology*, **3 (3)**, 270–280, doi:10.1038/s41564-017-0106-2.

[144] Lange, S., W. T. Franks, N. Rajagopalan, K. Döring, M. A. Geiger, A. Linden, B.-J. van Rossum, G. Kramer, B. Bukau, and H. Oschkinat, 2016: Structural analysis of a signal peptide inside the ribosome

tunnel by DNP MAS NMR. *Science advances*, **2 (8)**, e1600 379, doi: 10.1126/sciadv.1600379.

[145] Lange, S., A. H. Linden, Ü. Akbey, W. Trent Franks, N. M. Loening, B. J. V. Rossum, and H. Oschkinat, 2012: The effect of biradical concentration on the performance of DNP-MAS-NMR. *Journal of Magnetic Resonance*, **216**, 209–212, doi:10.1016/j.jmr.2012.01.002.

[146] LAURENT, T. C., E. C. MOORE, and P. REICHARD, 1964: Enzymatic Synthesis of Deoxyribonucleotides. IV. Isolation and Characterization of thioredoxin, the hydrogen donor from Escherichia coli B. *The Journal of biological chemistry*, **239 (10)**, 3436–44.

[147] Levitt, M., 1976: A simplified representation of protein conformations for rapid simulation of protein folding. *Journal of Molecular Biology*, **104 (1)**, 59–107, doi:10.1016/0022-2836(76)90004-8.

[148] Liang, J. N. and M. R. Pelletier, 1988: Destabilization of lens protein conformation by glutathione mixed disulfide. *Experimental Eye Research*, **47 (1)**, 17–25, doi:10.1016/0014-4835(88)90020-6.

[149] Loewen, P. C., 1979: Levels of glutathione in Escherichia coli. *Canadian Journal of Biochemistry*, **57 (2)**, 107–111, doi:10.1139/o79-013.

[150] Lovell, S. C., I. W. Davis, W. B. Arendall, P. I. De Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson, 2003: Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. *Proteins: Structure, Function and Genetics*, **50 (3)**, 437–450, doi:10.1002/prot.10286.

[151] Lu, J. and C. Deutsch, 2001: Pegylation: A method for assessing topological accessibilities in Kv 1.3. *Biochemistry*, **40 (44)**, 13 288–13 301, doi:10.1021/bi0107647.

[152] Lu, J. and C. Deutsch, 2005: Folding zones inside the ribosomal exit tunnel. *Nature Structural and Molecular Biology*, **12 (12)**, 1123–1129, doi:10.1038/nsmb1021.

[153] Lu, J. and C. Deutsch, 2005: Secondary structure formation of a transmembrane segment in Kv channels. *Biochemistry*, **44 (23)**, 8230–8243, doi:10.1021/bi050372q.

[154] Lu, J. and C. Deutsch, 2008: Electrostatics in the Ribosomal Tunnel Modulate Chain Elongation Rates. *Journal of Molecular Biology*, **384 (1)**, 73–86, doi:10.1016/j.jmb.2008.08.089.

[155] Lu, J., W. R. Kobertz, and C. Deutsch, 2007: Mapping the Electrostatic Potential within the Ribosomal Exit Tunnel. *Journal of Molecular Biology*, **371 (5)**, 1378–1391, doi:10.1016/j.jmb.2007.06.038.

[156] Lv, J. M., S. Q. Lü, Z. P. Liu, J. Zhang, B. X. Gao, Z. Y. Yao, Y. X. Wu, L. A. Potempa, S. R. Ji, M. Long, and Y. Wu, 2018: Conformational folding and disulfide bonding drive distinct stages of protein structure formation. *Scientific Reports*, **8 (1)**, 1–10, doi:10.1038/s41598-018-20014-y.

[157] Madej, E., L. K. Folkes, P. Wardman, G. Czapski, and S. Goldstein, 2008: Thiyl radicals react with nitric oxide to form S-nitrosothiols with rate constants near the diffusion-controlled limit. *Free Radical Biology and Medicine*, **44 (12)**, 2013–2018, doi:10.1016/j.freeradbiomed.2008.02.015.

[158] Mamathambika, B. S. and J. C. Bardwell, 2008: Disulfide-Linked Protein Folding Pathways. *Annual Review of Cell and Developmental Biology*, **24 (1)**, 211–235, doi:10.1146/annurev.cellbio.24.110707.175333.

[159] Mandal, K., S. K. Bose, B. Chakrabarti, and R. J. Siezen, 1987: Structure and stabilty of gamma-crystallins. II. Differences in microenvironments and spatial arrangements of cysteine residues. *Biochimica et biophysica acta*, **911**, 277–284.

[160] Manting, E. H. and A. J. Driessen, 2000: Escherichia coli translocase: The unravelling of a molecular machine. *Molecular Microbiology*, **37 (2)**, 226–238, doi:10.1046/j.1365-2958.2000.01980.x.

[161] Marino, J., G. von Heijne, and R. Beckmann, 2016: Small protein domains fold inside the ribosome exit tunnel. *FEBS Letters*, **590 (5)**, 655–660, doi:10.1002/1873-3468.12098.

[162] Mastronarde, D. N., 2005: Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology*, **152 (1)**, 36–51, doi:10.1016/j.jsb.2005.07.007.

[163] Matiollo, C., G. Ecco, A. C. O. Menegatti, G. Razzera, J. Vernal, and H. Terenzi, 2013: S-nitrosylation of Mycobacterium tuberculosis tyrosine phosphatase A (PtpA) induces its structural instability. *Biochimica et Biophysica Acta - Proteins and Proteomics*, **1834 (1)**, 191–196, doi:10.1016/j.bbapap.2012.10.007.

[164] Matsumura, M., W. J. Becktel, M. Levitt, and B. W. Matthews, 1989: Stabilization of phage T4 lysozyme by engineered disulfide bonds. *Proceedings of the National Academy of Sciences of the United States of America*, **86 (17)**, 6562–6566, doi:10.1073/pnas.86.17.6562.

[165] Matsumura, M., G. Signor, and B. W. Matthews, 1989: Substantial increase of protein stability by multiple disulphide bonds. *Nature*, **342 (6247)**, 291–293, doi:10.1038/342291a0.

[166] Mayr, E. M., R. Jaenicke, and R. Glockshuber, 1994: Domain interactions and connecting peptides in lens crystallins. *Journal of molecular biology*, **235 (1)**, 84–88, doi:10.1016/S0022-2836(05)80017-8.

[167] Mayr, E. M., R. Jaenicke, and R. Glockshuber, 1997: The domains in gammaB-crystallin: identical fold-different stabilities. *Journal of molecular biology*, **269 (2)**, 260–269, doi:10.1006/jmbi.1997.1033.

[168] McCarthy, A. A., P. W. Haebel, A. Törrönen, V. Rybin, E. N. Baker, and P. Metcalf, 2000: Crystal structure of the protein disulfide bond isomerase, DsbC, from Escherichia coli. *Nature Structural Biology*, **7 (3)**, 196–199, doi:10.1038/73295.

[169] McDermott, M. J., M. A. Gawinowicz-Kolks, R. Chiesa, and A. Spector, 1988: The disulfide content of calf γ-crystallin. *Archives of Biochemistry and Biophysics*, **262 (2)**, 609–619, doi:10.1016/0003-9861(88)90413-4.

[170] McNicholas, P., R. Salavati, and D. Oliver, 1997: Dual regulation of escherichia coli secA translation by distinct upstream elements. *Journal of Molecular Biology*, **265 (2)**, 128–141, doi:10.1006/jmbi.1996.0723.

[171] Mercier, E. and M. V. Rodnina, 2018: Co-translational Folding Trajectory of the HemK Helical Domain. *Biochemistry*, doi:10.1021/acs.biochem.8b00293.

[172] Merrifield, R. B., 1963: Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society*, **85 (14)**, 2149–2154, doi:10.1021/ja00897a025.

[173] Messens, J. and J.-F. Collet, 2006: Pathways of disulfide bond formation in Escherichia coli. *The International Journal of Biochemistry & Cell Biology*, **38 (7)**, 1050–1062, doi:10.1016/j.biocel.2005.12.011.

[174] Mobli, M., A. D. De Araújo, L. K. Lambert, G. K. Pierens, M. J. Windley, G. M. Nicholson, P. F. Alewood, and G. F. King, 2009: Direct visualization of disulfide bonds through diselenide proxies using77Se NMR spectroscopy. *Angewandte Chemie - International Edition*, **48 (49)**, 9312–9314, doi:10.1002/anie.200905206.

[175] Morgan, R. S., C. E. Tatsch, R. H. Gushard, J. M. Mcadon, and P. K. Warme, 1978: Chains of Alternating Sulfur and Π-Bonded Atoms in Eight Small Proteins. *International Journal of Peptide and Protein Research*, **11 (3)**, 209–217, doi:10.1111/j.1399-3011.1978.tb02841.x.

[176] Muto, H., H. Nakatogawa, and K. Ito, 2006: Genetically encoded but non-polypeptide prolyl-tRNA functions in the A site for SecM-mediated ribosomal stall. *Molecular cell*, **22 (4)**, 545–52, doi:10.1016/j.molcel.2006.03.033.

[177] Najmudin, S., V. Nalini, H. P. Driessen, C. Slingsby, T. L. Blundell, D. S. Moss, and P. F. Lindley, 1993: Structure of the bovine eye lens protein gammaB(gammaII)-crystallin at 1.47 A. *Acta crystallographica. Section D, Biological crystallography*, **49 (Pt 2)**, 223–33, doi:10.1107/S0907444992007601.

[178] Najmudin, S., V. Nalini, H. P. Driessen, C. Slingsby, T. L. Blundell, D. S. Moss, and P. F. Lindley, 1993: Structure of the bovine eye lens protein gammaB(gammaII)-crystallin at 1.47 A. *Acta crystallographica. Section D, Biological crystallography*, **49 (Pt 2)**, 223–233, doi:10.1107/S0907444992007601.

[179] Nakatogawa, H. and K. Ito, 2001: Secretion monitor, secM, undergoes self-translation arrest in the cytosol. *Molecular Cell*, **7 (1)**, 185–192, doi:10.1016/S1097-2765(01)00166-6.

[180] Nakatogawa, H. and K. Ito, 2002: The Ribosomal Exit Tunnel. *Cell*, **108 (5)**, 629–636.

[181] Nakatogawa, H., A. Murakami, and K. Ito, 2004: Control of SecA and SecM translation by protein secretion. *Current Opinion in Microbiology*, **7 (2)**, 145–150, doi:10.1016/j.mib.2004.01.001.

[182] Nakatogawa, H., A. Murakami, H. Mori, and K. Ito, 2005: SecM facilitates translocase function of SecA by localizing its biosynthesis. *Genes and Development*, **19 (4)**, 436–444, doi:10.1101/gad.1259505.

[183] Nelson, J. W. and T. E. Creighton, 1994: Reactivity and Ionization of the Active Site Cysteine Residues of DsbA, a Protein Required for Disulfide Bond Formation in Vivo. *Biochemistry*, **33 (19)**, 5974–5983, doi:10.1021/bi00185a039.

[184] Netzer, W. J. and F. U. Hartl, 1997: Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature*, **388 (6640)**, 343–349, doi:10.1038/41024.

[185] Nicola, A. V., W. Chen, and A. Helenius, 1999: Co-translational folding of an alphavirus capsid protein in the cytosol of living cells. *Nature Cell Biology*, **1 (6)**, 341–345, doi:10.1038/14032.

[186] Nilsson, O. B., R. Hedman, J. Marino, S. Wickles, L. Bischoff, M. Johansson, A. Müller-Lucks, F. Trovato, J. D. Puglisi, E. P. O'Brien, R. Beckmann, and G. von Heijne, 2015: Cotranslational Protein Folding inside the Ribosome Exit Tunnel. *Cell Reports*, **12 (10)**, 1533–1540, doi:10.1016/j.celrep.2015.07.065.

[187] Nilsson, O. B., A. Müller-Lucks, G. Kramer, B. Bukau, and G. Von Heijne, 2016: Trigger Factor Reduces the Force Exerted on the Nascent Chain by a Cotranslationally Folding Protein. *Journal of Molecular Biology*, **428 (6)**, 1356–1364, doi:10.1016/j.jmb.2016.02.014.

[188] Nilsson, O. B., A. A. Nickson, J. J. Hollins, S. Wickles, A. Steward, R. Beckmann, G. von Heijne, and J. Clarke, 2017: Cotranslational folding of spectrin domains via partially structured states. *Nature Structural & Molecular Biology*, **24 (3)**, 221–225, doi:10.1038/nsmb.3355.

[189] Nissen, P., P. Nissen, J. Hansen, N. Ban, and P. B. Moore, 2000: The Structural Basis of Ribosome Activity in Peptide Bond Synthesis. *Sciences*, **289 (8)**, 920–930, doi:10.1126/science.289.5481.920.

[190] Norledge, B., E.-M. Mayr, R. Glockshuber, O. Bateman, C. Slingsby, R. Jaenicke, and H. Driessen, 1996: The X-ray structures of two mutant crystallin domains shed light on the evolution of multi-domain proteins. *Nature Structural Biology*, **3 (3)**, 267–274, doi:10.1038/nsb0396-267.

[191] Nouwen, N., M. Piwowarek, G. Berrelkamp, and A. J. Driessen, 2005: The large first periplasmic loop of SecD and SecF plays an important role in SecDF functioning. *Journal of Bacteriology*, **187 (16)**, 5857–5860, doi:10.1128/JB.187.16.5857-5860.2005.

[192] Oh, E., A. H. Becker, A. Sandikci, D. Huber, R. Chaba, F. Gloge, R. J. Nichols, A. Typas, C. a. Gross, G. Kramer, J. S. Weissman, and B. Bukau, 2011: Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell*, **147 (6)**, 1295–1308, doi:10.1016/j.cell.2011.10.044.

[193] Onuchic, J. N., Z. Luthey-Schulten, and P. G. Wolynes, 1997: THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective. *Annual Review of Physical Chemistry*, **48 (1)**, 545–600, doi:10.1146/annurev.physchem.48.1.545.

[194] Ozenne, V., F. Bauer, L. Salmon, J. R. Huang, M. R. Jensen, S. Segard, P. Bernadó, C. Charavay, and M. Blackledge, 2012: Flexible-meccano: A tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. *Bioinformatics*, **28 (11)**, 1463–1470, doi:10.1093/bioinformatics/bts172.

