



University
of Glasgow

Cherry, Jon (2016) *A biochemical analysis into the co-translational folding of a G protein-coupled receptor; GPR35*. PhD thesis.

<http://theses.gla.ac.uk/7527/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

A biochemical analysis into the co-
translational folding of a G protein-coupled
receptor; GPR35

Jon Cherry
BSc, MRes

Submitted in fulfilment of the requirements
for the Degree of Doctor of Philosophy

Institute of Molecular, Cell and Systems Biology
College of Medical, Veterinary and Life Sciences

University of Glasgow

2016

Abstract

The folding and targeting of membrane proteins poses a major challenge to the cell, as they must remain insertion competent while their highly hydrophobic transmembrane (TM) domains are transferred from the ribosome, through the aqueous cytosol and into the lipid bilayer. The biogenesis of a mature membrane protein takes place through the insertion and integration into the lipid bilayer. A number of TM proteins have been shown to gain some degree of secondary structure within the ribosome tunnel and to retain this conformation throughout maturation. Although studies into the folding and targeting of a number of membrane proteins have been carried out to date, there is little information on one of the largest class of eukaryotic membrane proteins; the G-protein-coupled receptors (GPCRs).

This project studies the early folding events of the human ortholog of GPR35. To analyse the structure of the 1st TM domain, intermediates were generated and assessed by the biochemical method of pegylation (PEG-MAL). A structurally-similar microbial opsin (Bacterioopsin) was also used to investigate the differences in the early protein folding within eukaryotic and prokaryotic translation systems. Results showed that neither the 1st TM domain of GPR35 nor Bacterioopsin were capable of compacting in the ribosome tunnel before their N-terminus reached the ribosome exit point. The results for this assay remained consistent whether the proteins were translated in a eukaryotic or prokaryotic translation system.

To examine the communication mechanism between the ribosome, the nascent chain and the protein targeting pathway, crosslinking experiments were carried out using the homobifunctional lysine cross-linker BS³. Specifically, the data generated here show that the nascent chain of GPR35 reaches the ribosomal protein uL23 in an extended conformation and interacts with the SRP protein as it exits the ribosome tunnel. This confirms the role of SRP in the co-translational targeting of GPR35. Using these methods insights into the early folding of GPCRs has been obtained. Further experiments using site-directed mutagenesis to reduce hydrophobicity in the 1st TM domain of GPR35, highlighted the mechanisms by which GPCRs are targeted to the endoplasmic reticulum. Confirming that hydrophobicity within the signal anchor sequence is essential of SRP-dependent targeting.

Following the successful interaction of the nascent GPR35 and SRP, GPR35 is successfully targeted to ER membranes, shown here as dog pancreas microsomes (DPMs). Glycosylation of the GPR35 N-terminus was used to determine nascent chain structure as

it is inserted into the ER membrane. These glycosylation experiments confirm that TM1 has obtained its compacted state whilst residing in the translocon. Finally, a site-specific cross-linking approach using the homobifunctional cysteine cross-linker, BMH, was used to study the lateral integration of GPR35 into the ER. Cross-linking of GPR35 TM1 and TM2 could be detected adjacent to a protein of ~45kDa, believed to be Sec61 α . The loss of this adduct, as the nascent chain extends, showed the lateral movement of GPR35 TM1 from the translocon was dependent on the subsequent synthesis of TM2.

Acknowledgements

I would like to start by thanking the person who had the biggest influence on my PhD; my supervisor Dr. Cheryl Woolhead. From the moment I started, her guidance, knowledge and general support throughout this project has been invaluable to me. Her open door policy, relaxed demeanour and ever presence around the lab has made her a pleasure to work for and for that I will be ever thankful.

To all the members of the Woolhead lab both past and present, including Dr. Philip Robinson, Dr. Hazel Bracken, David Stewart and Donald Campbell. A special mention has to be made to Phil, who was forever willing to listen and help with any problems I may have had. Additionally, the members of the other groups in Lab 310, especially Prof. William Cushley, who is a true gent and always had time for a chat, and Dr. Phoebe Sharp who provided me with many a laugh.

To all my friends and family who have supported me throughout. Firstly, I would like to dedicate this thesis to my late grandfather, John Cherry. Without him none of this would have been possible and I hope I've made him proud. Secondly, my parents, who have been a constant pillar of support never once, questioning my decisions or motives. My brother Matthew and my flatmate Ad have been there throughout, always willing to lend a helping hand. Last but not least my girlfriend Emeline, who had to put up with me throughout, who had to read my entire thesis and whose support was undoubtedly what got me to the endpoint.

Finally, a big thank you must go to the EPSRC for funding this project. To the people who provided me with reagents throughout the 3 years; Graeme Milligan, Neil Bullied, Stephen High and Art Johnson, my assessors and finally, I would like to thank the University of Glasgow for giving me this opportunity.

Table of Contents

Abstract	2
Acknowledgements	4
Table of contents	5
List of Figures	10
List of Tables	13
Author’s declaration of originality	14
Abbreviations	15
1. Introduction	21
1.1. The Ribosome	21
1.1.1 Ribosome Overview	21
1.1.2. Ribosome structure.....	21
1.1.3 Ribosome function	24
1.1.3.1 Translation overview.....	24
1.1.3.2. Translation initiation	25
1.1.3.2. Translation elongation.....	29
1.1.3.4 Translation termination and recycling of ribosomes.....	32
1.1.4 The ribosome exit tunnel.....	34
1.1.4.1 The structure of the ribosome exit tunnel	34
1.1.4.2 Folding in the ribosome tunnel	36
1.1.4.3 Ribosome tunnel-nascent chain interactions.....	37
1.1.4.4 Nascent peptides at the ribosome exit site	38
1.2 Co-translationally-mediated targeting	39
1.2.1 Overview of SRP-dependant co-translational targeting.....	39
1.2.2 Components of the SRP system: A structural overview	40
1.2.3 The interaction between SRP and the RNC	43
1.2.4 RNC delivery to the membrane by SRP	45

1.2.5 Interaction between the ribosome and Sec translocon	46
1.3 Membrane protein insertion into the endoplasmic reticulum	48
1.3.1 Overview of the endoplasmic reticulum	48
1.3.2 The Sec61 translocon: structure and function	49
1.3.3 Sec61 accessory factors.....	51
1.3.4 Determinants of membrane proteins topology	54
1.3.5 Lateral movement of a TM domain from the Sec61 translocon	57
1.4. G protein-coupled receptors (GPCRs)	59
1.4.1 GPCR biogenesis: What is known?	59
1.4.2 G protein-coupled receptor 35: A model GPCR.....	60
1.5 Project Aims	62
2. Materials and Methods:	63
2.1 General Reagents:	63
2.2 General Buffers	66
2.3 <i>Escherichia coli</i> strains and plasmid vectors	69
2.3.1 <i>Escherichia coli</i> strains.....	69
2.3.2 Plasmid Vectors	70
2.4 General Methods	72
2.4.1 Preparation of LB plates:	72
2.4.2 Preparation of bacterial competent cell stocks:.....	72
2.4.3 Transformation of competent bacteria:	72
2.4.4 Small-scale preparation of plasmid DNA (MINI-PREP):	73
2.4.5 Large-scale preparation of plasmid DNA (MAXI-PREP)	73
2.4.6 Polymerase Chain Reaction (PCR).	74
2.4.6.1 DNA amplification by PCR	74
2.4.6.1 Site-directed mutagenesis by PCR.....	75
Single site directed mutagenesis	75

Multiple site directed mutagenesis	77
2.4.7 PCR purification	78
2.4.8 DNA agarose gel electrophoresis	78
2.4.9 DNA purification from agarose gel	79
2.4.10 Restriction Endonuclease Digestion	79
2.4.11 Ligation of DNA	80
2.4.12 DNA sequencing	81
2.4.13 <i>In vitro</i> transcription.....	81
2.5 <i>In vitro</i> translation systems.....	82
2.5.1 Preparation of <i>E. coli</i> S-30 extract	82
2.5.2 <i>In vitro</i> transcription/translation in <i>E. coli</i> S-30 Extract translation system.....	83
2.5.3 <i>In vitro</i> translation in Wheat Germ extract translation system	84
2.5.4 <i>In vitro</i> translation in Rabbit Reticulocyte Lysate translation system	85
2.5.5 TCA precipitation of translated peptides	86
2.5.6 Pegylation assay to assess compaction of nascent chain in the ribosome exit tunnel.	86
2.5.7 Chemical cross-linking of ribosome-bound nascent chains using BS ³	87
2.5.8 Chemical cross-linking of ribosome-bound nascent chains to Sec61 α using BMH	87
2.5.9 Immunoprecipitation of cross-linked products	88
2.5.10 Digestion assay assessing insertion into Dog Pancreas Microsomes (DPMs).....	89
2.6 Imaging	89
2.6.1 Gel electrophoresis and western blotting	89
2.6.2 Autoradiography	90
2.6.3 Immunoblotting.....	90
2.7 Statistical analysis	90
3. Investigating secondary structure formation in transmembrane domain 1 of hGPR35 and Bacterioopsin	91

3.1. Introduction	91
3.2. Results	95
3.2.1. Investigating the formation of secondary structure in TM1 of GPR35 and Bacterioopsin within the prokaryotic ribosome.	95
3.2.2. Investigating the formation of secondary structure in TM1 of GPR35 and Bacterioopsin within the eukaryotic ribosome.....	101
3.2.3. F ₀ c- pegylation of a bacterial membrane protein known to fold in the ribosome exit tunnel.....	105
3.2.4. Is the lack of isoleucines in the first TM of GPR35 and BO linked to the lack of secondary structure formation in the ribosome exit tunnel?	107
3.3. Discussion	111
4. The importance of secondary structure for SRP-mediated targeting?	116
4.1. Introduction	116
4.1. Results	116
4.2.1 Is the GPR35 nascent chain capable of interacting with SRP?.....	121
4.2.2. Does hydrophobicity drive the interaction between SRP and the first TM of GPR35?	127
4.2.3 Does SRP bind preferentially to a region of the first TM domain in GPR35?	130
4.3 Discussion	134
5: Targeting and integration of GPR35 to the Endoplasmic Reticulum.	139
5.1 Introduction	139
5.2 Results	146
5.2.1 Proteinase K digestion to determine successful targeting and integration of GPR35 to DPMs	146
5.2.3. Does a loss of hydrophobicity in the 1 st TM domain result in a loss of insertion into the translocon?	150
5.2.2 Assessing the secondary structure of GRP35 transmembrane domain 1 during translocation.	153

5.2.4. Analysis of GPR35 integration into the ER membrane by site-specific cross-linking.	157
5.3. Discussion	161
6. Final Discussion	166
6.1 Analysing secondary structure formation within the ribosome tunnel of TM 1 in a model GPCR.	166
6.2 Co-translational targeting and TM domain biogenesis in GPR35.	171
6.3 Conclusion	174
6.4 Future directions	176
Appendix	178
Appendix 1: Table of constructs	178
Appendix 2: Primers used for the generation of linear DNA intermediates.....	181
Appendix 3: Gene Sequences	186
Appendix 4: Western blot analysis of antibodies used in chapter 4	188
Bibliography	166

List of Figures:

Figure 1.1 Structure of the 70S prokaryotic ribosome.....	23
Figure 1.2 Prokaryotic translation initiation.	26
Figure 1.3 Schematic overview of eukaryotic translation initiation.	28
Figure 1.4 Schematic overview of translation elongation.....	30
Figure 1.5 Schematic overview of deacylated tRNA and peptidyl tRNA during elongation.....	31
Figure 1.6 Schematic overview of translation termination	33
Figure 1.7 The ribosome exit tunnel makes it way through the large ribosome subunit	35
Figure 1.8 Schematic overview of the signal recognition particle (SRP) targeting pathway.	39
Figure 1.9 Schematic representations of signal recognition particles from human and E.coli.	40
Figure 1.10 Crystal structure of core SRP domain and SRP bound to a nascent chain.	42
Figure 1.11 Structure of the prokaryotic and eukaryotic ribosome at the ribosome tunnel exit site.	44
Figure 1.12 Schematic diagram of Asparagine –linked (N-linked) glycosylation.....	49
Figure 1.13 Schematic diagram of co-translationally targeted membrane protein topology	57
Figure 1.11 Schematic representation of GPR35.....	61
Figure 3.1 Pegylation assay to measure the level of secondary structure formed within the ribosome exit tunnel.	94
Figure 3.2 Diagram to show how ribosome bound nascent chains (RNCs) are generated.	96
Figure 3.3. Analysis of the pegylation assay of GPR35 in the prokaryotic S-30 transcription/translation system.	97
Figure 3.4. Analysis of the pegylation assay of Bacterioopsin in the prokaryotic S-30 transcription/translation system.	99

Figure 3.5 Analysis of the pegylation assay of GPR35 in the eukaryotic translation system.....	103
Figure 3.6. Analysis of the pegylation assay of Bacterioopsin in the eukaryotic translation system.....	104
Figure 3.7. Pegylation assay control using F0c in the prokaryotic S-30 transcription/translation system.	106
Figure 3.8 Analysis of the pegylation assay of GPR35 Δ 4I in the eukaryotic translation system.....	110
Figure 4.1 Schematic diagram describing how integral membrane proteins are co-translationally targeted to the membrane.	117
Figure 4.2 Cross-linking assay to measure the level of interaction with components of the SRP targeting pathway.	119
Figure 4.3 Cross-linking assay of GPR35 to uL23 and Ffh in the prokaryotic S-30 transcription/translation system	122
Figure 4.4 Cross-linking assay of BO to uL23 and Ffh in the prokaryotic S-30 transcription/translation system.	124
Figure 4.5 Cross-linking assay of GPR35 to SRP54 in rabbit reticulocyte lysate system.....	126
Figure 4.6 Cross-linking assay of GPR35 Δ 4E mutant to uL23 and Ffh in the prokaryotic S-30 transcription/translation system.	128
Figure 4.7 Negatively charged amino acids reduce the hydrophobicity in TM domain 1.....	132
Figure 4.8 Cross-linking assay of GPR35 hydrophobic mutants to SRP54 in rabbit reticulocyte lysate system.	133
Figure 5.1 Sec61 translocon located in the mammalian ER membrane.	140
Figure 5.2 Experimental design of glycosylation and chemical cross-linking assays	145
Figure 5.3 Assay to assess insertion and orientation of GPR35 into DPMs.....	147
Figure 5.4 Glycosylation assay to assess translocation of GPR35 TM1 hydrophobic mutants.	151

Figure 5.5. Analysis GPR35 glycosylation in DPMs to detect the formation of secondary structure in the Sec61 translocon.	155
Figure 5.6. Integration GPR35 TM domain 1 into the DPM membrane	151
Figure 6.1. Schematic diagram of two possible scenarios describing the folding profile occurring in the 1st TM domain of GPR35 as it makes its way through the ribosome exit tunnel.....	170
Figure 6.2. Schematic diagram of our suggested model of biogenesis at the N-terminus of GPR35.....	103
Figure A4. Western blot analysis of components of the SRP mediated pathway in the S-30, WG and RRL extract.	103

List of Tables

Table 3.1. Transmembrane domains known to fold whilst in the ribosome tunnel.	108
Appendix A1.1. GPR35 construct mutations for pegylation assays	178
Appendix A1.2. BO and F _{0c} construct mutations for pegylation assays	178
Appendix A1.3. BO construct mutations for uL23 and SRP cross-linking assays	179
Appendix A1.4 GPR35 Δ NT, Δ CT and Δ 4E construct mutations for uL23 and SRP cross-linking assays	181
Appendix A1.5. GPR35 construct mutations for N-linked glycosylation assays	180
Appendix A2.1. Primers used in to generate GPR35 intermediates for pegylation assays	181
Appendix A2.2. Primers used in to generate BO intermediates for pegylation assays.....	183
Appendix A2.3. Primers used in to generate F _{0c} intermediates for pegylation assays.....	184
Appendix A2.4. Primers used in to generate GPR35 intermediates for SRP cross-linking assays	181
Appendix A2.5. Primers used in to generate BO intermediates for SRP cross-linking assays	183
Appendix A2.6. Primers used in to generate GPR35 intermediates for PK digestion and C-terminal N-linked glycosylation assays.....	184
Appendix A2.7. Primers used in to generate GPR35 intermediates for N-terminal N-linked glycosylation assays	184
Appendix A2.8. Primers used in to generate GPR35 intermediates for site specific cysteine cross-linking in TM1 and TM2 of GPR35.....	184

Author's declaration of originality

I declare that the work put forward for this thesis has been carried out by myself, unless otherwise stated. It is entirely of my own composition and has not been submitted, in whole or in part, for any other degree.

Jon Cherry

April 2016

Abbreviations

~	Approximately
°C	Degrees Celsius
Å	Angstroms
aa	amino acids
Ala(A)	Alanine
APS	Ammonium Persulphate
Arg (R)	Arginine
Asn (N)	Asparagine
Asp (E)	Aspartic Acid
ATP	Adenosine 5'-triphosphate
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
BMH	Bismaleimide
BO	Bacterioopsin
bp	Base Pair
BS³	Bis(sulfosuccinimidyl)suberate
BSA	Bovine Serum Albumin
CaCl₂	Calcium Chloride
cAMP	Adenosine 3', 5'-cyclic Monophosphate
Ci	Curie

CL	Cross-linking
cryo-EM	Cryo-Electron Microscopy
CTAB	Cetyltrimethylammonium Bromide
Cys (C)	Cysteine
Da	Dalton
ddH₂O	Double Distilled Water
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetracetic Acid
EF	Elongation Factor
Endo H	Endoglycosidase H
ER	Endoplasmic Reticulum
FRET	Fluorescence Resonance Energy Transfer
g	Gram
GDP	Guanosine 5' diphosphate
Gly (G)	Glycine
GPCR	G protein-coupled receptor
GTP	Guanosine-5'-Triphosphate

HCl	Hydrochloric Acid
HEPES	2-[4-(2-Hydroxyethyl)-1-Piperazine] Ethanesulphonic Acid
IF	Initiation Factor
Ile (I)	Isoleucine
IMP	Integral Membrane Protein
IP	Immunoprecipitation
IPTG	Isopropyl- β -D-thiogalactopyranoside
k	Kilo
K₂HPO₄	di-Potassium Hydrogen Phosphate
kb	Kilo Base Pair
KCl	Potassium Chloride
kD	Kilo Dalton
KGlu	Potassium Glutamate
KH₂PO₄	Potassium Dihydrogen Orthophosphate
KOAc	Potassium Acetate
KOH	Potassium Hydroxide
L	Litre
LB	Lysogeny Broth
Lys (K)	Lysine
m	Milli

M	Molar
Met (M)	Methionine
MgCl₂	Magnesium Chloride
MgOAc	Magnesium Acetate
MgSO₄	Magnesium Sulphate
mRNA	Messenger RNA
MW	Molecular Weight
MWCO	Molecular Weight Cut-off
n	Nano
Na₂HPO₄	Disodium Hydrogen Orthophosphate
NaCl	Sodium Chloride
NaOAc	Sodium Acetate
NH₄OAc	Ammonium Acetate
NMR	Nuclear Magnetic Resonance
OD	Optical Density
ORF	Open Reading Frame
OST	Oligosaccharyl transferase centre
p	Pico
PCR	Polymerase Chain Reaction
PEG-MAL	Methoxypolyethylene Glycol Maleimide

Phe (F)	Phenylalanine
PIC	Pre-Initiation Complex
PMSF	Phenylmethanesulphonyl Fluoride
Pro (P)	Proline
psi	Pounds Per Square Inch
PTC	Peptidyl Transferase Centre
RF	Release Factor
RNA	Ribonucleic Acid
RNaseA	Ribonuclease A
RNC	Ribosome-Nascent Chain
rNTP	Ribonucleotide Triphosphate
rpm	Rotations Per Minute
r-protein	Ribosomal protein
RRF	Ribosome Recycling Factor
r-RNA	Ribosomal RNA
s	Seconds
SD	Shine-Dalgarno
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SecM	Secretion Monitor

Ser (S)	Serine
SOB	Super Optimal Broth
SR	SRP receptor
SRP	Signal recognition particle
<i>T. thermophilus Thermus thermophilus</i>	
TAE	Tris Acetate EDTA
TCA	Trichloroacetic acid
TEMED	N,N,N',N'-Tetramethylethylenediamine
tmRNA	Transfer-messenger RNA
Tris	Tris (hydroxymethyl)aminoethane
tRNA	Transfer RNA
UTR	Untranslated Region
UV	Ultraviolet
v/v	Volume to Volume ratio
w/v	Weight to Volume Ratio
WT	Wild Type
α-helix	Alpha Helix
β-sheet	Beta Sheet
μ	Micro

1. Introduction

1.1. The Ribosome

1.1.1 Ribosome Overview

The ribosome is a multifunctional complex made up of ribosomal proteins (r-proteins) and ribosomal RNAs (r-RNAs) that combine to become the ‘intracellular protein nanofactory’ of the cell. Genomic DNA is transcribed into messenger RNA (mRNA) by the actions of RNA polymerase and in turn is translated by the ribosome, producing nascent polypeptides by the addition of amino acids (Green & Noller 1997). The rRNA moieties of the ribosome provide the catalytic activity to generate peptide bonds at the peptidyl transferase centre (PTC), which form at a 2×10^7 fold rate of enhancement (Sievers et al. 2004). The ribosome was often thought to be non-discriminating against the peptides it generated, but accumulating evidence suggests that is not the case, as more and more proteins have been identified to interact with the ribosome during translation. R-proteins and r-RNAs are believed to play vital roles in processes such as tRNA selection and binding, formation of secondary structure, translocation and interactions with targeting factors/chaperones.

1.1.2. Ribosome structure

All ribosomes consist of two subunits, made up of RNA and proteins. The 70S bacterial ribosomes contain a small subunit (30S), made up of one 16S r-RNAs and 21 r-proteins. The large subunit (50S) is made up of 5S and 23S rRNAs and 33 r-proteins. X-ray crystallographic structures of the individual subunits and complete 70S ribosome of *T. thermophilus* revealed the complex make-up and interactions that occur between r-proteins and r-RNA (Ban 2000). The small subunit 16S rRNA can be split into four domains and is the centre point to which r-proteins interact. The rRNA of the 50S subunit can be split into 7 domains and in contrast to the 30S is made up of interwoven RNA and r-proteins (Ban et al. 2000; Yusupova et al. 2001). Eukaryotic ribosomes, unlike their prokaryotic counterparts, are much larger and have a higher degree of complexity to their structure. The 80S ribosome made up of a small 40S subunit and large 60S subunit, containing 4 r-

RNAs (18S in the small subunit and 5S, 5.8S and 28S in the large subunit) and ~80 r-proteins, depending on the species.

Although the mass and structural complexities of the eukaryotic and prokaryotic ribosomes differ considerably, much of the core structure is highly conserved. In fact, early cryo-EM maps of the eukaryotic ribosome were built around the crystal structure of the prokaryotic 30S and 50S subunits (Wimberly et al. 2000; Ban et al. 2000) and enabled the identification of 46 eukaryotic r-proteins with bacterial and/or archaeal homologs (Spahn et al. 2001). One specific region that is conserved in all species of ribosome is the main catalytic and active site, which can be found at the interface between the large and small subunits; this region is known as the PTC (Figure 1.1). Through X-ray crystallographic experiments this region was shown to be made up solely of rRNA (Nissen et al 2000), with both ribosomal subunits contributing to the 3 tRNA binding sites, named the A (aminoacyl) site, which is required to accept incoming aminoacylated tRNA for synthesis; the P (peptidyl) site, which holds the tRNA molecule in place as the nascent chain is extended; and the E (exit) site, which holds the deacylated tRNA after dissociation with the nascent chain and prior to exit from the ribosome (Schmeing & Ramakrishnan 2009) (Figure 1.1 B). In each ribosome, the small subunit specifically binds mRNA and the anti-codon stem loops of the tRNA, ensuring the translation reaction is carried out effectively by monitoring the base pairing between the codon and anti-codons. The large subunit, catalyzes the peptide bond formation between the incoming amino acids on the A site tRNA and the nascent chain at the P site.

Although prokaryotic and eukaryotic ribosomes share a relatively conserved core, there are a number of variations within their composition. The differences between the two subsets of ribosomes mainly occur due to changes in sequence and lengths of r-proteins and rRNAs, known as expansion segments. Adaptations in cellular growth and stress conditions can be enough to vary rRNA elements, known as expansion bodies, proteins and add insertions to existing proteins (Jenner et al. 2012). Specifically, there were 32 extra r-proteins found within the ribosome crystal structure of yeast that showed no prokaryotic homology, with the majority of these found to be located on the solvent exposed surface of the ribosome. Such differences as these are believed to enable the eukaryotic ribosomes to interact with eukaryotic specific translation factors and chaperones (Ben-Shem et al. 2010)

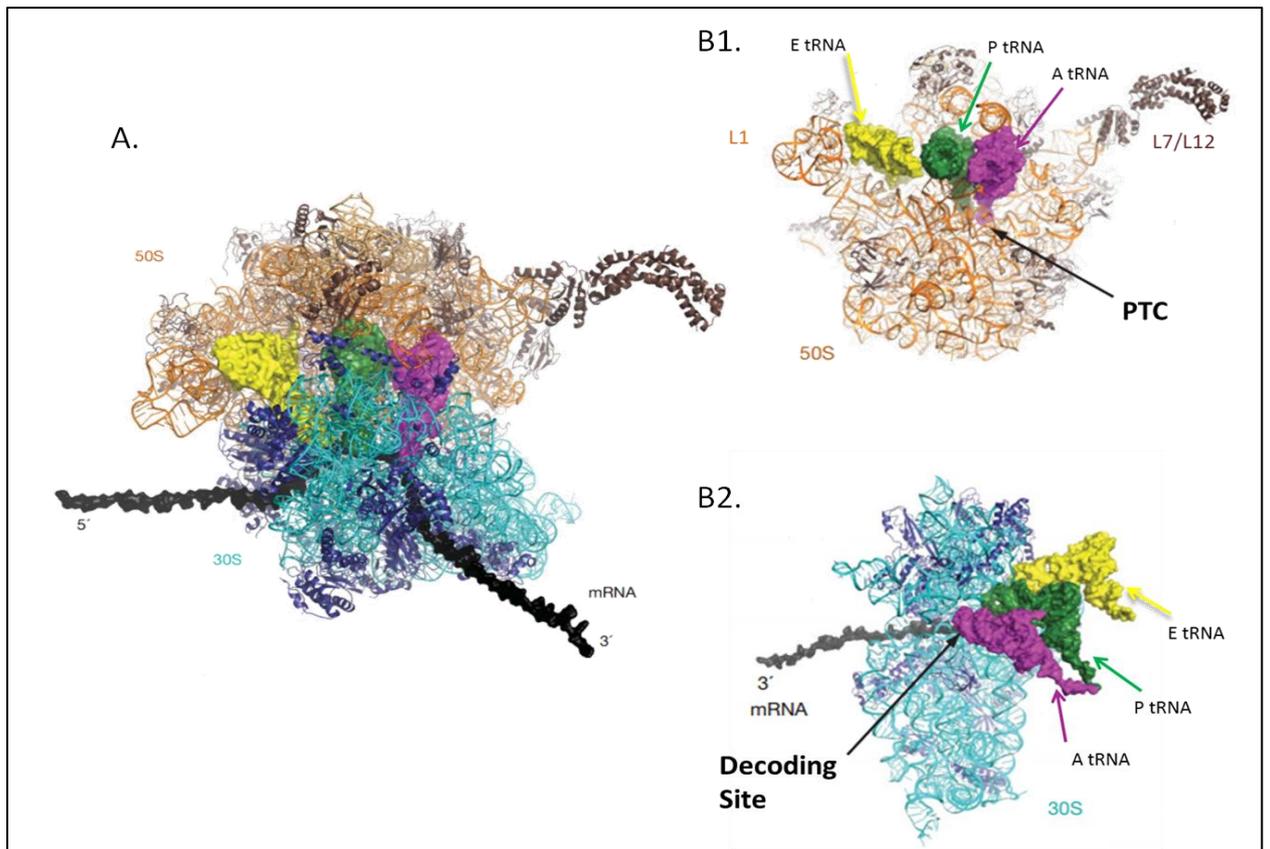


Figure 1.1 Structure of the 70S prokaryotic ribosome. A. An x-ray crystallographic structure of the 70S ribosome containing mRNA and displaying A-, P- and E-tRNA sites. B. Separated 30S (B1) and 50S subunits (B2). The 30S subunit is displayed in orange and the 50S subunit in blue, with the A-, P- and E-sites coloured magenta, green and yellow respectively. Figure reproduced from Schmeing and Ramakrishnan (2009).

The difference in complexity of the prokaryotic and eukaryotic ribosomes is believed to be due to the differences seen throughout the prokaryotic and eukaryotic cell as a whole. The initial assembly in eukaryotes commences in the nucleolus and progresses to the nucleoplasm where it requires >350 specialized factors to perform specific assembly steps on pre-ribosomal particles. Following this, the pre-ribosomes are exported to the cytoplasm and undergo final maturation before they can begin translation (Gerhardy et al. 2014). Subsequently, prokaryotic ribosomes do not undergo the same level of assembly due to their lack of nucleus; hence there is no compartmentalisation in bacterial cells. The number of maturation factors required is also much lower than in eukaryotes.

Eukaryotes themselves also contain a number of structurally distinct ribosome species. Alongside the ribosomes existing in the cytoplasm, the mitochondrion and chloroplasts of eukaryotes contain structurally different ribosomes. Mitochondrial ribosomes differ from

cytoplasmic and bacterial ribosomes, firstly in size and secondly in the number of proteins they synthesize. The ribosomes in the mitochondrion are made up from the large 39S subunit and the small 28S subunit (Greber et al. 2015). As the mitochondria only generate a small number of proteins themselves, the mito-ribosomes are only required to produce 13 peptides used for ATP synthesis or oxidative phosphorylation (Breiman et al. 2015). The chloroplastic ribosomes are very similar in composition and overall structure to eubacterial ribosomes. They too consist of a 30S and 50S subunit, generating a 70S ribosome. The ribosomes in a chloroplast differ slightly from bacterial ribosomes in that they contain some plastid-specific ribosomal proteins (PSRBs). These proteins are believed to have some structural relevance, yet this cannot be confirmed due to the lack of a high resolution crystal structure (Breiman et al. 2015).

1.1.3 Ribosome function

1.1.3.1 Translation overview

The ribosome is a large ribonucleoprotein with a primary function to synthesize proteins within the cell, using mRNA as a template and aminoacyl tRNAs as its substrates (Schmeing & Ramakrishnan 2009). Protein synthesis occurs at the catalytic PTC of the ribosome, which efficiently carries out its role in the transfer of amino acids to the growing nascent chain. The structure of the PTC is well adapted to deal with the wide range of amino acids that pass through the ribosome and is key to achieving a rate of synthesis 2×10^7 -fold quicker than that of an uncatalyzed bond. By lowering the entropy of activation, excluding water and optimally positioning substrates, the PTC is capable of enhancing the rate of peptide bond formation (Sievers et al. 2004). The translation process can be roughly split into 3 stages, initiation, elongation and termination, a continuous cycle that occurs due to the recycling of many components. Both eukaryotic and prokaryotic ribosomes carry out these 3 stages, with elongation and termination following a similar pattern. The process of initiation however differs between the two cell types and is partly due to the increasing complexity found within eukaryotes. The following sub-sections will describe in greater depth the role of each stage in the translation process, focusing on the simplified prokaryotic translation and describing any differences that exist in eukaryotic translation.

1.1.3.2. Translation initiation

Prokaryotic initiation

The initiation of translation in prokaryotes begins during transcription, as the two are a tightly coupled process within the cell. The formation of the initiation complex on the small ribosome subunit requires 3 intermediary initiation complexes involving formylated aminoacyl initiator tRNA (fMet-tRNA^{fmet}), the translated mRNA and 3 initiation factors (IF1, IF2 and IF3) (Laursen & Sørensen 2005). Firstly IF3 binds to the 30S subunit and promotes dissociation of the two ribosomal subunits, enabling IF1 to bind specifically to the A-site in the 30S subunit and directing the initiator tRNA to the ribosomal P-site (Petrelli et al. 2001). X-ray crystallography shows IF1 actively blocks the A site and allows the fMet-tRNA^{fmet} to associate with the P-site (Clemons Jr. et al. 2001). The actions of IF1 enables IF2, a small GTPase, to bind to the initiator tRNA and aid its binding at the P-site (Lockwood et al. 1971). Finally, IF3 binds to the E site of the small subunit to prevent its association with the 50S subunit (Dallas & Noller 2001). It is also believed to help selection of the initiator tRNA by destabilizing the binding of other tRNAs in the P site of the ribosome (Hartz et al. 1990).