[195] Pan, J. L. and J. C. Bardwell, 2006: The origami of thioredoxin-like folds. *Protein Science*, **15 (10)**, 2217–2227, doi:10.1110/ps.062268106.

[196] Pande, J., M. J. McDermott, R. H. Callender, and A. Spector, 1989: Raman spectroscopic evidence for a disulfide bridge in calf γII crystallin. *Archives of Biochemistry and Biophysics*, **269 (1)**, 250–255, doi:10.1016/0003-9861(89)90106-9.

[197] Patzelt, H., S. Rudiger, D. Brehmer, G. Kramer, S. Vorderwulbecke, E. Schaffitzel, A. Waitz, T. Hesterkamp, L. Dong, J. Schneider-Mergener, B. Bukau, and E. Deuerling, 2001: Binding specificity of Escherichia coli trigger factor. *Proceedings of the National Academy of Sciences*, **98 (25)**, 14 244–14 249, doi:10.1073/pnas.261432298.

[198] Pauli, J., M. Baldus, B. van Rossum, H. de Groot, and H. Oschkinat, 2001: Backbone and Side-Chain 13C and15N Signal Assignments of the α-Spectrin SH3 Domain by Magic Angle Spinning Solid-State NMR at 17.6 Tesla. *ChemBioChem*, **2 (4)**, 272–281, doi:10.1002/1439-7633(20010401)2:4<272::AID-CBIC272>3.0.CO;2-2.

[199] Pavlov, M. Y., R. E. Watts, Z. Tan, V. W. Cornish, M. Ehrenberg, and A. C. Forster, 2009: Slow peptide bond formation by proline and other N-alkylamino acids in translation. *Proceedings of the National Academy of Sciences*, **106 (1)**, 50–54, doi:10.1073/pnas.0809211106.

[200] Pecher, P. and U. Arnold, 2009: The effect of additional disulfide bonds on the stability and folding of ribonuclease A. *Biophysical Chemistry*, **141 (1)**, 21–28, doi:10.1016/j.bpc.2008.12.005.

[201] Peters, F. B., A. Brock, J. Wang, and P. G. Schultz, 2009: Photocleavage of the Polypeptide Backbone by 2-Nitrophenylalanine. *Chemistry and Biology*, **16 (2)**, 148–152, doi:10.1016/j.chembiol.2009.01.013.

[202] Peters, T. and L. K. Davidson, 1982: The biosynthesis of rat serum albumin. In vivo studies on the formation of the disulfide bonds. *The Journal of biological chemistry*, **257 (15)**, 8847–53.

[203] Petersen, M. T. N., P. H. Jonson, and S. B. Petersen, 1999: Amino acid neighbours and detailed conformational analysis of cysteines in proteins. *Protein Engineering*, **12 (7)**, 535–548, doi:10.1093/protein/12.7.535.

[204] Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, 2004: UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Com-*

putational Chemistry, **25 (13)**, 1605–1612, doi:10.1002/jcc.20084, `arXiv: 1011.1669v3`.

[205] Pollard, M. G., K. J. Travers, and J. S. Weissman, 1998: Ero1p: A novel and ubiquitous protein with an essential role in oxidative protein folding in the endoplasmic reticulum. *Molecular Cell*, **1 (2)**, 171–182, doi:10.1016/S1097-2765(00)80018-0.

[206] Qin, M., W. Wang, and D. Thirumalai, 2015: Protein folding guides disulfide bond formation. *Proceedings of the National Academy of Sciences*, **112 (36)**, 11 241–11 246, doi:10.1073/pnas.1503909112.

[207] Quinlan, R. A. and P. J. Hogg, 2018: γ-Crystallin redox–detox in the lens. *Journal of Biological Chemistry*, **293 (46)**, 18 010–18 011, doi:10.1074/jbc.H118.006240.

[208] R Core Team, 2019: *R: A Language and Environment for Statistical Computing*. Vienna, Austria, R Foundation for Statistical Computing.

[209] Rajapandi, T., K. M. Dolan, and D. B. Oliver, 1991: The first gene in the Escherichia coli secA operon, gene X, encodes a nonessential secretory protein. *Journal of Bacteriology*, **173 (22)**, 7092–7097, doi:10.1128/jb.173.22.7092-7097.1991.

[210] Ravera, E., B. Corzilius, V. K. Michaelis, C. Rosa, R. G. Griffin, C. Luchinat, and I. Bertini, 2013: Dynamic nuclear polarization of sedimented solutes. *Journal of the American Chemical Society*, **135 (5)**, 1641–1644, doi:10.1021/ja312553b.

[211] Robinson, P. J., S. Kanemura, X. Cao, and N. J. Bulleid, 2020: Protein secondary structure determines the temporal relationship between folding and disulfide formation. *Journal of Biological Chemistry*, **295 (8)**, 2438–2448, doi:10.1074/jbc.RA119.011983.

[212] Robinson, P. J., M. A. Pringle, C. A. Woolhead, and N. J. Bulleid, 2017: Folding of a single domain protein entering the endoplasmic reticulum precedes disulfide formation. *Journal of Biological Chemistry*, **292 (17)**, 6978–6986, doi:10.1074/jbc.M117.780742.

[213] Rodnina, M. V. and W. Wintermeyer, 2016: Protein Elongation, Co-translational Folding and Targeting. *Journal of Molecular Biology*, **428 (10)**, 2165–2185, doi:10.1016/j.jmb.2016.03.022.

[214] Rohou, A. and N. Grigorieff, 2015: CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, **192 (2)**, 216–221, doi:10.1016/j.jsb.2015.08.008.

[215] Rüdiger, S., A. Buchberger, and B. Bukau, 1997: Interaction of Hsp70 chaperones with substrates. *Nature Structural & Molecular Biology*, **4 (5)**, 342–349, doi:10.1038/nsb0597-342.

[216] Rudolph, R., R. Siebendritt, G. Nesslaŭer, a. K. Sharma, and R. Jaenicke, 1990: Folding of an all-beta protein: independent domain folding in gamma II-crystallin from calf eye lens. *Proceedings of the National Academy of Sciences of the United States of America*, **87 (12)**, 4625–4629, doi:10.1073/pnas.87.12.4625.

[217] Rutkowska, A., M. Beerbaum, N. Rajagopalan, J. Fiaux, P. Schmieder, G. Kramer, H. Oschkinat, and B. Bukau, 2009: Large-scale purification of ribosome-nascent chain complexes for biochemical and structural studies. *FEBS letters*, **583 (14)**, 2407–13, doi:10.1016/j.febslet.2009.06.041.

[218] Rutkowska, A., M. P. Mayer, A. Hoffmann, F. Merz, B. Zachmann-Brand, C. Schaffitzel, N. Ban, E. Deuerling, and B. Bukau, 2008: Dynamics of trigger factor interaction with translating ribosomes. *Journal of Biological Chemistry*, **283 (7)**, 4124–4132, doi:10.1074/jbc.M708294200.

[219] Sagemark, J., T. H. Elgán, T. R. Bürglin, C. Johansson, A. Holmgren, and K. D. Berndt, 2007: Redox properties and evolution of human glutaredoxins. *Proteins: Structure, Function, and Bioinformatics*, **68 (4)**, 879–892, doi:10.1002/prot.21416.

[220] Saio, T., X. Guan, P. Rossi, A. Economou, and C. G. Kalodimos, 2014: Structural basis for protein antiaggregation activity of the trigger factor chaperone. *Science*, **344 (6184)**, 590–591, doi:10.1126/science.1250494.

[221] Sanfelice, D., H. Koss, T. D. Bunney, G. S. Thompson, B. Farrell, M. Katan, and A. L. Breeze, 2018: NMR backbone assignments of the tyrosine kinase domain of human fibroblast growth factor receptor 3 in apo state and in complex with inhibitor PD173074. *Biomolecular NMR Assignments*, **12 (2)**, 231–235, doi:10.1007/s12104-018-9814-7.

[222] Schafer, F. Q. and G. R. Buettner, 2001: Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radical Biology and Medicine*, **30 (11)**, 1191–1212, doi:10.1016/S0891-5849(01)00480-4.

[223] Schanda, P., E. Kupĉe, and B. Brutscher, 2005: SOFAST-HMQC experiments for recording two-dimensional deteronuclear correlation spectra of proteins within a few seconds. *Journal of Biomolecular NMR*, **33 (4)**, 199–211, doi:10.1007/s10858-005-4425-x.

[224] Scheres, S. H., 2012: RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, **180 (3)**, 519–530, doi:10.1016/j.jsb.2012.09.006.

[225] Schmeing, T. M. and V. Ramakrishnan, 2009: What recent ribosome structures have revealed about the mechanism of translation. *Nature*, **461 (7268)**, 1234–1242, doi:10.1038/nature08403.

[226] Schmidt, B., L. Ho, and P. J. Hogg, 2006: Allosteric Disulfide Bonds †. *Biochemistry*, **45 (24)**, 7429–7433, doi:10.1021/bi0603064.

[227] Seckler, R., A. Fuchs, J. King, and R. Jaenicke, 1989: Reconstitution of the Thermostable Trimeric. *The Journal of biological chemistry*, **264 (20)**, 11 750–11 753.

[228] Seidelt, B., C. A. Innis, D. N. Wilson, M. Gartmann, J.-P. Armache, E. Villa, L. G. Trabuco, T. Becker, T. Mielke, K. Schulten, T. A. Steitz, and R. Beckmann, 2009: Structural Insight into Nascent Polypeptide Chain-Mediated Translational Stalling. *Science*, **326 (5958)**, 1412–1415, doi:10.1126/science.1177662.

[229] Seidelt, B., C. A. Innis, D. N. Wilson, M. Gartmann, E. Villa, L. G. Tra-buco, T. Becker, K. Schulten, T. a. Steitz, and R. Beckmann, 2009: Struc-trual insight into nascent polypeptide chain-mediated translational stalling. *Science*, **326 (5958)**, 1412–1415, doi:10.1126/science.1177662.Structural.

[230] Seras-Franzoso, J., R. Affentranger, M. Ferrer-Navarro, X. Daura, A. Villaverde, and E. García-Fruitósa, 2012: Disulfide bond formation and activation of Escherichia coli β-galactosidase under oxidizing con-ditions. *Applied and Environmental Microbiology*, **78**, 2376–2385, doi: 10.1128/AEM.06923-11.

[231] Serebryany, E., S. Yu, S. A. Trauger, B. Budnik, and E. I. Shakhnovich, 2018: Dynamic disulfide exchange in a crystallin protein in the human eye lens promotes cataract-associated aggregation. *Journal of Biological Chemistry*, **293 (46)**, 17 997–18 009, doi:10.1074/jbc.RA118.004551.

[232] Sevier, C. S., J. W. Cuozzo, A. Vala, F. Åslund, and C. A. Kaiser, 2001: A flavoprotein oxidase defines a new endoplasmic reticulum pathway for biosynthetic disulphide bond formation. *Nature Cell Biology*, **3 (10)**, 874–882, doi:10.1038/ncb1001-874.

[233] Sevier, C. S. and C. A. Kaiser, 2002: Formation and transfer of disulphide bonds in living cells. *Nature Reviews Molecular Cell Biology*, **3 (11)**, 836–847, doi:10.1038/nrm954.

[234] Sharma, a. K., V. Minke-Gogl, P. Gohl, R. Siebendritt, R. Jaenicke, and R. Rudolph, 1990: Limited proteolysis of gamma II-crystallin from calf eye lens. Physicochemical studies on the N-terminal domain and the intact two-domain protein. *European journal of biochemistry / FEBS*, **194 (2)**, 603–609.

[235] Sharma, D. and K. Rajarathnam, 2000: 13C NMR chemical shifts can predict disulfide bond formation. *Journal of Biomolecular NMR*, **18**, 165–171, doi:10.1023/A:1008398416292.

[236] Sharp, P. M., E. Cowe, D. G. Higgins, D. C. Shields, K. H. Wolfe, and F. Wright, 1988: Codon usage patterns in Escherichia coli, Bacillus sub-tilis, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Drosophila melanogaster and Homo sapiens; a review of the considerable within-species diversity. *Nucleic Acids Research*, **16 (17)**, 8207–8211, doi:10.1093/nar/16. 17.8207.

[237] Silva, G. M., L. E. Netto, K. F. Discola, G. M. Piassa-Filho, D. C. Pimenta, J. A. Bárcena, and M. Demasi, 2008: Role of glutaredoxin 2 and cytosolic thioredoxins in cysteinyl-based redox modification of the 20S proteasome. *FEBS Journal*, **275 (11)**, 2942–2955, doi:10.1111/j.1742-4658.2008.06441. x.

[238] Silvers, R., Schlepckow, K., Wirmer-Bartoschek, J. & Schwalbe, H., 2011: *NMR-Spectroscopic Investigation of Disulfide Dynamics in Unfolded States of Proteins*, Vol. 14. Springer Science+Business Media, 217–256 pp., `arXiv: 1011.1669v3`.

[239] Slingsby, C. and L. R. Miller, 1985: The reaction of glutathione with the eye-lens protein gamma-crystallin. *Biochem J*, **230 (1)**, 143–150, doi:10.1042/bj2300143.

[240] Slingsby, C., B. Norledge, A. Simpson, O. A. Bateman, G. Wright, H. P. Driessen, P. F. Lindley, D. S. Moss, and B. Bax, 1997: X-ray diffraction and structure of crystallins. *Progress in Retinal and Eye Research*, **16 (1)**, 3–29, doi:10.1016/S1350-9462(96)00018-3.

[241] Slingsby, C., B. Norledge, a. Simpson, O. a. Bateman, G. Wright, H. P. C. Driessen, P. F. Lindley, D. S. Moss, and B. Bax, 1997: X-ray diffraction and structure of crystallins. *Progress in Retinal and Eye Research*, **16 (1)**, 3–29, doi:10.1016/S1350-9462(96)00018-3.

[242] Smirnova, G. V., N. G. Muzyka, M. N. Glukhovchenko, and O. N. Oktyabrsky, 2000: Effects of menadione and hydrogen peroxide on glutathione status in growing Escherichia coli. *Free Radical Biology and Medicine*, **28 (7)**, 1009–1016, doi:10.1016/S0891-5849(99)00256-7.