At the point in which the 30S subunit binds to IF3, the ribosome is once again prepared to begin the initiation process (Figure 1.2A). At this point the subunit can then bind directly to a purine rich sequence within the mRNA known as the Shine Dalgarno sequence (5'-ACCUCCUUA-3'), found between 5-8 bases upstream of the AUG start codon (Shine & Dalgarno 1974)(Figure 1.2B). The Shine Dalgarno sequence base pairs to a complementary sequence within the 16S rRNA of the 30S subunit. Formation of the 30S initiation complex (30S-IC) is completed when IF1, IF2 and the fMet-tRNA^{fmet} join the 30S subunit (Figure 1.2C). Through GTP hydrolysis, IF2 binding promotes the release IF3 and the immediate joining of the 50S subunit(Figure 1.2D). As the 50S joins the 30S-IC to generate the 70S initiation complex (70S-IC) IF1, IF2 are released and the initiator tRNA moves into the PTC, readying the prokaryotic ribosome for elongation (Figure 1E).

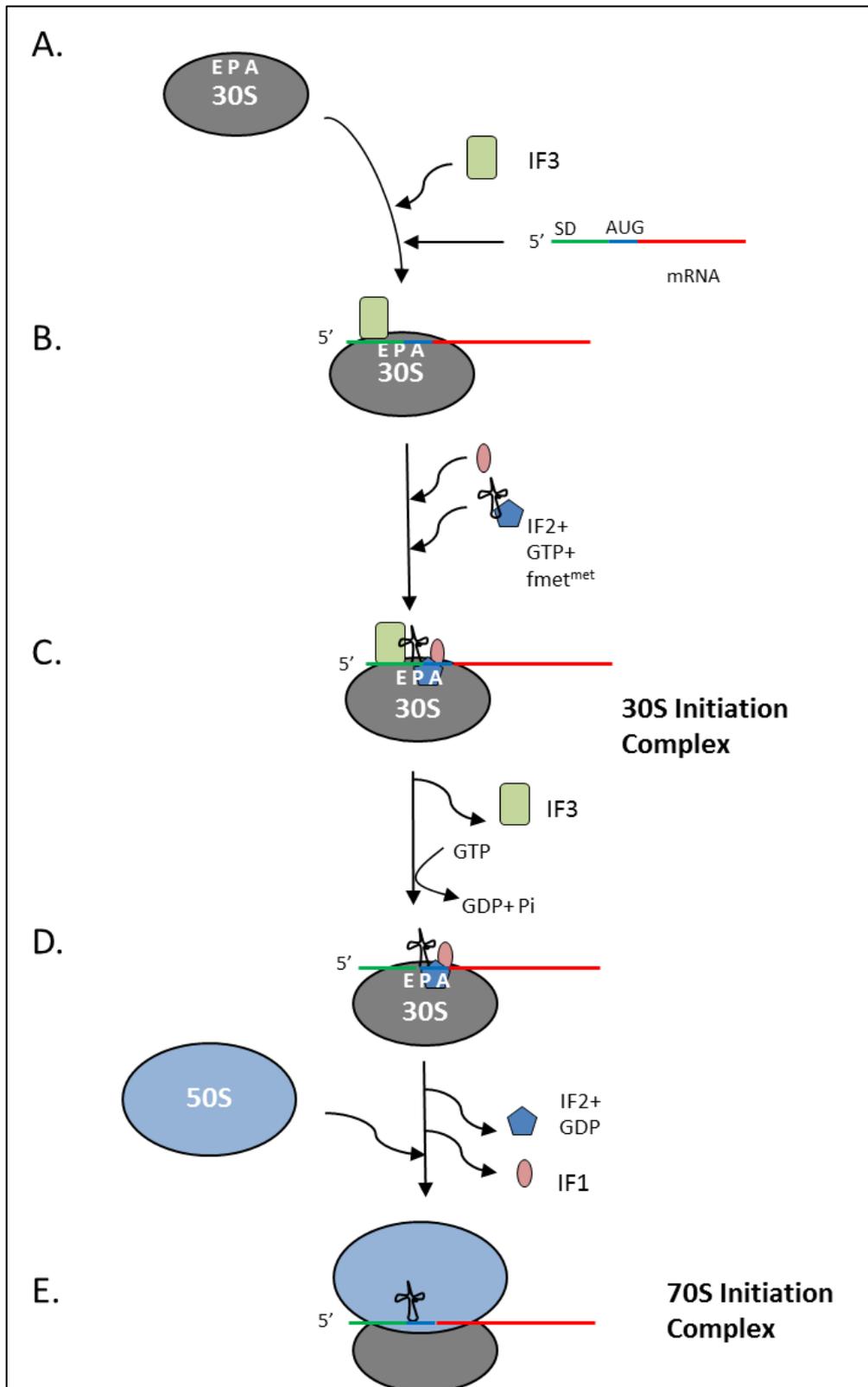


Figure 1.2 Prokaryotic translation initiation. (A) Initiation Factor (IF) 3 binds to the 30S ribosomal subunit. (B) The mRNA Shine-Dalgarno (SD) sequence interacts with and binds to the 30S subunit. (C) This is followed by the binding of initiator fMet-tRNA^{fMet}, IF1 and IF2, completing the 30S Initiation Complex. Hydrolysis of GTP by IF2 results in release of IF3 from the recruitment of the 50S subunit. (D) Subsequent release of the remaining Initiation Factors IF1 and IF2 enables the binding of the 50S subunit. (E) The 70S-IC is complete and the process of protein synthesis can continue into the elongation, termination and recycling stages.

Eukaryotic initiation

Eukaryotic initiation differs substantially to prokaryotic initiation, mainly due to actions of a multitude of initiation factors (Pestova & Hellen 2001). As in prokaryotes, the small 40S subunit acts as the centre point for the assembly of the pre-initiation complex (PIC) (Jackson et al. 2010; Jenner et al. 2012). The eukaryotic initiation process can be subdivided into the formation of 3 intermediary complexes: the first step is the formation of the 43S-PIC, which occurs when the 40S subunit, eIF3, eIF1 and eIF1A enter the initiation process following ribosome recycling. The multisubunit eIF3 prevents the association of the 40S and 60S subunits and aids the binding of eIF1/1A. Cryo-EM and kinetic studies have shown that together, eIF1A and eIF1, stabilize a conformational change on the 40S subunit, opening the mRNA channel and aiding the formation of a 43S-PIC (Pestova et al. 1998; Passmore et al. 2007). Formation of the 43S-PIC is completed when eIF2 and initiator tRNA (Met-tRNA_i; post-transcriptionally modified to distinguish it from elongation Met-tRNA (Sonenberg & Hinnebusch 2009)) bind to the 40S subunit (Figure 1.3A), generating a complex that is capable of scanning the incoming mRNA template.

Following the formation of the 43S-PIC, attachment to the 5' untranslated region (UTR) of the mRNA begins. To prepare for ribosomal attachment, the mRNA template is activated firstly by the attachment of a eukaryotic specific initiation factor complex, eIF4F. The eIF4F complex is made up of eIFs 4E, 4A and 4G and recruits the 43S-PIC to the mRNA template and to begin scanning (Pestova et al. 2001; Kolupaeva et al. 2007) (Figure 1.3B). Specifically, eIF4A and eIF4G have been shown structurally to undergo conformational changes to enable co-folding, which enhances binding to the m⁷G-cap (Gross et al. 2003). The crystal structure of yeast eIF4A provides a model for its role as an ATP-dependant helicase, interacting specifically with eIF4G to unwrap the mRNA and preparing it for an association with eIF4E (Schütz et al. 2008). At the 3' end of the mRNA template, 150 or more adenine nucleotides are found making up a region known as the poly A-tail (Sheets & Wickens 1989). At the poly-A tail, an interaction takes place between the poly(A)-binding protein (PABP) and the 5'm⁷G cap of the mRNA template, which results in a 'closed loop' mRNA conformation due a secondary interaction with eIF4G.

To prepare for ribosomal attachment to the mRNA template, eIF4F, eIF4B and eIF4A unwind the secondary structure of m⁷G cap in an ATPase dependant manner, enabling the

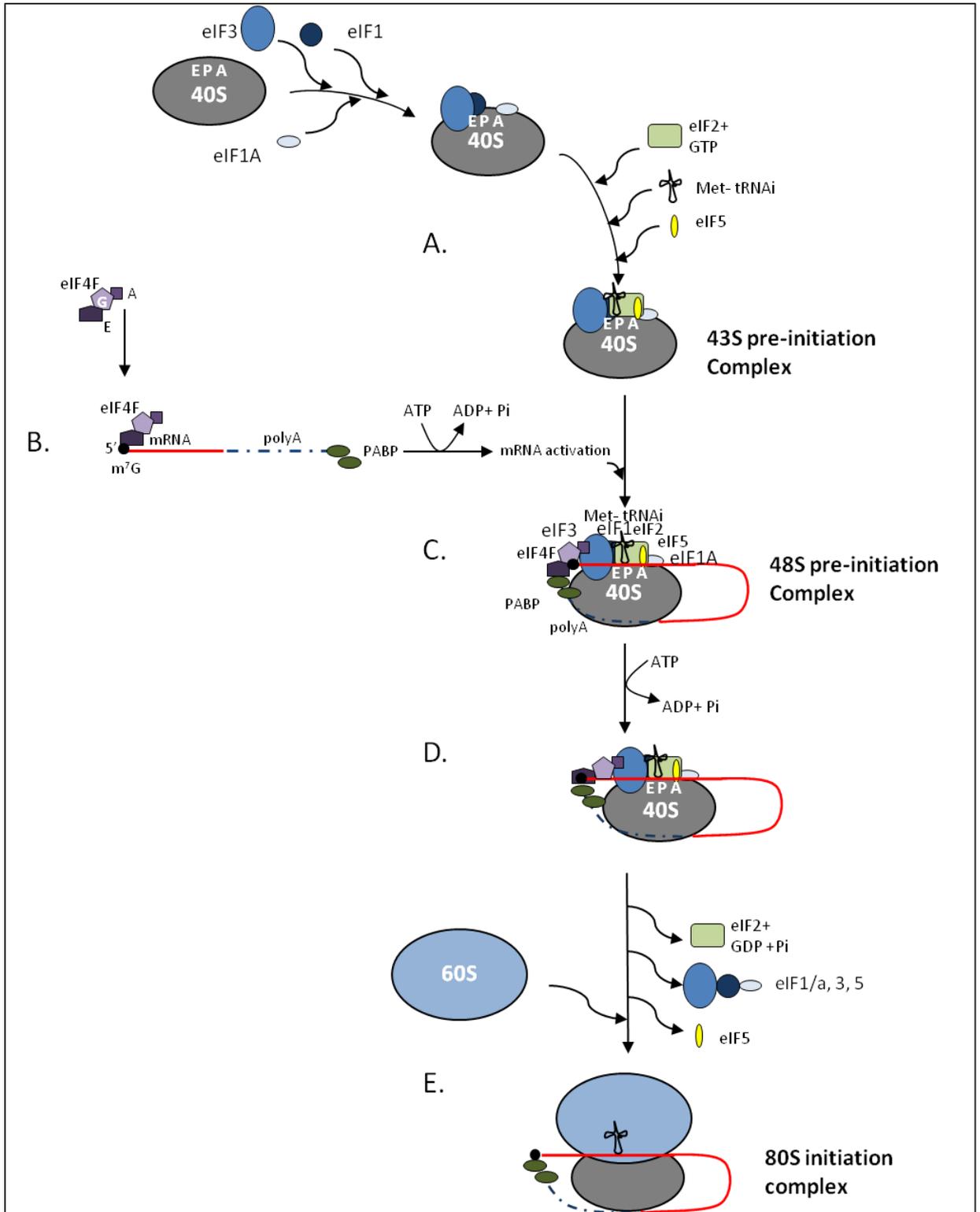


Figure 1.3 Schematic overview of eukaryotic translation initiation. (A) Initiation Factor (eIF) 1/1A and 3 binds to the 40S ribosomal subunit. This is followed by the binding of initiator Met-tRNA_i and eIF2 and eIF5, completing the 40S Initiation Complex. (B) The mRNA is primed with a 5' m⁷G cap and interacts with and binds to initiation factor complex eIF4F. (C) ATP hydrolysis enables the poly (A) binding protein and m⁷G cap to interact, activating the mRNA and completing the 43S complex. (D) Hydrolysis of GTP by eIF2 results in release of initiation factors and enables the binding of the 60S subunit. (E) The 70S-IC is complete and the process of protein synthesis can continue into the elongation, termination and recycling stages.

eukaryotic 43S-PIC to begin scanning from the 5' untranslated region (5'UTR) of the mRNA. Initiation factors eIF1/1A begin the search for a suitable AUG start codon (located within a Kozak Sequence GCC(A/G)CCAAUGG (Kozak 1986)) upon which codon-anticodon interaction can take place (Figure 1.3C). Eukaryotic PIC, in contrast to prokaryotic PIC, does not bind the incoming mRNA template directly at the AUG start codon; instead it is recruited to m⁷G-cap (post-transcriptionally modified GTP (Shatkin 1976)) found at the 5'-end of the mRNA. Interaction between the 40S, eIFs, initiator tRNA and mRNA template completes the formation of the 48S pre-initiation complex (Figure 1.3C).

Finally, following codon recognition, a step that commits the ribosome to the initiation process is required. Initiation factors, eIF1 and eIF1A (functionally similar to the IF3 and IF1 respectively in prokaryotes), promote irreversible GTP hydrolysis and aid the dissociation of incorrectly formed complexes from the mRNA template (Mitchell & Lorsch 2008). Upon forming the 48S complex correctly, subunit joining begins via the actions of eIF5 and eIF5b. The actions of eIF5 encourage the disassociation of initiation factors eIF1, eIF3 and residual eIF2-GDP (Figure 1.3E). The eIF5B has been shown to display GTPase activity and in doing so promoting the joining of the two subunits, whilst releasing itself from the 80S complex (Pestova et al. 2000).

1.1.3.2. Translation elongation

The elongation cycle consists of several steps by which amino acids are sequentially added to the polypeptide chain (Figure 1.4). In preparation for peptide synthesis, each individual amino acid is activated by ATP and paired to the correct tRNA in the cytoplasm, by the actions of a specific aminoacyl-tRNA synthase (Ibba et al. 1997). By the generation of aminoacyl-tRNAs, the initial decoding step of the elongation cycle can begin. Decoding ensures that the ribosome selects the correct aminoacyl-tRNA, as dictated by the mRNA codon, to take up its place in the vacant A-site (Schmeing & Ramakrishnan 2009). The delivery of the correct amino acid to the A-site occurs in a ternary complex made up of aminoacyl-tRNA, elongation factor Tu (EF-Tu) and GTP (Figure 1.4A). Complementary pairing of the codon-anticodon in the A-site is believed to occur at random, with the ribosome sampling multiple codons until there is a match (Figure 1.4B). However, the accuracy in codon selection at the A-site is high and is believed to be due to three universally conserved bases (A1492, A1493, and G530) situated in 16S RNA of the 30S

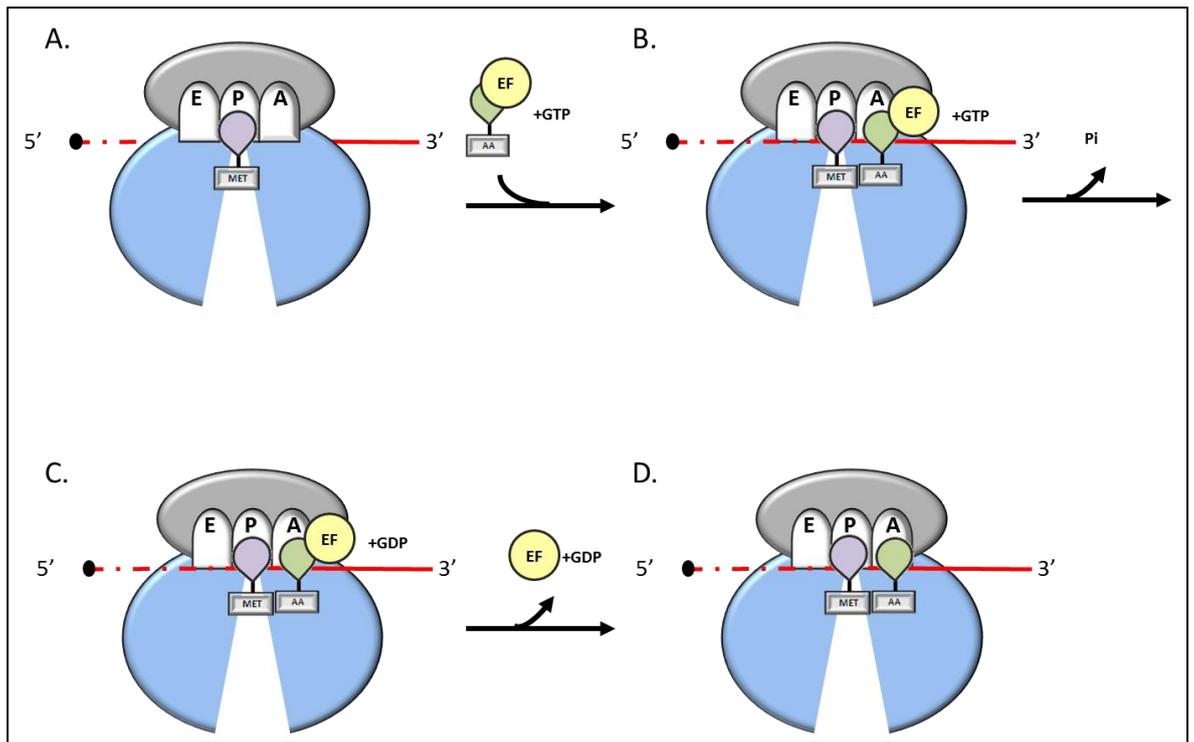


Figure 1.4 Schematic overview of translation elongation. (A). Elongating peptide makes its way through the ribosome exit tunnel bound to tRNA situated in the ribosome P-site. (B) A complex of aminoacyl tRNA, elongation factor Tu (EF) and GTP enter the ribosome A site. (C) Base pairing of the A-site tRNA with the mRNA results in hydrolysis of GTP by EF-Tu. (D) Dissociation of EF-Tu-GDP and accommodation of the aminoacyl-tRNA in the A-site in preparation for peptide bond synthesis.

subunit (Ogle et al. 2001). Once a codon-anticodon match occurs, a conformational change in the 30S subunit takes place, which stabilizes the ternary complex at the A-site. Stabilization of the ternary complex on the 30S subunit causes a distortion in the tRNA, which simultaneously leads to interactions occurring between the aminoacyl-tRNA, EF-Tu and the decoding site (Schmeing 2010). At this point EF-Tu can take up a GTPase activated state in which GTP hydrolysis occurs (Figure 1.4C), resulting in the movement of the aminoacyl-tRNA towards the PTC and the disassociation of EF-Tu-GDP (Figure 1.4D) (Rodnina et al. 1996; Blanchard et al. 2004).

The movement of the aminoacyl end of the tRNA into the PTC signals the start of the central catalytic event of protein synthesis, peptide bond formation. This reaction occurs when the α -amino group of the aminoacyl-tRNA nucleophilically attacks between the carboxylate group of the growing peptide chain and the 3'OH group of peptidyl-tRNA bound at the P-site (Beringer et al. 2005). At this point, the elongation cycle enters translocation, where the nascent peptide chain is moved to the A-site, leaving a deacylated

tRNA at the P-site. During translocation, the ribosome subunits move in a ratchet-like manner, during which time the mRNA is moved exactly 3 bases, whilst the tRNAs at the A and P-sites translocate to the P and E-sites respectively (Frank & Agrawal 2000; Blanchard et al. 2004; Cornish et al. 2008). To accommodate the movement of the nascent chain, the ribosome changes from a classical binding state (A/A and P/P) to a hybrid binding state (A/P and P/E) (Figure 1.5) (Frank and Agrawal 2000). Crystallographic structures of the ribosome trapped in these two states, verified previous research and discovered $\sim 9^\circ$ of relative rotation between the 30S and 50S subunits (Dunkle et al. 2011). Single molecule FRET assays demonstrated in real-time, how the spontaneous ratchet-like movement is capable of switching the ribosome between a ‘non-rotated’ and ‘rotated’ state after peptidyl transfer, until the point at which elongation factor G (EF-G) binds and stabilizes the ‘rotated’ conformation (Figure 1.5) (Cornish et al. 2008). The binding of EF-G catalyses the entire translocation reaction by the hydrolysis of GTP (Rodnina et al. 1997). Structural studies show that EF-G contains a GTPase domain which becomes activated upon binding the ribosome at the sarcin-loop (Connell et al. 2007). GTP hydrolysis generates the necessary conformational changes in the ribosome to translocate the mRNA and tRNA, with translocation being complete when the EF-G-GDP and Pi are released from the ribosome (Savelsbergh et al. 2003). Following translocation, the newly extended peptide is located in the P-site of the ribosome where another elongation cycle can proceed until synthesis is complete.

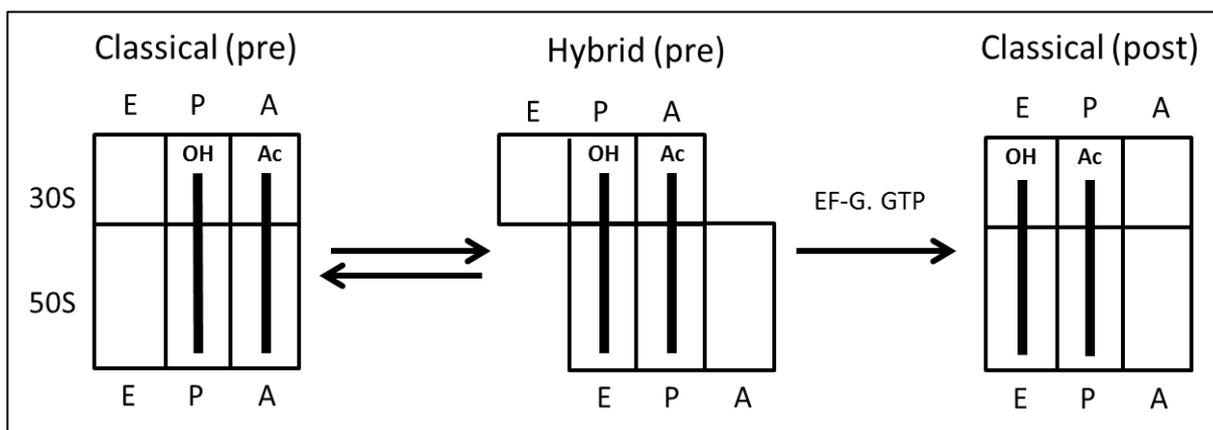


Figure 1.5 Schematic overview of deacylated tRNA and peptidyl tRNA during elongation. Deacylated tRNA is initially in the P-site and peptidyl tRNA is initially in the A-site in the classical binding state pre-translocation. Rotational movement of the 30S and 50S during translocation generates hybrid state and the actions of EF-G.GTP stabilises the ribosome, moving the deacylated tRNA into the E site.

1.1.3.4 Translation termination and recycling of ribosomes

Termination occurs when the synthesis of the nascent peptide is complete and a mRNA stop codon (UAA, UAG or UGA) moves into the A-site, signalling the end of translation. The stop codon is recognised by a class I release factor that is capable of cleaving the polypeptide chain from the P-site by hydrolysing the tRNA bond and releasing the protein from the ribosome. In prokaryotes, two class I release factors exist, RF1 and RF2. High resolution structures of RF1 and RF2 binding to the ribosome shows a conformational change in conserved bases (G530, A1492 and A1493) of 16S RNA at the decoding centre, generating stabilizing interactions between release factors, the stop codon and PTC, which allows the cleavage of the amide bond to take place (Figure 1.6A)(Laurberg et al. 2008; Weixlbaumer et al. 2008). Both factors are capable of recognising the UAA stop codon when bound to the ribosome by the conserved tripeptide motif GQQ. However, the UAG and UGA stop codons can only be recognised by RF1 motif PXT and RF2 motif SPF respectively (Ito et al. 2000). Following peptide release, the class II release factor RF3 binds to RF1/2 and disassociates them from the ribosome (Figure 1.6B) (Freistroffer et al. 1997). Through hydrolysis of GTP, a conformational change within RF3 enables an interaction with the ribosome at the P-site, destabilizing RF1/2 binding, accelerating their disassociation (Figure 1.6C)(Zavialov et al. 2001). Hydrolysis of GTP causes the release factor itself to dissociate from the ribosome, leaving the mRNA and deacylated tRNA bound to the P-site (Figure 1.6D)(Schmeing & Ramakrishnan 2009). At this point the recycling of the ribosomal subunits begins, enabling further rounds of protein synthesis. The process is carried out by the ribosome recycling factor (RRF) and EF-G (Hirashima & Kaji 1973), separating the 30S and 50S subunit through a mechanism that has yet to be fully understood. IF3 is then believed to engage the 30S subunit, releasing the associated tRNA and mRNA and preventing early reassembly (Zavialov et al. 2005).

In eukaryotes, termination is carried out in a similar fashion to prokaryotes, by release factors eRF1 (class I) and eRF3 (class II), neither of which are related to either RF1 or RF2 (Song et al 2000). eRF1 and eRF3 are believed to form a complex with GTP, in which eRF1 is responsible for recognising all 3 variations of the stop codon and eRF3 is required for GTP hydrolysis (Alkalaeva et al. 2006). Recycling of the ribosomal subunits, as in prokaryotes, follows the actions of the release factors. In mitochondria and chloroplasts, the mechanism follows closely that which was described in prokaryotes, with both containing homologs of the prokaryotic recycling factor (Rorbach et al. 2008). However, the method for recycling in cytoplasmic ribosomes differs, relying on the actions of

ABCE1, an ATPase which dissociates the large and small subunit (Pisarev et al. 2010). The release of mRNA and tRNA from the 30S subunit is believed to be carried out by a number of initiation factors, in preparation for a new round of protein synthesis.

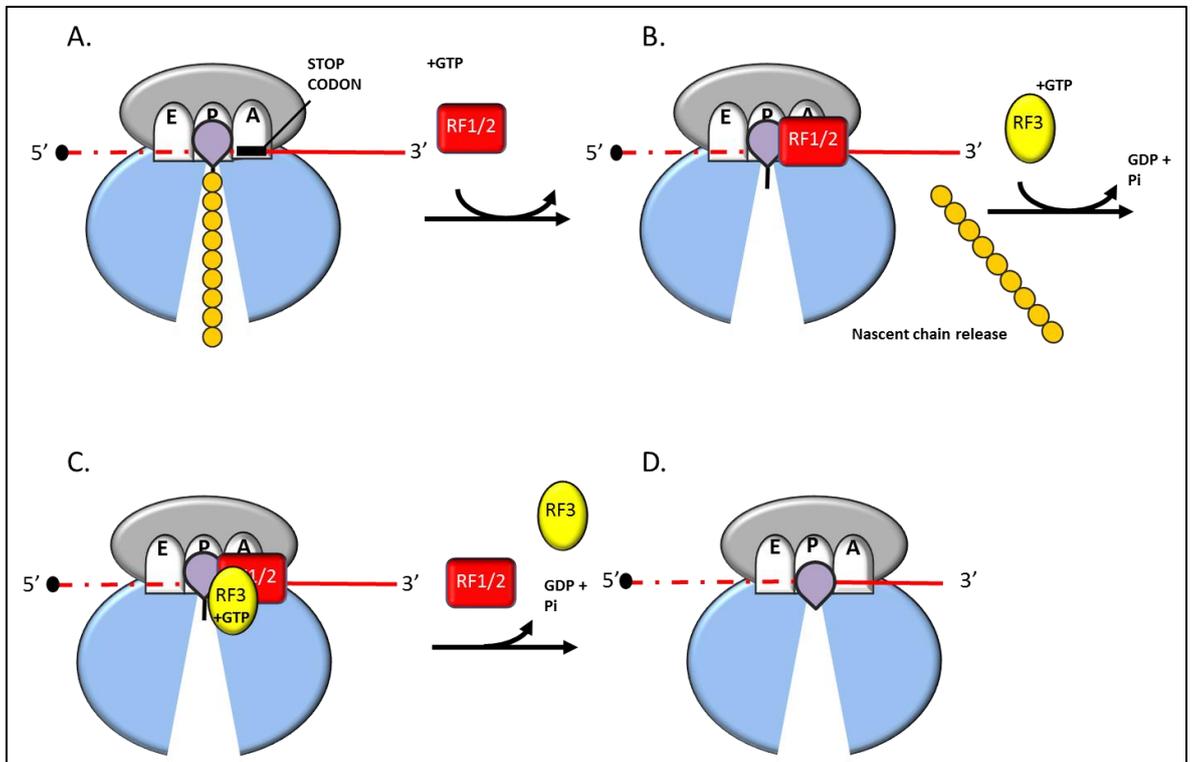


Figure 1.6 Schematic overview of translation termination. (A) Termination begins as stop codon enters the A-site (B) Release Factor binding (RF1 and RF2) and GTP hydrolysis triggers the release of the nascent chain. (C) GTP bound release factor RF3 binds to RF1/2. (D) GTP hydrolysis of RF3 drives the dissociation of release factors and deacyl-tRNA .

1.1.4 The ribosome exit tunnel

1.1.4.1 The structure of the ribosome exit tunnel

As elongation continues and the nascent peptide increases in length, the ever expanding polypeptide follows a route from the PTC, through the large subunit and out into the cytoplasm. The existence of a tunnel within the large subunit of the ribosome was suggested many decades ago through biochemical and imaging studies. Protease digestion assays carried out by Malikin and Rich (1967), showed the first signs of the ribosome offering protection to the translating nascent chain, suggesting an exit tunnel was present. Following this, the first image reconstruction of the large ribosomal subunit was produced, providing a reliable model of an exit tunnel, ~100-120Å length, making its way from the PTC through the large subunit (Yonath et al. 1987). The first studies using cryo-Electron Microscopy (cryo-EM) confirmed the existence of the tunnel (Frank et al. 1995), with more recent high resolution structures containing a nascent peptide removing any doubt (Figure 1.7B) (Bhushan et al. 2010; Seidelt et al. 2009)

Numerous 3D-structures of different ribosomes have provided us with a detailed view of the ribosome exit tunnel (Ban 2000; Ben-Shem et al. 2010; Jenner et al. 2012). The structure and environment of the tunnel in cytoplasmic ribosomes is phylogenetically well conserved. The size of the tunnel from PTC to the point of exit is ~100Å in length and its diameter ranges from 10Å at its narrowest point to 20Å at the vestibule near its exit (Figure 1.7)(Ban et al. 2000; Voss et al. 2006; Bhushan et al. 2010). The tunnel itself is shaped by the conserved r-RNA 23S and multiple r-proteins. In all ribosome structures, r-proteins uL4 and uL22 create a constriction point ~25-30Å from the PTC, the point at which the tunnel is narrowest. At the bottom of the tunnel is the uL23 protein, which lines the exit interface and is believed to play a key role in signalling to cytoplasmic ribosome associated factors (Figure 1.7A)(Woolhead et al. 2004; Robinson et al. 2012). Specific to bacteria is a finger-loop of uL23 which protrudes from the tunnel wall. In eukaryotes this is replaced by the r-protein eL39 (Ben-Shem et al. 2010; Anger et al. 2013). Originally the ribosome tunnel was believed to possess 'teflon-like' qualities allowing the nascent chain to move through relatively unhindered (Ban et al. 2000). However, increasing evidence is now suggesting that the tunnel plays a key role in interacting with the elongating nascent chain as it makes its way out of the ribosome. Although the tunnel is predominantly hydrophilic, patches of hydrophobicity do exist and together with the r-proteins may generate sites where the translating peptide and tunnel can interact. The solvent face of the ribosome exit tunnel also shows a considerable degree of conservation and plays a key role in harbouring

co-translational factors that are required for the capture of an exiting nascent chain. Prokaryotic ribosomes have a total of 6 r-proteins located around the rim of the tunnel, were as eukaryotes, with their greater complexity, have 10.

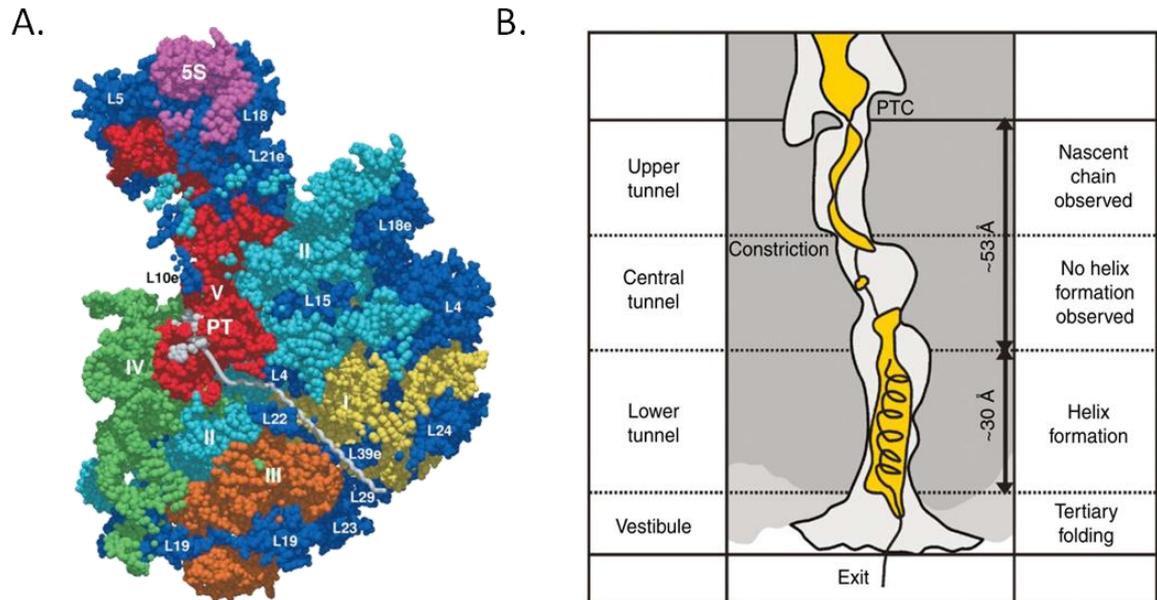


Figure 1.7 The ribosome exit tunnel makes it way through the large ribosome subunit

A. Eukaryotic large ribosomal subunit, indicating the path of the nascent chain (grey line) through the exit tunnel. The ribosomal proteins uL4 and uL22 form the narrowest part of the exit tunnel known as the constriction site. eL39 and uL23 form contacts within the tunnel at the distal of the tunnel close to the exit site. Nomenclature of r-proteins is not up to date with renaming that occurred in 2014 (Ban et al., 2014). Structures reproduced from (Nissen 2000) B. Schematic cross-section of the 80S ribosome tunnel highlighting regions in which secondary and tertiary structure may form. Diagram reproduced from (Bhushan et al., 2010)

1.1.4.2 Folding in the ribosome tunnel

Over the last two decades it has become clear that the ribosome exit tunnel is much more than a passive channel, with many studies highlighting its ability to promote a certain degree of folding within the nascent chain. The dimensions of the tunnel make it difficult to imagine the folding of a complex structure, but increasing evidence has built up to suggest that secondary structures can be formed within the ribosome (Lu & Deutsch 2005a; Bhushan et al. 2010; Woolhead et al. 2004). Indeed a number of biochemical studies, using pegylation and cross-linking assays, and structural studies have shown that the ribosome is capable of housing nascent chains up to 40aa in length, which is more than expected to be required to traverse $\sim 100\text{\AA}$ in an extended form (Lu & Deutsch 2005b; Houben et al. 2005; Bhushan et al. 2010).

The ribosome tunnel, as described in Section 1.1.4.1, does not offer structural uniformity throughout, with certain regions believed to accommodate secondary structure more than others. Indeed a number of studies have shown that the ribosome contains distinct ‘folding zones’ within the tunnel (Figure 1.7B). The first study to directly show compaction taking place within a nascent peptide used the biophysical technique fluorescence resonance energy transfer (FRET) (Woolhead et al. 2004). These experiments carried out on an α -helical transmembrane (TM) domain, highlighted an area just after the PTC in the upper tunnel as a possible ‘folding zone’. Following this, a biochemical assay using an engineered peptide with a high α -helical propensity, indicated that folding could also take place in the lower regions of the ribosome tunnel (Lu & Deutsch 2005b). These experiments were later followed by cryo-EM data to confirm two distinct regions where folding may be possible; one close to the PTC, before the constriction point and a second after the constriction point as you near the exit site (Figure 1.7B)(Bhushan et al. 2010).

In recent years, further interest has explored the possibility of tertiary structure forming within the ribosome. Due to size restrictions within the tunnel, the only viable region that could possibly house such structures would be the vestibule, which is $\sim 20\text{\AA}$ in diameter. Computational studies agreed that this was the most likely region of the tunnel to see tertiary interactions and suggested that transient tertiary structures could form in the final 20\AA of the tunnel (O’Brien et al. 2010). Subsequent biochemical studies have since shown the presence of a helical hairpin structure forming between two TM domains, in the vestibule of the tunnel, whilst the nascent peptide is stably bound to the ribosome (Tu et al 2014). Recently, the complete folding of small protein domains (Nilsson et al. 2015;

Marino et al. 2016), such as the ADR1a zinc finger, within the lower ribosome tunnel has been observed, suggesting the folding capacity of the ribosome could be even greater than imagined.

1.1.4.3 Ribosome tunnel-nascent chain interactions

The concept of the ribosome tunnel having a passive role on the nascent chain during elongation is has been dispelled. Mounting evidence has built up over the last decade and a half, suggesting that interactions between the ribosome and the nascent chain are essential during protein biogenesis. Interactions between the ribosome and nascent chain have been postulated to occur for a number of reasons: firstly, the ribosome acts as a sensory organelle, preparing for when the nascent chain is about to leave the tunnel. The ribosomal proteins uL23 and uL22 proteins are both suggested to have finger-loop domains that protrude into the ribosome tunnel. These loops are believed to sense the elongating nascent chain and prepare the ribosome for the recruitment of targeting factors, chaperones and possible interaction with the membrane translocon (Woolhead et al. 2004; Bornemann et al. 2008; Liao et al. 1997). Another possible reason, as discussed above, is to aid the formation of secondary structure. Some suggest that the formation of secondary structure may be required for an interaction between the nascent chain and targeting factors such as the signal recognition particle (SRP) (Robinson et al. 2012). The presence of secondary structure within the ribosome may also increase the efficiency of insertion into the ER membrane (Tu et al. 2014). Finally, a major tunnel-nascent chain interaction that has been identified in certain nascent peptides, is to induce co-translational stalling from within the tunnel. Proteins such as SecM for example, contain specific arrest motifs that selectively position amino acids in regions of the tunnel where interactions between the nascent chain and uL4 can cause translational arrest (Nakatogawa & Ito 2002). These ‘translation arrest peptides’ have been identified in both eukaryotic and prokaryotic genomes and act to regulate gene expression by sensing cellular conditions (Cruz-Vera et al. 2011).

1.1.4.4 Nascent peptides at the ribosome exit site

As newly-synthesised proteins approach the exit of the ribosome, a number of mechanisms are in place to ensure proper protein homeostasis occurs. At the exit site, a number of protein biogenesis machineries meet and compete to gain access to the translating nascent peptide (Zhang & Shan 2012). As the nascent chain begins to leave the tunnel, there is a high chance that it will come into contact with chaperones associated with the ribosome. Interactions with these chaperones prevent misfolding and aggregation of the nascent chain. In prokaryotes, a high number of polypeptides come in to contact with the ribosome-associated chaperone trigger factor (TF), which has been shown to provide stability, prevent aggregation and even unfold misfolded nascent chains (Kaiser et al. 2006; Hoffmann et al. 2012). In eukaryotes, most nascent chains again encounter chaperones, such as heat shock proteins and the nascent polypeptide-associated complex (NAC), which offer protection from aggregation similar to prokaryotic chaperones (Wang et al. 1995; Yam et al. 2008).

After initially binding ribosome associated chaperones, there are a number of routes that a nascent chain can potentially take. The first major route is for proteins that are cytosolic or destined for cellular organelles. They can be passed onto post-translational chaperones or post-translational targeting factors enabling them to be managed after they have become detached from the ribosome. Post-translational chaperones, such as DnaK/DnaJ in prokaryotes and the TRiC/CCT complex in eukaryotes provide favourable conditions for the peptide to begin folding into their tertiary structure (Calloni et al. 2012). Post-translational targeting factors prepare the nascent chain for translocation, in doing so maintaining their loosely folded state and directing them to the appropriate translocon (Cross & High 2009).

Peptides that are destined for cellular membranes or the secretory pathway require a second route, carried out by co-translational machineries. The most well studied example of the co-translational targeting pathways is through SRP (discussed at length in section 1.2). SRP is found in both eukaryotic and prokaryotic cells and is essential for the targeting of many nascent peptides to their target membranes. The SRP machinery enables an interaction with the appropriate translocon (Sec61 and SecYEG in eukaryotes and prokaryotes respectively) and aids the unloading of the nascent peptide (Akopian et al. 2013).

1.2 Co-translationally-mediated targeting

1.2.1 Overview of SRP-dependant co-translational targeting

The correct localization of a nascent peptide to its cellular destination is an essential step in protein biogenesis. As approximately 30% of the proteome is initially destined for the ER or plasma membrane in prokaryotes, it is widely recognised that many are targeted by the signal recognition particle (SRP) (Figure 1.8). SRP interacts with the nascent chain co-translationally as the peptide begins to emerge from the ribosome exit tunnel. Through a methionine rich M-domain, SRP can recognise a specialised N-terminal sequence within the nascent chain. This sequence provides a signal for the SRP to bind to the ribosome and generate a complex that can be targeted to the membrane. Upon binding to the ribosome-nascent chain complex (RNC), SRP interacts with the SRP receptor (SR) and delivery to the translocon occurs. At the translocon the RNC is transferred from the SRP and insertion of the nascent chain into the membrane can begin.

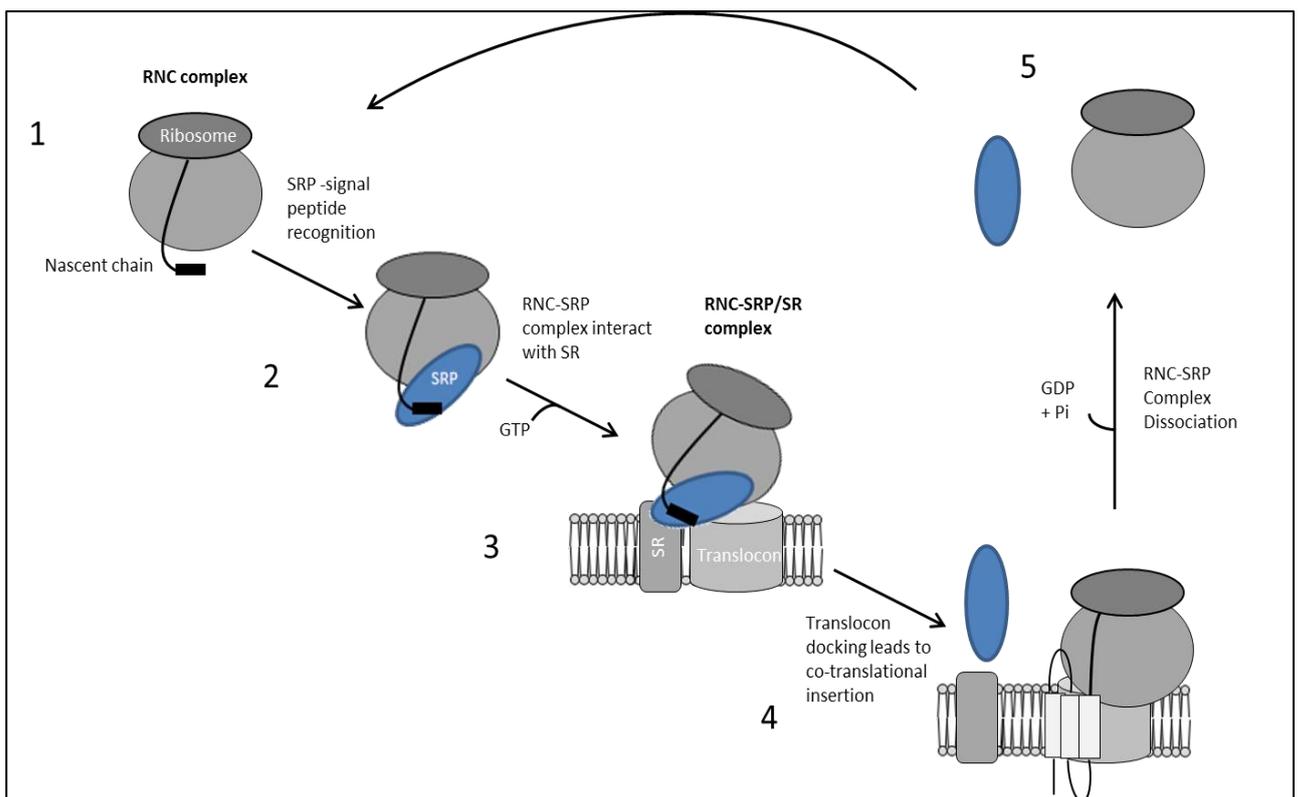


Figure 1.8 Schematic overview of the signal recognition particle (SRP) targeting pathway.

1. SRP binds to the signal sequence (Black square) within the nascent chain as it emerges from the ribosome. 2. SRP and SRP receptor (SR) bind and dock to the translocon when bound to GTP. 3. Docking enables co-translational insertion into the membrane. GTP hydrolysis allows the disassociation of SRP and the SR from the RNC-translocon complex. 5. Insertion of polypeptide triggers release and recycling of ribosome for further rounds of translation.

1.2.2 Components of the SRP system: A structural overview

The SRP system plays a vital role in the proper localization of secretory and membrane proteins to the endoplasmic reticulum (ER) membrane. Throughout all kingdoms SRP has a highly conserved function, yet a number of organisms have evolved producing variability in the SRP structure (Figure 1.9)(Pool 2005). The eukaryotic SRP system (Figure 1.9A), even though it is the most complex, is possibly the best characterised and is centred around two conserved proteins, SRP54 (SRP protein, homologous to the prokaryotic Ffh) and the SRP receptor (SR α subunit, homologous to the prokaryotic FtsY).

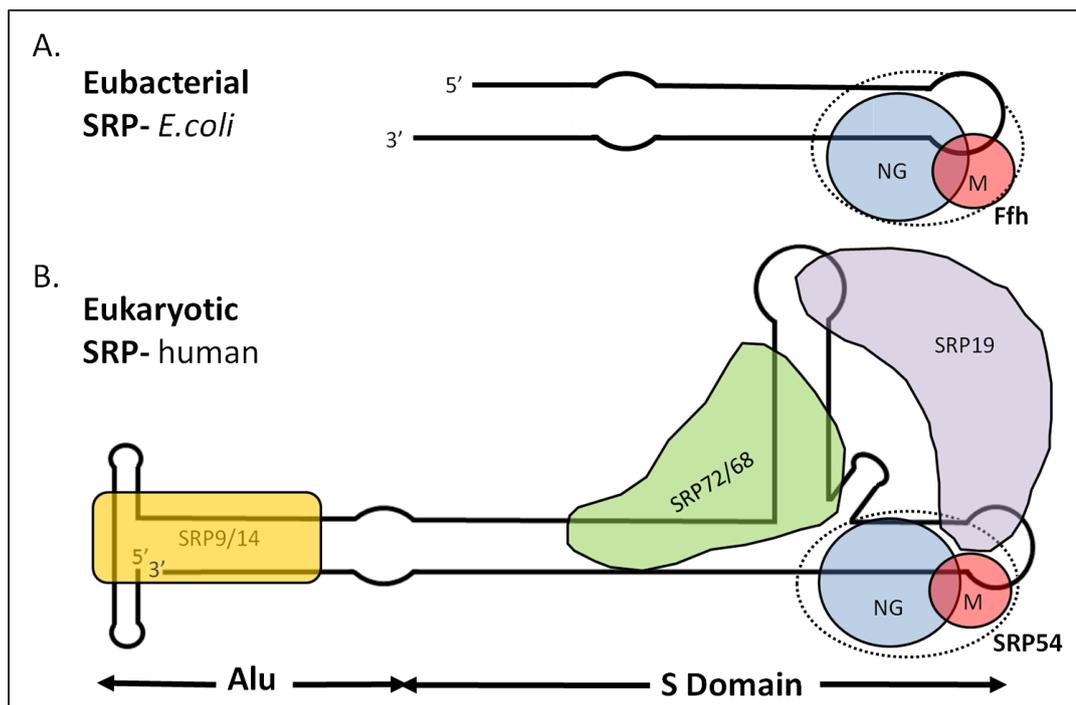


Figure 1.9 Schematic representations of signal recognition particles from human and *E.coli*. Both human and *E.coli* contain M and N/G domains (red and blue circles respectively) which are required for SR binding and signal peptide recognition. The eukaryotic SRP contains higher levels of complexity, with human SRP containing the protein domains SRP19 (purple), SRP72/68 (green) and SRP9/14 (yellow).

Eukaryotic SRP is known to be made up of 6 protein components and a 7S RNA. In higher eukaryotes the six protein domains are named according to their molecular weights; SRP9, 14, 19, 54, 68 and 72 (LUIRINK 2004). Assembly of SRP takes place in two compartments of the cell; the SRP RNA is transcribed at the nucleus and the proteins are synthesised in the cytoplasm before being imported into the nucleus. A pre-SRP complex is assembled by the formation of the SRP proteins (except SRP54) around the RNA component. Export of the pre-SRP complex and binding with SRP54 in the cytoplasm completes the assembly of mature SRP (Politz et al. 2000; Grosshans et al. 2001). Limited nuclease assays were carried out and show that mature SRP can be split in two major

domains; the Alu-domain and the S-domain (Figure 1.9) (Gundelfinger et al. 1983). The Alu-domain has been proven to be an essential requirement for translational arrest (discussed further in section 1.2.4) in co-translationally targeted RNCs (Thomas et al. 1997). Crystal structures of the Alu-domain show it is comprised of a SRP9 and SRP14 heterodimer, as well as the 3' and 5' end of the 7S RNA (Figure 1.9) (Weichenrieder et al. 2001). The S-domain is composed of the remaining for protein components and the central core of the RNA domain. The S-domain is believed to contain the site for SRP receptor (SR) binding and the site required to interact with the signal peptide. The 54 kD subunit found within the S domain is recognised as one of the most conserved domains throughout all kingdoms and it is this that has been shown to play a key role in recognition of signal peptides (Keenan & Freymann 2001).

The core, and only components present in all SRPs are the SRP54 (Ffh in bacteria) and helix 8 of SRP RNA. Helix 8 is essential for facilitating the binding of SRP54, in doing so providing a base for SRP54 assembly (Oubridge et al. 2002). The SRP54 protein is made of a number of components and can be split into: N, G and M domains (Figure 1.9B)(LUIRINK 2004). The N-domain and G-domain are found in the centre and at the N-terminus respectively and together make up the SRP GTPase. Structural studies show that the N-domain is formed of a 4-helix bundle, which serves as a platform for the GTPase of the G-domain. The two domains specifically communicate through 3 motifs to carry out the GTPase activity: the ALLEADV motif in the N-domain, and the GQ and DARGG motifs in the G domain (Grudnik et al. 2009). A strikingly similar NG-domain also exists in the SR, which associates with the NG-domain in SRP54.

The final protein component that makes up SRP54 is a methionine rich M-domain found at the C-terminus (Figure 1.9 and 1.10A). Through structural studies in both prokaryotic and eukaryotic SRP, it has been shown that the methionine rich M-domain, due to the presence a hydrophobic groove, can interact with the SRP RNA and the signal peptide of emerging nascent chains from the ribosome (Figure 1.10B) (Keenan et al. 1998; Halic et al. 2006b). Crystal structures have also revealed the presence of a helix-turn-helix motif in this region, which plays a key role in the interaction with helix 8 in the SRP RNA (Keenan et al. 1998; Batey et al. 2000). The M-domain itself is connected to the NG-domain via a flexible linker, allowing the SRP54 (Ffh) protein to make structural rearrangements that are key in the selection of appropriate nascent chains (Keenan et al. 1998; Rosendal et al. 2004).

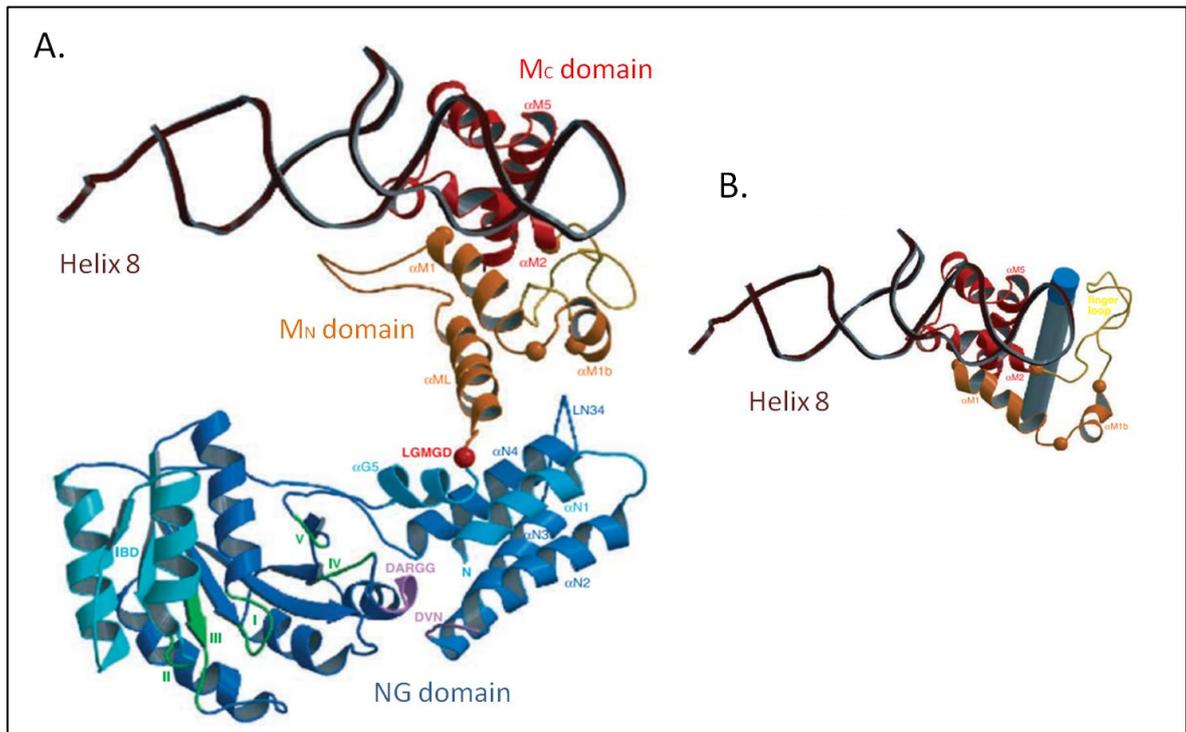


Figure 1.10 Crystal structure of core SRP domain and SRP bound to a nascent chain. (A) The SRP core consisting of SRP54 and helix 8 of the SRP RNA is shown from the crenarchaeon *S. solfataricus*. The SRP54 domains are coloured in blue, the rigid part of the M domain is in red (MC domain), the flexible part (MN domain) is in orange and in yellow is the finger loop. The RNA helix is brown. (B) Model for SRP-nascent chain binding based on the closed structure of *S. solfataricus*. The nascent peptide (grey and blue) is bound between the rigid and flexible regions of the M domain. Figure adapted from (LUIRINK 2004)

The final core SRP structure is the 7S RNA domain, a 114-nucleotide-long molecule that is made up of the most evolutionary conserved part of the SRP RNA, helix 8. High resolution structures of the 7S RNA (and also the 4.5S RNA in bacterial) have proved very informative, with the RNA believed to have the ability to stabilize the M-domain and promote GTPase activity between the SRP and SR complex (Batey et al. 2000).

The SRP receptor is final component that make up the SRP system. As described previously, these proteins belong to a specific family of SRP-type GTPases and confer a similar activity as that described in the NG-domain of SRP54. In eukaryotes, SR is made up of 2 subunits: SR α and SR β . The SR α subunit is essential for the interaction with SRP54 and also carries out a GTP dependant interaction with the SR β subunit. SR β is bound to the cytosolic face of the membrane and is believed to add another level of specificity to eukaryotic targeting (Bacher et al. 1996).

1.2.3 The interaction between SRP and the RNC

The structure of the main SRP system components, as described above, highlights the importance of their interactions in targeting the nascent chain to the membrane co-translationally. However, co-translational targeting does not occur through the actions of the SRP components alone, in fact a multitude of extra SRP interactions are known to be essential such as those with the ribosome, the emerging nascent chain and membrane proteins (Bibi 2011). The ribosome itself has been proven to be a critical part of the co-translational targeting machinery and can often be seen as the initiator of the process.

The ribosome has been shown to weakly bind SRP whilst in a non-translating state (Flanagan et al. 2003), but as translation occurs and a nascent peptide makes its way through the ribosome tunnel, decisions are made as to whether the SRP targeting system is required. Biochemical and cryo-EM data from prokaryotic ribosomes has shown that the uL23 protein, located near the exit site of the large subunit, acts as a docking site for the N-domain in bacterial Ffh (Ullers et al. 2003; Halic et al. 2006a). Cross-linking experiments also showed that uL23 could specifically interact with the emerging nascent chain through a finger-loop domain which protruded into the ribosome tunnel. This interaction with the nascent chain plays a sensory role in detecting the proximity of potential SRP cargo to the exit site of the ribosome. If the nascent chain contains structural or chemical properties that require co-translational targeting, the ribosome responds by altering the conformation of the globular domain on uL23 and in doing so increasing its affinity for SRP (Bornemann et al. 2008). The ribosome has also, through biochemical and structural studies, been shown to interact with SRP through a number of other r-proteins and rRNAs. uL29, also found close to the exit site is believed to make contact with N-domain of Ffh, whereas the SRP M-domain is believed to interact with the rRNAs, uL22 and uL24 (Halic et al. 2006a; Schaffitzel et al. 2006; Yosef et al. 2010). In some cases, these interactions have been postulated to play a role in altering the conformational shape of the ribosome upon SRP binding, possibly impacting the interactions between the tunnel and the nascent chain or effecting future targeting events such as translocon docking (Halic et al. 2006b; Yosef et al. 2010). After an interaction with the ribosome at the exit site, SRP binds the nascent chain through certain signals located within its N-terminus. Generally, peptides that are required to be co-translationally targeted by SRP contain an N-terminal signal sequence that directs the RNC-SRP complex to the appropriate destination. SRP is highly promiscuous in its choice of signal sequence, with a

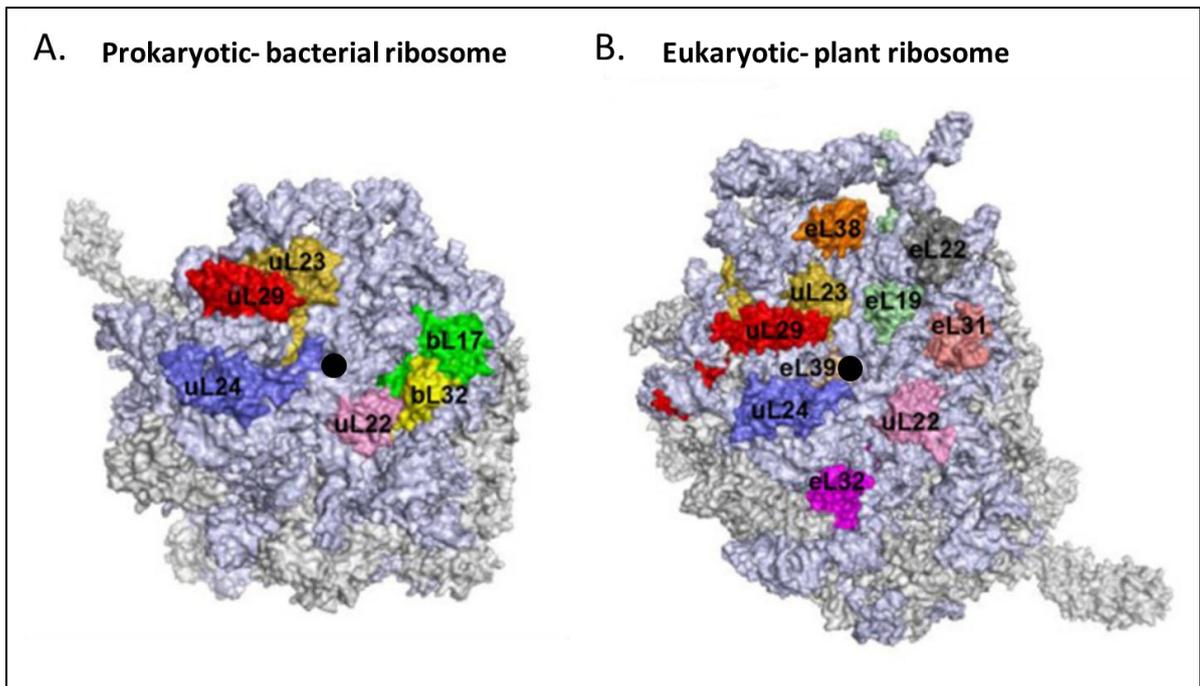


Figure 1.11 Structure of the prokaryotic and eukaryotic ribosome at the ribosome tunnel exit site. The ribosome exit site is marked (black circle), with the position of surrounding r-proteins highlighted by different colours. Conservation of the protein is denoted by the use of the same colour. (A) Cryo-EM structure of the bacterial ribosome. (B) Cryo-EM structure of the plant cytoplasmic ribosome. Adapted from (Breiman et al. 2015b)

peptide containing a hydrophobic core of 8-18aa usually being adequate (von Heijne 1985; Akopian et al. 2013). High levels of diversity have been shown to exist between signal sequences, with length, shape and sequence content all varying. This enables SRP to interact with a variety of domains that can act as a signal sequence, such as short signal peptides or larger, integral TM domains (signal anchors). Cross-linking experiments have shown that the hydrophobic groove in the M-domain of SRP is responsible for the recognition and binding of the signal sequence (Zopf et al. 1990), with structural data suggesting that changes in conformational shape whilst bound to the RNC allows efficient capture of the nascent chain (Wild et al. 2004; Halic et al. 2006b). However, evidence also suggests that both the NG-domain and SRP RNA may also play some part (Bradshaw et al. 2009; Clérico et al. 2009).

1.2.4 RNC delivery to the membrane by SRP

Following the recognition of the signal sequence by SRP, the RNC-SRP complex is targeted to the Sec translocase machinery. In prokaryotes, SRP targets membrane proteins to the plasma membrane but in eukaryotes, it targets both luminal and membrane proteins to the ER. Upon recognising the presence of a signal sequence, SRP is capable of slowing down elongation of the nascent chain to ensure correct protein biogenesis; a process known as translation arrest. Cryo-EM and crystal structures have confirmed the exact interactions taking place between the ribosome and SRP to enable this process to occur (Weichenrieder et al. 2001; Blau et al. 2006; Bousset et al. 2014). Elongation arrest primarily requires the actions of the Alu domain, more specifically the first 48 residues and the SRP9/14 heterodimer (Halic et al. 2004; Lakkaraju et al. 2008). The binding of these domains to the ribosome, at sites known to be required for elongation factor binding, determines elongation arrest is carried out through direct competition between SRP and EFs. *In vivo* experiments using human cells containing SRP14 lacking arrest activity, showed an inefficient delivery of the nascent chain to the ER and cellular defects, thus confirming its role in proper protein biogenesis (Lakkaraju et al. 2008).

Following translation arrest, the RNC-SRP complex is recruited to the membrane via an interaction that takes place between SRP and SR (FtsY in prokaryotes). Upon binding the signal sequence at the ribosome exit site, ribosomal components prime SRP54 for an interaction with GTP (Bacher et al. 1996; Buskiewicz et al. 2009). At this point GTP is not stably bound to either SRP54, nor SR α , therefore this interaction is not enough to drive the GTPase activity and they remain in their free state (or open conformation). The GTP-bound state however, is essential in initiating the interaction between SRP and its receptor, in which both GTPases are said to form an almost identical dimer, capable of activating each other (Powers & Walter 1995).