[243] Smirnova, G. V. and O. N. Oktyabrsky, 2005: Glutathione in bacteria. *Biochemistry (Moscow)*, **70 (11)**, 1199–1211, doi:10.1007/s10541-005-0248-3.

[244] Smith, B. C. and M. A. Marletta, 2012: Mechanisms of S-nitrosothiol formation and selectivity in nitric oxide signaling. *Current Opinion in Chemical Biology*, **16 (5-6)**, 498–506, doi:10.1016/j.cbpa.2012.10.016.

[245] Song, H., P. Mugnier, A. K. Das, H. M. Webb, D. R. Evans, M. F. Tuite, B. A. Hemmings, and D. Barford, 2000: The crystal structure of human eukaryotic release factor eRF1 - Mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. *Cell*, **100 (3)**, 311–321, doi:10.1016/s0092-8674(00)80667-4.

[246] Spyrou, G., E. Enmark, A. Miranda-Vizuete, and J. Å. Gustafsson, 1997: Cloning and expression of a novel mammalian thioredoxin. *Journal of Biological Chemistry*, **272 (5)**, 2936–2941, doi:10.1074/jbc.272.5.2936.

[247] Srinivasan, N., R. Sowdhamini, C. Ramakrishnan, and P. Balaram, 1990: Conformations of disulfide bridges in proteins. *International Journal of Peptide and Protein Research*, **36 (2)**, 147–155, doi:10.1111/j.1399-3011.1990.tb00958.x.

[248] Srivastava, S. K. and E. Beutler, 1973: Cleavage of lens protein-GSH mixed disulfide by glutathione reductase. *Experimental Eye Research*, **17 (1)**, 33–42, doi:10.1016/0014-4835(73)90165-6.

[249] Su, T., J. Cheng, D. Sohmen, R. Hedman, O. Berninghausen, G. von Heijne, D. N. Wilson, and R. Beckmann, 2017: The force-sensing peptide VemP employs extreme compaction and secondary structure formation to induce ribosomal stalling. *eLife*, **6**, 1–17, doi:10.7554/eLife.25642.

[250] Takahashi, Y. H., K. Inaba, and K. Ito, 2004: Characterization of the menaquinone-dependent disulfide bond formation pathway of Escherichia coli. *Journal of Biological Chemistry*, **279 (45)**, 47 057–47 065, doi:10.1074/jbc.M407153200.

[251] Tidor, B. and M. Karplus, 1993: The contribution of cross-links to protein stability: A normal mode analysis of the configurational entropy of the native state. *Proteins: Structure, Function, and Bioinformatics*, **15 (1)**, 71–79, doi:10.1002/prot.340150109.

[252] Toth, E. A., C. Worby, J. E. Dixon, E. R. Goedken, S. Marqusee, and T. O. Yeates, 2000: The crystal structure of adenylosuccinate lyase from Pyrobaculum aerophilum reveals an intracellular protein with three disulfide bonds 1 1Edited by I. A. Wilson. *Journal of Molecular Biology*, **301 (2)**, 433–450, doi:10.1006/jmbi.2000.3970.

[253] Tsai, A., G. Kornberg, M. Johansson, J. Chen, and J. Puglisi, 2014: The dynamics of SecM-induced translational stalling. *Cell Reports*, **7 (5)**, 1521–1533, doi:10.1016/j.celrep.2014.04.033.

[254] Tu, B. P., S. C. Ho-Schleyer, K. J. Travers, and J. S. Weissman, 2000: Biochemical basis of oxidative protein folding in the endoplasmic reticulum. *Science*, **290 (5496)**, 1571–1574, doi:10.1126/science.290.5496.1571.

[255] Tu, D., G. Blaha, P. B. Moore, and T. A. Steitz, 2005: Structures of MLSBK antibiotics bound to mutated large ribosomal subunits provide a structural explanation for resistance. *Cell*, **121 (2)**, 257–270, doi:10.1016/j.cell.2005.02.005.

[256] Tu, L. W. and C. Deutsch, 2010: A Folding Zone in the Ribosomal Exit Tunnel for Kv1.3 Helix Formation. *Journal of Molecular Biology*, **396 (5)**, 1346–1360, doi:10.1016/j.jmb.2009.12.059.

[257] Ulrich, E. L., H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C. F. Schulte, D. E. Tolmie, R. Kent Wenger, H. Yao, and J. L. Markley, 2008: BioMagResBank. *Nucleic Acids Research*, **36 (SUPPL. 1)**, 402–408, doi:10.1093/nar/gkm957.

[258] Vajpai, N., A. K. Schott, M. Vogtherr, and A. L. Breeze, 2014: NMR backbone assignments of the tyrosine kinase domain of human fibroblast growth factor receptor 1. *Biomolecular NMR Assignments*, **8 (1)**, 85–88, doi:10.1007/s12104-013-9458-6.

[259] Vázquez-Laslop, N., H. Ramu, D. Klepacki, K. Kannan, and A. S. Mankin, 2010: The key function of a conserved and modified rRNA residue in the ribosomal response to the nascent peptide. *EMBO Journal*, **29 (18)**, 3108–3117, doi:10.1038/emboj.2010.180.

[260] Voss, N. R., M. Gerstein, T. A. Steitz, and P. B. Moore, 2006: The Geometry of the Ribosomal Polypeptide Exit Tunnel. *Journal of Molecular Biology*, **360 (4)**, 893–906, doi:10.1016/j.jmb.2006.05.023.

[261] Wang, C.-C., W.-C. Lai, and W.-J. Chuang, 2016: Predicting the redox state and secondary structure of cysteine residues in proteins using multidimensional classification of NMR chemical shifts. *Journal of Biomolecular NMR*, **66 (1)**, 55–68, doi:10.1002/prot.20875.

[262] Waudby, C. a., H. Launay, L. D. Cabrita, and J. Christodoulou, 2013: Protein folding on the ribosome studied using NMR spectroscopy. *Progress in*

nuclear magnetic resonance spectroscopy, **74**, 57–75, doi:10.1016/j.pnmrs. 2013.07.003.

[263] Waugh, D. S., 1996: Genetic tools for selective labeling of proteins with. *Journal of biomolecular NMR*, **8**, 184–192.

[264] Weissman, J. S. and P. S. Kim, 1991: Reexamination of the Predominance Intermediates. *Science*, **253 (1986)**, 1386–1393.

[265] Welker, E., W. J. Wedemeyer, M. Narayan, and H. A. Scheraga, 2001: Coupling of conformational folding and disulfide-bond reactions in oxidative folding of proteins. *Biochemistry*, **40 (31)**, 9059–9064, doi:10.1021/ bi010409g.

[266] Wilson, D. N. and J. H. D. Cate, 2012: The Structure and Function of the Eukaryotic Ribosome. *Cold Spring Harbor Perspect Biol Perspect Biol*, **4**, 1–18, doi:10.1101/cshperspect.a011536.

[267] Wingert, R. A., *et al.*, 2005: Deficiency of glutaredoxin 5 reveals Fe-S clusters are required for vertebrate haem synthesis. *Nature*, **436 (7053)**, 1035–1039, doi:10.1038/nature03887.

[268] Wistow, G., B. Turnell, L. Summers, and C. Slingsby, 1983: X-ray analysis of the eye lens protein gamma-II crystallin at 1.9 resolution. *J. Mol. Biol.*, **170**, 175–202.

[269] Wolynes, P. G., J. N. Onuchic, and D. Thirumalai, 1995: Navigating the folding routes. *Science*, **267**, 1619–1620, doi:10.1126/science.7886447.

[270] Woolhead, C. A., A. E. Johnson, and H. D. Bernstein, 2006: Translation Arrest Requires Two-Way Communication between a Nascent Polypeptide and the Ribosome. *Molecular Cell*, **22 (5)**, 587–598, doi:10.1016/j.molcel. 2006.05.021.

[271] Woolhead, C. a., P. J. McCormick, and A. E. Johnson, 2004: Nascent Membrane and Secretory Proteins Differ in FRET-Detected Folding Far inside the Ribosome and in Their Exposure to Ribosomal Proteins. *Cell*, **116 (5)**, 725–736, doi:10.1016/S0092-8674(04)00169-2.

[272] Woycechowsky, K. J. and R. T. Raines, 2000: Native disulfide bond formation in proteins. *Current Opinion in Chemical Biology*, **4 (5)**, 533–539, doi:10.1016/S1367-5931(00)00128-9.

[273] Xia, B., A. Vlamis-Gardikas, A. Holmgren, P. E. Wright, and H. J. Dyson, 2001: Solution structure of Escherichia coli glutaredoxin-2 shows similarity to mammalian glutathione-S-transferases. *Journal of Molecular Biology*, **310 (4)**, 907–918, doi:10.1006/jmbi.2001.4721.

[274] Yachdav, G., E. Kloppmann, L. Kajan, M. Hecht, T. Goldberg, T. Hamp, P. Hönigschmid, A. Schafferhans, M. Roos, M. Bernhofer, L. Richter, H. Ashkenazy, M. Punta, A. Schlessinger, Y. Bromberg, R. Schneider, G. Vriend, C. Sander, N. Ben-Tal, and B. Rost, 2014: PredictProtein - An open resource for online prediction of protein structural and functional features. *Nucleic Acids Research*, **42 (W1)**, 337–343, doi:10.1093/nar/gku366.

[275] Ying, B. W., H. Taguchi, M. Kondo, and T. Ueda, 2005: Co-translational involvement of the chaperonin GroEL in the folding of newly translated polypeptides. *Journal of Biological Chemistry*, **280 (12)**, 12 035–12 040, doi:10.1074/jbc.M500364200.

[276] Yu, C. H., Y. Dang, Z. Zhou, C. Wu, F. Zhao, M. S. Sachs, and Y. Liu, 2015: Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Molecular Cell*, **59 (5)**, 744–754, doi:10.1016/j.molcel.2015.07.018.

[277] Yu, N.-T., D. C. DeNagel, and C. Slingsby, 1989: Raman spectroscopy of calf lens γ-II crystallin: Direct evidence for the formation of mixed disulfide bonds with 2-mercaptoethanol and glutathione. *Experimental Eye Research*, **48 (3)**, 399–410, doi:10.1016/S0014-4835(89)80008-9.

[278] Zaman, S., M. Fitzpatrick, L. Lindahl, and J. Zengel, 2007: Novel mutations in ribosomal proteins L4 and L22 that confer erythromycin resistance in Escherichia coli. *Molecular Microbiology*, **66 (4)**, 1039–1050, doi:10.1111/j.1365-2958.2007.05975.x.

[279] Zhang, H., S. Neal, and D. S. Wishart, 2003: RefDB: A database of uniformly referenced protein chemical shifts. *Journal of Biomolecular NMR*, **25 (3)**, 173–195, doi:10.1023/A:1022836027055.

[280] Zhang, J., X. Pan, K. Yan, S. Sun, N. Gao, and S.-F. Sui, 2015: Mechanisms of ribosome stalling by SecM at multiple elongation steps. *eLife*, **4 (DECEMBER2015)**, 1–25, doi:10.7554/eLife.09684.

[281] Zheng, S. Q., E. Palovcak, J. P. Armache, K. A. Verba, Y. Cheng, and D. A. Agard, 2017: MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods*, **14 (4)**, 331–332, doi:10.1038/nmeth.4193.

[282] Zhou, M., J. Guo, J. Cha, M. Chae, S. Chen, J. M. Barral, M. S. Sachs, and Y. Liu, 2013: SUPPLEMENT: Non-optimal codon usage affects expression, structure and function of clock protein FRQ. *Nature*, **495 (7439)**, 111–5, doi:10.1038/nature11833.

[283] Zivanov, J., T. Nakane, B. O. Forsberg, D. Kimanius, W. J. Hagen, E. Lindahl, and S. H. Scheres, 2018: New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife*, **7**, 1–38, doi:10.7554/elife.42166.

5   Appendices

5.1   Amino acid sequences and DNA sequences of RNCs and isolated GBC and amio acid sequence of the U32 peptides.

Table 5.1: *Amino acid sequence of RNC constructs, GBC full-length and NTD protein and peptides from SPPS.*

Construct name	Amino acid sequence
U32SecM	MGHHHHHHHH HHMGKITFYE DRGFQGHCYE CSSDCPNLQP YFSRCFSTPV WISQAQGIRA GP
U32SecM C18 C22	MGHHHHHHHH HHMGKITFYE DRGFQGHAYE CSSDCPNLQP YFSRAFSTPV WISQAQGIRA GP
U32SecM C15 C18	MGHHHHHHHH HHMGKITFYE DRGFQGHCYE CSSDAPNLQP YFSRAFSTPV WISQAQGIRA GP
U32SecM C18	MGHHHHHHHH HHMGKITFYE DRGFQGHAYE CSSDAPNLQP YFSRAFSTPV WISQAQGIRA GP
U43SecM	MGHHHHHHHH HHMGKITFYE DRGFQGHCYE CSSDCPNLQP YFSRCNSIRV DSGCWMFSTP VWISQAQGIR AGP
U78SecM	MGHHHHHHHH HHMGKITFYE DRGFQGHCYE CSSDCPNLQP YFSRCNSIRV DSGCWMLYER PNYQGHQYFL RRGDYPDYQQ WMGFNDSIRS CFSTPVWISQ AQGIRAGP
U174 full-length	MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA MNAKVGSLRR VMDFYHHHHH H
U/H83 (NTD)	MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR LIPHHHHHH
U/H88-174 (CTD)	MGTFRMRIYE RDDFRGQMSE ITDDCPSLQD RFHLTEVHSL NVLEGSWVLY EMPSYRGRQY LLRPGEYRRY LDWGAMNAKV GSLRRVMDFY HHHHHH
Peptide C15	MGKITFYEDR GFQGHCYECS SDAPNLQPYF SRA
Peptide C18	MGKITFYEDR GFQGHAYECS SDCPNLQPYF SRA
Peptide C22	MGKITFYEDR GFQGHAYECS SDCPNLQPYF SRA

Table 5.2: *DNA sequence of RNC constructs, GBC full-length and NTD protein. Sequences include a 5' NcoI and 3' XhoI restriction site.*

Construct name	DNA sequence
U32SecM	CCATGGGGCA TCATCACCAC CATCACCATC ATCACCATAT GGGGAAGATC ACTTTTTACG AGGACCGGGG CTTCCAGGGC CACTGCTACG AGTGCAGCAG TGACTGCCCC AACCTGCAGC CCTATTTCAG CCGCTGTTTC AGCACGCCCG TCTGGATAAG CCAGGCGCAA GGCATCCGTG CTGGCCCTTG ACTCGAG
U32SecM C18 C22	CCATGGGGCA TCATCACCAC CATCACCATC ATCACCATAT GGGGAAGATC ACTTTTTACG AGGACCGGGG CTTCCAGGGC CACGCTTACG AGTGCAGCAG TGACTGCCCC AACCTGCAGC CCTATTTCAG CCGCGCTTTC AGCACGCCCG TCTGGATAAG CCAGGCGCAA GGCATCCGTG CTGGCCCTTG ACTCGAG
U32SecM C15 C18	CCATGGGGCA TCATCACCAC CATCACCATC ATCACCATAT GGGGAAGATC ACTTTTTACG AGGACCGGGG CTTCCAGGGC CACTGCTACG AGTGCAGCAG TGACGCTCCC AACCTGCAGC CCTATTTCAG CCGCGCTTTC AGCACGCCCG TCTGGATAAG CCAGGCGCAA GGCATCCGTG CTGGCCCTTGA CTCGAG
U32SecM C18	CCATGGGGCA TCATCACCAC CATCACCATC ATCACCATAT GGGGAAGATC ACTTTTTACG AGGACCGGGG CTTCCAGGGC CACGCTTACG AGTGCAGCAG TGACGCTCCC AACCTGCAGC CCTATTTCAG CCGCGCTTTC AGCACGCCCG TCTGGATAAG CCAGGCGCAA GGCATCCGTG CTGGCCCTTG ACTCGAG
U43SecM	CCATGGGGCA TCATCACCAC CATCACCATC ATCACCATAT GGGGAAGATC ACTTTTTACG AGGACCGGGG CTTCCAGGGC CACTGCTACG AGTGCAGCAG TGACTGCCCC AACCTGCAGC CCTATTTCAG CCGCTGTAAC TCCATCCGCG TGGACAGCGG CTGCTGGATG TTCAGCACGC CCGTCTGGAT AAGCCAGGCGC AAGGCATCCG TGCTGGCCCT TGACTCGAG

U78SecM	CCATGGGGCA TCATCACCAC CATCACCATC ATCACCATAT GGGGAAGATC ACTTTTTACG AGGACCGGGG CTTCCAGGGC CACTGCTACG AGTGCAGCAG TGACTGCCCC AACCTGCAGC CCTATTTCAG CCGCTGTAAC TCCATCCGCG TGGACAGCGG CTGCTGGATG CTGTATGAGC GCCCCAACTA CCAGGGCCAC CAGTACTTCC TGCGGCGCGG CGACTACCCC GACTACCAGC AGTGGATGGG CTTCAACGAC TCCATCCGCT CCTGCTTCAG CACGCCCGTC TGGATAAGCC AGGCGCAAGG CATCCGTGCT GGCCCTTGAC TCGAG
U174 full-length	CCATGGGGAA GATCACTTTT TACGAGGACC GGGGCTTCCA GGGCCACTGC TACGAGTGCA GCAGTGACTG CCCCAACCTG CAGCCCTATT TCAGCCGCTG TAACTCCATC CGCGTGGACA GCGGCTGCTG GATGCTGTAT GAGCGCCCCA ACTACCAGGG CCACCAGTAC TTCCTGCGGC GCGGCGACTA CCCCGACTAC CAGCAGTGGA TGGGCTTCAA CGACTCCATC CGCTCCTGCC GCCTCATCCC GCAACACACC GGCACTTTCA GAATGAGAAT CTATGAGAGA GATGACTTCA GAGGACAGAT GTCAGAGATC ACAGACGATT GTCCCTCTCT TCAAGACCGC TTCCACCTCA CTGAGGTTCA CTCCCTCAAC GTGCTGGAGG GTTCCTGGGT CCTCTATGAG ATGCCAAGCT ACAGGGGAAG GCAGTACCTG CTGAGGCCAG GGGAGTACAG GAGATATCTT GACTGGGGGG CAATGAATGC CAAAGTTGGT TCTTTAAGAC GGGTGATGGA TTTTTATCAT CATCATCATC ATCATTGACT CGAG
U83 (NTD)	CCATGGGGAA GATCACTTTT TACGAGGACC GGGGCTTCCA GGGCCACTGC TACGAGTGCA GCAGTGACTG CCCCAACCTG CAGCCCTATT TCAGCCGCTG TAACTCCATC CGCGTGGACA GCGGCTGCTG GATGCTGTAT GAGCGCCCCA ACTACCAGGG CCACCAGTAC TTCCTGCGGC GCGGCGACTA CCCCGACTAC CAGCAGTGGA TGGGCTTCAA CGACTCCATC CGCTCCTGCC GCCTCATCCC GCATCATCAT CATCATCATT GACTCGAG

H83 (NTD)	CCATGGGAAA AATCACTTTC TATGAAGATC GCGGTTTTCA GGGTCATTGT TATGAATGTA GCAGCGATTG TCCGAACCTG CAGCCGTACT TTAGCCGCTG TAACAGCATC CGCGTTGATA GCGGTTGTTG GATGCTGTAC GAACGCCCGA ACTATCAGGG TCATCAGTAT TTTCTGCGCC GCGGTGATTA TCCGGATTAT CAGCAGTGGA TGGGTTTTAA CGATAGCATC CGCAGCTGTC GCTTAATCCC ACATCATCAT CATCATCATT GACTCGAG
U88-174 (CTD)	CCATGGGCAC TTTCAGAATG AGAATCTATG AGAGAGATGA CTTCAGAGGA CAGATGTCAG AGATCACAGA CGATTGTCCC TCTCTTCAAG ACCGCTTCCA CCTCACTGAG GTTCACTCCCT CAACGTGCTG GAGGGTTCCT GGGTCCTCTA TGAGATGCCA AGCTACAGGG GAAGGCAGTA CCTGCTGAGG CCAGGGGAGT ACAGGAGATA TCTTGACTGG GGGGCAATGA ATGCCAAAGT TGGTTCTTTA AGACGGGTGA TGGATTTTTA TCATCATCAT CATCATCATT GACTCGAG
H88-174 (CTD)	CCATGGGTAC TTTTCGGATG CGGATCTACG AACGGGACGA TTTTCGGGGA CAGATGTCCG AAATCACTGA TG ACTGTCCG TCATTGCAAG ATCGCTTTCA TTTAACTGAA GTACATAGCT TAAACGTTCT GGAAGGGAGC TGGGTCTTAT ACGAAATGCC GAGCTATCGC GGACGCCAGT ATCTGCTGCG CCCGGGAGAA TATCGCCGGT ACTTGGATTG GGGAGCTATG AACGCAAAGG TAGGGTCACT ACGGCGCGTT ATGGACTTCT ACCACCATCA TCACCACCAC TAATGACTCGAG

5.2   Lists of buffers and media

5.2.1   For electrophoresis and Western-blotting

2x SDS reducing sample buffer
0.25 M Tris/HCl (pH 6.8)
4% (w/v) SDS
20% (v/v) glycerol
10% (v/v) 2-mercaptoethanol
0.25% (w/v) bromphenol blue

2x SDS non-reducing sample buffer
0.25 M Tris/HCl (pH 6.8)
4% (w/v) SDS
20% (v/v) glycerol
0.25% (w/v) bromphenol blue

6x SDS reducing sample buffer
375 mM Tris/HCl(pH 6.8)
10% (w/v) SDS
50% (v/v) glycerol
10% (v/v) 2-mercaptoethanol
0.03% (v/v) bromphenol blue

6x SDS non-reducing sample buffer
375 mM Tris/HCl(pH 6.8)
10% (w/v) SDS
50% (v/v) glycerol
0.03% (v/v) bromphenol blue

6x DNA sample buffer
60% (v/v) glycerol
0.25% (w/v) bromphenol blue
0.25% (w/v) xylene cyanol FF

SDS-PAGE running buffer
25 mM Tris
250 mM glycine
1% (w/v) SDS

5% stacking gel
0.25 M Tris/HCl (pH 6.8)
5% (v/v) acrylamide/bisacrylamide
(29:1)
0.1% (w/v) APS
0.1% (w/v) SDS

15% resolving gel
0.375 M Tris/HCl (pH 8.8)
15% (v/v) acrylamide/bisacrylamide
(29:1)
0.1% (w/v) APS
0.1% (w/v) SDS

Coomassie Blue staining solution
10% (v/v) ethanol
5% (v/v) acetic acid (glacial)
0.0025% (w/v) Coomassie Brilliant
Blue G250
0.0025% (w/v) Coomassie Brilliant
Blue R250

TAE buffer
40 mM Tris (pH 8.0)
0.1% (v/v) acetic acid (glacial)
1 mM EDTA

1% agarose gel
1% agarose (w/v)
TAE buffer

Anode buffer (pH 7.4)
75 mM Tris/HCl
20% (v/v) methanol

Cathode buffer (pH 9.0)
25 mM Tris/HCl
40 mM aminocaproic acid
20% (v/v) methanol

PBS buffer (pH 7.4)
10 mM Na_2HPO_4
2 mM K_2HPO_4
137 mM NaCl
2.7 mM KCl

PBS+M
5% (w/v) milk powder
PBS buffer (pH 7.4)

PBST
0.05% (v/v) Tween-20
PBS buffer (pH 7.4)

5.2.2   Media for protein expression

M9 minimal medium (pH 7.4)
42 mM Na_2HPO_4
22 mM KH_2PO_4
8.5 mM NaCl
0.2 mM $CaCl_2$
2 mM $MgSo_4$
1x Vitamin mix
1x Trace elements
1g/L $^{15}NH_4Cl$
10 g/L ^{12}C-glucose
or 2g/L ^{13}C-glucose
100 µg/ml ampicillin

5000x Trace elements
50 mM $FeCL_3$
20 mM $CaCl_2$
10 mM $MnCl_2$
10 mM $ZnSO_4$
2 mM $CoCl_2$
2 mM $CuCl_2$
2 mM $NiSO_4$
2 mM Na_2MoO_4
2 mM Na_2SeO_3
2 mM H_3BO_3

2000x Vitamin mix
100 mM Thiamine (vitamin B_1)
100 mM Biotin (vitamin H)
100 mM Niacin (vitamin B_3)
10 mM Cobalamin (vitamin B_{12})

1000x MDG Trace elements
50 mM $FeCl_2$ in 0.1M HCl
20 mM $CaCl_2$
1 mM $MnCl_2 \cdot 4\ H_2O$
1 mM $ZnSO_4 \cdot 7\ H_2O$
2 mM $CoCl_2 \cdot 6\ H_2O$
2 mM $CuCl_2 \cdot 2\ H_2O$
2 mM $NiCl_2 \cdot 6\ H_2O$
2 mM $Na_2MoO_4 \cdot 2\ H_2O$
2 mM $Na_2SeO_3 \cdot 5\ H_2O$
2 mM H_3BO_3

5x M9 salts +aa -Cys
2.5 g/L Ala
2.0 g/L Arg
2.0 g/L Asp
2.0 g/L Gln
3.25 g/L Glu
2.75 g/L Gly
0.5 g/L His
1.15 g/L Ile
1.15 g/L Leu
2.1 g/L Lys
1.25 g/L Met
0.65 g/L Phe
0.5 g/L Pro
10.5 g/L Ser
1.15 g/L Thr
0.85 g/L Tyr
1.15 g/L Val
2.5 g/L Adenosine
1.2 g/L Cytosine
3.25 g/L Guanosine
1.0 g/L Thymine
2.5 gŁUracile
2.5 g NH_4Cl
2.5 g NaCl
15g KH_2PO_4
42.29 g $Na_2HPO_4 \cdot 2\ H_2O$

MDG medium (pH 7.4)
25 mM Na_2HPO_4
25 mM KH_2PO_4
50 mM NH_4Cl
5 mM Na_2SO_4
2 mM $MgSO_4$
0.4% (w/v) glucose
0.2% (w/v) L-aspartic acid, pH 7.0
1x MDG Trace elements
30 µg/ml kanamycin
100 µg/ml ampicillin

5x M9 salts (pH 7.4)
2.5 g NH_4Cl
2.5 g NaCl
15g KH_2PO_4
42.29 g $Na_2HPO_4 \cdot 2\ H_2O$

1x M9+aa +^{13}C, ^{15}N Cys final medium
1x M9+aa -Cys (pH 7.4)
100 mg/L Trp
100 mg/L NAD
50 mg/L Thiamine
20 mg/L ^{13}C, ^{15}N Cys
40 µg/ml Biotin
10 g/L glucose
20 mM $MgSO_4$
100 mM $CaCl_2$
100 µg/ml ampicillin
30 µg/ml kanamycin

Lysogeny broth medium after Miller (LB)
10 g/L tryptone
5 g/L yeast extract
10 g/L NaCl

Mg^{2+}/Ca^{2+} solution
3.25 g $MgCl_2 \cdot 6\ H_2O$
0.6 g $CaCl_2 \cdot 2\ H_2O$
add up to 200 ml with H_2O

5.2.3   Buffers for protein purification

GBC lysis/NTA washing buffer (pH 8.0)
50 mM sodium phosphate
500 mM NaCl

GBC NTA elution buffer (pH 8.0)
50 mM sodium phosphate
500 mM NaCl
250 mM imidazole

20 mM imidazole

GBC IEX A buffer (pH 9.0)
50 mM Tris/HCl

GBC IEX B buffer (pH 6.0)
50 mM Tris/HCl
50 mM NaCl

GBC SEC buffer (pH 8.0)
50 mM Tris/HCl
200 mM NaCl

GBC NMR buffer (pH 8.0)
50 mM Tris/HCl
200 mM NaCl
10% (v/v) D_2O
0.1% (w/v) DSS

JE28 lysis buffer (pH 7.6)
20 mM Tris/HCl
10 mM $MgCl_2$
150 mM KCl
20 mM NH_4Cl

Je28 elution buffer (pH 7.6)
20 mM Tris/HCl
10 mM $MgCl_2$
150 mM KCl
30 mM NH_4Cl
250 mM imidazole