GTPase activation occurs over a number of steps, details of which have only been provided through studies in prokaryotic systems. However, the conserved nature of the interaction between SRP and its receptor suggest the eukaryotic system may function in a similar manner (Nyathi et al. 2013). Initially an intermediate complex is formed between Ffh and FtsY, due to the binding of GTP to Ffh. Both Ffh and FtsY at this point remain in their open conformations, forming an intermediate that is highly unstable as it interacts primarily through electrostatic attractions between the N-domains (Zhang et al. 2008). Once the early intermediate complex has formed, extensive rearrangements in both

molecules drive the more significant interactions between the NG-domains, generating a stabilized closed conformation (Shan et al. 2004; Zhang et al. 2008). At this point the RNC plays an important role in ‘pausing’ GTP hydrolysis, ensuring it does not occur before the nascent chain is ready to be translocated through the Sec translocon (Shen et al. 2012). Finally, GTPase activity occurs upon the rearrangement of helix-turn-helix motifs within the RNA component of SRP, which must be brought in close proximity to the molecules of GTP to enable complex activation (Akopian et al. 2013). Mutagenesis experiments have highlighted that each individual rearrangement step acts as essential checkpoint in the targeting of membrane proteins. Compromised GTPase activation in either SRP or the receptor prevents the structural rearrangements required in the complex to allow protein translocation (Shan et al. 2007; Grudnik et al. 2009). These structural rearrangements are critical to promote the release of the signal sequence by SRP and in doing so enabling the binding of the RNC to the Sec translocon (Sec61 in eukaryotes and SecYEG in prokaryotes).

The delivery of the RNC-SRP-SR complex to the membrane is required before signal peptide release from the complex can begin. Upon arriving at the membrane, the RNC-SRP-SR forms a quaternary complex by interacting with the SecYEG translocon. The rearrangement of the cytosolic loops of the translocon, C4 and C5, have been shown to be essential in triggering GTPase activation between SRP and SR (Shen et al. 2012; Akopian et al. 2013). Finally, the activation of the GTPase in SRP and SR enables the hydrolysis of GTP, hence releasing the signal peptide and allowing the RNC to bind to the translocon, and SRP-SR dissociation to occur.

1.2.5 Interaction between the ribosome and Sec translocon

The release of the RNC by the SRP-SR complex enables the ribosome to bind to the translocon in preparation for protein insertion. As discussed above, the interaction between the RNC-SRP-SR complex and the translocon is co-ordinated, ensuring the constant shielding of the translating signal peptide or TM domain. Ribosome-translocon binding sites are evolutionarily conserved regions, with the proteins close to the exit site of the ribosome (uL23, uL24 and uL29 (Figure 1.11)), as well as the rRNA region, providing the greatest number of contacts with the translocon (Becker et al. 2009; Frauenfeld et al. 2011). Cryo-EM data available shows the binding sites that are presented by the ribosome remain accessible to the translocon regardless as to whether it is in a translating or non-

translating state (Gogala et al. 2014). During translation the ribosome undergoes major structural changes, which play a role in the binding of the translocon. Contacts occur between the ribosome and the cytosolic loops of Sec61 α , C4 and C5 (Frauenfeld et al. 2011). Upon binding, the translocon is also believed to undergo conformational changes, most likely causing the opening of the channel by the displacement of the plug (Berg et al. 2004). Many models suggest that the Sec61 translocon is capable of switching between an ‘open’ and ‘closed’ conformation during the insertion of the nascent peptide into the ER membrane. However, based on the comparison with a non-translocating, Sec61 channel, recent cryo-EM models suggest that upon docking of the ribosome the translocon channel remains constitutively opened (Pfeffer et al. 2015). A constitutively opened channel would therefore, as suggested in a number of studies, require a number of accessory factors to maintain membrane permeability especially during the lateral integration of membrane proteins.

A multitude of components have been suggested to interact with the RNC-translocon complex as it forms. On the cytosolic side of the ER, a gap of $\sim 15\text{\AA}$ between the ribosome and the Sec translocon is believed to exist (Berg et al. 2004). The role of the gap is poorly understood, with some suggesting that it may allow the nascent chain to form secondary structures that were not possible in the ribosome (Conti et al. 2015). The size of the gap is small enough to prevent the movement of large molecules through the pore, but in the absence of a gating system would require assistance maintain membrane permeability. Proteins such as the translocating-chain associated membrane protein (TRAM) or translocating-associated protein (TRAP) have been shown in structural studies to form at the interface between the ribosome and Sec61 translocon, suggesting a functional role in maintaining an ion-tight channel (Pfeffer et al. 2014).

1.3 Membrane protein insertion into the endoplasmic reticulum

1.3.1 Overview of the endoplasmic reticulum

The endoplasmic reticulum is the largest site for the synthesis of proteins within eukaryotic cells and processes the majority of secreted and integral membrane proteins made by the cell. Targeting signals within nascent chains of newly-synthesised proteins direct them to the ER, via the co-translational targeting pathway, where they interact with specialised translocon machinery and make their way into the ER lumen. Due to the large amount of protein traffic entering the ER, it follows that it is an important site for post-translational modifications and protein folding.

Although the targeting of membrane and secretory proteins to the ER follows the same route, the maturation of membrane proteins at the ER membrane differs substantially. Insertion of the nascent chain into the ER and its exposure to the ER lumen enables the interaction of the peptide with a variety of enzymes that catalyse post-translational modifications of the amino acid sequence. An integral membrane protein will only partially translocate through the ER membrane, with integral TM segments becoming integrated laterally, hence anchoring the protein in the ER membrane. Upon integration of the proteins into the membrane, the ER assists with the formation of the protein's native structure and this in turn may be influenced by post-translational modifications, or vice-versa (Braakman & Bulleid 2011).

Multiple post-translational modifications can occur in the ER lumen such as N-linked glycosylation, disulphide bond formation and the cleavage of signal sequences. In the case of N-linked glycosylation, nascent chains are post-translationally modified by the addition of a carbohydrate group to the amide (NH) group on asparagine residues. The addition of carbohydrate groups is catalysed by a specific complex found in close proximity to the translocon, on the luminal side of the ER. This complex, the oligosaccharyltransferase (OST) complex, seeks out the consensus sequence Asn-X-Ser/Thr within the nascent chain, as the site of N-linked glycosylation (Figure 1.12). N-linked glycosylation aids the proper folding, enhances stability and ensures correct topology in membrane proteins (Burda & Aebi 1999).

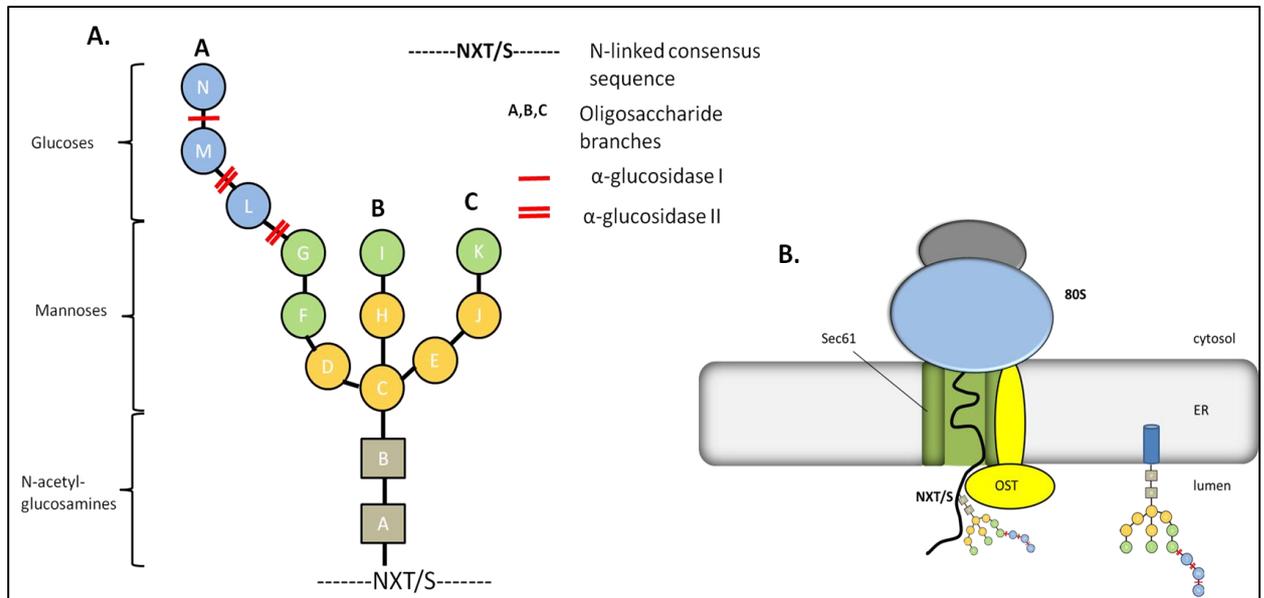


Figure 1.12 Schematic diagram of Asparagine-linked (N-linked) glycosylation. (A) Synthesis of N-linked oligosaccharide occurs in 2 phases; firstly in the cytosol, 2 N-acetylglucosamines are added to a dolichol molecule, before the addition of 5 manose residues. Secondly, after being flipped into the ER lumen by a flippase, 3 addition mannose residues and 3 glucose residues are added. (B) The oligosaccharyltransferase (OST) complex then attaches glycans to a protein at the NXT/S consensus motif.

A second form of post-translational modification that is facilitated by the ER is disulphide bond formation. By providing a slightly oxidising environment, the ER enables a family of protein disulphide isomerases (PDIs) to assist folding by the formation, rearrangement and removal of disulphide bonds. Folding within the ER is also aided by a number of chaperones such as BiP, a luminal Hsp70, which interacts with membrane proteins to ensure the production of a correctly folded structure (discussed further in section 1.3.3).

1.3.2 The Sec61 translocon: structure and function

The protein-conducting channel that enables the translocation of both membrane and secretory proteins is well conserved throughout the kingdoms. The Sec61/SecYEG complexes allow proteins to be translocated across or inserted into the ER of eukaryotes and the plasma membrane of prokaryotes. Early screens in *E. coli* and *Saccharomyces cerevisiae* identified a number of proteins are now known to make up the Sec translocon and to be involved in translocation of proteins at the plasma membrane (Emr et al. 1981; Rothblatt et al. 1989). The heterotrimeric Sec61 complex that makes up the translocon in the eukaryotic ER, is believed to have evolved from the SecYEG complex, associated with

the plasma-membrane in prokaryotes (Rapoport 2007). The Sec61 heterotrimer is made up of two well conserved subunits, the α and γ subunits, with a third and less well conserved subunit, the β subunit, completing the complex (Nyathi et al. 2013).

The major component of the Sec61 complex is the Sec61 α subunit, which forms the central, aqueous pore of the translocon (Görlich & Rapoport 1993; Crowley et al. 1994). The structure of the Sec61 α subunit was initially modelled on the crystal structure of the *Methanococcus jannaschii* SecY translocon, (Berg et al. 2004). The SecY structure provided evidence that the α -subunit could be split into two domains, joined by a linker, with one half containing TM helices 1-5 and the other 6-10. The linker between the two domains of α -subunit is believed to act as a hinge which is capable of opening the translocon to the lipid bilayer, between TM2 and 7, allowing for membrane protein integration (Berg et al. 2004). The two domains of the α -subunit form an hourglass shaped pore with two funnels, one opening onto the cytosolic face and the other onto the ER lumen. At the centre of the pore is a ring of hydrophobic residues (such as 6 isoleucines in the SecY translocon (Berg et al. 2004)) that form a constriction site. The pore is $\sim 6\text{\AA}$ in diameter when in its 'closed' conformation, but would be required to increase to $\sim 20\text{\AA}$ in the presence of an unfolded or α -helical protein. Fluorescent quenching assays used probes inserted into translocating nascent chain to map the diameter of the translocon. These experiments suggested that the diameter of the channel could increase to between 40 and 60\AA (Hamman et al. 1998).

The γ -subunit, a small 8 kDa tail-anchored membrane protein, is the second phylogenetically conserved protein that makes up the Sec61 complex. It contains a diagonal TM domain that traverses the ER bilayer and is in contact with both halves of the α -subunit. The γ -subunit specifically contacts the α -subunit at the back of the channel, near the opening of the two TM domains. Mutagenesis experiments in yeast were carried out to remove the Sec61 γ /Sec61p subunits, with results suggesting a possible role in the lateral movement of the nascent chain into the ER membrane (Wilkinson et al. 2010). In SecY, proteolysis experiments cleaving the hinge region showed the channel remained in working order in the presence of Sec61 γ /SecE, again suggesting a role in opening the channel laterally (Lycklama A Nijeholt et al. 2013).

Finally the third subunit which makes up the Sec61 complex is the 12 kDa β -subunit. Like Sec61 γ , it is a tail anchored membrane protein however its role in translocation remains relatively unclear. The β -subunit has been shown to interact with SR and signal peptidase

in yeast, suggesting Sec61 β may span the membrane (Helmers et al. 2003; Kalies et al. 1998). *In vivo* studies have also shown yeast homologues, Sbh1p and Sbh2p, to be non-essential, but in the absence of the β -subunit translocation rates are dramatically reduced (Finke et al. 1996).

On the luminal side of the Sec61 α pore a small helix, referred to as the plug, occupies the centre of the channel when the translocon is in a resting state. The plug is postulated to influence the opening and closing of the translocon by sealing the pore, generating a selective gate system allowing substrate translocation (Egea et al. 2005). However, experiments have suggested that the existence of a plug is non-essential and residues from neighbouring subunits can replace its function, without the same substrate selectivity (Li et al 2007). Recent cryo-EM models have also suggested that a translocating channel does not alter between the 'open' and 'closed' conformation, instead is constitutively 'open' which suggest the need for accessory factors (Pfeffer et al. 2015). Early studies of the Sec translocon suggested the chaperone BiP, played a role in closing the Sec61 translocon and maintaining membrane permeability. BiP bound to ADP was associated with the 'closed' conformation of the translocon and upon binding ATP; opening of the channel could occur (Alder & Johnson 2004). Whether the translocon gating mechanism is carried out intrinsically by the plug domain or aided by BiP, maintaining permeability of the ER membrane during translocation is essential.

1.3.3 Sec61 accessory factors

The Sec61 complex is the core machinery involved in the insertion of membrane proteins into the ER membrane. Although this is the case, a number of accessory components are required to assist with the translocation and integration of integral membrane proteins. Accessory factors could aid membrane protein biogenesis in various ways; firstly, they could directly impact on translocation, interacting with the translocon to assist with insertion or lateral movement into the ER membrane. Secondly, accessory factors may aid the folding and assist with the orientation of a nascent chain. Finally, others may have an effect on the lipid bilayer surrounding the translocon, into which some proteins will be integrated.

Sec62/63 complex

The Sec62/63 complex aids Sec61-dependant translocation in a 1:1 ratio with the Sec61 translocon. The complex is made up from two conserved membrane proteins Sec62 and Sec63, yet in yeast it also known to contain two extra subunits Sec71 and Sec72 (Deshaies et al. 1991). Mutagenesis experiments in yeast show that Sec71 and 72 are non-essential, although do increase the rate of protein translocation (Fang & Green 1994).

Sec62 is a 32 kDa protein with two-membrane spanning domains and a positively charged N-terminus, whilst the Sec63 protein has a molecular weight of 73kDa and has 3 membrane spanning domains (Deshaies et al. 1991). Both protein components are said to interact directly with the Sec61 complex, as well as interacting with each other, through the N-terminus of the Sec62 and the C-terminus of Sec63. The roles of the two components differ dramatically, with Sec62 believed to be largely involved in post-translational translocation, whereas Sec63 is involved in both co-translational and post-translational translocation (Lang et al. 2012). The Sec62 pathway, through cross-linking studies, has been shown to interact specifically with β -subunit of Sec61 (Jadhav et al. 2015). This interaction is absent in the presence of RNC-SRP complexes, further confirming what was previously believed to be an SRP independent pathway (Jadhav et al. 2015). Recent studies in mammalian systems have also shown Sec63 to play a non-essential role in co-translational translocation, with its function being recovered by the Hsp40 protein, Erj1 (Kroczyńska et al. 2004). Experiments in mammalian cells have also discovered the existence of a ribosome-binding site on the cytosolic face of the TM subunit, suggesting an evolutionary change providing a role in co-translational translocation (Muller et al. 2010).

TRAM (translocating chain associated membrane protein)

TRAM is a 37 kDa integral membrane protein, with 8 TM domains that span the ER membrane. The interactions between TRAM and the components of the Sec61 translocon has been well studied, identifying it as a key partner of many secretory proteins (Görlich et al. 1992) and a required component for the insertion of TM proteins in proteoliposomes (Görlich & Rapoport 1993). Cross-linking studies show TRAM interacts with both signal sequences and TM domains within the nascent chain (Görlich et al. 1992; McCormick et al. 2003), directly aiding their insertion and also enhancing the integration of domains with low levels of hydrophobicity into the bilayer (Voigt et al. 1996; Heinrich et al. 2000). Cross-linking data also provides evidence that TM domains can interact with both Sec61 α

and TRAM simultaneously, hence TRAM may function to assemble multiple TMs at the lateral gate before allowing their integration into the bilayer (Sadlish et al. 2005).

TRAP (translocon associated protein complex)

The TRAP complex is made up of 4 membrane spanning subunits (α , β , γ , δ) that are capable of interacting with the Sec61 complex (Hartmann et al. 1993). Initially assays using the detergent Nonidet P-40, suggested the complex only consisted of the α and β -subunits of TRAP (then known as the signal sequence receptor (SSR)). This was later discovered not to be the case when assays carried out with milder detergents discovered γ and δ -subunits also comprised part of the complex (Gorlich et al 1990; Hartmann et al 1993).

The precise function of TRAP is yet to be determined, but studies have suggested that it may interact with weakly hydrophobic signal sequences and aid their insertion into the ER membrane (Fons et al. 2003). More recent *in vivo* studies using mutagenesis and silencing RNAs have also suggested that it may aid membrane protein topogenesis, with the topologies of proteins found to be mixed in the absence of TRAP (Sommer et al. 2013).

BiP (Binding immunoglobulin protein)

A member of the heat shock protein (Hsp) 70-family of chaperones, the 74 kDa protein BiP, is found in the lumen of the ER. BiP is found to interact with the Sec61 translocon indirectly via the chaperones Sec63 and Erj1 (Blau et al. 2005; Lang et al. 2012). BiP has been described to have a wide variety of functions during ER transport; it has been shown to act as a seal for the Sec61 channel, helping with the gating mechanism and preventing the permeability of the membrane being disrupted (Alder et al. 2005; Schäuble et al. 2012). BiP has also been suggested to be essential for the efficient translocation of proteins into the ER, acting with a ratcheting mechanism on the translocating peptide to aid its movement into or across the ER membrane (Tyedmers et al. 2003a). Experiments using microsomes washed out and lacking luminal chaperones failed to translocate with optimal efficiency, something that could be restored upon the reconstitution of luminal proteins (Nicchitta & Blobel 1993; Tyedmers et al. 2003b)

OST (oligosaccharyl transferase)

Asparagine-linked (N-linked) glycosylation is one of the most common post-translational modifications in eukaryotic cells and is highly conserved in bacteria also. Approximately

70% of proteins that pass through the ER of eukaryotes have the potential for glycoylation, with errors leading to increased risk of congenital disease (Denks et al. 2014). N-linked glycosylation occurs at the OST (as described in section 1.3.1), a heteromeric complex that is embedded into the ER membrane. The catalytic subunit of the OST, first discovered in yeast, is the STT3 subunit and is highly conserved throughout all kingdoms (Burda & Aebi 1999). The OST complex has been shown to interact with a number of components regarded as essential for translocation. Co-translational N-linked glycosylation sees an interaction between the ribosome, Sec61 translocon and the OST complex, with TRAM also having been shown to interact directly with the OST (Shibatani et al. 2005; Harada et al. 2009; Pfeffer et al. 2014)

1.3.4 Determinants of membrane proteins topology

The insertion of membrane proteins into the lipid bilayer of the ER is mediated by the Sec61 translocon. During the initial stages of insertion the topology of many polypeptides is decided; with the composition of their amino acid structure being the major influence. Membrane protein topology is usually dictated by the first ‘stop-transfer’ or ‘signal anchor’ sequence that enters the translocon. These sequences can take up one of two orientations; a type I orientation, with the N-terminus of the peptide existing in the ER lumen or a type II orientation, with the N-terminus facing out into the cytosol (Figure 1.13). These topologies are generally determined by the positive-inside rule; which relies on existence of positively charged amino acids (lysines or arginines) flanking the TM domain. The positively charged flank of the TM domain is thereby positioned at the cytosolic membrane, whilst the more negatively charged region of the TM domain exists nearer the luminal side (Hartmann et al. 1989; von Heijne 1990). This rule was originally observed for bacterial proteins however, a similar charge bias exists for targets of the ER, with the more positively charged flank remaining cytoplasmic (von Heijne 1992). Further studies, using site-directed mutagenesis, have shown that the topology of a protein can be reversed by simply altering the position of the positively charged residues in relation to the TM domain (Harley et al. 1998). An interaction between the positively charged flanking residues of the nascent chain and the positively charged plug region of the Sec61 translocon may generate repellent electrostatic forces between the two. Therefore, if the positively charged residues within the nascent chain were N-terminal, it is likely that a reorientation from a type I to type II topology would occur.

Single-spanning membrane proteins

Single-spanning membrane proteins are seen as the least complex of the membrane proteins, with the positive-inside rule impacting highly on their topology (Figure 1.13). Although this is the case, studies have highlighted additional factors which may impact on the topology of a membrane protein during insertion (Goder & Spiess 2001). One such factor is the length and structure of the N-terminal sequence that becomes exposed to the cytosol following elongation. If folding takes place in this region, it becomes less favourable for the Sec61 translocon to feed the N-terminus through the channel, therefore a type II orientation with a cytosolic N-terminus is generated (Beltzer et al. 1991). Experiments removing the long and structured N-terminus of a single-spanning membrane receptor (truncations reducing the length or mutations to remove the structure within the N-terminus) were able to reverse the topology and produce a type I orientation.

To avoid a type II topology, yet maintain a long and structured N-terminus, some membrane proteins required the services of a cleavable signal sequence. Cleavable signal sequences are found downstream of the first TM anchor and are orientated so their N-terminus faces the cytosol. The topology of the signal sequence commits the first TM domain to a type I topology by forcing its N-terminus through the Sec61 translocon. The constraints placed on the TM domain whilst in the translocon channel prevent it from reverting back to a type II orientation (Shao & Hegde 2011).

The hydrophobicity of a nascent chain also contributes to the topology of a membrane domain. *In vivo* experiments, in which polypeptide chains contained a stretch of between 7-25 leucine residues, proved higher hydrophobicity or a longer TM domain containing more hydrophobic residues could orientate themselves differently than a short, less hydrophobic domain (Wahlberg and Spiess 1998). The possibility that large hydrophobic TM domains could rapidly integrate into the lipid bilayer as they are inserted into the Sec61 channel may also occur. Integration during translocation would fix membrane domains in a type I topology as the N-terminus is fed through into the lumen whilst the TM segment enters the translocon channel (Hessa et al 2005).

Polytopic membrane proteins

Complex multi-spanning membrane proteins (type III membrane proteins) were originally believed to insert into the lipid bilayer in a linear fashion. This 'linear insertion model' proposed that the topology of TM domains within a polytopic membrane protein were defined by the topology of the most N-terminal signal anchor (Blobel 1980). This would

therefore require the remaining TM segments to alternate between signal and stop transfer signal sequences (Wessels & Spiess, 1988). Although this model of insertion is feasible, it would require a highly unusual degree of regularity throughout the individual domains of the membrane protein. Certainly in eukaryotic proteins, which have been shown to be highly diverse, this is not expected to be the case. Indeed experiments carried out in a 12 TM domain glucose transporter (Glut1), in which mutations were carried out to alter the charge difference across TM1, showed inverting the topology of TM1 had no effect on the overall topology of the downstream TM domains. The mutations did generate aberrant TM segments in domains 1 and 2 that failed to insert, however, the remaining TM domains inserted in the correct orientation suggesting downstream domains were not affected by TM1 topology (Sato et al. 1998).

Although polytopic membrane proteins follow the positive-inside rule to some extent, less stringently in eukaryotes than in bacteria, there are other factors which may help guide TM domain topology that cannot be seen in single-spanning membrane proteins. One example is Aquaporin-1, which undergoes a reorientation of 3 TM domains and two connecting peptide loops after translocon and integration (Foster et al. 2000). Post-translational reorientation of TM3 through 180° from a type I topology to a type II topology, leads to a repositioning of TM2 and TM4 from the ER lumen into the lipid bilayer, thus generating a mature form of the protein (Lu et al. 2000). This suggests that the translocation of subsequent domains can alter the topology of domains previously translocated through or inserted into the ER membrane. A second example using the Glutamate transporter from *Pyrococcus horikoshii*, shows that a TM domains can be repositioned during post-translational folding and oligomerization in the ER membrane. This suggests that polytopic membrane domains can alter their topology prior to integration into the lipid bilayer (Kauko et al. 2010). Finally, glycosylation of a membrane protein can be a determining factor in topogenesis. Watson et al (2013) generated truncated peptides of the potassium channel, TASK-1, and displayed evidence that a number of nascent chains could vary their topology. The peptides were capable of reorientation and a mixed population of topologies could be detected. N-linked glycosylation was believed to sterically hinder reorientation and trap the protein in the topology where the glycosylation site was in the lumen of the ER (Watson et al. 2013).

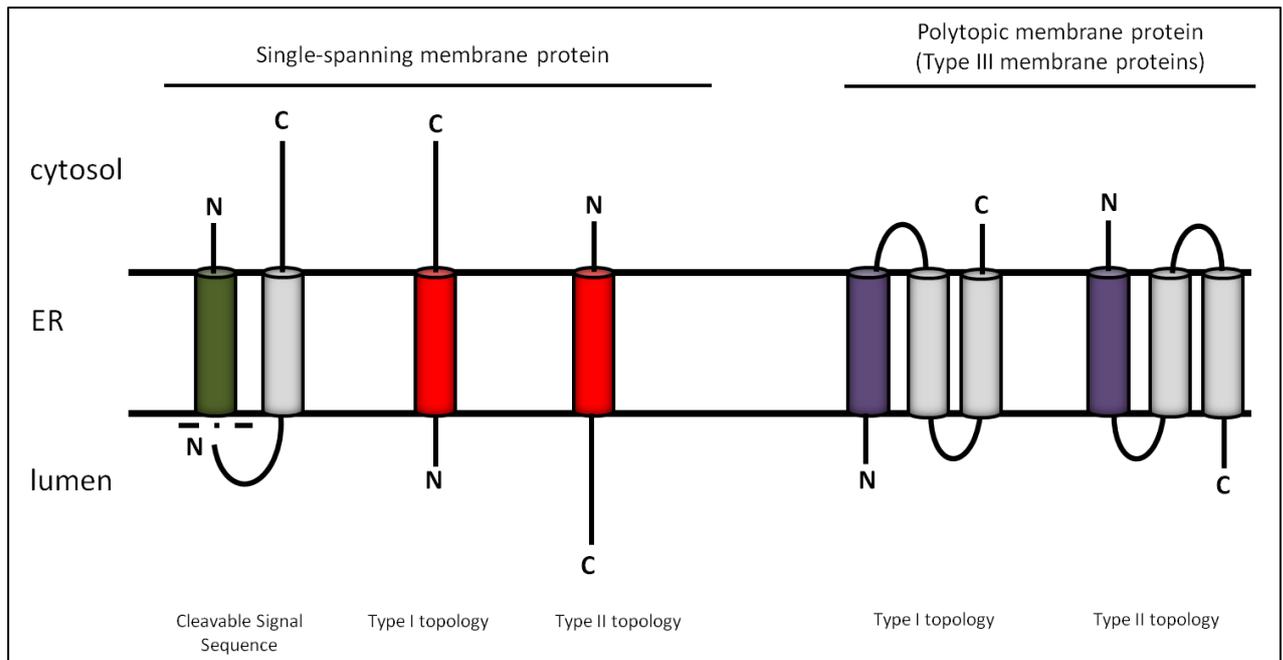


Figure 1.13 Schematic diagram of co-translationally targeted membrane protein topology. Single-spanning and polytopic membrane proteins are capable of taking up different topologies within the ER membrane.

1.3.5 Lateral movement of a TM domain from the Sec61 translocon

The Sec61 translocon not only plays a key role in the movement of proteins through the ER membrane, but also allows the integration of TM domains into the lipid phase of the bilayer. Recently both the cryo-EM and crystal structures of the Sec translocon have emerged whilst engaged with a nascent chain (Gogala et al. 2014; Li et al. 2016). This provided clear evidence of the opening of the lateral gate in which the hydrophobic nascent chain requires to be partitioned into the membrane bilayer. Cross-linking experiments had previously shown large stretches of the nascent chain can interact with the Sec61 channel and the lipid bilayer simultaneously (McCormick et al. 2003), which both these models confirmed by repositioning the nascent chain in a groove outside the lateral gate and in contact with the ER membrane (Gogala et al. 2014; Li et al. 2016). This lateral movement of the nascent chain from the Sec61 channel into the bilayer is highly dependent on the intrinsic hydrophobicity of the sequence (Hessa et al. 2005; Hessa et al. 2007). The Sec translocon is believed to have a hydrophobicity threshold, therefore recognising domains destined to be integrated in to the ER bilayer. Studies have shown that the constriction

point in the Sec61 channel may be the key site for recognising a translocating TM segment (Hessa et al. 2005).

The lateral movement of membrane proteins into the ER is unlikely to be a solo act carried out by the Sec61 complex. In fact a number of photo cross-linking experiments have shown that a variety of other components may be involved. One model suggests that events taking place as early as the ribosome tunnel, at the uL22 protein and may be responsible for co-ordinating the gating mechanism involved in releasing the nascent chain from the translocon channel. The opening and closing of the translocon in this model is driven by BiP, suggesting the chaperone plays a key role in the integration of membrane proteins (Tyedmers et al. 2003b; Alder et al. 2005). Previous experiments have also shown TM domains within the nascent chain to interact with the TRAM/ Sec61 complexes simultaneously during integration. TRAM has been shown to restrict the movement of both single-spanning and polytopic membrane proteins into the phospholipid bilayer (Hung et al. 1996; McCormick et al. 2003). Specifically in polytopic membrane proteins, experiments using single-site cross-linking to Sec61 α and TRAM prove that TM domains can move to different locations within the translocon (Meacock et al. 2002; McCormick et al. 2003). This, coupled with studies suggesting different TM domains can be simultaneously cross-linked to Sec61, supports the theory that the Sec61 channel is dynamic enough to hold more than one TM domain at a time (Sadlish et al. 2005; McCormick et al. 2003).

The ability for Sec61 to house multiple TM domains simultaneously suggests the possibility that more than one model of integration exists in polytopic membrane proteins. The simplest model for the integration of multiple TM segments is the sequential model. This mode of integration was suggested to occur due to the inability to isolate TM segments from the membrane, via alkali and urea extraction, before translation had been terminated (Mothes et al. 1997). Whilst the sequential integration of polytopic membrane proteins is highly possible, subsequent experiments have shown that it is not the only mode of integration. A second method relies on the movement of multiple TM domains from the Sec61 channel. Ismail et al (2007 and 2008) show that TM segments from the G protein-coupled receptor, Opsin, can move out of the translocon in both a singular and multimeric form. TM1 shows a delayed release from the translocon, coinciding with the translocation and exit of TM2 (Ismail 2006), whereas TM5-7 move into the bilayer in a trimer of TM domains (Ismail 2006). Finally, TM domains with unusually low hydrophobicity may require the help of the next domain to become integrated into the membrane. If the

hydrophobicity is low and the TM domain is not recognised as such by the translocon, it may transiently enter the ER lumen and await the following TM domain to integrate it into the bilayer (Öjemalm et al. 2012).

1.4. G protein-coupled receptors (GPCRs)

1.4.1 GPCR biogenesis: What is known?

The GPCR family is one of the largest and most versatile groups of transmembrane proteins, consisting of 4 major classes, with over 800 different genes encoding for GPCRs in humans. GPCRs are comprised of seven transmembrane helical domains, essential for their function of signal transduction from extra- to intracellular. They play a crucial role in regulating numerous physiological functions, as well as being the target for ~40% of all drugs, making them an area of ever increasing interest in research.

Although the structure of a GPCR directly impacts its function, relatively little research has been carried out on the events leading up to and resulting in their biogenesis at the ER membrane. Early studies using opsin as a GPCR model protein were capable of determining, like many integral membrane proteins, GPCRs could be targeted to the ER in a SRP-dependant manner (Audigier et al. 1987). These experiments used SRP-depleted microsomes to confirm that only in the presence of SRP could Opsin be effectively targeted to the ER. This was later confirmed by Laird and High (1997), who went on to use Opsin as a model to study integration into the ER membrane (Laird & High 1997). This indicated that Opsin integrated into the ER membrane through the Sec61 translocon, as cross-linking to components such as Sec61 α , β and associating factor TRAM confirmed. More recent experiments by the same group, provided information on the mode of integration of individual TM domains within Opsin (Meacock et al. 2002; Ismail 2006; Ismail et al. 2008). In particular Ismail et al (2006 and 2008), brought together two pieces of work which produced a complete model for the integration of Opsin. This provided evidence of variability in the behaviour of individual TM domains whilst integrating into the lipid bilayer. Site-specific cross-linking assays produced results indicative of TM domains exiting the Sec61 translocon individually, as pairs and as larger complexes (Ismail et al. 2008). As this is the only GPCR model of integration available it is impossible to say whether it could be applied to the entire family of proteins.

The correct biogenesis at the ER membrane is essential for all IMPs, and it is during the processes of insertion and integration that their mature structure begins to form. Incorrect folding at this point could lead to serious aberrations in the mature structure of a membrane protein. In GPCRs, one such event occurs frequently due to the loss of a specific disulphide bond forming between the extracellular loops 1 and 2, located in the ER lumen during biogenesis. This has been shown to occur in a number of family members of GPCRs, leading to impaired trafficking from the ER and poor receptor function (Peeters et al. 2012)

1.4.2 G protein-coupled receptor 35: A model GPCR

G protein-coupled receptor 35 (GPR35) is an orphan, 7-transmembrane domain G protein-coupled receptor (GPCR) first identified more than 15 years ago and is yet to be well characterized (Figure 1.14)(O'Dowd et al. 1998). The human orthologue of GPR35 was found to be located on chromosome 2, region q37.3 and corresponded to a protein of 309 amino acids in length (O'Dowd et al. 1998; Milligan 2011). The same sequence was also identified to contain an N-terminal extension of 31 amino acids, producing a differentially spliced isoform known as the long GPR35 isoform (GPR35b) (Figure 1.14) (Okumura et al. 2004). GPR35 shares a sequence similarity with a number of different GPCRs and has been found to be related to the purinergic receptor LPA4 (32%), the hydroxycarboxylic acid binding receptors HCA2 and HCA3 (30%), and the cannabinoid and lysophosphatidylinositol-binding GPR55 receptor (30%) (Mackenzie & Milligan 2015).

The short (GPR35a) and long (GPR35b) isoforms of GPR35 have been shown to play a role in a number of different disease physiologies. Originally, GPR35a was discovered to exist in a wide variety of tissue types (e.g. small intestine, colon and stomach), with GPR35b originally discovered in human gastric cancer cells (Milligan 2011). Although the two different isoforms exist, very little is known about how they are differentially regulated or how they differ functionally. One experiment in human gastric cells found that GPR35a was located in the tumorous regions of the cancer at low levels, whereas GPR35b could be found at higher levels, in both the tumorous and non-tumorous regions (Okumura et al. 2004). Both isoforms also show little variation in how they respond to GPR35 agonists. However, it has been suggested that the 31aa extension in GPR35b, although not causing any pharmacological differences, may provide different protein-protein interactions as it extends the N-terminus (MacKenzie et al. 2014). In humans, both forms of the GPR35 gene show polymorphic variations in the amino acid sequence that may

result in disease phenotypes (Figure 1.14). Single nucleotide polymorphisms have been shown to enhance the risk of early onset inflammatory bowel disease and colitis. However, due to the poor characterisation of the receptors and the lack of information regarding their expression, the role of GPR35 is often overlooked in many associated disease phenotypes (Mackenzie & Milligan 2015). Rat and mouse models have provided some evidence that GPR35 overexpression in the dorsal route ganglion may result in elevated pain and inflammatory response.

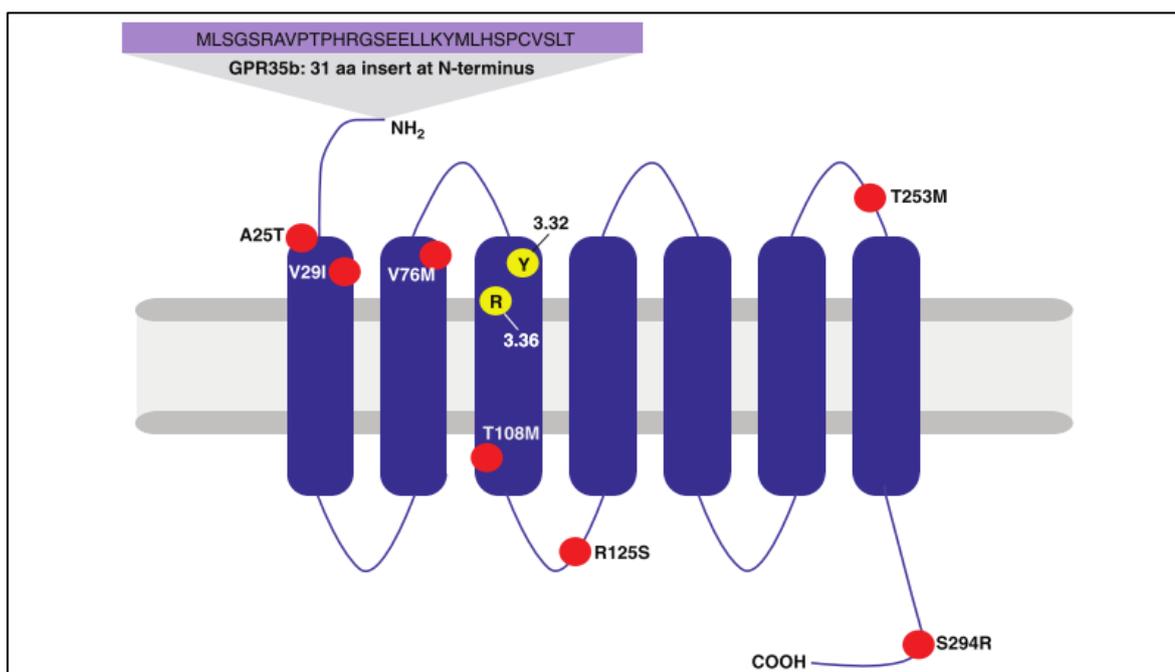


Figure 1.14 Schematic representation of GPR35. GPR35 exists in two isoforms differing by a 31 amino acids (purple sequence bar). Non-synonymous mutations (red circles) and Arg/Tyr residues that play a role in ligand interactions (yellow circles) identified. Diagram adapted from Milligan et al., 2000.

GPR35 is a plasma membrane receptor, found in a wide variety of tissue and cell types. Much of the interest around GPR35 has focused on functional and pharmacological research into the receptor. No work to date has been carried out to investigate the early folding events and the biogenesis of the GPR35 at the ER membrane. Therefore, the work presented in this thesis will attempt to determine the folding profile of TM1 during protein synthesis, the role the first TM domain plays in co-translational targeting and investigate the biogenesis of GPR35 at the ER membrane.

1.5 Project Aims

The aims of this project were to investigate the folding profile of the N-terminus of GPR35 as it is synthesised, co-translationally targeted and integrated into the ER membrane.

Chapter 3 of this thesis investigates the folding profile of the 1st TM domain of GPR35 as it is elongated through the ribosome exit tunnel. Using an *in vitro* translation and an assay known as pegylation, this enables me to directly measure the compaction within the nascent chain of GPR35 intermediates containing the first 1st TM domain. These pegylation assays will be carried out in both prokaryotic and eukaryotic translation systems to investigate any environmental differences that may occur between the ribosomes of the two systems.

Chapter 4 of this study investigates the mechanism of co-translational targeting, identifying the driving force behind the interaction between the principle targeting component SRP and GPR35 nascent chain. Initially, we set out to confirm an interaction with components of the prokaryotic SRP targeting pathway uL23 and Ffh. Using the chemical cross-linker BS³ we were capable of confirming the previously acquired folding profile of GPR35 TM1 as it made its way through the tunnel. Following this, mutations to reduce the hydrophobicity of TM1 were engineered to enable investigations into the importance of hydrophobicity within the nascent chain for driving an SRP-dependant mode of targeting. This was also carried out in the eukaryotic *in vitro* translation system to investigate any differences existing between the prokaryotic Ffh and eukaryotic SRP54.

Finally, Chapter 5 of this project focused on the insertion and integration of GPR35 into the ER membrane. Using N-linked glycosylation as a marker for insertion, I investigated the effect of a lack of hydrophobicity in the 1st TM domain of GPR35 on the translocation of the N-terminus across the ER bilayer. Glycosylation assays were also used as a marker for investigating the folding profile of TM1 as it was inserted into the membrane, following the model described by Whitley et al (1996). Finally the mode of integration of the 1st and 2nd domains of GPR35 was investigated using a site-specific cross-linking assay and the homobifunctional cysteine cross-linker BMH.

2. Materials and Methods:

2.1 General Reagents:

Agilent Technologies UK Ltd., Wokingham, Berkshire, UK

PfuTurbo DNA Polymerase; QuikChange Multi Site-Directed Mutagenesis Kit;
QuikChange Site-Directed Mutagenesis Kit

BioRad Laboratories Ltd., Hemel Hempstead, Hampshire, UK

Agarose

Fisher Scientific UK Ltd., Loughborough, Leicestershire, UK

Ammonium persulphate (APS); Ethylenediaminetetraacetic acid (EDTA); Glucose;
Glycerol; Glycine; Methanol; Potassium hydroxide (KOH); Sucrose; Tris base

Formedium Ltd., Hunstanton, Norfolk, UK

Bacterial Agar; Tryptone; Yeast Extract Powder

Integrated DNA Technologies, Leuven, Belgium

Oligonucleotide primers

Invitrogen Ltd., Paisley, UK

PureLink PCR Purification Kit; SeeBlue Pre-stained Standard

Kodak, Hemel Hempstead, Hertfordshire, UK

X-ray film

Melford Laboratories Ltd., Chelworth, Ipswich, UK

Dithiothreitol (DTT); Isopropyl- β -D-thiogalactopyranoside (IPTG)⁶³

Merck Millipore, Billerica, MA, USA

Amicon Ultra 0.5 ml Centrifugal Filters (10K); Cetyltrimethylammonium bromide; di-Potassium hydrogen phosphate (K₂HPO₄)

New England Bioscience (UK) Ltd., Hitchin, Hertfordshire, UK

100 mM dNTPs (dATP, dCTP, dGTP, dTTP); 10x T4 DNA Ligase Buffer; T4 DNA Ligase; Prestained Protein Marker, Broad Range (7-175 kDa)

PerkinElmer, Cambridge, UK

EXPRESS Protein Labelling Mix (35S Met)

Polysciences Incorporated, Eppelheim, Germany

polyethyleneimine (PEI) (linear, mw 2500)

Promega, Southampton, UK

1 kb DNA Ladder; 100 bp DNA Ladder; 100 mM rNTPs (rATP, rCTP, rGTP, rUTP); Bovine Serum Albumin (BSA); Dithiothreitol (DTT); E. coli S30 Extract System for Linear Templates; *HindIII*; *KpnI*; *MfeI*; Recombinant RNasin Ribonuclease Inhibitor; RNA Polymerase Transcription Buffer; SP6 RNA Polymerase; T7 RNA Polymerase; Wheat Germ Extract; *XbaI*

Qiagen Ltd., Crawley, West Sussex, UK

Nuclease-free Water; QIAquick Gel Extraction Kit; QIAprep Spin Miniprep Kit

Roche Diagnostic Ltd., Burgess Hill, UK

Creatine kinase; Dpn1; tRNA from E. coli MRE 600

Severn Biotech Ltd., Kidderminster, Worcestershire, UK

30 % Acrylamide [Acrylamide: Bis-acrylamide ratio 37.5:1]

Sigma-Aldrich Ltd., Gillingham, Dorset, UK

Adenosine 3',5'-cyclic monophosphate (cAMP); Adenosine 5'-triphosphate (ATP); Ammonium persulphate; Ampicillin; β -mercaptoethanol; Brilliant blue; Bromophenol blue; Calcium chloride (CaCl₂); Creatine phosphate; Dulbecco's modification of Eagle's medium (DMEM); EDTA; Ethidium bromide; Folinic acid; Hydrochloric acid; Isopropanol; L-Amino acids; L-Glutathione oxidised; Lysozyme; Magnesium acetate (MgOAc); Methoxypolyethylene glycol maleimide (PEG-MAL); N,N,N',N'-Tetramethylethylenediamine (TEMED); 1X penicillin/streptomycin mixture ; PIPES; Phosphoenol pyruvate; Poly(ethylene glycol) MW 8000; poly-D-lysine hydrobromide; Potassium Glutamate (KGlu); Puromycin; Pyruvate kinase from Bacillus stearothermophilus; Phenylmethylsulphonyl fluoride (PMSF); Ribonuclease A from bovine pancreas; Tricine

Takara Bio Europe, Saint-Germain-en-Laye, France

Ex Taq DNA Polymerase; Ex Taq DNA Polymerase Buffer

Thermo Fisher Scientific Inc., Waltham, Massachusetts, USA

Slide-A-Lyzer Dialysis Cassette 3,500 MWCO

tRNA probes, College Station, Texas, USA

Non radioactive ϵ ANB-Lys-tRNA^{amb}

VWR International Ltd., Lutterworth, Leicestershire, UK

Acetone; Ammonium acetate (NH₄OAc); Disodium hydrogen orthophosphate (Na₂HPO₄); Ethanol; Glacial acetic acid; 2-(4-(2-Hydroxyethyl) -1-piperazinyl)-ethanesulphonic acid (HEPES); Magnesium chloride (MgCl₂); Potassium acetate (KOAc); Potassium dihydrogen orthophosphate (KH₂PO₄); Potassium chloride (KCl); Sodium acetate (NaOAc); Sodium chloride (NaCl); Sodium dodecyl sulphate (SDS); Trichloroacetic acid (TCA); 30 mm glass cover slips

2.2 General Buffers

Competent cell buffer

60 mM CaCl₂, 15% (v/v) glycerol, 10 mM PIPES (pH 7).

Gel fixing solution

40% (v/v) methanol, 7% (v/v) Glacial acetic acid.

6x DNA loading buffer

30% (v/v) glycerol, 0.25% (w/v) bromophenol blue.

LB media

1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl. LB agar 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 1.5% (w/v) agar.

PEG buffer

20 mM HEPES (pH 7.2), 100 mM NaCl, 5 mM MgCl₂.

RNC (Ribosome Nascent Chain buffer)

20 mM HEPES (pH 7.5), 14 mM MgOAc, 100 mM KOAc.

Run-out premix

0.75 mM HEPES (pH 7.5), 7.5 mM DTT, 21.3 mM MgOAc, 75 μM each amino acid, 6 mM ATP, 20 mg/mL phosphoenol pyruvate, 0.14 mg/mL pyruvate kinase.

2x SDS PAGE Sample

125 mM Tris-HCl (pH 6.8), 20% (v/v) glycerol, 4% (w/v) Buffer SDS, 5% 2-mercaptoethanol, 0.04 % (w/v) bromophenol blue.

SDS-PAGE resolving gel

375 mM Tris-HCl (pH 8.8), 12.5% Acrylamide, 0.1% (w/v) SDS, solution 0.05% (v/v) ammonium persulphate, 0.005% (v/v) TEMED

10x SDS-PAGE running buffer

0.25 M Tris, 1.92 M glycine, 1% (w/v) SDS. buffer SDS-PAGE stacking gel 125 mM Tris-HCl (pH 6.8), 6% Acrylamide, 0.1% (w/v) SDS, solution 0.05% (v/v) ammonium persulphate, 0.005% (v/v) TEMED.

S-30 extract Buffer 1

20 mM HEPES (pH 7.5), 14 mM MgOAc, 100 mM KCl, 6 mM β -mercaptoethanol, 0.5 mM PMSF.

S-30 extract Buffer 2

20 mM HEPES (pH 7.5), 14 mM MgOAc, 100 mM KCl, 1 mM DTT, 0.5 mM PMSF.

S-30 extract Buffer 3

20 mM HEPES (pH 7.5), 14 mM MgOAc, 100 mM KOAc, 1 mM DTT, 0.5 mM PMSF.

SOC media

2 % (w/v) Tryptone, 0.5 % (w/v) Yeast Extract, 0.05 % (w/v) NaCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM Glucose.

50x TAE buffer

2 M Tris, 5.71 % (v/v) glacial acetic acid, 0.05 M EDTA (pH 8.0).

2.5x Translation premix

137.5 mM HEPES (pH 7.5), 520 mM KGlu, 68.75 mM NH₄OAc, 48.25 mM MgOAc, 4.25 mM DTT, 3 mM ATP, 2 mM each rNTPs, 625 μ g/mL creatine kinase, 200 mM

creatine phosphate, 444 $\mu\text{g/mL}$ E. coli tRNA, 2 mM IPTG, 60 mg/mL PEG 8000, 170 μM folinic acid, 1.6 mM cAMP.

10x Tricine gel Anode running buffer

1 M Tris, 1 M tricine, 1% (w/v) SDS.

10x Tricine gel Cathode

2 M Tris-HCl (pH 8.9). running buffer

Tricine separating gel solution

496 mM Tris-HCl (pH 8.45), 8.4% Acrylamide, 6.5% (v/v) solution glycerol, 0.07% (v/v) ammonium persulphate, 0.004% (v/v) TEMED.

Tricine spacer gel

1 M Tris-HCl (pH 8.45), 10% Acrylamide, 0.07% (v/v) ammonium persulphate, 0.004% (v/v) TEMED.

Tricine stacking gel solution

750 mM Tris-HCl (pH 8.45), 4% Acrylamide, 0.07% (v/v) ammonium persulphate, 0.004% (v/v) TEMED.

2.3 *Escherichia coli* strains and plasmid vectors

2.3.1 *Escherichia coli* strains

The following strains of *E. coli* were used at various points throughout this project:

C41 competent cells: F⁻ *ompT hsdSB (rB- mB-) gal dcm* (DE3)

DH5 α competent cells: F⁻ Φ 80*lacZ* Δ M15 Δ (*lacZYA-argF*)

U169 *recA1 endA1 hsdR17(r_k⁻, m_k⁺) phoA supE44thi-1 gyrA96 relA1 λ ⁻*

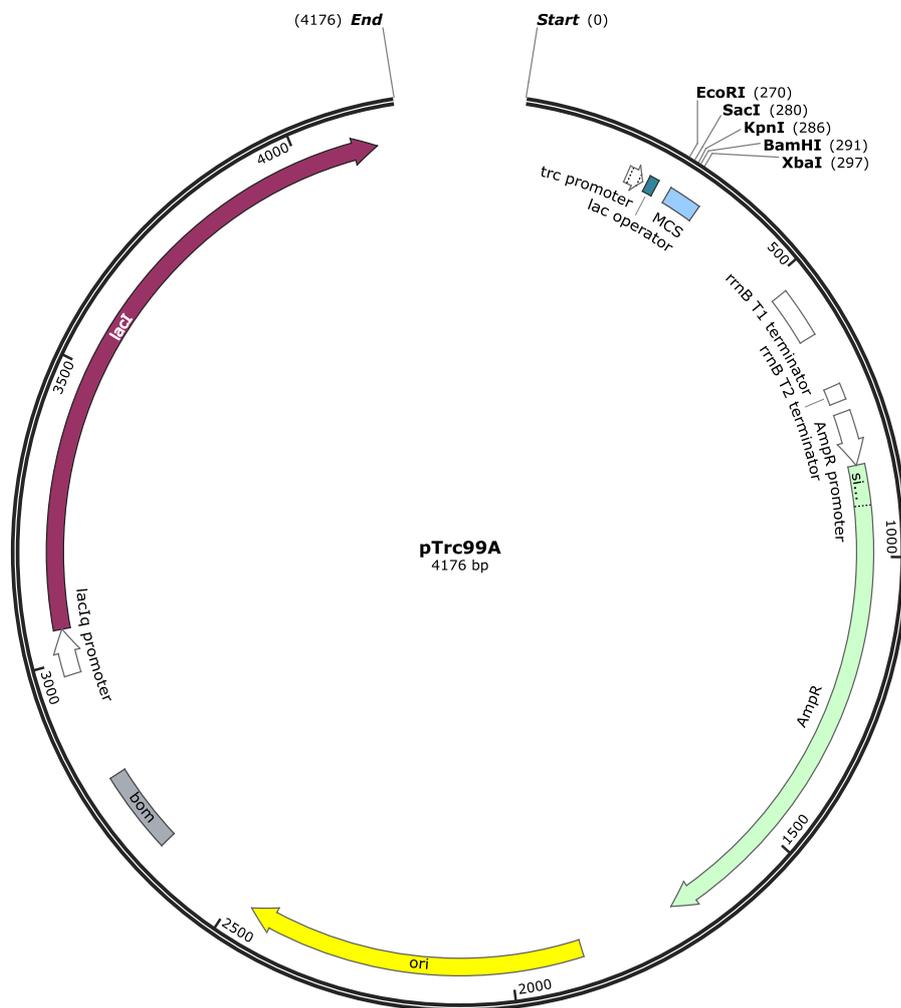
XL1 Blue competent cells: *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacIq Z Δ M15 Tn10 (Tetr)]*.

XL10 Gold competent cells: *TetrD(mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacIqZDM15 Tn10 (Tetr) Amy Camr]*

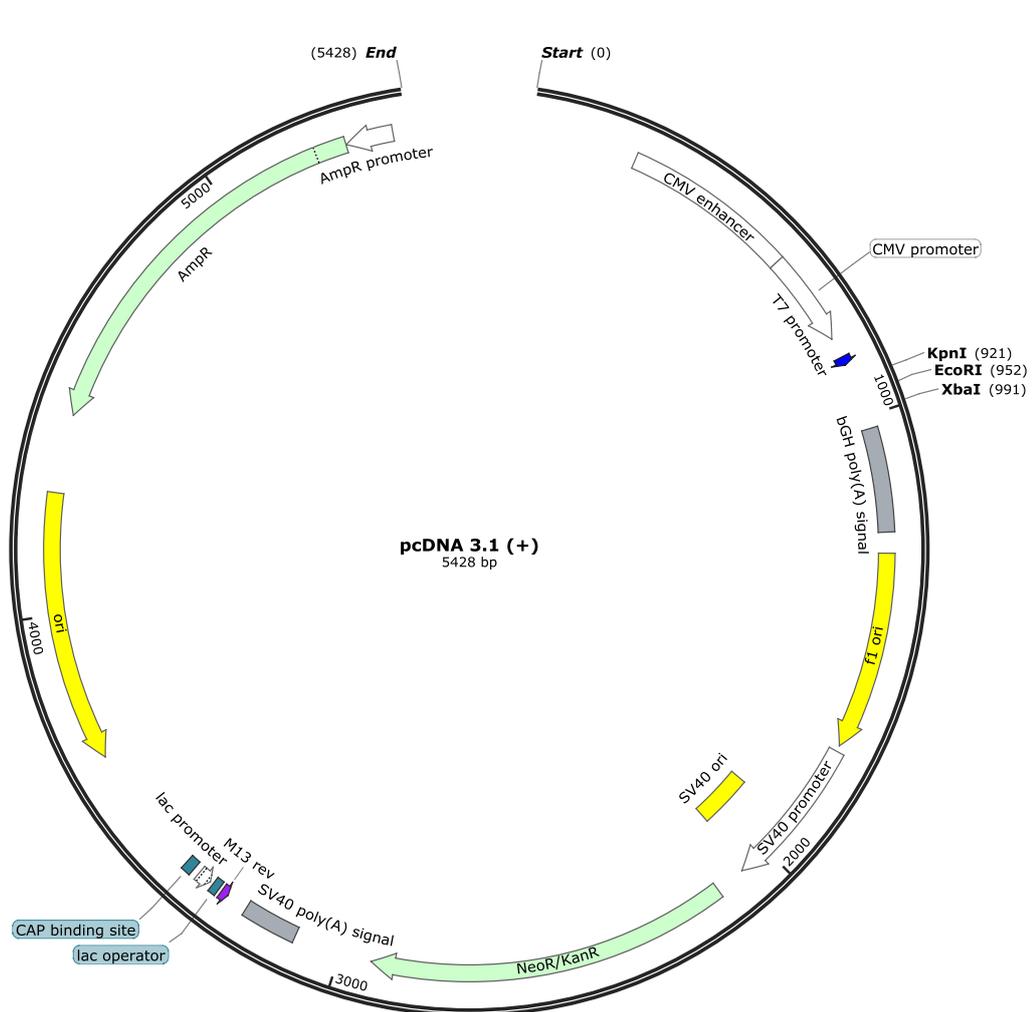
2.3.2 Plasmid Vectors

The following plasmid vectors were used throughout this project:

pTrc99A- a bacterial expression vector with a lacI promoter chemically induced by isopropyl- β -D-thiogalactopyranoside (IPTG). It contains an ampicillin resistance gene which acts as a selectable marker and its multiple cloning site is useful for restriction enzyme cloning. The restriction enzymes shown are those used during cloning experiments in this plasmid vector.



pcDNA 3.1- a high level expression vector for mammalian cell lines. Genes of interest are placed under the control of the T7 promoter. It contains an ampicillin resistance which acts as a selectable marker and its multiple cloning site is useful for restriction enzyme cloning. The restriction enzymes shown are those used during cloning experiments in this plasmid vector.



2.4 General Methods

2.4.1 Preparation of LB plates:

LB-Agar was autoclaved to 126°C for 30 minutes and allowed to cool before ampicillin was added at a final concentration of 100 µg/mL. The LB-Agar was mixed well and poured into plates under sterile conditions.

2.4.2 Preparation of bacterial competent cell stocks:

Glycerol stocks of competent bacterial cells, XL-1 Blue or DH5α, were streaked out onto LB-Agar plates and incubated overnight (~16 hours) at 37°C. A single colony was picked and used to inoculate 5 mL of LB media and incubated overnight at 37°C in a shaking incubator. 500µL of the overnight culture was then used to inoculate 50mL of sterile LB broth which was then incubated with moderate agitation (~220 rpm) until an Optical Density (OD) of 0.375 at 595 nm had been reached. Upon reaching the required OD, cells were chilled for 10 minutes before being transferred to sterile 50 mL falcon tubes and centrifuged at 2,500 rpm for 10 minutes at 4°C to pellet the bacteria. The supernatant was removed and the pellet re-suspended in 10 mL competent cell buffer (see section 2.2 for details) before being centrifuged again at 2,500 rpm for 5 minutes at 4°C. The supernatant is again removed and the pellet re-suspended in 10 mL of competent cell buffer, incubated on ice for 30 minutes and then centrifuged at 2,500 rpm for 5 minutes at 4°C. Finally, the pellet is re-suspended in 2 mL of competent cell buffer and the bacterial cells are aliquoted into sterile eppendorf tubes, frozen with liquid nitrogen and stored at -80°C.

2.4.3 Transformation of competent bacteria:

Transformations of well-established plasmids were carried out in DH5α cells, whilst XL1-Blue and XL10-Gold cells were used for the transformation of site-directed mutagenesis plasmids. A 50 µL aliquot of competent cells were thawed on ice and 1 µL of the relevant DNA plasmid was added. The cells were then incubated on ice for a further 30 minutes before a heat shock at 42°C was carried out for 1 minute and returned to ice for 5 minutes. 500 µL of LB broth was then added and samples were incubated at 37°C for 1 hour in the

shaking incubator. Following this, the entire transformation sample was split and spread over two LB-ampicillin plates and incubated overnight (~16 hours) at 37°C.

2.4.4 Small-scale preparation of plasmid DNA (MINI-PREP):

A single colony is selected from a bacterial transformation and used for the inoculation of 5 mL LB media containing 100 µg/mL of ampicillin which was then incubated overnight at 37°C in the shaking incubator. The overnight culture was centrifuged at 6000 x g for 10 minutes at 4°C to pellet the cells. The plasmid DNA was then purified from the bacterial cells using the Qiagen QIAprep Spin Miniprep Kit. Firstly, the pellet was re-suspended in 250 µL of Buffer P1 and transferred to a sterile eppendorf tube. Cell lysis was then carried on the addition of 250 µL of Buffer P2. The cell lysis was then neutralised after 5 minutes by the addition of 350 µL of Buffer N3. The sample was centrifuged at 15871 x g for 10 minutes and the supernatant transferred to a QIAprep spin column. The column was centrifuged at 15871 x g for 1 minute and the flow-through was discarded. The column was then washed with 750 µL of Buffer PE and centrifuged at 15871 x g for 1 minute with the flow-through again discarded. To remove any residual PE buffer the column was spun again for 2 minutes. The plasmid DNA was then eluted into a sterile eppendorf by adding 50 µL of nuclease-free water to the column and centrifuging for 2 minutes at 15871 x g. Plasmid DNA is then stored at -20°C

2.4.5 Large-scale preparation of plasmid DNA (MAXI-PREP)

After successful small-scale preparation of plasmid DNA, larger quantities required for transfection of cells was required. This was prepared using the Qiagen Maxi Plasmid Kit, using the materials and reagents supplied and following the manufacturer's instructions.

300 µL bacterial cultures in LB broth containing 0.01 mg/mL ampicillin, were grown overnight at 37°C in the shaking incubator. The overnight cultures were centrifuged at 6000 x g for 10 minutes at 4°C to pellet the cells. The pellet was then re-suspended in 10 mL of P1 solution, followed by the addition of P2 solution to lyse the cells and incubated at room temperature for 5 minutes. 7.5 mL of Buffer P3 was then added to neutralise the lysate, before 20 minute incubation on ice, followed by two centrifugation steps at 20000 x g for 30 minutes at 4°C. The supernatant was then loaded onto a Qiagen Maxi column and

under gravity was allowed to flow through with the DNA remaining bound to the column. Three wash steps in QC buffer followed and plasmid DNA elution was carried out using buffer QF. The eluted DNA was then precipitated by adding 10 mL of isopropanol, centrifuged at 15000 x g for 30 minutes at 4°C and washed twice with 70% (v/v) ethanol. The pellet was subsequently air-dried and the DNA was re-suspended in 300 µL of ddH₂O. The DNA concentration was determined by spectrometry and stored at -20°C.

2.4.6 Polymerase Chain Reaction (PCR).

2.4.6.1 DNA amplification by PCR

DNA sequences of interest were amplified by specifically designed forward and reverse oligonucleotide primers (refer to Appendix 2). Forward and reverse primers also included any appropriate restriction sites required for restriction digests. All primers used in this study were synthesised by Integrated DNA technologies. PCR samples were set up on ice in thin-walled PCR tubes, using plasmid DNA as the template. PCR reactions were set up to follow the general protocol below:

10x DNA polymerase buffer	10 µL
dNTP mix (2.5 mM each: dATP ,dCTP ,dGTP, dTTP)	8 µL
Forward primer (100 pmol)	1 µL
Reverse primer (100 pmol)	1 µL
Template DNA (~100 ng/µL)	1 µL
Ex Taq DNA polymerase (5 units/µL)	0.5 µL
ddH ₂ O	Final volume 100 µL

PCR was carried out in the Applied Biosciences 2720 thermal cycler using the following standard conditions:

Temp	Step	Time	# Cycles
94°C	Initial Denaturing	1 minute	1 cycle
94°C	Denaturing	30 seconds	} 30 cycles
56°C	Annealing	30-60 seconds	
72°C	Elongation	1 minute/kb	
72°C	Final Elongation	8 minutes	1 cycle
4°C		Hold	

The standard protocol for PCR can be altered to account for the different melting temperatures of individual primers. Optimal annealing temperature is ~5°C lower than the lowest melting temperature of the set of primers.

Upon completion of PCR, the products were purified by PCR clean-up (see Section 2.4.6) and run on an agarose gel (see Section 2.4.8) to confirm the obtained product was of the correct base pair size.