JE28 overlay buffer (pH 7.6)
20 mM Tris/HCl
60 mM NH_4Cl
5.25-10 mM Mg acetate
0.25 mM EDTA

JE28 sucrose gradient buffer (pH 7.6)
20 mM Tris/HCl
60 mM NH_4Cl
5.25 mM Mg acetate
0.25 mM EDTA
10% (w/v) and 40% (w/v) sucrose

RNC lysis buffer (pH 7.5)
50 mM HEPES/KOH
500 mM K acetate
12 mM Mg acetate

RNC elution buffer (pH 7.5)
50 mM HEPES/KOH
500 mM K acetate
12 mM Mg acetate
500 mM imidazole

RNC sucrose gradient buffer (pH 7.5)
50 mM HEPES/KOH
1M K acetate 12 mM Mg acetate
5 mM EDTA
10% and 40% (w/v) sucrose

Tico buffer (pH 7.5)
10 mM HEPES/KOH
30 mM NH_4Cl
10 mM $MgCl_2$
1 mM EDTA

5.3   Parameters of proteins

Table 5.3: *Biophysical parameters of GBC proteins and peptides and GBC RNC constructs.*

Protein	ε at 280 nm (M^{-1} cm^{-1})	pI	Sequence length (aa)	Molecular weight (kDa)
U/H83 (NTD)	reduced 22920 oxidized 23295	7.07	89	10.82
U174	reduced 44350 oxidized 44725	7.16	181	21.92
U/H88-174 (CTD)	reduced 21430 oxidized 21430	7.10	96	11.70
U32 C15A C32A (peptide)	reduced 4470 oxidized 4495	5.46	33	3.82
U32SecM	-	7.18	61	7.14
U32SecM C15 C18	-	7.19	61	7.08
U32SecM C18 C22	-	7.19	61	7.08
U32SecM C18	-	7.20	61	7.05
U43SecM	-	7.17	73	8.49
U78SecM	-	7.19	107	12.78

5.4   Additional figures
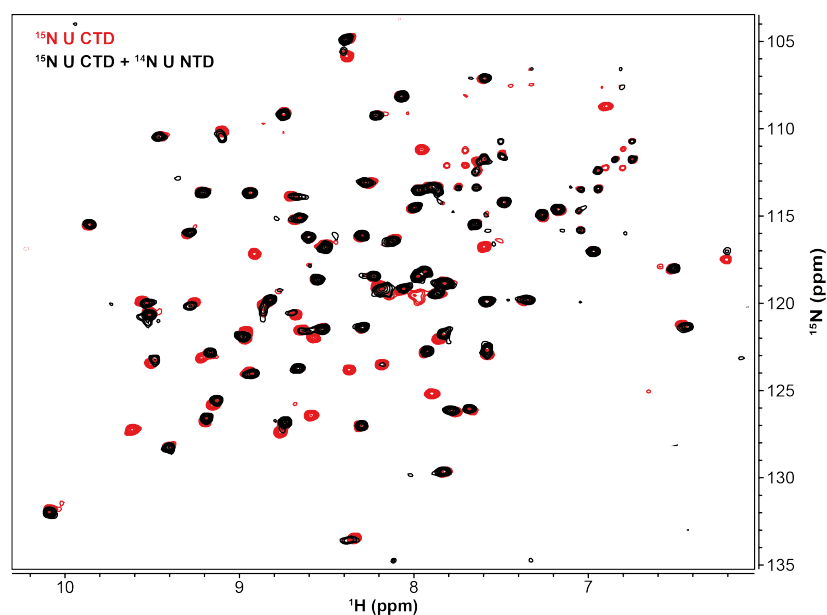
Figure 5.1: *Interaction betwenn the CTD and NTD of GBC determined by NMR spectroscopy.* *2D backbone ^1H-^{15}N-BEST-TROSY NMR spectrum of ^{15}N labeled CTD (red) and ^{15}N labeled CTD with unlabeled NTD of U. Protein concentrations: 100 µM CTD, 300 µM NTD in 50 mM Tris, 200 mM NaCl, 10% D_2O, 100 µM DSS, pH 8.0. Acquisition parameter: NS=64, TD=1024, 256, T=298K.*
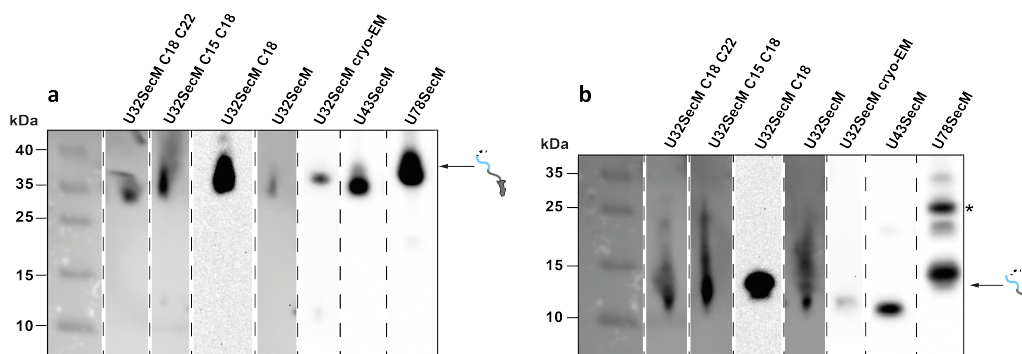
Figure 5.2: *Western blot of RNC GBC constructs with variable polypeptide chain lengths.* *a)* *Western blot analysis of purified ribosome-bound RNC.* *b)* *Western blot analysis of RNCs after treatment with RNase A. NC was detected with a poly-histidine antibody (Merck, Darmstadt, Germany). The western-blots were carried out on different gels after protein purification and were aligned to the protein marker. U78SecM showed dimer formation after release from the ribosome even under reducing conditions (*). Variations of the background are a result of an update of the imaging software.*
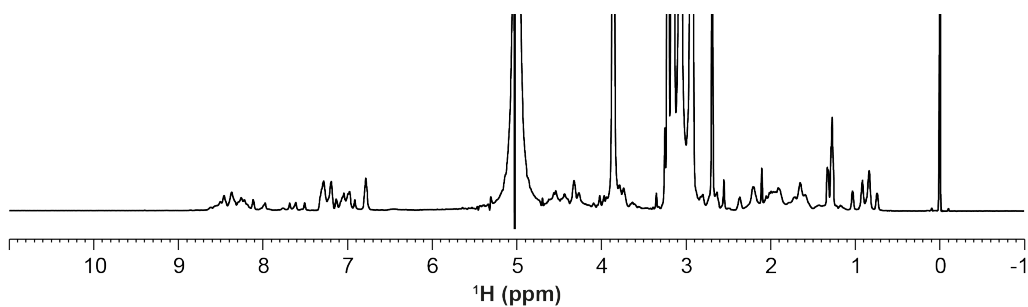
Figure 5.3: *1D ^1H NMR spectrum of $^{13}C,^{15}N$ cysteine C18 labeled U32. Protein concentration approx 164 μM in Tico buffer, 10% D_2O, 100 μM DSS, pH 7.2. Acquisition parameter: NS=64, TD=16384, T=277K.*
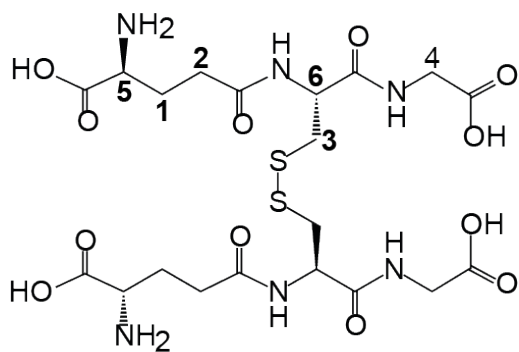
Figure 5.4: *Assignment of GSSG. Carbon chemical shifts: 1=29.230 ppm, 2= 33.924 ppm; 3=40.954 ppm, 4=46.108 ppm; 5=56.904 ppm, 6=55.041 ppm).*
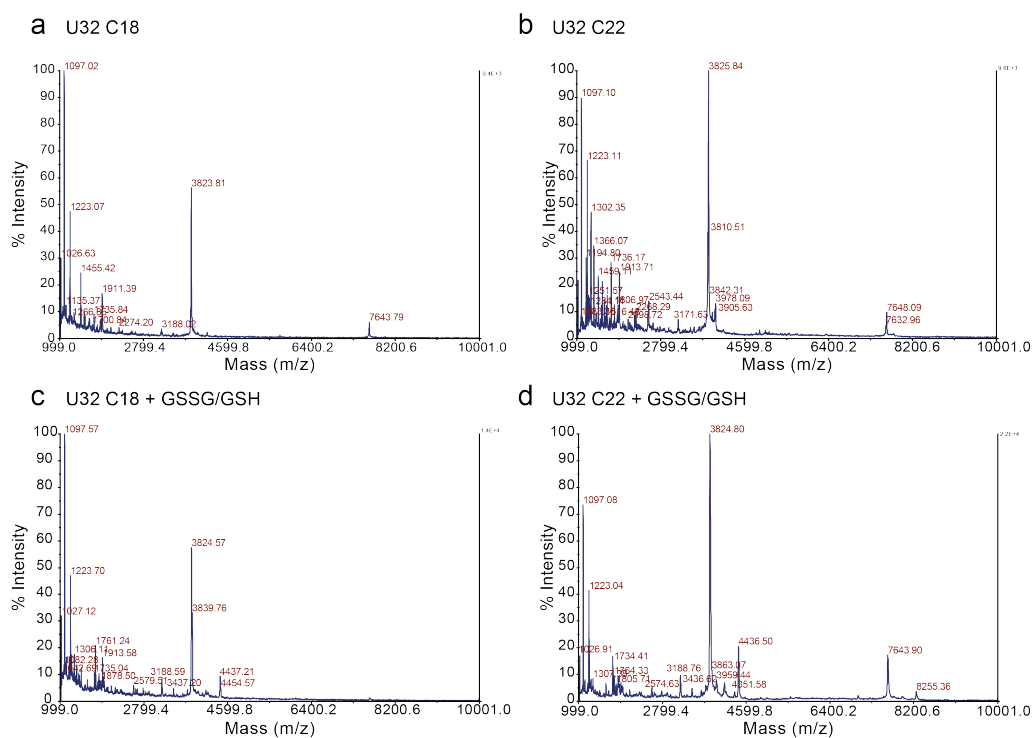
Figure 5.5: ***MALDI spectrum of U32 peptides before and after addition of GSSG/GSH.a)*** *MALDI spectrum of U32 C18.* ***b)*** *MALDI spectrum of U32 C22. Calculated mass of U32 peptides: 3824 Da. Dimer: 7648 Da.* ***c)*** *MALDI spectrum of U32 C18 peptide after addition of GSSG/GSH (final concentration 5.77 mM GSSG, 0.57 mM GSH).* ***d)*** *MALDI spectrum of U32 C22 peptide after addition of GSSG/GSH (final concentration 5.77 mM GSSG, 0.57 mM GSH). Calculated mass with GSH (307 Da) bound: 4131 Da (1 GSH bound), 4438 Da (2 GSH bound).*

5.5   Backbone assignment tables for U174, U174 after oxidation with Cu(II) and the reduced NTD H83.

Table 5.4: *Backbone assignment of full–length GBC U (not oxidized, naturally in reduced conformation).*

No.	Amino acid	^{N}H (ppm)	N (ppm)	Cα (ppm)	Cβ (ppm)	C'$^{-1}$ (ppm)
2	Lys	8.260	122.300	57.460	35.080	174.060
3	Ile	8.700	129.330	59.110	41.370	172.340
4	Thr	8.000	120.580	61.230	70.030	170.530
5	Phe	8.460	126.120	56.400	40.380	171.870
6	Tyr	8.810	118.570	56.790	39.190	174.430
7	Glu	8.910	118.740	59.440	30.710	177.700
8	Asp	7.900	115.320	52.360	43.690	177.020
9	Arg	8.680	118.810	56.860	30.580	175.720
10	Gly	8.780	105.770	47.200		177.800
11	Phe	8.110	114.500	56.000	35.020	174.110
12	Gln	6.180	116.750	53.680	31.970	174.510
13	Gly	8.270	104.920	43.950		174.270
14	His	8.760	119.280	57.920	31.240	174.200
15	Cys	7.290	118.090	53.750	33.100	173.590
16	Tyr	8.330	119.220	57.590	43.490	172.650
17	Glu	7.700	128.790	54.480	32.040	172.470
18	Cys	9.080	124.210	57.720	30.120	173.490
20	Ser	8.250	116.000	57.520	66.330	173.150
21	Asp	8.180	117.640	55.670	40.580	172.060
22	Cys	9.520	122.500	56.530	29.660	177.770
24	Asn	7.660	116.620	53.290	37.730	175.360
25	Leu	8.580	125.920	54.680	42.890	174.000
26	Gln	7.650	118.880	59.570	27.600	176.330
28	Tyr	7.300	114.710	58.050	38.860	176.440
29	Phe	7.160	112.720	57.390	38.590	175.820
30	Ser	9.770	114.700	59.970	65.270	174.070
31	Arg	8.090	119.490	54.940	30.980	174.140
32	Cys	9.660	120.010	58.380	28.000	173.040
33	Asn	8.620	121.600	53.220	43.620	172.280
34	Ser	8.420	109.170	60.430	64.400	174.290
35	Ile	9.090	118.460	60.630	45.080	172.490
36	Arg	8.780	125.210	55.730	33.030	175.490

37	Val	9.580	128.930	62.160	31.840	175.940
38	Asp	8.380	129.940	56.730	43.490	176.000
39	Ser	8.130	107.780	58.120	63.910	175.820
40	Gly	8.920	111.830	43.890		174.660
41	Cys	7.980	116.830	58.180	29.790	171.380
42	Trp	8.870	123.250	55.670	32.240	172.520
43	Met	9.680	122.810	53.150	34.350	174.240
44	Leu	10.030	127.350	52.820	43.950	174.840
45	Tyr	8.410	115.200	56.330	39.780	175.290
46	Glu	9.300	124.150	56.260	34.880	176.760
49	Asn	10.010	113.880	54.210	36.610	175.420
50	Tyr	7.770	112.860	56.260	33.760	173.850
51	Gln	6.220	116.270	53.750	32.570	175.160
52	Gly	8.600	105.950	44.280		174.840
55	Tyr	8.920	117.170	55.470	42.560	175.750
56	Phe	9.160	125.350	56.600	39.720	174.260
57	leu	9.320	132.480	55.870	47.060	175.810
58	Arg	7.030	116.140	53.820	33.360	173.480
59	Arg	8.870	119.990	58.120	31.110	175.400
60	Gly	8.630	112.590	44.480		175.880
61	Asp	8.270	119.490	53.680	42.560	170.640
64	Asp	7.150	115.180	52.950	43.820	173.580
65	Tyr	8.170	118.670	57.590	38.130	176.440
66	Gln	8.100	123.390	58.710	26.880	178.060
67	Gln	7.920	117.990	57.190	27.540	180.320
68	Trp	7.030	115.660	56.930	27.470	176.990
69	Met	7.610	113.610	56.660	29.130	174.230
70	Gly	7.710	104.930	45.610		176.210
72	Asn	7.440	112.740	51.500	37.660	175.210
73	Asp	7.780	114.500	52.950	39.050	171.660
74	Ser	7.990	112.860	57.590	63.880	175.710
75	Ile	8.850	130.360	59.380	38.190	175.190
76	Arg	8.860	121.240	56.260	33.630	174.790
77	Ser	8.310	111.650	60.370	64.800	175.880
78	Cys	9.420	112.240	55.730	32.830	173.540
79	Arg	10.100	121.000	54.210	35.080	171.450
80	Leu	8.910	126.670	53.550	41.110	173.720
81	Ile	8.880	131.930	58.510	38.130	176.210
88	Phe	9.250	127.210	57.120	40.310	174.170

89	Arg	9.280	119.590	56.790	35.150	175.000
90	Met	9.040	125.370	54.540	36.870	174.480
91	Arg	9.030	121.330	54.610	34.020	172.290
92	Ile	8.600	115.800	59.640	40.510	174.710
93	Tyr	8.200	118.330	57.260	40.510	173.540
94	Glu	8.030	118.940	58.380	32.900	176.460
95	Arg	7.970	113.400	53.950	33.430	176.650
96	Asp	8.550	118.430	53.420	40.310	175.030
97	Asp	9.280	115.870	55.400	40.250	175.270
98	Phe	8.220	109.100	54.150	33.560	173.720
99	Arg	6.510	117.910	54.870	32.570	175.300
100	Gly	8.380	104.790	44.020		175.550
101	Gln	8.820	119.660	57.790	29.460	173.550
102	Met	8.300	127.010	53.550	34.490	175.010
103	Ser	8.270	121.200	58.650	65.200	176.680
104	Glu	8.760	126.710	55.400	32.100	171.370
105	Ile	9.060	125.160	60.630	41.040	174.850
106	Thr	8.860	113.060	60.570	71.160	175.520
107	Asp	7.930	122.840	51.700	43.420	172.990
108	Asp	8.110	116.550	55.730	40.840	173.580
109	Cys	9.360	120.510	56.130	28.930	176.440
112	Leu	8.420	133.780	59.310	41.570	175.390
113	Gln	8.530	121.680	59.510	28.270	178.930
114	Asp	8.000	118.440	56.730	40.580	178.140
115	Arg	7.170	114.470	57.520	31.570	176.940
116	Phe	7.900	113.270	56.930	41.110	176.490
117	His	7.640	115.530	57.120	27.340	174.740
118	Leu	6.980	116.960	54.010	45.540	174.390
119	Thr	8.760	109.160	61.560	69.970	176.230
120	Glu	7.600	119.830	54.680	34.420	174.290
121	Val	8.500	121.340	62.090	34.820	175.270
122	His	7.280	119.700	57.320	34.490	173.410
123	Ser	8.070	108.120	59.900	64.900	176.600
124	Leu	9.840	115.220	55.540	44.350	171.900
125	Asn	9.480	120.510	51.300	40.910	174.300
126	Val	9.410	128.360	63.410	30.780	175.240
127	Leu	7.860	129.330	57.390	42.360	175.000
128	Glu	8.380	116.290	56.530	34.550	177.240
129	Gly	8.130	111.900	44.150		175.940

130	Ser	6.860	108.480	59.040	64.240	169.660
131	Trp	8.900	121.880	55.730	33.360	173.330
132	Val	9.530	120.040	61.820	34.090	173.410
133	Leu	8.900	129.610	54.280	44.550	173.770
134	Tyr	8.860	121.290	56.400	39.120	174.560
135	Glu	9.470	124.400	58.850	33.630	177.470
136	Met	8.170	110.670	52.360	33.360	174.790
138	Ser	9.510	110.470	58.580	61.030	175.160
139	Tyr	7.850	113.850	55.400	34.090	172.990
140	Arg	6.230	116.160	53.820	34.020	174.530
141	Gly	8.140	106.360	44.480		174.840
142	Arg	8.770	120.100	58.450	31.380	173.870
143	Gln	7.730	118.530	53.820	31.110	174.490
144	Tyr	8.690	115.660	56.260	40.640	175.400
145	Leu	8.860	126.400	56.130	40.780	174.790
146	Leu	9.770	128.650	52.890	43.620	176.300
147	Arg	7.750	121.540	53.090	28.600	175.460
149	Gly	9.170	113.610	44.080		176.560
150	Glu	8.280	115.930	57.320	32.100	171.000
151	Tyr	9.140	122.430	56.860	38.330	177.050
152	Arg	8.680	123.720	59.380	30.250	175.070
153	Arg	8.000	114.330	54.150	32.830	174.000
154	Tyr	7.920	119.360	57.990	37.070	176.950
156	Asp	7.920	117.890	57.190	41.040	180.600
157	Trp	7.280	114.570	57.190	27.930	178.080
158	Gly	7.570	106.910	45.470		175.940
159	Ala	6.370	121.330	51.170	22.310	173.670
160	Met	8.600	114.840	55.010	33.230	176.550
161	Asn	7.480	114.030	51.500	39.050	175.720
162	Ala	7.830	118.810	51.370	18.340	175.880
163	Lys	7.870	118.810	58.180	31.770	175.880
164	Val	7.590	122.150	61.890	35.610	177.040
165	Gly	9.100	111.220	45.410	35.810	174.240
166	Ser	8.300	112.790	58.650	64.740	174.040
167	Leu	8.780	118.470	54.870	47.920	172.570
168	Arg	9.530	121.680	54.280	36.080	174.850
169	Arg	8.500	124.690	55.730	29.590	174.230
170	Val	8.340	123.180	63.810	29.260	176.750
171	Met	7.170	125.910	54.210	36.470	172.740

| 172 | Asp | 8.450 | 122.170 | 52.950 | 41.830 | 173.680 |
| 173 | Phe | 8.390 | 120.720 | 57.920 | 40.180 | 175.650 |

Table 5.5: *Backbone assignment of full–length Cu(II)-oxidized GBC U .*

No.	Amino acid	^{N}H (ppm)	N (ppm)	Cα (ppm)	Cβ (ppm)
2	Lys	8.250	122.300	57.460	35.080
3	Ile	8.700	129.330	59.110	41.370
4	Thr	8.000	120.580	61.230	70.030
5	Phe	8.460	126.160	56.400	40.380
6	Tyr	8.810	118.570	56.790	39.190
7	Glu	8.910	118.740	59.440	30.710
8	Asp	7.900	115.320	52.360	43.690
9	Arg	8.680	118.810	56.860	30.580
10	Gly	8.780	105.770	47.200	
11	Phe	8.110	114.500	56.000	35.020
12	Gln	6.180	116.700	53.680	31.970
13	Gly	8.270	104.920	43.950	
14	His	8.760	119.280	57.920	31.240
15	Cys	7.460	116.960	53.750	33.360
16	Tyr	8.330	119.220	57.590	43.490
17	Glu	7.700	128.800	54.280	32.100
18	Cys	9.080	124.250	57.720	30.120
22	Cys	9.470	123.190	61.100	38.390
24	Asn	7.690	116.750	53.290	37.660
25	Leu	8.620	125.940	54.680	42.890
26	Gln	7.650	118.880	59.570	27.600
28	Tyr	7.300	114.710	58.050	38.860
29	Phe	7.160	112.720	57.390	38.590
30	Ser	9.770	114.720	59.970	65.270
31	Arg	8.090	119.490	54.940	34.090
32	Cys	9.660	119.990	58.450	
33	Asn	8.420	109.120	60.430	64.400
34	Ser	9.090	118.460	60.630	45.080
35	Ile	8.780	125.210	55.730	33.030
36	Arg	9.570	128.680	62.290	32.100
37	Val	8.380	129.930	56.730	43.490
38	Asp	8.000	107.250	58.210	63.820

39	Ser	8.920	111.780	43.890	
40	Gly	7.910	116.490	58.310	29.850
41	Cys	8.790	122.740	55.670	33.100
42	Trp	9.450	122.120	53.090	34.620
43	Met	10.030	127.350	52.820	43.950
44	Leu	8.410	115.200	56.330	39.780
45	Tyr	9.300	124.150	56.260	34.880
46	Glu	10.010	113.860	54.210	36.610
50	Tyr	7.830	112.480	55.800	34.090
51	Gln	6.220	116.240	53.750	32.570
52	Gly	8.600	105.950	44.280	
55	Tyr	8.920	117.170	55.470	42.560
56	Phe	9.160	125.350	56.600	39.720
57	leu	9.320	132.480	55.870	47.060
58	Arg	7.030	116.140	53.820	33.360
59	Arg	8.870	119.990	57.920	31.110
60	Gly	8.580	112.250	44.350	
61	Asp	8.170	119.340	53.620	42.560
64	Asp	7.150	115.180	52.950	43.820
65	Tyr	8.170	118.670	57.590	38.130
66	Gln	8.100	123.390	58.710	26.880
67	Gln	7.920	117.990	57.190	27.540
68	Trp	7.030	115.660	56.930	27.470
69	Met	7.610	113.590	56.660	29.130
70	Gly	7.710	104.930	45.610	
72	Asn	7.440	112.740	51.500	37.660
73	Asp	7.780	114.500	52.950	39.050
74	Ser	7.990	112.860	57.590	63.880
75	Ile	8.850	130.360	59.380	38.190
76	Arg	8.860	121.240	55.800	33.630
77	Ser	8.240	111.120	59.900	65.460
78	Cys	9.260	111.850	56.790	32.300
79	Arg	10.130	122.120	56.590	34.880
88	Phe	9.250	127.210	57.120	40.310
89	Arg	9.280	119.590	56.790	35.150
90	Met	9.040	125.370	54.540	36.870
91	Arg	9.030	121.330	54.610	34.020
92	Ile	8.600	115.800	59.440	40.510
93	Tyr	8.200	118.330	57.390	40.580

94	Glu	8.030	118.940	58.380	32.900
95	Arg	7.970	113.400	53.950	33.430
96	Asp	8.550	118.430	53.420	40.180
97	Asp	9.280	115.870	55.400	40.310
98	Phe	8.220	109.100	54.150	33.560
99	Arg	6.510	117.910	54.870	32.570
100	Gly	8.380	104.790	44.020	
101	Gln	8.820	119.660	57.790	29.460
102	Met	8.300	127.010	53.550	34.490
103	Ser	8.270	121.200	58.650	65.200
104	Glu	8.760	126.710	55.400	32.100
105	Ile	9.060	125.160	60.630	41.040
106	Thr	8.860	113.100	60.570	71.160
107	Asp	7.930	122.840	51.700	43.420
108	Asp	8.110	116.550	55.600	40.840
109	Cys	9.360	120.490	56.130	28.930
112	Leu	8.420	133.780	59.310	41.570
113	Gln	8.530	121.680	59.510	28.270
114	Asp	8.000	118.440	56.730	40.580
115	Arg	7.170	114.470	57.520	31.570
116	Phe	7.900	113.240	56.930	41.110
117	His	7.640	115.530	57.120	27.340
118	Leu	6.980	116.990	54.010	45.540
119	Thr	8.760	108.950	61.560	69.970
120	Glu	7.600	119.830	54.680	34.420
121	Val	8.500	121.340	62.090	34.820
122	His	7.280	119.700	57.320	34.490
123	Ser	8.070	108.120	59.900	64.900
124	Leu	9.840	115.220	55.540	44.350
125	Asn	9.480	120.510	51.300	40.910
126	Val	9.410	128.360	63.410	30.780
127	Leu	7.860	129.360	57.390	42.360
128	Glu	8.380	116.290	56.530	34.550
129	Gly	8.130	111.870	44.150	
130	Ser	6.860	108.500	59.040	64.240
131	Trp	8.900	121.920	55.730	33.360
132	Val	9.530	120.040	61.820	34.090
133	Leu	8.900	129.670	54.280	44.550
134	Tyr	8.860	121.370	56.400	39.120

135	Glu	9.470	124.400	58.850	33.630
136	Met	8.170	110.670	52.360	33.360
138	Ser	9.510	110.480	58.580	61.030
139	Tyr	7.850	113.850	55.400	34.090
140	Arg	6.230	116.160	53.820	34.020
141	Gly	8.170	106.280	44.610	
142	Arg	8.770	120.100	58.450	31.380
143	Gln	7.730	118.530	53.820	31.110
144	Tyr	8.690	115.660	56.260	40.640
145	Leu	8.860	126.440	56.130	40.780
146	Leu	9.770	128.650	52.890	43.620
147	Arg	7.750	121.540	53.090	28.600
149	Gly	9.170	113.610	44.080	
150	Glu	8.280	115.930	56.730	32.100
151	Tyr	9.140	122.430	56.860	38.330
152	Arg	8.680	123.720	59.380	30.250
153	Arg	8.000	114.330	54.150	32.830
154	Tyr	7.920	119.360	57.990	37.070
156	Asp	7.920	117.890	57.190	41.040
157	Trp	7.280	114.570	57.190	27.930
158	Gly	7.570	106.910	45.470	
159	Ala	6.370	121.330	51.170	22.310
160	Met	8.600	114.840	55.010	33.230
161	Asn	7.480	114.030	51.500	39.050
162	Ala	7.830	118.810	51.370	18.340
163	Lys	7.870	118.810	58.180	31.770
164	Val	7.590	122.150	61.890	35.610
165	Gly	9.100	111.200	45.410	35.810
166	Ser	8.300	112.790	58.650	64.740
167	Leu	8.780	118.470	54.870	47.920
168	Arg	9.530	121.730	54.280	36.080
169	Arg	8.500	124.690	55.730	29.590
170	Val	8.340	123.180	63.810	29.260
171	Met	7.180	125.900	54.210	36.470
172	Asp	8.460	122.120	52.950	41.830
173	Phe	8.390	120.540	57.920	40.180