2.4.6.1 Site-directed mutagenesis by PCR

Single site directed mutagenesis

Single point mutations were generated in constructs using the Quick Change Site Directed Mutagenesis kit (Agilent). Forward and reverse primers were designed to incorporate a desired mutation (refer to Appendix 1). PCR samples were set up on ice, in thin-walled PCR tubes and the following protocol was followed:

10x DNA polymerase buffer	5 μ L
dNTP mix (10 mM each: dATP ,dCTP ,dGTP, dTTP)	1 μ L
Forward primer (100 pmol)	1 μ L
Reverse primer (100 pmol)	1 μ L
Template DNA (~100 ng/ μ L)	1 μ L
Pfu Turbo DNA polymerase (2.5 units/ μ L)	1 μ L
ddH ₂ O	Final volume 50 μ L

PCR was carried out in the Applied Biosciences 2720 thermal cycler using the following standard conditions:

Temp	Step	Time	# Cycles
94°C	Initial Denaturing	30 seconds	1 cycle
94°C	Denaturing	30 seconds	} 18 cycles
58°C	Annealing	1 minute	
72°C	Elongation	1 minute/kb	
72°C	Final Elongation	10 minutes	1 cycle
4°C		Hold	

Upon completion of single site directed mutagenesis, products were treated with 1 μ L of the restriction enzyme Dpn1 for 1 hour at 37°C. The Dpn1 enzyme recognizes methyl groups on the backbone of the parental DNA molecules and removes them via digestion. 1 μ L of the final reaction is used to transform 50 μ L of competent XL-1 blue cells (see Section 2.4.2). Plasmid DNA is isolated through small-scale DNA preparation (see Section 2.4.4) and was subsequently sent for sequencing to confirm whether the mutation had been successfully incorporated.

Multiple site directed mutagenesis

Multiple point mutations were incorporated into the same construct simultaneously by using the QuickChange Multi Site-Directed Mutagenesis kit. On this occasion, only forward primers containing the desired mutations were designed and used, with care taken so that primers did not overlap. PCR samples were set up on ice in thin-walled PCR tubes and the following protocol was followed:

10x DNA polymerase buffer	2.5 μ L
dNTP mix (10 mM each: dATP ,dCTP ,dGTP, dTTP)	1 μ L
Forward primer (100 ng 1-3 primers or 50 ng 4-5 primers)	1 μ L of each
QuickChange Quick Solution	0.75 μ L
Template DNA (~100 ng/ μ L)	1 μ L
QuickChange Multi enzyme blend (2.5 units/ μ L)	1 μ L
ddH ₂ O	Final volume 25 μ L

PCR was carried out in the Applied Biosciences 2720 thermal cycler using the following standard conditions:

Temp	Step	Time	# Cycles
94°C	Initial Denaturing	1 minute	1 cycle
94°C	Denaturing	1 minute	} 30 cycles
55°C	Annealing	1 minute	
65°C	Elongation	2 minutes/kb	
4°C		Hold	

Upon completion of single site directed mutagenesis, products were treated with 1 μ L of the restriction enzyme Dpn1 for 1 hour at 37°C. The Dpn1 enzyme recognizes methyl groups on the backbone of the parental DNA molecules and removes them via digestion. 1 μ L of the final reaction is used to transform 50 μ L of competent XL-1 blue cells (see Section 2.4.2). Plasmid DNA was isolated through small-scale DNA preparation (see

Section 2.4.4) and was subsequently sent for sequencing to confirm whether the mutation had been successfully incorporated.

2.4.7 PCR purification

Products produced by PCR were purified using the PureLink PCR Purification Kit (Invitrogen). 1 volume of the PCR sample (100 μ L) was mixed with 4 volumes of Binding Buffer (B2) in an eppendorf and added to a PureLink Spin Column provided. The sample was then centrifuged at 15871 x g for 1 minute to bind the DNA to the column. The flow-through was discarded and the column was washed with 650 μ L of Wash Buffer and centrifuged at 15871 x g for 1 minute. The flow-through was once again discarded and the column centrifuged at 15871 x g for a further 2 minutes to remove any residual Wash Buffer. The column was then transferred to a fresh eppendorf and the PCR product was eluted from the column by applying 50 μ L of ddH₂O and centrifuging at 15871 x g for 1 minute. The PCR product was analyzed on an agarose gel (see Section 2.4.8) to confirm both the PCR and purification were successful.

2.4.8 DNA agarose gel electrophoresis

Between 1-2% (w/v) agarose powder was dissolved in 1x TAE buffer (see Section 2.2) by heating in the microwave. The molten gel was allowed to cool before adding 0.5 μ g/mL ethidium bromide. The molten gel was then poured into a mould containing a loading comb and allowed to set. The gel was then placed in a gel tank, the comb removed and immersed in 1x TAE buffer. The DNA samples were prepared by adding 6x Loading dye (see section 2.2) before being loaded into individual wells on the gel. Either a 100bp or 1kb DNA ladder (Promega) was run alongside the samples as a marker for size. Electrophoresis was carried out at 80 volts for ~ 40 minutes and the gel was visualised, with images being recorded, using the BioRad Molecular Imager ChemiDoc XRS+ System.

2.4.9 DNA purification from agarose gel

DNA samples were resolved by gel electrophoresis and purified using the QIAquick Gel Extraction kit (Qiagen). DNA samples resolved within a gel were visualised on a UV light box and bands corresponding the correct size were excised using a scalpel and placed into sterile eppendorfs. The gel was weighed and 3 gel volumes (1 g: 300 μ L) of Buffer QG were added and incubated at 50 $^{\circ}$ C for 10 minutes until the gel was completely dissolved. 1 volume of isopropanol was added to the sample, mixed and transferred to a QIAquick spin column and centrifuged at 15871 x g for 1 minute. The flow-through was discarded and the column was washed with 750 μ L of PE Buffer and centrifuged at 15871 x g for 1 minute. The flow-through was once again discarded and the column centrifuged at 15871 x g for a further 2 minutes to remove any residual Wash Buffer. The column was the transferred to a fresh eppendorf and the PCR product was eluted from the column by applying 50 μ L of ddH₂O and centrifuging at 15871 x g for 1 minute. The products were collected and stored at -20 $^{\circ}$ C.

2.4.10 Restriction Endonuclease Digestion

Plasmid DNA or PCR products containing specific restriction sites were digested by pairs of selectively chosen restriction enzymes. Samples for restriction digestions were set up as follows:

Component	Volume
Restriction enzyme buffer	10 μ L
Plasmid DNA/PCR product (100 ng/ μ L)	1.5 μ L
Restriction enzyme 1 (20 units/ μ L)	0.5 μ L
Restriction enzyme 2 (20 units/ μ L)	0.5 μ L
10x BSA (as recommended by manufacturer)	0.15 μ L
ddH ₂ O	Final Volume 15 μ L

Reactions were incubated at a temperature and a time in accordance to the manufacturer's guidelines. Upon completion, enzymes were heat inactivated at 65 °C for 10 minutes. Samples were resolved on an agarose gel and purified by gel extraction (see Section 2.4.9). Samples were stored at -20 °C.

2.4.11 Ligation of DNA

PCR products with 5' and 3' ends cleaved by restriction enzymes underwent a ligation reaction to insert the product into a vector prepared by digestion with the same enzymes. The ligation protocol is as follows:

Components	Volume
10x T4 DNA ligase buffer	2 µL
Insert DNA- PCR product (~ 100 ng/µL)	6 µL
Plasmid vector DNA (~ 100 ng/µL)	1 µL
T4 DNA Ligase (400 units/µL)	1 µL
ddH ₂ O	Final Volume 20 µL

Reactions were incubated at room temperature for 1 hour or at 16 °C in the cold room overnight. 5 µL of the reaction mix was used to transform DH5α competent cells (see Section 2.4.2). Plasmid DNA was isolated through small-scale DNA preparation (see Section 2.4.4) and was subsequently sent for sequencing to confirm whether the ligation was successful.

2.4.12 DNA sequencing

Plasmid DNA required for sequencing was prepared via small-scale DNA preparation and diluted to a concentration of 20 ng/ μ L. Samples were sent to the Sequencing Service at the School of Life Sciences, University of Dundee. Results were analysed and compared to known nucleotide sequences using BLAST and the Swiss Institute of Bioinformatics (SIB) ExPASy translate online software.

2.4.13 *In vitro* transcription

Using Linear DNA from PCR, RNA was transcribed by the following protocol. Reactions were set up in thin-walled PCR tubes and incubated in the Applied Biosciences 2720 thermal cycler for 2 hours at 37°C.

Components	Volume
5x Transcription buffer	10 μ L
dNTP mix (25 mM each: dATP ,dCTP ,dGTP, dTTP)	6 μ L
Linear DNA (~ 100 ng/ μ L)	10 μ L
DTT (1M)	2 μ L
RNasin® Ribonuclease Inhibitor (40 units/ μ L)	1 μ L
T7 RNA polymerase	2 μ L
ddH ₂ O	Final Volume 50 μ L

Synthesised RNA was used directly in the Rabbit Reticulocyte Lysate (RRL) *in vitro* translation system, but required the following clean-up measures when used in the Wheat Germ (WG) *in vitro* translation system. RNA was precipitated by adding 0.1 volumes of 3M NaOAc (pH 5.2) and 3 volumes of 100% ethanol and incubating on ice for 10 minutes. The mix was then centrifuged at 19071 x g at 4°C for 10 minutes. The supernatant was discarded and the remaining pellet was washed with 200 μ L of 70% (v/v) ethanol and

centrifuged at 19071 x g at 4°C for a further 10 minutes. The supernatant was again discarded and the pellet was resuspended in 50 µL of ddH₂O. At this point, 2 µL of the RNA was resolved on an agarose gel (see Section 2.4.7), confirming the RNA was the correct size and adequate quality. Aliquoted samples were stored at -80 °C.

2.5 *In vitro* translation systems

2.5.1 Preparation of *E. coli* S-30 extract

C41 cells from glycerol stocks were streaked onto an LB-agar plate and incubated overnight at 37°C. From the plate, two individual colonies were picked and used to inoculate separate 5 mL LB-broth cultures overnight. The following day, the two 5 mL cultures were used to inoculate 500 mL cultures of SOC media (see Section 2.4.2). The cells were incubated at 37°C in a shaking incubator and grown to mid-log phase ($A_{600}=0.8$). At this point, 1 litre of ice was added to the cultures and then centrifuged at 6732 x g for 10 minutes at 4°C. The supernatant was discarded and the pellet re-suspended in 100 mL of Buffer 1 and re-centrifuged at 6732 x g for 10 minutes at 4°C. The supernatant was again discarded and the pellet re-suspended in 100 mL of Buffer 1 and then centrifuged again at 6732 x g for 10 minutes at 4°C. After removal of the supernatant, the wet cell mass of the pellet was weighed and re-suspended in Buffer 2 to a concentration of 0.5 g/mL. Lysosyme (100 mg/mL) was added to a final concentration of 1 mg/mL before the cell suspension was passed twice through a French Press at 8,000 psi. The extract was then centrifuged at 29994 x g for 30 minutes at 4°C. The supernatant was retained and transferred to a fresh tube and centrifuged again at 29994 x g for 30 minutes at 4°C. The supernatant was then incubated with 0.15 volumes of Run-out premix at 26°C for 70 minutes, before the extract was dialysed 3 times using a Slide-A-Lyzer Dialysis Cassette 3500 MWCO, in Buffer 3 for a minimum of 1 hour each. Following dialysis, the extract was centrifuged at 19071 x g at 4°C for 10 minutes before being aliquoted and then snap frozen to be stored at -80°C.

2.5.2 *In vitro* transcription/translation in *E. coli* S-30 extract translation system

Translation reactions in the *E. coli* transcription/translation system were set up as follows in a standard 25 μL volume (volumes can be adjusted accordingly).

Components	Volume
Translation premix	10 μL
1mM each L-amino acid (minus methionine)	2.5 μL
Linear DNA (~ 100 ng/ μL)	2.5 μL
S-30 extract	7.5 μL
[³⁵ S] methionine	10 μCi
5 $\mu\text{g}/\mu\text{L}$ of anti-ssrA oligonucleotide (5'-TTAAGCTGCTAAAGCGTAGTTTTTCGTCGTTTGCGACTA-3')	1 μL
	Final Volume 25 μL

Translation reactions were incubated at 37°C for 30 minutes and were terminated by incubating on ice for 5 minutes. SDS-PAGE analysis of results was carried out as described in Section 2.6.1

2.5.3 *In vitro* translation in Wheat Germ extract translation system

Translation reactions in the Wheat Germ Extract translation system were set up as follows in a standard 50 μL volume (volumes can be adjusted accordingly).

Components	Volume
Wheat Germ Extract	25 μL
1mM each L-amino acid (minus methionine)	4 μL
RNA substrate in ddH ₂ O (~500 ng/ μL)	5 μL
Potassium Acetate (1M)	6.5 μL
[³⁵ S] methionine	2.5 μL
RNasin® Ribonuclease Inhibitor (40 u/ μL)	1 μL
ddH ₂ O	Final Volume 50 μL

Translation reactions were incubated at 26°C for 30 minutes and were terminated by incubating on ice for 5 minutes. SDS-PAGE analysis of results was carried out as described in Section 2.6.1

2.5.4 *In vitro* translation in Rabbit Reticulocyte Lysate translation system

Translation reactions in the Flexi® Rabbit Reticulocyte Lysate translation system were set up as follows in a standard 25 μL volume (volumes can be adjusted accordingly).

Components	Volume
Rabbit Reticulocyte Lysate	16.5 μL
1mM each L-amino acid (minus methionine)	1 μL
RNA substrate in ddH ₂ O (~100 ng/ μL)	1 μL
Potassium Chloride (2.5 M)	0.4 μL
DTT (1M)	0.2 μL
[³⁵ S] methionine	1 μL
RNasin® Ribonuclease Inhibitor (40 units/ μL)	1 μL
ddH ₂ O	Final Volume 25 μL

Translation reactions were incubated at 30°C for 30 minutes and were terminated by incubating on ice for 5 minutes. SDS-PAGE analysis of results was carried out as described in Section 2.6.1.

2.5.5 TCA precipitation of translated peptides

Translation reactions were carried out as described in Section 2.5.2.4 and following the 5 minute incubation on ice, 1% (w/v) final volume of TCA was added to the reaction mix. The samples were incubated on ice for 15 minutes and then centrifuged at 19071 x g at 4°C for 10 minutes. The supernatant was discarded and the pellet was twice washed with 1 mL of 100% cold acetone and incubated on ice for 10 minutes, before being centrifuged at 19071 x g at 4°C for 10 minutes. The supernatant was removed and the precipitated pellets were dried in the speedy vacuum for 15 minutes. Pellets were re-suspended in 2x Sample Buffer and SDS-PAGE analysis of results was carried out as described in Section 2.6.1.

2.5.6 Pegylation assay to assess compaction of nascent chain in the ribosome exit tunnel.

Translation reactions were carried out as described in Section 2.5.2 and 2.5.3 at a volume of 50 µL and following the 5 minutes incubation on ice, the translation product was overlaid onto 100 µL sucrose cushion (0.5M sucrose in RNC buffer). Ultracentrifugation of the samples in a Beckman TLA-100 rotor took place at 436000 x g for 10 minutes at 4°C to isolate ribosome bound nascent chains (RNCs). The supernatant was discarded and pellets were re-suspended on ice in 60 µL PEG Buffer. The samples were split in half (30 µL) and added to either 30 µL PEG Buffer containing 2mM PEG-MAL or a control containing 30 µL of PEG Buffer only. The reactions were incubated on ice for 2 hours. To terminate the reaction, 100 mM DTT was added and incubated at room temperature for 10 minutes. The samples were then CTABr precipitated by adding 10 volumes of 0.5M NaOAc (pH 4.7) and 2% (w/v) CTABr (600 µL of each) and incubated on ice for 15 minutes. They were then centrifuged at 15871 x g for 15 minutes at room temperature before discarding the supernatant. The remaining pellets were twice washed with 1 mL of 100% cold acetone and incubated on ice for 10 minutes, before being centrifuged at 19071 x g at 4°C for 10 minutes. The supernatant was removed and the precipitated pellets were dried in the speedy vacuum for 15 minutes. The dried pellets were re-suspended in 15 µL 1mg/mL RNaseA in ddH₂O and incubated at room temperature for 10 minutes, before the addition of 2x Sample Buffer and analysis of results by SDS-PAGE as described in Section 2.6.1.

2.5.7 Chemical cross-linking of ribosome-bound nascent chains using BS³

Translation reactions (100 μ L) were carried out as previously described in Section 2.5.2 and 2.5.4 to generate nascent peptides of desired lengths. After the 5 minute incubation on ice, a 7 μ L portion of the reaction was overlaid onto a 50 μ L sucrose cushion (0.5M sucrose in RNC buffer) (tube A) while the remainder is overlaid onto a 100 μ L sucrose cushion (tube B). Ultracentrifugation of the samples in a Beckman TLA-100 rotor took place at 436000 x g for 10 minutes at 4°C to isolate ribosome bound nascent chains (RNCs). Pellet A was re-suspended in 8 μ L of RNC buffer, 100 μ g/mL RNase A and 5 mM EDTA at incubated at 26°C for 10 minutes. The sample was the heated in 2 X Sample Buffer at 95°C for 5 minutes before the analysis of results by SDS-PAGE as described in Section 2.6.1.

Pellet B was re-suspended in 88 μ L of BS³ buffer (RNC buffer and 1 mM BS³) and incubated on ice for 2 hours. The reactions are quenched by the addition of 5 mL of 1M Tris (pH 8.0) and incubated at room temperature for 15 minutes. 7 μ L of the sample was added to 100 μ g/ mL RNase A and 5 mM EDTA at incubated at 26°C for 10 minutes and heated in 2 X sample buffer at 95°C for 5 minutes before the analysis of results by SDS-PAGE as described in Section 2.6.1. The remainder of the samples were incubated with 1% (w/v) final volume of TCA and placed on ice for 15 minutes before centrifugation at 19071 x g at 4°C for 10 minutes. The supernatant was discarded and the pellets were used for immunoprecipitation experiments.

2.5.8 Chemical cross-linking of ribosome-bound nascent chains to Sec61 α using BMH

Translation reactions (100 μ L) were carried out as previously described in Section 2.5.2 and 2.5.4 to generate nascent peptides of desired lengths. For cross-linking assays to Sec61 α , the translation mix also contained 1 μ L (<80 equivalents/ μ L) of dog pancreas microsomes (DPMs) and 1 μ L (0.5 pmol/ μ L) of signal recognition particle (tRNA probes). After the translation reaction was terminated, ultracentrifugation in a Beckman TLA-100 rotor at 436000 x g for 10 minutes at 4°C is required to pellet the DPMs. The DPM pellet was then washed with 3M potassium acetate and centrifuged at 19071 x g at 4°C for 10 minutes. The pellet was resuspended in 50 μ L of RNC buffer and split into (A) a 5 μ L

sample and (B) a 45 μL sample. Sample A was treated with 100 $\mu\text{g}/\text{mL}$ RNase A and 5 mM EDTA at incubated at 26°C for 10 minutes. The sample was then heated in 2 X Sample Buffer at 95°C for 5 minutes before the analysis of results by SDS-PAGE as described in Section 2.6.1.

5 μL of BMH buffer (DMSO and 20 mM BMH) was added to sample B and incubated 30 °C for 30 minutes. To terminate the reaction, 100 mM DTT was added and incubated at room temperature for 10 minutes. 6 μL of the sample was added to 100 $\mu\text{g}/\text{mL}$ RNase A and 5 mM EDTA and incubated at 26°C for 10 minutes and heated in 2 X sample buffer at 95°C for 5 minutes before the analysis of results by SDS-PAGE as described in Section 2.6.1. The remainder of the samples were centrifuged at 19071 x g at 4°C for 10 minutes to isolate the DPM pellet. The supernatant was discarded and the pellets were used for immunoprecipitation experiments

2.5.9 Immunoprecipitation of cross-linked products

TCA precipitated pellets from cross-linking experiments were re-suspended in 50 μL Solubilisation Buffer (see Section 2.2 for all buffers used in the following protocol) and heated to 95°C for 5 minutes. The sample was then incubated in 1 mL of ice cold RIPA buffer on ice for 5 minutes, before centrifugation at 19071 x g at 4°C for 10 minutes. The supernatant was retained and transferred to a fresh tube to be incubated on ice for 2 hours with the desired primary antibody. 30 μL Protein- A sepharose beads (0.2 g in 1 mL of RIPA buffer) were added to each sample and incubated for 30 minutes at 4°C. They were then centrifuged at 2655 x g for 1 minute at 4°C and the supernatant discarded. The beads were then washed in High Strength (HS) RIPA buffer and again centrifuged at 2655 x g for 1 minute at 4°C and the supernatant discarded. The HS-RIPA buffer wash was repeated once before a wash with RIPA buffer and re-centrifuged at 2655 x g for 1 minute at 4°C. The supernatant was discarded and 15 μL of 1x Sample Buffer was added to the beads before they were heated to 95°C for 5 minutes. SDS-PAGE analysis of results was carried out as described in Section 2.6.1.

2.5.10 Digestion assay assessing insertion into Dog Pancreas Microsomes (DPMs)

Translation reactions (50 μ L) were carried out as previously described in Section 2.5.4 to generate peptides of desired length. For insertion assays the translation mix also contained 1 μ L (<80 equivalents/ μ L) of dog pancreas microsomes (DPMs) and 1 μ L (0.5 pmol/ μ L) of signal recognition particle (tRNA probes). After the translation reaction was terminated, ultracentrifugation in a Beckman TLA-100 rotor at 436000 x g for 10 minutes at 4°C is required to pellet the DPMs. The DPM pellet was then washed with 3M potassium acetate and centrifuged at 19071 x g at 4°C for 10 minutes. The pellet was re-suspended in 20 mM HEPES buffer (pH 7.4) and Proteinase K is added at a concentration of 0.2 mg/mL (in the presence or absence of 1% (v/v) Triton X-100) and incubated at room temperature for 30 minutes. The DPMs were once again centrifuged at 19071 x g at 4°C for 10 minutes. The supernatant was removed and the DPMs were re-suspended in 1x Sample Buffer and analysed by SDS-PAGE gel.

2.6 Imaging

2.6.1 Gel Electrophoresis and western blotting

Polyacrylamide separating gels of the appropriate percentage were prepared as described by Laemmli (1970) for the separation of large proteins with a MW > 15 kDa. To resolve smaller proteins with a MW <15 kDa tricine gels were used (see Section 2.2 for the solutions used to make separating, stacker and spacer (tricine gel only) components of each gel). Prepared samples were added to the gel, alongside the SeeBlue Pre-stained Marker (Invitrogen) and electrophoresed at 120 volts until through the stacking gel. When the samples entered the resolving gel, the voltage was increased to 200 volts and the gel was run until complete. Following completion, the gel apparatus was removed, the stacking gel cut off and the separating retained. If the results of the gel were to be visualised by autoradiography (see Section 2.1), the gel was placed in destain solution (see Section 2.2) for 1 hour whilst shaking and dried for 2 hours at 60°C on the gel drier. If the gel was to be analysed by immunoblotting, it transferred to an Invitrogen Xcell II wet blotting apparatus. The proteins were then transferred to a nitrocellulose membrane in transfer buffer (see Section 2.2) at 100 mA for 1 hour.

2.6.2 Autoradiography

Following gel electrophoresis, destaining and drying, the protein samples radiolabelled with ^{35}S -methionine were visualised by autoradiography. Dried gels were placed into an autoradiographic cassette and exposed to Kodak X-ray film for a certain period of time before being processed by the Kodak X-Omat 100 processor.

2.6.3 Immunoblotting

Following western blotting, the nitrocellulose membrane was incubated for 1 hour at 4°C with 10 (w/v) Marvel skimmed milk powder in order to block non-specific binding of antibodies. The membrane was washed in TBS (see Section 2.2 for all solution used in the following protocol) and incubated with the appropriate antibody (prepared at 1:1000 in TBS) at 4°C overnight. Three 10 minute wash steps using TBS/0.1% (v/v) Tween 20 followed by a 1 hour incubation with the secondary antibody (prepared at 1:20000 in TBS). Three further wash steps in TBS were carried out, before the membrane was developed using Amersham™ ECL western blotting reagent. Equal volumes of the two ECL solutions were mixed and poured over the nitrocellulose membrane and left to incubate for 2 minutes. The excess ECL was then dabbed off and the membrane was placed into an autoradiographic cassette. The membrane was exposed to Kodak X-ray film for a certain period of time before being processed by the Kodak X-Omat 100 processor.

2.7 Statistical analysis

Group data are expressed as mean \pm SD. Groups were compared using one-way ANOVA with Tukey's *post hoc* analysis. Analysis was carried out using GraphPad Prism version 5.0 (GraphPad Software, Inc., USA). Significance was accepted at $p < 0.05$.

3. Investigating secondary structure formation in transmembrane domain 1 of hGPR35 and Bacterioopsin

3.1. Introduction

Like all other proteins, a GPCR starts its life as a nascent peptide in the ribosome tunnel. GPCRs are comprised of seven transmembrane (TM) helical domains, which are vital in how the receptor functions. Although these TM segments are critical to both the structure and function of GPCRs, little is known about their biogenesis and the mechanisms that are intrinsic to obtaining the secondary, tertiary and quaternary structure of these domains. Much research has been carried out investigating the mature structure, indicating the helical nature of the individual TM domains, but little is known as to when this helicity first arises. Potentially, this could occur during peptide synthesis and may involve the ribosome itself.

Many theories suggest that the tertiary structure of integral membrane proteins (IMPs) is established as they are co-translationally inserted into the membrane. This therefore asks the question, when does secondary structure form and can it occur before the nascent chain has left the ribosome? The large subunit of the ribosome is approximately 100Å in length when measured from the PTC to the base of the exit tunnel (Lu & Deutsch 2005b). The diameter of the tunnel also ranges from 10-20Å, suggesting the ribosome tunnel could tolerate the formation of an α -helix within a nascent chain during translation (Mingarro et al. 2000; Ban 2000; Kramer et al. 2001). The ribosome tunnel has been shown to be both structurally and biochemically diverse, as well as being favourable towards the formation of α -helices (Ziv et al. 2005). For example, structural studies have provided evidence of ribosomal proteins not only lining, but protruding into the tunnel, generating zones of helix stabilization (Lu & Deutsch 2005b). Biochemical studies have also suggested that these zones could be key in promoting peptide folding (Woolhead et al. 2004; Bhushan et al. 2010; Robinson et al. 2012).

The size restraints and dimensions of the tunnel lead us to believe only helices can form in the upper and middle regions of the tunnel, leaving other secondary structures such as β -sheets and tertiary structures such as β -hairpins to develop in the distal regions of the ribosome exit tunnel (Kosolapov & Deutsch 2009; Conti et al. 2014; Lu & Deutsch 2014; Marino et al. 2016). Although the environment of the tunnel is favourable towards the formation of secondary structure, examples such as TM1 in leader peptidase (Lep) and

TM2 in Kv1.3 choose not to take up their folded conformation in preference of an extended conformation whilst fully in the tunnel (Houben et al. 2005; Tu & Deutsch 2010). The amino acid sequence of a nascent chain is believed to be equally as important as the ribosome environment in relation as to whether secondary structure forms in the tunnel. The tunnel acts to decode certain motifs within the amino acid sequence but it is the two in tandem which control such events as elongation, folding and termination (Kramer et al. 1999; Woolhead et al. 2006).

The formation of secondary structure within TM domains of IMPs has been frequently remarked upon, with several studies showing the ability of the ribosome to recognise a TM segment as it makes its way through the tunnel. Experiments using a single-spanning membrane protein produced results suggesting that upon recognition of the TM segment, the ribosome was capable of aiding its folding (Liao et al. 1997; Woolhead et al. 2004). Later experiments identified this region of the ribosome to contain the uL22 protein and linked the folding of TM domains within the tunnel to a signalling pathway for ER gating (Woolhead et al. 2004). This prediction was later verified by cross-linking data suggesting the appearance and folding of the TM domains within the tunnel triggered a structural rearrangement of the Sec61 translocon, priming it for TM integration (Pool 2009). The importance of co-translational targeting of RNCs has lead us to question the importance of TM domain secondary structure for SRP. Actively translating ribosomes have been shown to increase their affinity for SRP as the nascent chain enters the distal regions of the tunnel (Berndt et al. 2009; Flanagan et al. 2003) which in some cases has been enhanced by TM domain folding (Tu et al. 2000).

Not only has secondary structure formation within the tunnel been shown to play a pivotal role in the targeting of TM proteins to the membrane, there is some evidence that it also aids the assembly of multimeric membrane spanning proteins in the ER. Studies by Kosolapov and Deutsch (2003) reveal that the folding of the 1st TM domain of a voltage gated potassium (Kv) channel, whilst ribosome bound, is essential for acquiring its tertiary and quaternary structure within the ER membrane. Folding within TM 1 and later TM segments of the Kv channels was assayed in further experiments using a technique known as pegylation. Pegylation will be utilized in this chapter to analyze the co-translational folding of the N-terminal TM 1 of GPR35.

The pegylation assay has been used by several groups to demonstrate nascent peptide compaction within the ribosome tunnel (Lu & Deutsch 2005b; Lu & Deutsch 2005a; Conti

et al. 2014). The pegylation assay firstly requires an *in vitro* translation to produce stalled RNCs. The assay relies on mass-tagging the nascent peptide with a molecule of methoxy-polyethylene glycol maleimide (PEG-MAL), to cysteine residues, which increases the molecular weight of a nascent peptide by 5 kD when separated on an SDS-PAGE gel. Each nascent chain will contain a single, specifically-placed cysteine residue (the marker cysteine), which when exposed to the PEG-MAL molecule will become mass-tagged and detectable by a gel shift assay (Figure 3.1B). Due to the size of the PEG-MAL molecule, RNCs containing a cysteine buried within the ribosome tunnel will be protected from pegylation (Figure 3.1 C1). As the cysteine emerges from the ribosome it is no longer protected and becomes available for pegylation (Figure 3.1 C2). As the ribosome tunnel is known to be approximately 100Å, from the PTC to its exit, a fully extended peptide at 3-3.4Å/ amino acid would require approximately 30 amino acids to traverse the tunnel, where as an α -helix at 1.5Å/ amino acid would require approximately 67 amino acids (Lu et al. 2007). From these studies, nascent chains within the ribosome exit tunnel have been shown to exist in different forms, with different levels of compaction within the tunnel (Figure 3.1A). Hence, it is difficult for a pegylation to discriminate between tight compaction in one region of the nascent chain and full extension in another. Therefore, using pegylation we can measure the point at which the nascent peptide first emerges from the ribosome exit tunnel and by calculating the nascent chain length detect whether secondary structure has formed within the ribosome.

The aim of this chapter is therefore, to use a pegylation assay to investigate the folding profile of the N-terminal region of the human orthologue of GPR35 (a GPCR), specifically assessing if secondary structure is apparent within the first TM domain. The folding profile of a structurally similar microbial Opsin (Bacterioopsin) was also assessed and used to investigate the differences between prokaryotic and eukaryotic *in vitro* translation systems. Using two different systems offered the opportunity to compare two different subsets of ribosomes, in turn providing us with the chance to investigate if the two proteins of interest produced the same co-translational folding profile in the eukaryotic and prokaryotic systems. This would therefore enable us to establish whether ribosome environment or sequence specificity within the nascent peptide had a greater impact on generating secondary structure within a nascent chain.