Table 5.6: *Backbone assignment of the reduced NTD of GBC H82.*

No.	Amino acid	^{N}H (ppm)	N (ppm)	Cα (ppm)	Cβ (ppm)	C'$^{-1}$ (ppm)
2	Lys	8.340	122.540	57.370	35.110	173.080
3	Ile	8.700	129.160	59.010	41.240	172.310
4	Thr	8.020	120.650	61.120	69.880	170.530
5	Phe	8.420	126.070	56.380	40.450	171.890
6	Tyr	8.810	118.780	56.770	39.130	174.420
7	Glu	8.930	118.780	59.140	30.700	177.650
8	Asp	7.910	115.390	52.300	43.670	176.940
9	Arg	8.700	118.860	56.380	30.370	175.720
10	Gly	8.820	105.870	47.160		177.730
11	Phe	8.090	114.480	55.920	35.050	174.120
12	Gln	6.170	116.700	53.480	32.080	174.510
13	Gly	8.270	104.870	44.000		174.290
14	His	8.770	119.150	57.890	31.030	174.240
15	Cys	7.310	118.190	53.480	33.200	173.580
16	Tyr	8.350	119.200	57.300	43.210	172.590
17	Glu	7.760	128.830	54.400	31.950	172.540
18	Cys	9.100	124.540	57.700	30.040	173.430
19	Ser	8.910	119.160	57.630	65.330	172.870
20	Ser	8.230	116.260	57.300	66.190	173.270
21	Asp	8.160	118.010	55.590	40.380	171.970
22	Cys	9.520	122.540	56.450	29.710	177.470
24	Asn	7.650	116.390	53.350	37.810	175.300
25	Leu	8.540	125.820	54.600	43.080	174.010
26	Gln	7.660	118.780	59.340	27.800	176.330
28	Tyr	7.290	114.630	58.030	38.730	176.440
29	Phe	7.140	112.680	57.300	38.540	175.860
30	Ser	9.720	114.640	59.870	65.140	173.930
31	Arg	8.060	119.400	54.870	31.100	174.140
32	Cys	9.690	120.020	58.220	28.130	173.000
33	Asn	8.650	121.630	53.150	43.470	172.250
34	Ser	8.410	109.000	60.330	64.350	174.270
35	Ile	9.090	118.430	60.670	45.060	172.510
36	Arg	8.780	125.040	55.850	33.010	175.500
37	Val	9.610	129.040	62.110	31.890	175.990
38	Asp	8.280	129.670	56.770	43.610	176.000
39	Ser	8.180	108.160	57.960	63.950	176.040

40	Gly	8.180	108.660	43.940	63.750	176.840
41	Cys	7.550	115.640	57.700	29.380	171.510
42	Trp	8.610	122.150	55.850	32.410	172.660
43	Met	9.500	121.150	52.820	33.800	174.460
44	Leu	9.510	125.290	52.490	44.260	174.700
45	Tyr	8.400	115.780	56.120	39.850	175.150
46	Glu	9.120	123.770	57.760	33.660	176.630
47	Arg	8.110	112.120	53.550	30.830	175.480
49	Asn	10.010	113.540	54.080	36.500	175.400
50	Tyr	7.750	112.660	55.790	33.660	173.750
51	Gln	6.140	115.770	53.610	32.680	174.880
52	Gly	8.490	105.800	44.070		174.730
53	His	8.220	119.470	58.550	31.230	173.900
54	Gln	7.720	116.760	52.630	32.350	173.330
55	Tyr	8.960	122.660	55.190	41.700	174.880
56	Phe	8.590	124.300	56.580	41.240	173.720
57	leu	9.130	130.540	55.590	46.640	174.710
58	Arg	7.330	117.270	53.550	32.810	174.470
59	Arg	8.380	120.310	57.830	30.640	176.080
60	Gly	8.670	113.140	44.530		175.790
61	Asp	8.250	119.290	53.810	42.620	170.500
62	Tyr	9.240	119.280	54.730	39.390	175.850
64	Asp	7.130	115.310	52.820	43.800	173.580
65	Tyr	8.100	118.770	57.430	38.080	176.370
66	Gln	8.120	123.530	58.550	27.010	177.980
67	Gln	7.890	117.770	57.240	27.740	180.080
68	Trp	7.020	115.640	56.970	27.540	176.750
69	Met	7.630	114.510	56.450	28.590	174.570
70	Gly	7.750	105.130	45.710		176.360
71	Phe	7.810	118.900	59.140	40.510	173.860
72	Asn	7.430	112.890	51.310	37.550	175.150
73	Asp	7.750	114.390	52.560	39.000	171.710
74	Ser	7.920	112.890	57.560	63.950	175.860
75	Ile	8.850	130.140	59.540	38.540	175.140
76	Arg	8.720	121.010	55.850	33.530	174.810
77	Ser	8.250	111.510	60.130	64.810	175.890
78	Cys	9.400	112.510	55.660	33.010	173.520
79	Arg	10.030	120.280	55.000	34.980	171.660
80	Leu	8.690	126.410	53.680	41.430	173.600

| 81 | Ile | 8.490 | 128.920 | 57.500 | 37.680 | 176.350 |

5.6   R script for generation of Cα–Cβ chemical shift plots for RNCs.

```r
### environment ----

## working directory
setwd("E:/PhD projects/RNCs/DNP/Cys_chemical_shifts_R")

## packages
# install.packages("foreach")
# install.packages("extrafont")
library(ggplot2)
library(foreach)
library(grid)
library(extrafont)
# font_import()
loadfonts(quiet = TRUE)



## functions
insert_minor <- function(major_labs, n_minor) {
  labs = c(sapply(major_labs, function(x) c(x, rep("", 4))))
  labs[1:(length(labs)-n_minor)]
}



#### data processing ----

### new own data
w32 <- read.csv2("W33SecM_final_assignment.csv", skip=1,
header=TRUE)
w32[, 1] = as.numeric(as.character(w32[, 1]))
w32[, 2] = as.numeric(as.character(w32[, 2]))
w32[, 3] = as.numeric(as.character(w32[, 3]))
w43 <- read.csv2("W44SecM_final_assignment.csv", skip=1,
header=TRUE)
w43[, 1] = as.numeric(as.character(w43[, 1]))
w43[, 2] = as.numeric(as.character(w43[, 2]))
w43[, 3] = as.numeric(as.character(w43[, 3]))
w78 <- read.csv2("W79SecM_final_assignment.csv", skip=1,
header=TRUE)
w78[, 1] = as.numeric(as.character(w78[, 1]))
w78[, 2] = as.numeric(as.character(w78[, 2]))
w78[, 3] = as.numeric(as.character(w78[, 3]))
w32c15ac32a <- read.csv2("W32SecM_C15A_C32A_final_assignment2.csv",
 skip=1, header=TRUE)
w32c15ac32a[, 1] = as.numeric(as.character(w32c15ac32a[, 1]))
w32c15ac32a[, 2] = as.numeric(as.character(w32c15ac32a[, 2]))
w32c15ac32a[, 3] = as.numeric(as.character(w32c15ac32a[, 3]))
```

```
w32c22ac32a <- read.csv2("W32SecM_C22A_C32A_2.csv", skip=1,
header=TRUE)
w32c22ac32a[, 1] = as.numeric(as.character(w32c22ac32a[, 1]))
w32c22ac32a[, 2] = as.numeric(as.character(w32c22ac32a[, 2]))
w32c22ac32a[, 3] = as.numeric(as.character(w32c22ac32a[, 3]))

W32C18 <- read.csv2("W32C18_assignment.csv", skip=1, header=TRUE)
W32C18[, 1] = as.numeric(as.character(W32C18[, 1]))
W32C18[, 2] = as.numeric(as.character(W32C18[, 2]))
W32C18[, 3] = as.numeric(as.character(W32C18[, 3]))

### wang 2016 ----

## read text file
lns = readLines("Data_Wang_2016.txt", warn = FALSE)

## define and find groups in file
lbl = c(paste("Oxi", c("H", "E", "C"), sep = ", ")
        , paste("Red", c("H", "E", "C"), sep = ", "))

ids = sapply(lbl, function(i) grep(i, lns))

## reformat data for each group separately
bnd = do.call(rbind, lapply(1:length(ids), function(i) {

  # identify start and end position of current group
  st = ids[i]
  nd = ifelse(i < length(ids), ids[i+1] - 1, length(lns))
  sbs = lns[st:nd]

  # find rows with actual data
  rws = grep("([[:digit:]]|[A-Z]|[[:punct:]]){5,}", sbs,
  value = TRUE)
  rws = gsub(" $", "", rws)

  # split by space ('\\s'), ...
  lst = strsplit(rws, "\\s+")

  if (!all(sapply(lst, length) == 9)) {
    stop("Longer or shorter row identified, please check.")
  }

  # ... column bind, ...
  mat = do.call(cbind, lapply(1:9, function(j) {
    sapply(lst, "[[", j)
  }))

  # ... and return together with name of current group
  cbind(rep(names(ids)[i], nrow(mat)), mat)
}))
```

```r
## transform columns into numerics and factors
(required for grouping in ggplot)
dat = data.frame(bnd, stringsAsFactors = FALSE)
for (j in 4:ncol(dat)) {
  dat[, j] = as.numeric(dat[, j])
}

dat[, 1] = factor(dat[, 1], levels = names(ids))

## set column names
names(dat) = c("Group", "PDB chain", "BMRB", "Num"
               , "Co", "Ca", "Cb"
               , "Hn", "Ha", "Nh")


### visualize data ----

## colors, point symbols, and line types
clr = c(rep("black", 3), rep("darkslateblue", 3))
pts = rep(c(16, 17, 8), 2)
lty = rep(c("dashed", "solid", "dotted"), 2)
names(lty) = names(pts) = names(clr) = levels(dat[, 1])

## create standalone plot
sbs1 = subset(dat, Group %in% c("Oxi, H", "Oxi, E", "Red, H",
"Red, E"))

p1 = ggplot(aes(x = Cb, y = Ca, shape = Group, color = Group
                , linetype = Group), data = sbs1) +
  # geom_point(size = 2.5) + # point module
  stat_ellipse(level = .9, lwd = 1) + # ellipse module
  scale_color_manual("", values = clr) +
  scale_shape_manual("", values = pts) +
  scale_linetype_manual("", values = lty) +
  scale_x_reverse(breaks = 20:55
                    , labels = insert_minor(seq(20, 55, by = 5),
                      4)) +
  scale_y_reverse(breaks = 50:70
                    , labels = insert_minor(seq(50, 70, by = 5),
                      4)) +
  theme_bw() +
  theme(panel.grid = element_blank()) +
  #theme(text=element_text(family="Calibri", size=12)) +
  labs(x = expression("C"[beta] ~ "Chemical Shift (ppm)")
       , y = expression("C"[alpha] ~ "Chemical Shift (ppm)"))

## -"- with colored points
jnk = foreach(own = list(w32, w43, w78, w32c15ac32a,  w32c22ac32a,
 W32C18 ),
 nms = list("w32", "w43", "w78", "w32c15ac32a",  "w32c22ac32a",
  "W32C18"))
```

```r
%do% {

own = data.frame(own, Group = "Oxi, E")

p2 = ggplot(aes(x = Cb, y = Ca, shape = Group, color = Group),
data = sbs1) +
  # geom_point(size = 2.5) +
  geom_polygon(aes(x=col1, y=col2),
  data=data.frame(col1=c(35,35,50,50),
   col2=c(50,67,67,50)),
                inherit.aes = FALSE, color="#99BFA4",
                fill= "#99BFA4")+
  geom_polygon(aes(x=col3, y=col4), data=data.frame(col3=c(24,24,
  33,33),
   col4=c(50,67,67,50)),
                inherit.aes = FALSE, color="#918FB2",
                fill= "#918FB2")+
  stat_ellipse(aes(linetype = Group), level = .9, lwd = 0.7) +
  scale_color_manual("", values = clr) +
  scale_shape_manual("", values = pts) +
  scale_linetype_manual("", values = lty) +
  scale_x_reverse(breaks = 20:55
                  , labels = insert_minor(seq(20, 55, by = 5),
                   4)) +
  scale_y_reverse(breaks = 50:70
                  , labels = insert_minor(seq(50, 70, by = 5),
                   4)) +
  geom_point(aes(x = Cb, y = Ca), data = own, shape = 19
             , size = 3.5, color = "black") +
  geom_text(aes(x = Cb, y = Ca, label = number), data = own,
  hjust = -.75, vjust = -.75
             , show.legend = FALSE, fontface = 2, inherit.aes = FALSE
             , color = "black") +
  theme_bw() +
  theme(panel.grid = element_blank()
        , legend.key.width = unit(5, "line")) +
  theme(legend.position = "bottom") +
  #theme(text=element_text(family="Calibri", size=12)) +
  labs(x = expression("C"[beta] ~ "Chemical Shift (ppm)")
        , y = expression("C"[alpha] ~ "Chemical Shift (ppm)"))

## write to disk
tiff(paste0("prediction_wang_", nms, ".tiff"), width = 5.5,
height = 6.5
     , units = "in", res = 300, compression = "lzw")
#postscript(paste0("3prediction_wang_", nms, ".eps"), width = 7.8,
height = 5.5)
grid.newpage()
print(p2, newpage = FALSE)
dev.off()
```

```r
  return(p2)
}

## -"- with coils and colored points
jnk = foreach(own = list(w32, w43, w78, w32c15ac32a, w32c22ac32a,
 W32C18)
               , nms = list("w32", "w43", "w78", "w32c15ac32a",
               "w32c22ac32a", "W32C18")) %do% {
  own = data.frame(own, Group = "Oxi, E")
  p3 = ggplot(aes(x = Cb, y = Ca, shape = Group, color = Group),
  data = dat) +
    #geom_point(size = 2.5) +
    geom_polygon(aes(x=col1, y=col2), data=data.frame(col1=c(35,35,
    50,50),
     col2=c(50,67,67,50)),
                  inherit.aes = FALSE, color="#99BFA4",
                  fill= "#99BFA4")+
    geom_polygon(aes(x=col3, y=col4), data=data.frame(col3=c(24,24,
    33,33),
     col4=c(50,67,67,50)),
                  inherit.aes = FALSE, color="#918FB2",
                  fill= "#918FB2")+
    stat_ellipse(aes(linetype = Group), level = .9, lwd = .7) +
    scale_color_manual("", values = clr) +
    scale_shape_manual("", values = pts) +
    scale_linetype_manual("", values = lty) +
    scale_x_reverse(breaks = 20:55
                     , labels = insert_minor(seq(20, 55, by = 5),
                      4)) +
    scale_y_reverse(breaks = 50:70
                     , labels = insert_minor(seq(50, 70, by = 5),
                      4)) +
    geom_point(aes(x = Cb, y = Ca), data = own, shape = 19
                , size = 3.5, color = "black") +
    geom_text(aes(x = Cb, y = Ca, label = number), data = own,
    hjust = -.75,  vjust = -.75
                , show.legend = FALSE, fontface = 2,
              inherit.aes = FALSE, color = "black") +
    theme_bw() +
    theme(panel.grid = element_blank()
           , legend.key.width = unit(5, "line")
           , legend.position = "bottom", legend.direction="horizontal"
          # , text = element_text(family = "Calibri", size = 14)) +
           , text = element_text(size = 14)) +
    labs(x = expression("C"[beta] ~ "Chemical Shift (ppm)")
         , y = expression("C"[alpha] ~ "Chemical Shift (ppm)")) +
    guides(color = guide_legend(nrow = 2, byrow = TRUE))

  ## write to disk
  tiff(paste0("4prediction_wang_coil_", nms, ".tiff"), width = 5.5,
  height = 6.5 , units = "in", res = 300,
```

```
  compression = "lzw")
  #setEPS()
  #postscript(paste0("3prediction_wang_coil_", nms, ".eps")
            #, width = 5.5, height = 6.5)
  grid.newpage()
  print(p3, newpage = FALSE)
  dev.off()

  return(p3)
}
```