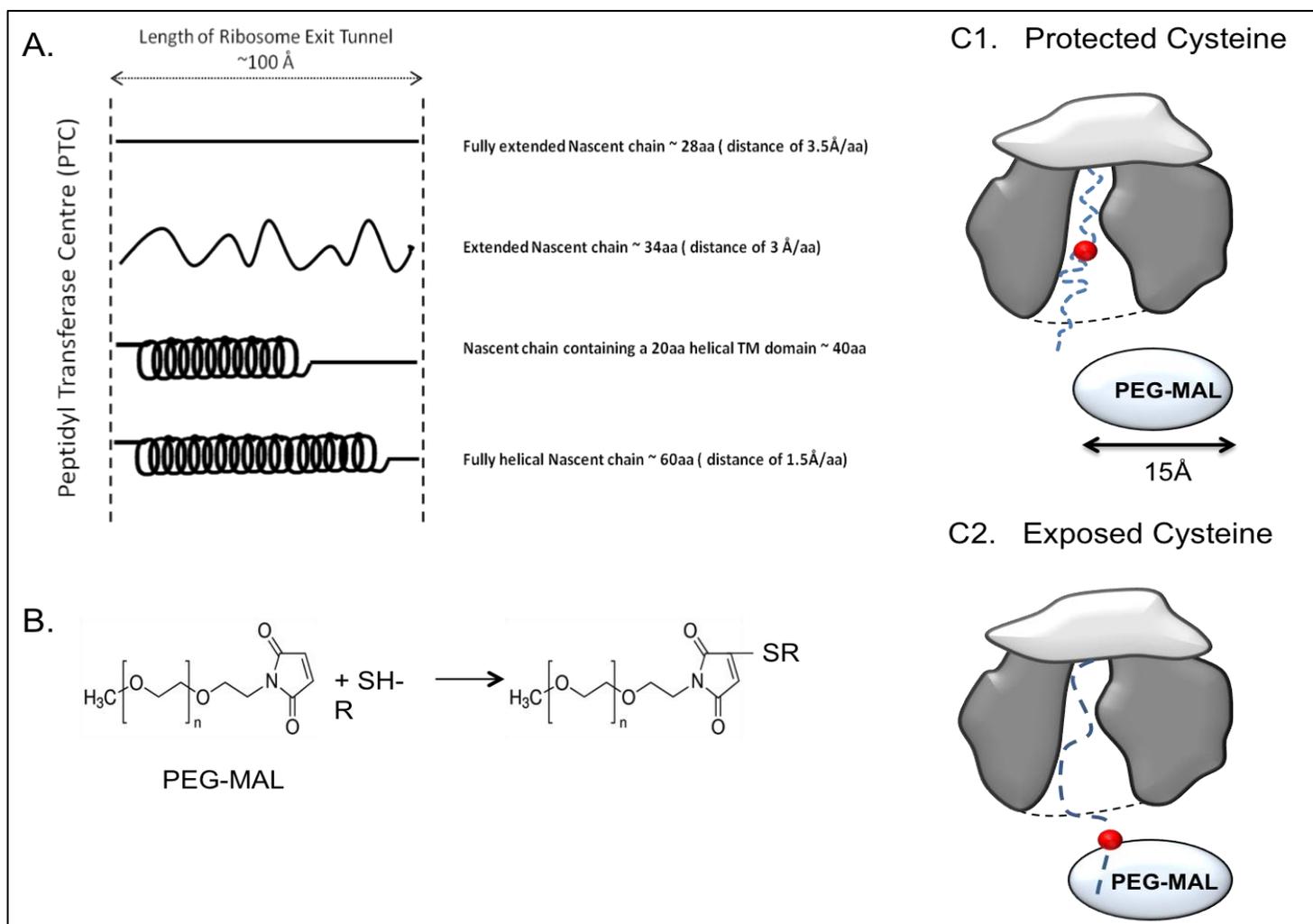


Figure 3.1 Pegylation assay to measure the level of secondary structure formed within the ribosome exit tunnel. (A) PEG-MAL molecule forms a thiol bond between maleimide head group and free cysteine residues, attaching polyethylene glycol chains to other molecules. (B) Schematic representing the length of nascent chain required to traverse the ribosome exit tunnel, depending on the degree of compaction. (C1) Schematic diagram representing the protection offered to the marker cysteine (red circle) from a molecule of PEG-MAL by the ribosome tunnel and (C2) interaction between PEG-MAL and the cysteine when exposed from the ribosome exit tunnel.

3.2. Results

3.2.1. Investigating the formation of secondary structure in TM1 of GPR35 and Bacterioopsin within the prokaryotic ribosome.

To measure the compaction of the first TM domain of GPR35 and BO within the ribosome exit tunnel, we first had to prepare the constructs to be used in the pegylation assay. For pegylation assays, we require a single cysteine residue in the N-terminus of our genes of interest, upstream of the first TM domain. Figures 3.3A and 3.4A show the sequence of the N-terminal regions of both hGPR35 and BO respectively. GPR35 contains a native cysteine at residue 8 (C8) which was calculated to be too far upstream from the first TM domain to determine if compaction occurred when using pegylation as an assay. Therefore, by site-directed mutagenesis (see section 2.4.6.1), C8 was substituted for an alanine and a cysteine residue was placed at amino acid position 15 (C15). The C15 residue is located 10 amino acids upstream of TM1 and will become a marker for pegylation, hence it will be known as the marker cysteine (Figure 3.3A).

To determine if secondary structure was forming within TM1 of GPR35 whilst ribosome bound, we generated translation intermediates of various lengths ranging from 25-50aa from the peptidyl transferase centre (PTC) to the marker cysteine. Individual intermediates were generated from linear DNA, which lacked a stop codon in the reverse primer and hence produced stable RNCs. The length each intermediate increased by 5aa within the range and were chosen to provide a detailed analysis of the movement of the TM domain through the ribosome tunnel. RNCs were generated using the coupled transcription/translation S-30 expression system. Following pegylation, the RNCs were precipitated using cetyltrimethylammonium bromide (CTABr) was carried out to isolate the nascent chains that remained bound to the ribosome and to discriminate against those that may have become disengaged from the PTC. Pegylation of C15 was detectable by a single 5 kDa shift of translation product on a tricine gel.

The results of the pegylation assay to assess the compaction of TM1 in GPR35 are shown in Figure 3.3. The intermediate length of 25aa at which the marker cysteine is 25 amino acids from the PTC shows clearly what is expected when pegylation does not occur. At this length there is no change in the size of the translation product, with C15 expected to be found within the ribosome tunnel and hence, should be unavailable to interact with the PEG-MAL molecule. This agrees with previous studies suggesting that PEG-MAL is inaccessible to the mid-tunnel of the ribosome

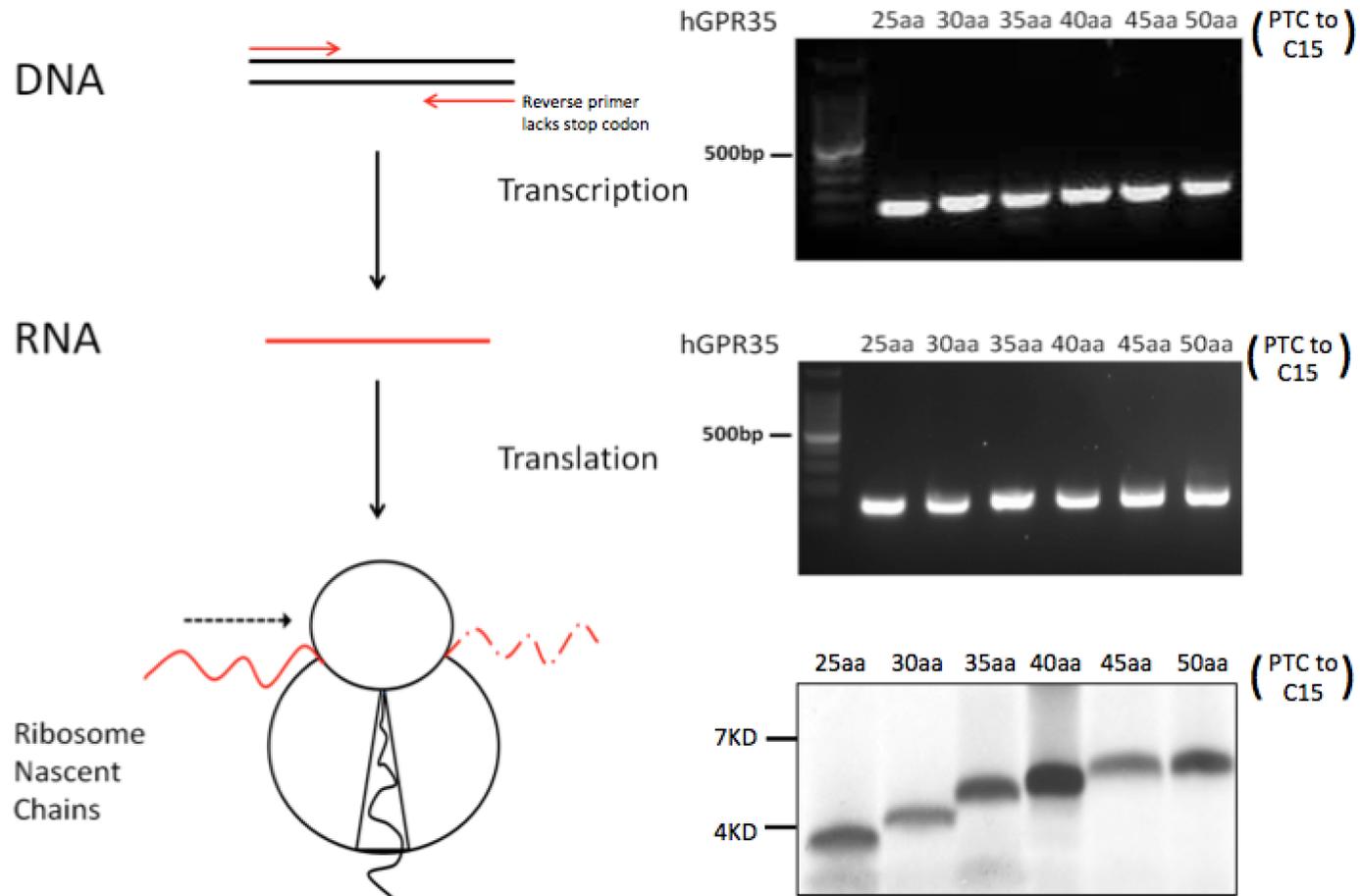


Figure 3.2 Diagram to show how ribosome bound nascent chains (RNCs) are generated. Firstly, linear DNA is generated from PCR reactions using a plasmid DNA template. The reaction specifically contains a reverse primer without a stop codon to prevent the release of the nascent peptide from the ribosome. Transcription of linear DNA to messenger RNA is preformed either manually or as part of a coupled transcription/translation reaction. Finally, within in vitro translation systems generate nascent peptides attached to the ribosome, which can be used to measure secondary structure within the ribosome tunnel.

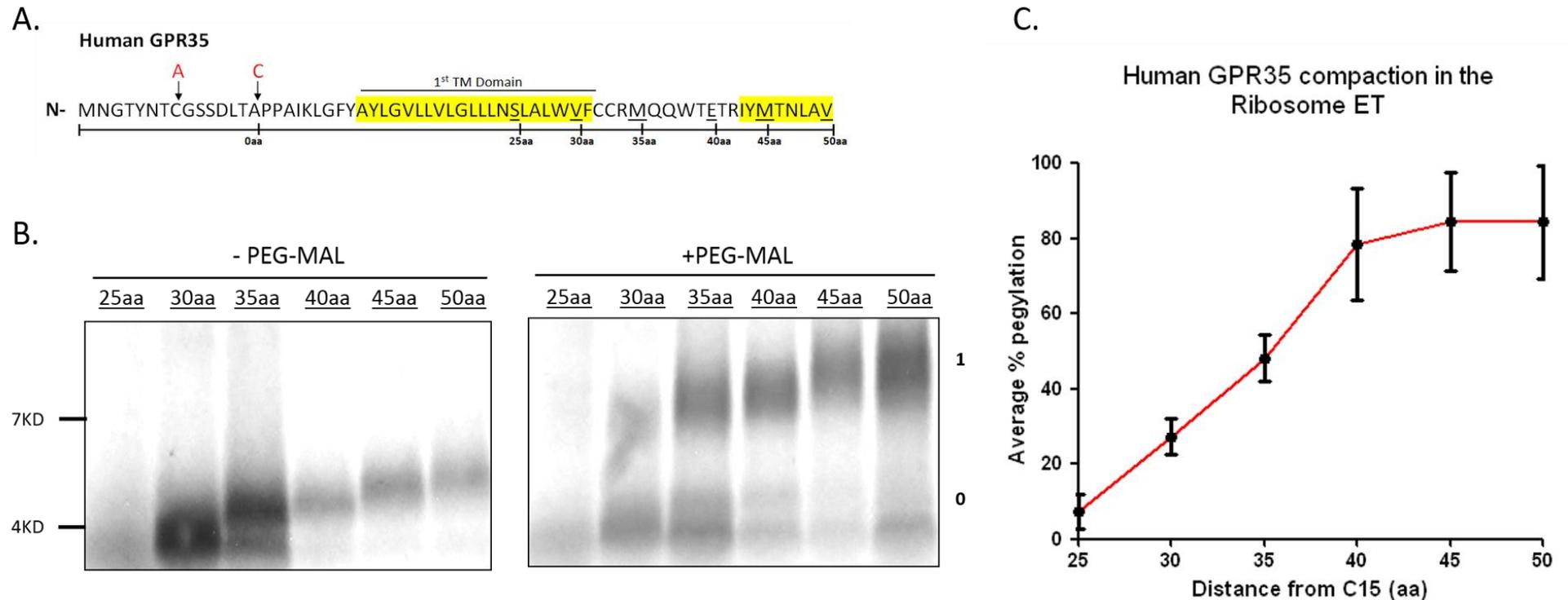


Figure 3.3. Analysis of the pegylation assay of GPR35 in the prokaryotic S-30 transcription/translation system. (A) Schematic diagram of the GPR35 gene highlighting the substitution mutations made by SDM to place a single cysteine (marker cysteine) at position 15 for pegylation and remove the native cysteine at position 8. From the marker cysteine, the position of the reverse primers generating various lengths of intermediates can be seen. (B) Intermediate lengths of GPR35 were expressed in the S-30 coupled transcription/translation reaction. The reaction was split into two, with one half being incubated with 1mM PEG-MAL and the other half incubated in buffer as a control. CTAB precipitation was carried out to ensure selection of ribosome-bound nascent chains and exclude those which had dissociated from the PTC. A representative gel displays how the pegylated (1) and non-pegylated (0) samples were resolved by SDS-PAGE, measuring if pegylation had taken place by a gel-shift of the translation product by ~ 5kDa. (C) The average % pegylation of individual intermediates shows pegylation of cysteine 15 occurs as early as 30 amino acids from the PTC. The % pegylation of intermediates shows an increase between 30 and 40 amino acids before it plateaus. All % pegylation values were calculated using $[\text{pegylated}/(\text{unpegylated} + \text{pegylated})]$ adjusted for background. Average percentage pegylation is calculated from an n of 3. Error bars indicate standard deviation.

(Lu & Deutsch 2005b; Lu & Deutsch 2005a). As the nascent chain increases in length to 30aa, a small fraction of the translation product, approximately 30%, becomes pegylated resulting in a gel shift (Figure 3.3 B and C). At this point the marker cysteine is believed to be in the vestibule of the exit tunnel where the PEG-MAL molecule can interact with cysteine residues in an inefficient manner. As the distance of the marker cysteine increases to 35-40aa from the PTC, we see a greater rate of pegylation increasing from 45% to 75%, correlating with the ability of the PEG-MAL molecule to efficiently bond to the C15 residue at the N-terminus of GPR35. Further lengths of 45aa and 50aa show pegylation at levels of approximately 80%, little change from that seen at 40aa, suggesting the marker cysteine had fully exited the ribosome tunnel and was completely exposed at that point to the PEG-MAL molecule (Figure 3.3).

The findings above suggest that TM1 of GPR35 exists in an extended conformation as it makes its way through the prokaryotic ribosome tunnel. With the marker cysteine located 10 amino acids downstream from TM1 showing signs of pegylation as early as 30aa from the PTC, and pegylation increasing thereafter, it seems unlikely that compaction of the nascent chain is taking place while the N-terminus of TM1 is in the lower regions of the tunnel. However, as the prokaryotic ribosome would not be considered a native environment for the expression of a eukaryotic GPCR, this may affect how the components of ribosome tunnel interact with the GPR35 nascent chain. To investigate a possible prokaryotic specific interaction, the folding profile of a structurally similar protein, BO, was examined in this translation system.

To enable the assessment of secondary structure in BO by pegylation, SDM was carried out to introduce a marker cysteine at position 15, to make the system comparable to the results for GPR35. As BO had no native cysteines upstream of the first TM domain, C15 (the marker cysteine) would be the sole site for pegylation. The pegylation assay was carried out as with GPR35 using identically sized intermediates. In every assay carried out with BO, neither the 25 nor 30aa intermediate could be expressed, possibly due to prokaryotic ribosome specific instability of these lengths of BO at the PTC. The first length to express is the intermediate presenting the marker cysteine 35aa from the PTC. At this length PEG-MAL already shows signs of binding with ~60% efficiency, generating a shift in a high portion of the translated product. This, as with the 35aa intermediate of GPR35, would suggest the marker cysteine now resides outside exit site of the ribosome tunnel and is available for efficient pegylation of the nascent chain.

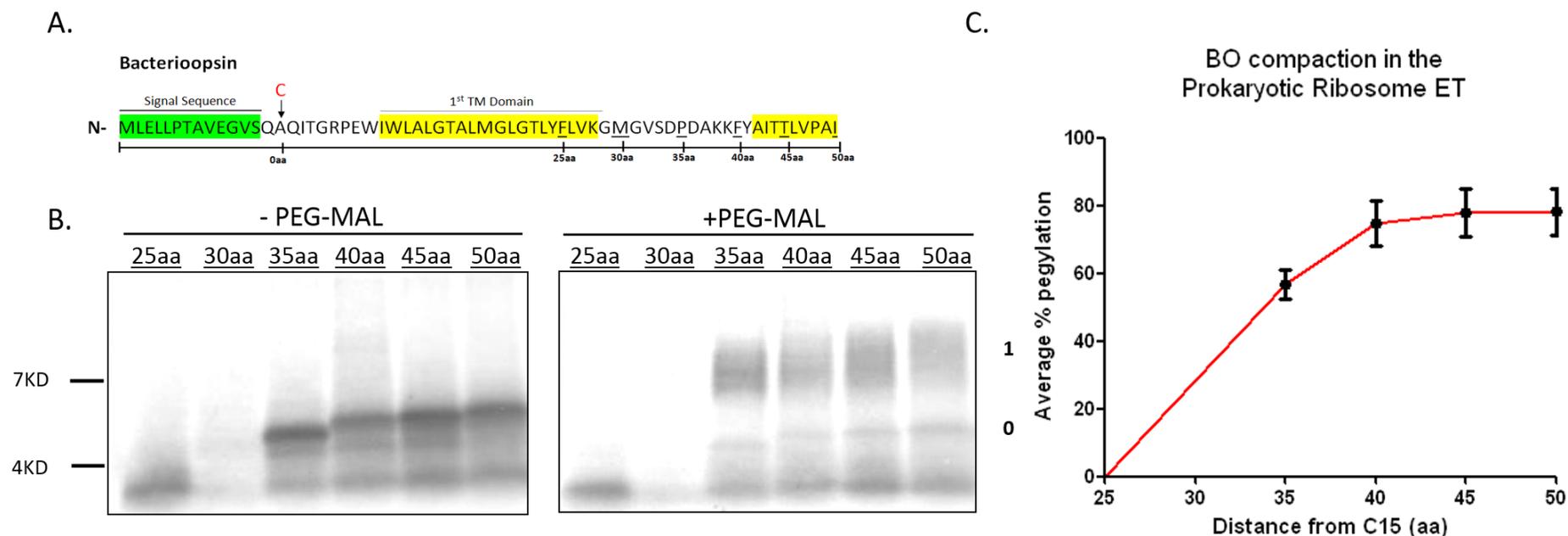


Figure 3.4. Analysis of the pegylation assay of Bacterioopsin in the prokaryotic S-30 transcription/translation system. (A) Schematic diagram of the BO gene highlighting the substitution mutations made by SDM to place a single cysteine (marker cysteine) at position 15 for pegylation. From the marker cysteine, the position of the reverse primers generating various lengths of intermediates can be seen. (B) Intermediate lengths of BO were expressed in the in vitro S-30 couple transcription/translation reaction. The reaction was split into two, with one half being incubated with 1mM PEG-MAL (+PEG-MAL) and the other half incubated in buffer as a control (-PEG-MAL). The representative gels display how the pegylated (1) and non-pegylated (0) samples were resolved by SDS-PAGE, measuring if pegylation had taken place by a gel-shift of the translation product by ~5kDa. (C) The average % pegylation of individual intermediates shows pegylation of cysteine 15 occurs from 35 amino acids from the PTC. The % pegylation of intermediates shows an increase between 35 and 40 amino acids before it plateaus in the subsequent intermediates. Both the 25 and 30aa BO constructs failed to express in the prokaryotic system. All % pegylation values were calculated using $[\text{pegylated}/(\text{unpegylated} + \text{pegylated})]$ adjusted for background. Average percentage pegylation is calculated from an n of 3. Error bars indicate standard deviation.

(Figure 3.4 B and C). All subsequent intermediate lengths of BO were pegylated with upwards of 70% efficiency, suggesting the nascent chain was fully available to PEG-MAL.

The above results of the pegylation assay indicate that the first TM domain of BO, like GPR35, is absent of secondary structure and remains in an extended conformation whilst its N-terminus is in the lower regions of the ribosome tunnel. This suggests that either the sequence coding for the nascent peptide in both proteins does not favour the generation of secondary structure or that the environment of the *E. coli* ribosome does not support the co-translational folding of these nascent chains. To test this theory, the next set of experiments will be assessing the capability of GPR35 and BO to compact within eukaryotic ribosome exit tunnel. This will enable us to investigate the impact the ribosome tunnel environment has on the translating nascent chain and whether the change in environment produces a change in the folding profile of GPR35 and BO.

3.2.2. Investigating the formation of secondary structure in TM1 of GPR35 and Bacterioopsin within the eukaryotic ribosome.

Following the investigation of the formation of secondary structure within the 1st TM domain of GPR35 and BO in the prokaryotic ribosome, the next step was to assay if secondary structure was capable of forming within the eukaryotic ribosome exit tunnel. The ribosome tunnel environment has been shown to play a vital role in the compaction of nascent peptides whilst being co-translationally expressed (Woolhead et al. 2006; Kosolapov et al. 2004; Robinson et al. 2012). With this mind, the pegylation assay previously carried out in the S-30 coupled transcription/translation system will be attempted in the eukaryotic wheat germ (WG) extract translation system. A eukaryotic system will enable us to assess how critical the change in ribosome tunnel environment is in the formation of secondary structure within TM1 of GPR35 and BO.

Again the pegylation assay assessed RNC lengths ranging from 25-50aa from the PTC to the marker cysteine at residue 15, in both model proteins. Levels of pegylation of the RNCs in the eukaryotic system follow closely the results previously seen the prokaryotic system. In GPR35, we again witness no pegylation at the 25aa intermediate as expected with the marker cysteine buried deep within the ribosome exit tunnel and inaccessible to PEG-MAL. As the cysteine is moved 5aa further from the PTC we begin to see pegylation occur, highlighted by the ~30% gel shift of the translation product. Following the trend of the prokaryotic system, we once again see pegylation rates steadily increase to ~80% of the translation product in the intermediate which places C15 40aa from the PTC (Figure 3.5).

The results using the same assay with BO show an almost identical outcome. In the eukaryotic system however, both the 25 and 30aa intermediates were expressed, something that was not possible in the prokaryotic system. The results show, a pegylation shift in every intermediate length except for the 25aa construct, which is expected. Unlike the prokaryotic S-30 system, the 30aa intermediate is expressed and becomes pegylated with ~30% efficiency. The quantified rate of pegylation follows closely the trends set in the prokaryotic assay with the subsequent intermediates. The 35aa length of BO becomes pegylated with ~50% efficiency, before the later lengths between 40-50aa from the PTC reach maximal rates of ~75% (Figure 3.6). The results shown above, coupled with those seen in section 3.2.1 suggest that the first TM domain of both GPR35 and BO exist in an extended conformation and hence show no sign of the formation of secondary structure, whilst traversing the ribosome tunnel. The results obtained agree with previous studies

suggesting an extended peptide (3.4Å per aa) would require ~30aa to traverse the ribosome tunnel, whereas a fully compacted peptide would require ~65aa (1.5Å/aa) to travel the same distance (Lu & Deutsch 2005b). As we know from both the sequence and the structure of the mature GPR35 and BO proteins, this region of the nascent will contain a combination of TM domains and loops. Therefore, we would expect that if secondary structure was to form whilst in the ribosome exit tunnel, pegylation would occur with the marker cysteine ~45aa from the PTC. As PEG-MAL binding is seen at C15 when it is ~30-35aa from the PTC, the existence of significant secondary structure within the exit tunnel can be ruled out.

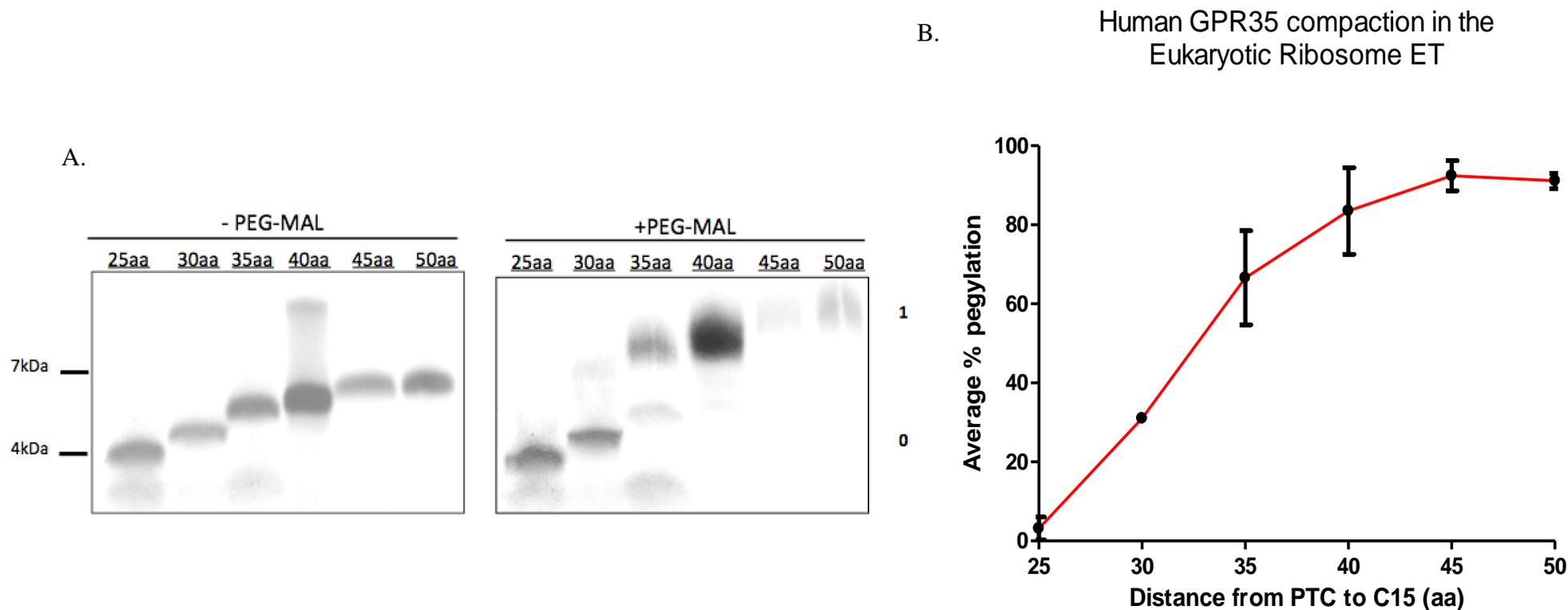


Figure 3.5 Analysis of the pegylation assay of GPR35 in the eukaryotic translation system. (A) Intermediate lengths of GPR35 were expressed in the *in vitro* eukaryotic WG translation reaction. The reaction was split into two, with one half being incubated with 1mM PEG-MAL and the other half incubated in buffer as a control. Representative gel displays how the pegylated (1) and non-pegylated (0) samples were resolved by SDS-PAGE, measuring if pegylation had taken place by a gel-shift of the translation product by ~ 5kDa. (B) The average % pegylation of individual intermediates shows pegylation of cysteine 15 occurs as early as 30 amino acids from the PTC. The % pegylation of intermediates shows an increase between 30 and 40 amino acids before it plateaus. All % pegylation values were calculated using $[\text{pegylated}/(\text{unpegylated} + \text{pegylated})]$ adjusted for background. Average percentage pegylation is calculated from an n of 3. Error bars indicate standard deviation.

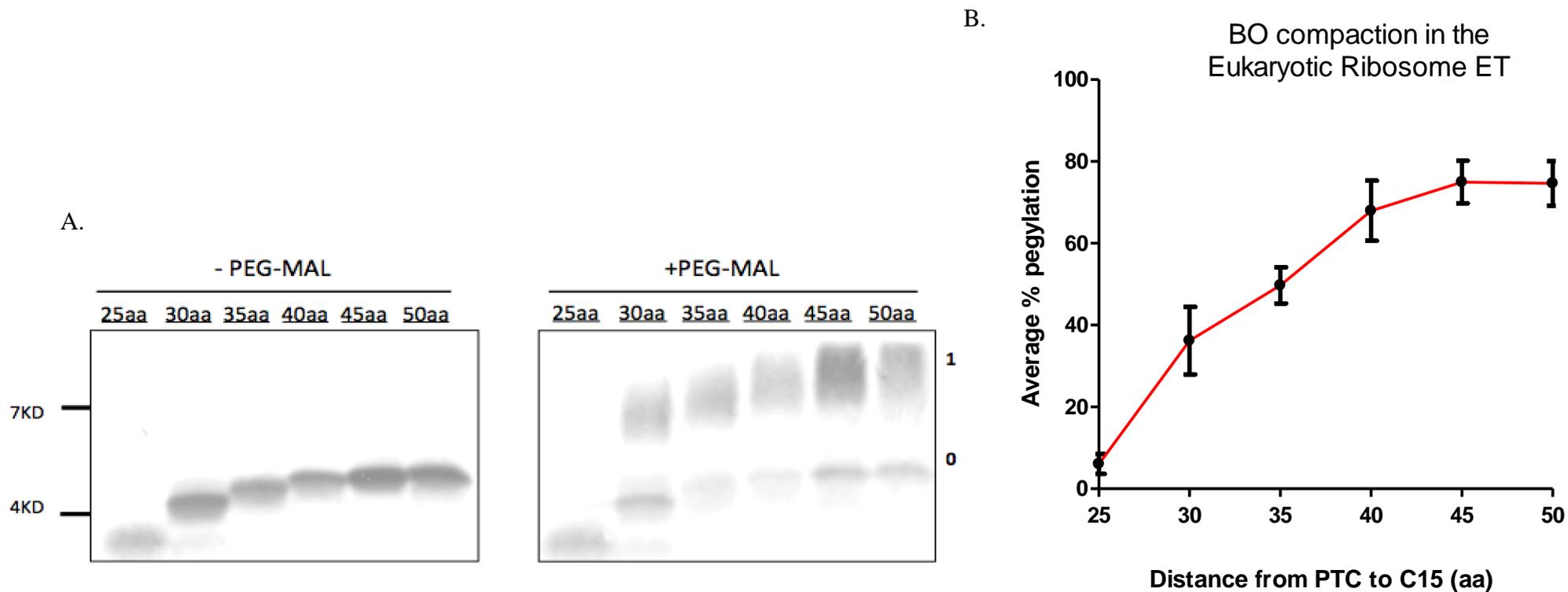


Figure 3.6. Analysis of the pegylation assay of Bacterioopsin in the eukaryotic translation system. (A) Intermediate lengths of BO were expressed in the *in vitro* eukaryotic WG translation reaction. The reaction was split into two, with one half being incubated with 1mM PEG-MAL and the other half incubated in buffer as a control. Representative gel displays how the pegylated (1) and non-pegylated (0) samples were resolved by SDS-PAGE, measuring if pegylation had taken place by a gel-shift of the translation product by ~ 5kDa. (B) The average % pegylation of individual intermediates shows pegylation of cysteine 15 occurs as early as 30 amino acids from the PTC. The % pegylation of intermediates shows an increase between 30 and 40 amino acids before it plateaus. All % pegylation values were calculated using [pegylated/(unpegylated + pegylated)] adjusted for background. Average percentage pegylation is calculated from an n of 3. Error bars indicate standard deviation

3.2.3. F₀c- pegylation of a bacterial membrane protein known to fold in the ribosome exit tunnel.

The investigation of secondary structure using the pegylation assay has so far shown the lack of significant α -helical structure within the first TM domain of two large integral membrane proteins; a GPCR, GPR35 and a prokaryotic opsin, BO. As both had previously unknown co-translational folding profiles, it was decided to use a membrane protein shown previously to fold whilst traversing the ribosome exit tunnel to ensure the assay was valid. This protein, F₀c, is a bacterial membrane protein, specifically subunit c of the F_o component of the ATP synthase. F₀c is a 79 amino acid protein and contains two TM domains both known to be inserted into the bacterial periplasm (Van Der Laan et al. 2004). FRET and cross-linking studies carried out by Robinson et al (2012) show the first TM domain of F₀c undergoes compaction in the distal regions of the ribosome tunnel. Cross-linking to the ribosomal protein uL23 determined that F₀c required ~48aa to traverse the length of the ribosome tunnel.