Acknowledgements

First of all, I would like to thank **Prof. Dr. Harald Schwalbe** for giving me the opportunity to write my PhD thesis in his group. It was a very pleasant working atmosphere and you gave me the chance to learn new techniques and develop myself. Nevertheless, you were always very supporting with this project.

Additionally, I would like to thank my collaboration partners **Dr. Jiafei Mao** and **Prof. Dr. Clemens Glaubitz**. Without the solid–state NMR experiments, I would not have been able to record NMR spectra on such a huge complex and this project would just not exist. Jiafei, you were allways available for helpful discussions for which I am very thankful.

For the cryo-EM experiments, I would like to thank **Julian Reitz**, **Victor-Valentin Hordinau** and **Prof. Achilleas S. Frangakis**. You all helped me preparing the cryo-EM samples, taking micrographs on the microscope and getting the final electron density. This was a new exciting field to me and the outcome was amazing. I also do not want to forget **Dr. Anja Seybert**, **Dr. Margot Frangakis**, **Diana Grewe** as well as **Utz Ermel**, who helped with my first steps to get the RNC sample on the grids and also sometimes on the microscope, if something was not working as planned.

Without **Dr. Denis Kudlinzki**, a model of the nascent chain within the ribosome would not exist. Thank you so much for helping me with this part of my project. Thanks to you, I learned to handle COOT and PHENIX and was able so see the nascent chain in 3D.

I would like to thank **Jakob Meier-Credo**, **Imke Wüllenweber** and **Dr. Julian Langer** for the LC–MS/MS experiments. Thanks to the finding of the cysteine modifications, all other experiments started to make sense and my hypothesis of S-glutathionylation was confirmed.

Regarding the solid-phase peptide synthesis and subsequent HPLC-purification, I would like to thank **Stephan Rawer** and **Kerstin Witt**. It was a lot of fun working with both of you and it worked out just fine.

For the support during the subgroup meetings, project days and mentoring meetings, I thank **Prof. Dr. Wachtveitl**, **Dr. Boris Fürtig**, **Dr. Sridhar Sreeramulu**, **Dr. Krishna Saxena**, **Dr. Martin Hengesbach**, **Dr. Anna Wacker** and **Vanessa de Jesus**. At this point, **Dr. Florian Buhr**, **Dr. Christian Richter**, **Dr. Nusrat Qureshi** and **György Pinter** should not be forgotten for their scientific input, corrections and help at the NMR spectrometer. Special thanks go to **Dr. Sridhar Sreeramulu** for his help with the manuscript, mentoring my thesis and discussing the results when I had a hard time interpreting them.

I also would like to thank my interns **Anja Dölle, Beatriz Sanchez Batalla, Keerthihan Thiyagarajah, Lena Sierks, Robyn Köhne** and my bachelor student **Vanessa Schlidt**, providing samples or checking the assignment of the proteins.

Publication

<u>L. Schulte</u>, J. Mao, J. Reitz, S. Sreeramulu, D. Kudlinzki, V.-V. Hodirnau, J. Meier-Credo, K. Saxena, F. Buhr, J. Langer, M. Blackledge, A. S. Frangakis, C. Glaubitz, H. Schwalbe (2020): Cysteine oxidation and disulfide formation in the ribosomal exit tunnel. *resubmitted.*

Author Contributions

Linda Schulte prepared the NMR and cryo-EM samples. Linda Schulte recorded and analyzed liquid-state NMR data. Jiafei Mao recorded solid-state NMR data and analyzed the data together with Linda Schulte. Valentin-Viktor Hodirnau supervised cryo-EM sample preparation. Valentin-Viktor Hodirnau (Institute for Biophysics, Buchmann Institute for Molecular Life Science, Goethe University Frankfurt, Frankfurt, Germany) and Linda Schulte recorded cryo-EM micrographs. Julian Reitz (Institute for Biophysics, Buchmann Institute for Molecular Life Science, Goethe University Frankfurt, Frankfurt, Germany) and Linda Schulte processed the cryo-EM electron map. Denis Kudlinzki, Linda Schulte, Julian Reitz and Achilleas S. Frangakis (Institute for Biophysics, Buchmann Institute for Molecular Life Science, Goethe University Frankfurt, Frankfurt, Germany) built the structural model and refined it. Martin Blackledge conducted the flexible-meccano simulations. Jakob Meier-Credo and Julian Langer (Max Planck Institute of Biophysics, Frankfurt, Germany) conducted mass spectrometric experiments.

Except where stated otherwise by reference or acknowledgment, the work presented was generated by myself under the supervision of my advisors during my doctoral studies. All contributions from colleagues are explicitly referenced in the thesis. The material listed below was obtained in the context of collaborative research:

Figure 3.28: **Simulation of U32SecM within the ribosomal exit tunnel of *E. coli* using flexible-meccano**, Prof. Dr. Martin Blackledge (Institut de Biologie Structurale, Grenoble, France) simulated the nascent chain inside the ribosomal exit tunnel and prepared this figure together with Linda Schulte.

Figure 3.29: **Ramachandran plots of the simulated U32SecM NC in the ribosomal exit tunnel and free in solution**, Prof. Dr. Martin Blackledge (Institut de Biologie Structurale, Grenoble, France) simulated the nascent chain inside the ribosomal exit tunnel and prepared this figure. Linda Schulte updated the axis and the description.

Figure 3.30: **Simulation of U32SecM within the ribosomal exit tunnel of *H. sapiens* using flexible-meccano**, Prof. Dr. Martin Blackledge (Institut de Biologie Structurale, Grenoble, France) simulated the nascent chain inside the ribosomal exit tunnel and prepared this figure together with Linda Schulte.

Figure 3.38: **LC-MS/MS investigation of RNC fragments**, Jakob-Meier-Credo (Max Planck Institute of Biophysics, Frankfurt, Germany) measured the RNC fragments using LC-MS/MS and prepared this figure together with Linda Schulte.

Figure 3.39: **Cryo-EM structure of U32SecM**, Julian Reitz processed the electron denisty and built the model of U32SecM into the nascent chain. The rest of the ribosome was built by Denis Kudlinzki and Linda Schulte. The ChimeraX file for this figure was provided by Julian Reitz (Institute for Biophysics, Buchmann Institute for Molecular Life Science, Goethe University Frankfurt, Frankfurt, Germany) and Prof. Dr. Achilleas S. Frangakis (Institute for Biophysics, Buchmann Institute for Molecular Life Science, Goethe University Frankfurt, Frankfurt, Germany). The Figure was prepared by Linda Schulte.

Figure 3.41: **Cryo-EM structure and structural analysis of U32SecM**, the NC in panel **c)** was simulated using flexible meccano by Prof. Dr. Martin Blackledge (Institut de Biologie Structurale, Grenoble, France). The figure was prepared by Linda Schulte.

Whenever a figure, table or text is identical to a previous publication, it is stated explicitly in the thesis that copyright permission and/or co-author agreement has been obtained.

Conference contributions

L. Schulte, F. Buhr, J. Mao, S. Jha, A. Komar, C. Glaubitz, H. Schwalbe 'Effects of synonymous codon mutations on γB crystallin within the ribosomal exit tunnel', **Poster presentation**, EMBO Ribosome Structure and Function, Strasbourgh (France), July 2016.

L. Schulte, F. Buhr, J. Mao, S. Jha, A. Komar, C. Glaubitz, H. Schwalbe 'Effects of synonymous codon mutations on γB crystallin within the ribosomal exit tunnel', **Poster presentation**, Faltertage, Halle (Saale) (Germany), October, 2016.

L. Schulte, F. Buhr, J. Mao, S. Jha, A. Komar, J. Langer, A. Frangakis, C. Glaubitz, H. Schwalbe 'Folding of γB crystallin within the ribosomal exit tunnel', **Poster presentation**, Design and Light Control, Niedernberg (Germany), August 2017.

L. Schulte, F. Buhr, J. Mao, S. Jha, A. Komar, J. Langer, A. Frangakis, C. Glaubitz, H. Schwalbe 'Folding ofγB crystallin within the ribosomal exit tunnel', **Poster presentation**, New Approaches for Investigating Nacent Peptide Folding, Cambridge (UK), December 2017.

L. Schulte, J. Mao, J. Reitz, V. V. Hordinau, K. Saxena, F. Buhr, A. Frangakis, M. Blackledge, C. Glaubitz, H. Schwalbe 'Investigation of co-translational protein folding', **Poster presentation**, 4th iNEXt Annual User Meeting Lund (Sweden), June 2019.

Curriculum Vitae

Personal Data

Name	**Linda Schulte**
Place of Birth	**Darmstadt**
Date of Birth	**12.06.1989**

Education

Since 06/15	**Goethe University of Frankfurt** *PhD student at the institute for organic chemistry and chemical biology (Prof. Schwalbe)* • Influence of synonymous codon usage on protein conformation • Co-translational folding on the ribosome determined by solid-state NMR enhanced by DNP, liquid-state NMR and cryo-EM
10/2011 – 08/2014	**University of Bayreuth** Field of study — Biochemistry and Molecular Biology Qualification — Master of Science (M.Sc.) Topic: NMR studies concerning the authenticity of natural products
10/2008 – 10/2011	**University of Bayreuth** Field of study — Biochemistry Qualification — Bachelor of Science (B.Sc.) Topic: Expression and purification of Phytochelatin synthase – Enzymes that tie and cleave peptide bonds

Internships and Activities

10/2013 – 03/2014	**ALNuMed GmbH** *Student employee* • Extraction and analysis of plant metabolites (NMR) • Machine learning-based classification of plant characteristics (country of origin, year of harvest, plant type)
07/2013	**University of Bayreuth** *Research assistant (HIWI) at the department of biopolymers* • Supervision of student practical laboratory courses • Recording NMR spectra

08/2012 – 1/2013	**University of Leiden, Netherlands**
	Research project at the protein chemistry group (Prof. Ubbink)
	• Development of steady state and stopped-flow assays for Cytochrome c Peroxidase
03/2011 – 07/2011	**University of Bayreuth**
	Research assistant (HIWI) at the department of experimental physics IV
	• Expression and purification of recombinant protein
08/2011 – 10/2011	**University of California, Davis, USA**
	Internship at the metabolic lab (Prof. Fiehn)
	• Extraction of *Arabidopsis* metabolites and validation of gas chromatography–mass spectrometry
04/2008 – 05/2008	**BASF AG, Ludwigshafen**
	Internship
	• Implementation of daily analytics • Collaboration to approach a testing plant

List of academic teachers

Prof. Dr. Wulf Blankenfeldt, Prof. Dr. Breu, Prof. Dr. Klaus Ersfeld, Prof. Dr. Andreas Fery, Dr. Frank Fischer, Prof. Dr. Andreas Golembiowski, Dr. Stephan Heidmann, Prof. Dr. Thomas Hellweg, Prof. Dr. Klaus H. Hoffmann, Prof. Dr. Ortwin Meyer, Prof. Dr. Paul Rösch, Prof. Dr. Ernst Rößler, Prof. Dr. Franz. X. Schmid, Dr. Holger Schmidt, Prof. Dr. Rainer Schobert, Prof. Dr. Wolfgang Schumann, Prof. Dr. Stephan Schwarzinger, Dr. Kristian Schweimer, Prof. Dr. Jürgen Senker, Prof. Dr. Mathias Sprinzel, Prof. Dr. Clemens Steegborn, Dr. Steinle-Neumann, Prof. Dr. Olaf Stemmann, Prof. Dr. Matthias Ullmann, Prof. Dr. Benedikt Westermann, Prof. Dr. Birgitta Wöhrl

Erklärung

Ich erkläre an Eides statt, dass ich meine Dissertation

Investigation of co–translational protein folding using cryo–EM and solid-state NMR enhanced by DNP

selbstständig und ohne unerlaubte Hilfe angefertigt und mich dabei keinerlei anderer als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch an keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Frankfurt am Main, den

Linda Schulte