The F₀c pegylation assay also required a specifically placed single cysteine residue downstream of the first TM domain. A construct was designed previously with a cysteine substituted at position 5 by SDM. Therefore, intermediates of F₀c will be named by size from cysteine 5 (the marker cysteine) to the PTC. As F₀c has previously been shown to fold in the tunnel (Robinson et al. 2012), intermediate lengths between 35-70aa were generated with the 35aa intermediate expected to act as a negative control for pegylation and the 70aa acting as a positive control. As predicted, there was no sign of a shift in the translation product at 35aa, a length at which both GPR35 and BO were shown to interact strongly with PEG-MAL (Figure 3.7). Using Robinson et al (2012) as a guide, the first length at which we expected the marker cysteine to emerge from the ribosome tunnel was ~45aa from the PTC. Upon pegylation of this length we saw a translation shift correlating with GPR35 and BO intermediates of 35aa, with ~50% pegylation showing the marker cysteine was outwith the tunnel and capable of interacting with PEG-MAL (Figure 3.7). Further intermediate lengths of 50, 55 and 70aa were expressed and showed higher pegylation levels of ~75%, similar to the latter intermediates of 40, 45 and 50aa in GPR35 and BO (Figure 3.3 and 3.4). This indicates that the marker cysteine outside the ribosome exit tunnel and was capable of interacting with PEG-MAL efficiently.

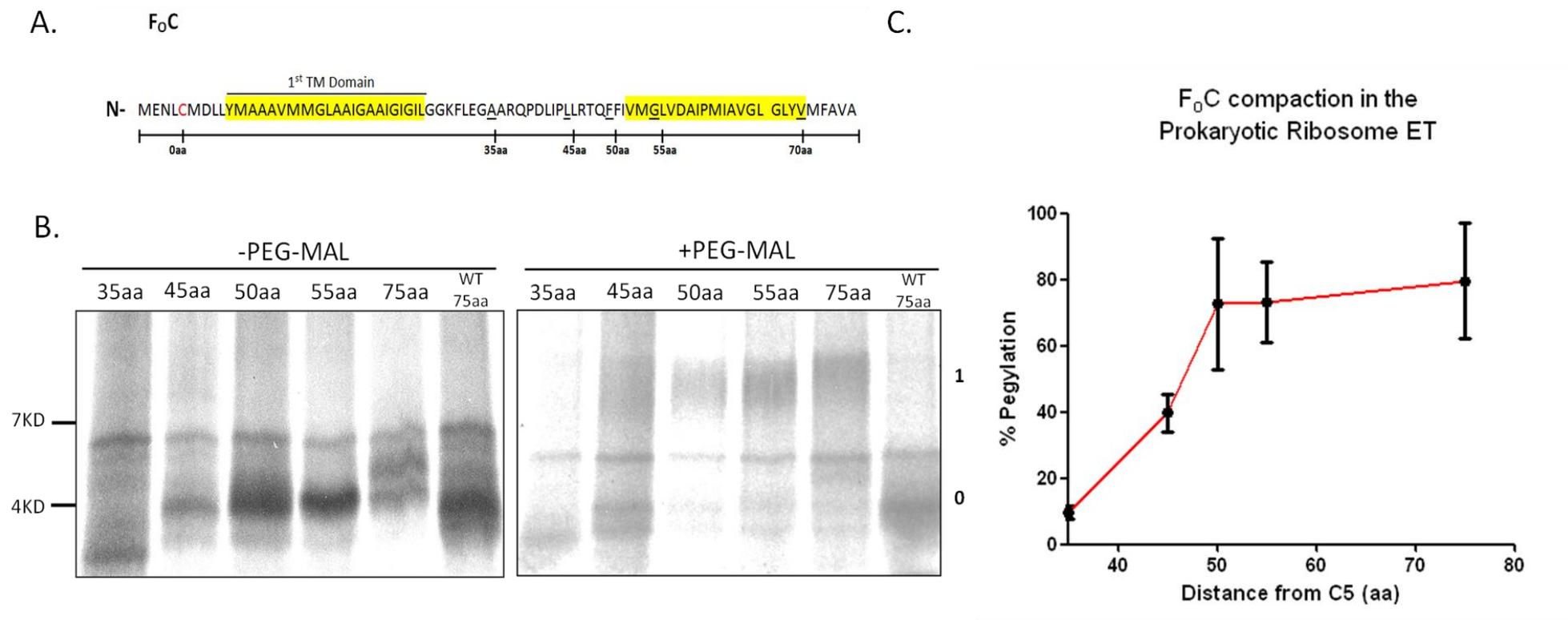


Figure 3.7. Pegylation assay control using F_0C in the prokaryotic S-30 transcription/translation system. (A) Schematic diagram of the F_0C gene highlighting the substitution mutations made by SDM to place a single cysteine (marker cysteine) at position 5 for pegylation. From the marker cysteine, the position of the reverse primers generating various lengths of intermediates can be seen. TM domains are indicated (yellow), with emphasis on the 1st TM of which compaction will be measured by pegylation. (B) Intermediate lengths of F_0C were expressed in the in vitro S-30 couple transcription/translation reaction. The reaction was split into two, with one half being incubated with 1mM PEG-MAL and the other half incubated in buffer as a control. Representative gel displays how the pegylated (1) and non-pegylated (0) samples were resolved by SDS-PAGE, measuring if pegylation had taken place by a gel-shift of the translation product by ~ 5kDa. (C) The average % pegylation of individual intermediates shows pegylation of cysteine 5 occurs at 45 amino acids from the PTC. The % pegylation of intermediates shows an increase between 35 and 50 amino acids before it plateaus. All % pegylation values were calculated using $[\text{pegylated}/(\text{unpegylated} + \text{pegylated})]$ adjusted for background. Average percentage pegylation is calculated from an n of 3. Error bars indicate standard deviation.

The results of the pegylation assay of the F_{0c} membrane protein indicates that PEG-MAL is a tool capable of detecting compaction within the TM domains of integral membrane proteins. Not only does it confirm that F_{0c} , as described in Robinson et al. (2012), exists in a compacted conformation within the ribosome ET, but it also confirms that pegylation is an adequate tool assessing structure in the TM domains of GPR35 and BO were there appears to be none (see section 3.2.1 and 3.2.2).

3.2.4. Is the lack of isoleucines in the first TM of GPR35 and BO linked to the lack of secondary structure formation in the ribosome exit tunnel?

As previously discussed in this chapter, pegylation assays determined that the first TM domains of GPR35 and BO are not capable of forming secondary structure within the ribosome ET. Pegylation assays in both eukaryotic and prokaryotic ribosomes were carried out to determine the importance of the tunnel environment in forming secondary structure. Whilst many studies have indeed confirmed the ribosome tunnel environment was essential for the formation of structure within the nascent peptide (Woolhead et al. 2006; Lu & Deutsch 2005b; Lu & Deutsch 2005a), it does not seem to have an effect on the folding profile of the two model proteins. This raises the question as to what can make some domains fold early in the ribosome tunnel, while others remain extended until they exit the ribosome.

As the sequences of our model proteins seemed to favour an extended nascent peptide in the ribosome tunnel, it was decided to analyse the sequences of nascent peptides deemed to be ‘known folders’ in the ribosome tunnel and investigate whether any sequence similarities or differences could be detected. A search of the literature was carried out to obtain the sequences of α -helical domains previously shown to fold in the ribosome exit tunnel (Table 3.1). The compacted peptides were selected from a wide range of non-synthetic domains (transmembrane segments, signal peptides and stalling peptides).

A comparison of the individual sequences highlighted a clear lack of sequence complementarity, even within segments of the same protein and varying degrees of hydrophobicity and size. A noticeable similarity, however, between all the domains shown in Table 3.1, was the high number of amino acids deemed to have a high helical propensity (Pace & Scholtz 1998). At this point it, was noticed, that although TM1 of both GPR35

Domain Type	Name	Sequence (Folded domain red)	Reference
Signal Peptide	EspP(1-25)	<u>MNKIYSLKYHITGGGLIAVSELSGRVSSR</u>	(Peterson et al. 2010)
TM Domain			
	F ₀ C	AAAV <u>MMGLAAIGAAIGILGGKFLE</u> GAAR	(Robinson et al. 2012)
	111p (VSVG)	NELDR <u>SSIASFFIIKLIIGLFLVLR</u> FR LQ	(Woolhead et al. 2004)
	S1 Kv 1.3	GPAR <u>GIAIVSVLVILISIVIFCL</u> ETLP	(Kosolapov et al. 2004)
	S3 Kv 1.3	SRNI <u>MNLIDIVAIPYFTITLGT</u> ELAE RQGN	(Tu & Deutsch 2010)
	S4 Kv1.3	GQQAMSLA <u>LRVIRLVRFRIFKLSHRH</u> SKGLQI	(Tu & Deutsch 2010)
	S5 Kv1.3	KASM <u>RELGLLIFFLFIGVILFSSAVYFAE</u> ADDP	(Tu & Deutsch 2010)
	S6 KV 1.3	GGKI <u>VGSLCAIAGVLTIALPVPVIVSN</u> FNYF	(Lu & Deutsch 2005a)
	Bovine Opsin TM2	HPLN <u>WILVNLAIDLAETIISTISVVNQ</u> MYGYF	(Lin et al. 2011)
	Bovine Opsin TM3	CVVE <u>GTVSLCGITGLWSLAISW</u> ERWM	(Lin et al. 2011)
Stalling Peptide	SecM	PQAK <u>FSTPVWISQAQGRAGP</u> QRLT	(Woolhead et al. 2006)
	CGS1	SIKAR <u>RNCNIGVAQIVAAK</u> WS	(Onoue et al . 2011)
Model Proteins	GPR35 TM1	LGFY <u>AYLGVLLVLGLLNSLALWVF</u> CCRM	This study
(Non-folders)	BO TM1	GRPE <u>WILWALGTALMGLGTLYFLVK</u> GMGV	This study

Table 3.1. Transmembrane domains known to fold whilst in the ribosome tunnel. A literature review highlighted numerous domains capable of folding in the ribosome tunnel. The sequence of individual domains are shown, with regions that form secondary structure highlighted in red.

and BO are made up primarily of residues with a high helical propensity, unlike the other domains in Table 3.1, they lack central isoleucine residues. Although isoleucine residues do not have the greatest helical propensity, it was queried as to whether this residue and other important factors in helix formation, such as the hydrophobicity of environment (Blaber et al. 1994), steric contacts in the helix (Hermans et al. 1992) and favourable side chain-helix VDW interactions (Wang & Purisima 1996; Lu & Deutsch 2005b) could alter the formation of a helix.

To investigate if isoleucine residues generated a nascent peptide more favourable for the formation of an α -helix within the ribosome tunnel, SDM within the first TM domain was carried out to substitute leucine 27,31,36 and 40 to isoleucine residues (GPR35 Δ 4I). Although leucines have a greater helical propensity than isoleucines (Pace & Scholtz 1998), the change was made on the basis that they are isomers and only differ in structure due to the position of their methyl groups. Upon creation of the new construct, a pegylation assay was carried out as before in the WG extract translation system. Intermediates increasing in size by 5aa from 25-50aa were generated and the pegylation results were analysed by gel-shift on an SDS-PAGE gel.

As with previous assays, the 25aa intermediate was used as the negative control for pegylation as the marker cysteine would be inaccessible to PEG-MAL regardless of compaction in the tunnel. From 30aa onwards, the pegylation profile of Δ 4I was shown to occur in a similar manner as the WT peptide (Figure 3.8). Pegylation beginning when C15 was 30aa from the PTC suggests that once again we have a TM domain that lacks secondary structure and the substitutions from Leu to Ile had no effect on the ability of TM1 to fold whilst traversing the ribosome tunnel. Although many domains of large TM proteins have been shown to fold whilst making their way through the ribosome tunnel (Table 3.1), the results throughout the chapter consistently suggest this is not the case for GPR35.

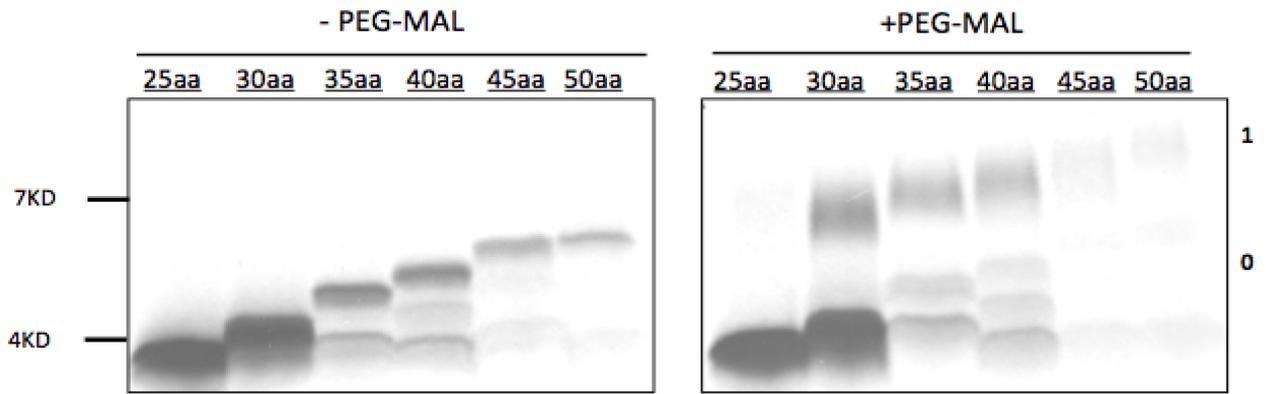


Figure 3.8 Analysis of the pegylation assay of GPR35 Δ 4I in the eukaryotic translation system. Intermediate lengths of GPR35 Δ 4I were expressed in the *in vitro* eukaryotic WG translation reaction. The reaction was split into two, with one half being incubated with 1mM PEG-MAL and the other half incubated in buffer as a control. Representative gel displays how the pegylated (1) and non-pegylated (0) samples were resolved by SDS-PAGE, measuring if pegylation had taken place by a gel-shift of the translation product by \sim 5kDa.

3.3. Discussion

The formation of secondary structure within TM domains of integral membrane proteins has been shown to occur in several other studies (Kosolapov & Deutsch 2003; Woolhead et al. 2004; Lu & Deutsch 2005b). However, until now, there have been no studies into the early folding events, which occur while the nascent chain makes its way through the ribosome tunnel, of the largest class of eukaryotic integral membrane proteins; GPCRs. In the mature protein of both GPR35 and BO, the first TM domain contains secondary structure in the form of an α -helix. As this is the case, it is possible that secondary structure which makes up the mature protein structure begins to form within the ribosome tunnel as witnessed in other peptides (Woolhead et al. 2004; Lu & Deutsch 2005b; Robinson et al. 2012). Previous research carried out in proteins with a similar TM domain secondary structure suggests that the first TM domain of the chosen model proteins could be likely to form secondary structure within the exit tunnel of the ribosome. Previously, Lu and Deutsch (2005a) used cysteine modification by PEG-MAL to detect conformational changes in the domains of a K⁺ channel. In this chapter, we used a similar pegylation assay investigate the compaction within TM1 of GPR35 and BO as it traversed the ribosome exit tunnel.

The pegylation results in chapter 3.2.1 reveal that the first TM domain in both BO and GPR35 show no sign of forming structure in the prokaryotic ribosome tunnel. Two intermediate lengths placing the marker cysteine at 25aa and 50aa acids from the PTC provided us with a negative and positive control respectively. In a fully extended form of both GPR35 and BO, the marker cysteine when 25aa from the PTC would be expected to be buried deep within the ribosome tunnel and hence inaccessible to becoming mass-tagged by the PEG-MAL molecule. If compaction of the TM domains were to occur, the intermediate presenting the marker cysteine 50aa from the PTC should provide enough length to place C15 out with the ribosome tunnel and fully accessible to PEG-MAL (approximately 80% pegylation is has been shown to be maximal (Kosolapov et al. 2004)). The 25aa intermediate of both model proteins fails to pegylate in either the eukaryotic or prokaryotic expression system, thus proving the cysteine was indeed located within the ribosome tunnel and inaccessible to PEG-MAL (Figure 3.3 -3.6) (BO intermediates of 25aa and 30aa lengths, however, failed to express in the S-30 system, but did express in the eukaryotic WG system). As the nascent chain was extended by 5aa, placing C15 30aa from the PTC, pegylation of the translated product could first be recognised at relatively low levels (~30% in comparison to the 50aa construct which was ~80%) (Figure 3.3 and 3.4).

At this point, the first TM of both GPR35 and BO would be passing the constriction point or residing in the upper regions of the lower tunnel suggesting compaction of the nascent chain in the upper tunnel, as witnessed previously by Woolhead et al (2004), had not occurred. The low efficiency of pegylation was believed to be due to the marker cysteine residing in the distal regions of the ribosome tunnel, hence making the interaction between the SH group of the cysteine and the PEG-MAL molecule difficult. Although a molecule of PEG-MAL has a radius of $\sim 15\text{\AA}$ and could theoretically protrude into the vestibule of the ribosome, the 8\AA backbone of the nascent chain makes this increasingly difficult, hence the low efficiency of pegylation. An increase in distance of the marker cysteine from the PTC by 5aa sees an instant increase in pegylation to $\sim 50\%$, suggesting steric exclusion from the tunnel may reflect the changes in pegylation (Kosolapov & Deutsch 2003). As the marker cysteine is moved 5aa further from the PTC, pegylation rates increase to $\sim 75\%$, levels matching the maximal rate of pegylation shown to occur in the 50aa intermediate (Kosolapov et al. 2004; Lu & Deutsch 2005a). At intermediate lengths of 35aa and 40aa, the TM domain would be residing in the lower and distal regions of the tunnel respectively. If compaction of the TM domain had have occurred at this point, as has been discussed in a number of studies (Lu and Deutsch 2005a; Bhushan et al 2010), we would have expected to see little to no pegylation of the nascent chain. Hence, we suggest that until the point at which the N-terminus of GPR35 TM1 reaches the lower tunnel, secondary structure is not present.

With the first TM domains of both GPR35 and BO producing pegylation results indicative of an extended peptide, something that is often not seen in large TM proteins (Tu & Deutsch 2010), a control membrane protein F_0c known to fold in the ribosome tunnel was used to ensure pegylation was a valid assay to determine secondary structure within the ribosome. The F_0c results confirmed previous studies by Robinson et al (2012), with comparative pegylation results to other large TM domains shown to compact within the ribosome tunnel (Kosolapov et al. 2004; Tu & Deutsch 2010). Unlike GPR35 and BO intermediates, at 35aa from the PTC there is no PEG-MAL shift in the translation product (Figure 3.7). Not until 45aa from the PTC do we see pegylation of nascent chain at rates comparable to the GPR35 and BO 35aa constructs of $\sim 40\%$ and also in agreement with previous CL assays carried out by Robinson et al (2012). Maximal pegylation at $\sim 80\%$ of the F_0c intermediates occurs with the marker cysteine 50aa from the PTC highlighting the difference between an extended peptide and a peptide containing a TM domain capable of folding in the ribosome tunnel.

What is clearly stated in many studies involving co-translationally folded peptides is the numerous factors believed to be involved to enable nascent peptides to fold in the ribosome exit tunnel; none more important than the roles played by the ribosome tunnel and nascent chain itself. The molecular environment that makes up the ribosome exit tunnel (rRNA, ribosomal proteins, water and ions) is specialised for housing the nascent chain and in doing so executing functions (such as peptide elongation, folding, nascent chain recognition and targeting) essential in the biogenesis of proteins. The nascent chain is also specialised, containing specific sequences whose precise order acts like a barcode, transmitting important information that enables the correct biogenesis of the mature protein (Gong & Yanofsky 2002; Cruz-Vera et al. 2011; Lin et al. 2011). With these two important factors in mind we explored the critical role two different populations of ribosome have on the folding profile of GPR35 and BO, analysing whether ribosome environment or nascent chain sequence plays a more significant role in the formation of secondary structure.

Comparisons between the prokaryotic and eukaryotic ribosome shows much evolutionary conservation, as well as numerous areas with not such a high degree of conservation (Ban et al 2000, Yuspov et al). The eukaryotic ribosome is believed to be approximately 40% bigger than that of the prokaryotic ribosome, with its core structure made up of eukaryotic specific proteins (Ben-Shem et al 2011). Although this is the case, the ribosome tunnel, a region essential for the formation of secondary structure is believed to be highly conserved in terms of size and make-up (Ban et al 2000, Nissan et al 2000, Ben-Shem et al 2011). Indeed, multiple studies have shown that folding within the nascent chain can occur in both subsets of ribosome (Woolhead et al 2004, Robinson et al 2010). This conservation between the two tunnel environments may go some way to explaining the results seen in section 3.2.2. determining that both GPR35 and BO intermediates exist as extended nascent chains, whilst ribosome bound in the prokaryotic translation system and show an identical folding profile in the eukaryotic WG system (Figure 3.5 and 3.6). Therefore, even though the prokaryotic ribosome tunnel was not viewed as a native environment for the expression of a eukaryotic GPCR and hence postulated to be less favourable for the formation of structure within the nascent peptide, this seems not to be the case. With this in mind, it could be suggested that a nascent chain seen to fold in the prokaryotic ribosome should be fully capable of interacting in the same way within the eukaryotic ribosome, and vice-versa.

As we have seen from the results in section 3.2.2. the differences between prokaryotic and eukaryotic ribosomes have little impact on the formation of secondary structure within

TM1 of GPR35 and BO. With this being the case, it would seem only natural to suggest that the sequence within the nascent chain of the two model proteins plays a substantial role in how they are structured within the ribosome. Reviewing numerous domains capable of folding in the ribosome tunnel led us to a wide variety of proteins; signal sequences (Peterson et al 2010), TM domains (signal anchors) (Kosolapov et al 2004, Woolhead et al 2004, Lu and Deutsch 2005, Robinson et al 2010, Lin et al 2011) and stalling peptides (Woolhead et al 2006, Onoue 2011). Analysis of these sequences showed neither sequence complementarity, nor consistency in size, but did show consistency with high levels of hydrophobicity and residues of a high α -helical propensity, two important factors (Pace and Scholtz 1998). The ability of signal sequences and TM signal anchors to form α -helical structures has also been suggested as a possible factor when targeting proteins to their desired organelle. Specifically, SRP-dependant targeting is believed to require a short region of α -helical peptide (~10aa in length) to bind and target nascent peptides to the ER membrane. This suggests forming secondary structure within the ribosome tunnel may be a necessity in the case of some proteins.

When comparing peptides that were 'known folders' in the ribosome tunnel, an observation was made concerning the lack of isoleucines in the two model proteins; GPR35 and BO (Table 3.1). Isoleucine has a helical propensity of 0.41 kcal/mol (Pace and Scholtz 1998) and although it does not have the greatest propensity to form a helix (eg. Ala 0, Leu 0.21 kcal/mol), it is possible that interactions between the aliphatic side chains of the residue and the ribosome tunnel may be favourable in the formation of structure within a nascent chain. With little else to go on, it was decided to introduce 4 isoleucine residues in place of leucines, generating the GPR35 4 Δ I construct. As before, a pegylation assay was carried out on 25-50aa intermediates of 4 Δ I, assessing if the change in the substitutions of isoleucine residues had any effect on the folding of TM1. Results suggested that it did not, with a folding profile very similar to GPR35 WT seen once again (Figure 3.5). This highlights the complexity within the sequence of the nascent chain and suggests it is very unlikely that single amino acids are essential for the formation of secondary structure within the ribosome. In fact Pace and Scholtz (1998) suggest that conformational entropy of residues within a peptide is key for formation of helical structures, alluding to the fact that a stretch of amino acids favouring a helical formation may be more essential than just one. Alongside this, other factors such as hydrophobicity, environment and side-chain interactions all may impact on how structure forms within the nascent chain (Hermans et al 1992, Blaber et al 1994, Wang and Purisma 1996). Together, the studies in section 3.2.2

and 3.2.4 reinforce the idea that if the sequence within the nascent chain does not code for the formation secondary structure within the ribosome tunnel, then it will not be present.

In summary, the work carried out in this section has showed that the first TM domain of two large TM proteins, GPR35 and BO, exists in an extended conformation as they make their way through the ribosome tunnel. Using the biochemical method of pegylation, both model proteins were shown to take up an extended conformation in the prokaryotic and eukaryotic ribosome tunnel, until the point at which the N-terminus of TM1 is in the lower tunnel. Using two different species of ribosome enabled us to determine that both ribosome environment and nascent chain sequence are essential for generating secondary structure within the TM domains of an IMP.

4. The importance of secondary structure for SRP-mediated targeting?

4.1. Introduction

Approximately 30 % of all proteins synthesised are IMPs, which are destined to be transported to either the eukaryotic ER or bacterial plasma membrane (PM). Although these IMPs may be structurally diverse, they share a common feature; their hydrophobic TM domains, structures that need to reside within the lipid bilayer (White and von Heijne 2005). This poses a problem for IMPs as insertion into biological membranes requires transport from the aqueous cytosol where they are synthesized, to an environment where they are energetically stable. One major route of IMP targeting is carried out by a specialized system known as the signal recognition particle (SRP) pathway that has been shown to play a vital role in the protection, targeting and integration of TM domains into the lipid bilayer (Bibi 2011).

The SRP pathway is made up from a number of distinct steps (as previously described in section 1.2 in more detail) (Wild et al. 2004; Egea et al. 2005) (Figure 4.1). In the eukaryotic pathway, the first step begins as the SRP docks to the ribosome and binds the nascent chain, creating a SRP-Ribosome nascent chain complex (RNC), in turn inducing a translation arrest. Following this, the complex is targeted to the ER membrane, where GTP binding to SRP triggers an interaction between itself and the SRP receptor (SR), creating the SRP-SR complex. At this point, the RNC can then be transferred to the Sec61 translocon, where translation process is re-engaged and translocation can begin. The final step in the SRP pathway sees dissociation of SRP from the SR via GTP hydrolysis, recycling SRP back into the cytosol (Figure 4.1)

The initial stages of SRP-mediated targeting are believed to begin in a co-translational manner whilst the nascent chain makes its way through the ribosome exit tunnel. The exit tunnel, once believed to be simply a passive water-filled channel, is now known to be key in many co-translational interactions involving the translating peptide (Woolhead et al. 2004; Lu et al. 2007). In the initial stages of SRP mediated targeting, the tunnel is believed to have a sensory role. Studies have shown as the nascent chain nears the exit of the ribosome, the ribosomal protein L39e in eukaryotes and the tunnel loop domain of uL23 in the prokaryotic ribosome interact specifically with hydrophobic membrane segments

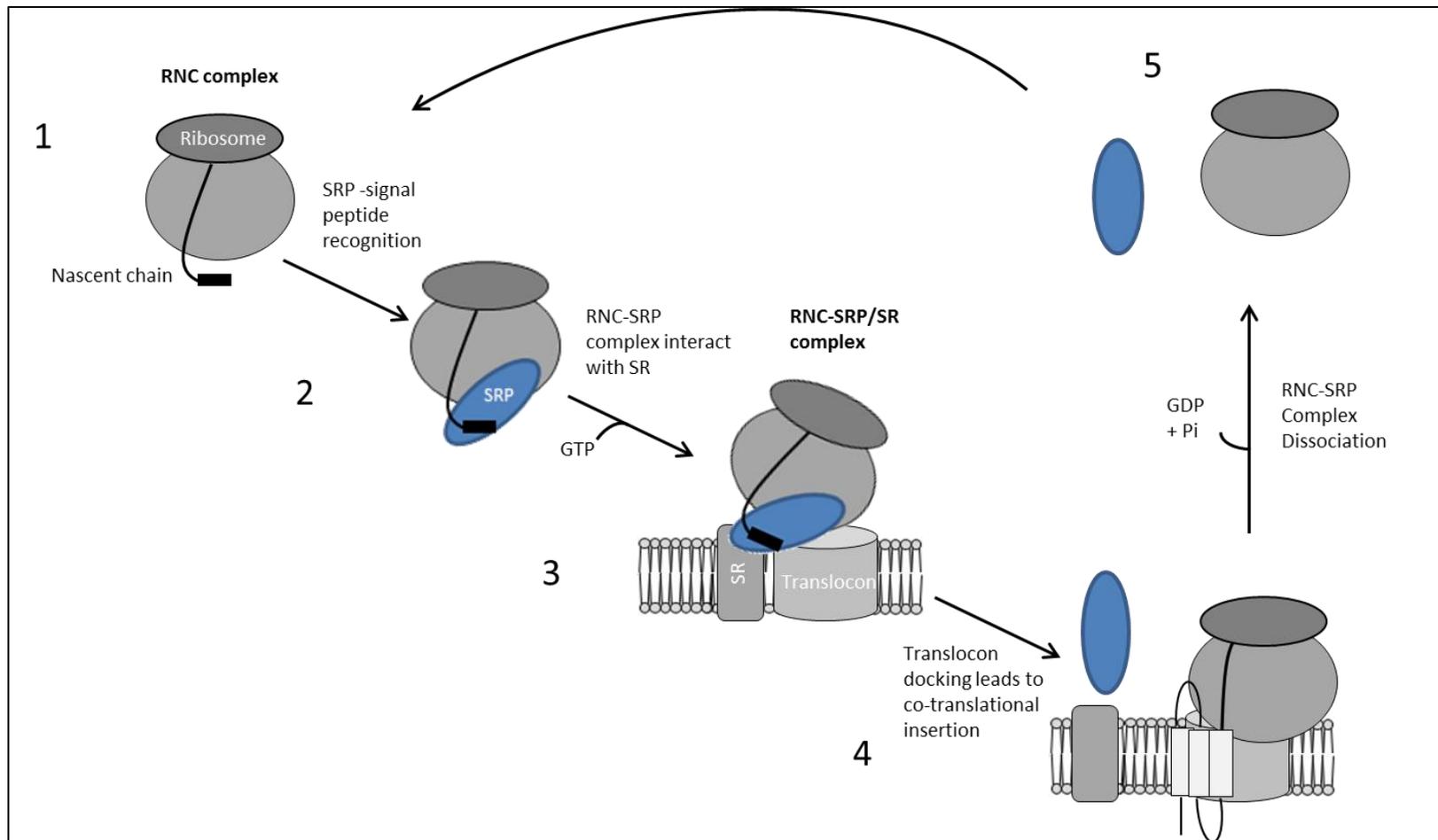


Figure 4.1 Schematic diagram describing how integral membrane proteins are co-translationally targeted to the membrane. (1) Protein transport begins as nascent chain containing a signal peptide emerges from the ribosome. (2) The signal peptide is recognised by SRP and the SRP/ribosome-nascent-chain complex is formed. (3) The SRP/ribosome nascent chain complex then binds to the SRP receptor (SR). For the formation of a stable SRP/SR complex, GTP has to be present. (4) The SRP/RNC complex docks to the translocon, where the signal peptide is transferred from SRP to the translocation channel. (5) GTP hydrolysis in both, SRP and SR, leads to the dissociation of the SRP/SR complex.

(Woolhead et al. 2004; Robinson et al. 2012). These interactions within the tunnel of the ribosome are believed to begin SRP recruitment to the exit site, positioning it for the capture and protection of the emerging TM domain. Experiments have shown that translation of a nascent chain drives conformational changes within the large subunit of the ribosome and increases the affinity for an interaction with SRP (Flanagan et al. 2003; Bornemann et al. 2008). Specifically, a hydrophobic segment within a nascent peptide, known as the signal sequence enables the translating ribosome to enter the SRP pathway. These hydrophobic stretches of amino acids usually span between 10-20 residues and often make up cleavable signal sequences or signal anchor domains.

SRP interaction with the ribosome begins as the nascent chain makes its way through the tunnel and nears the exit point. SRP, a large ribonucleoprotein, composed of RNA and protein (6 protein subunits in eukaryotic SRP and one in bacteria) has a high affinity for the translating ribosome. It binds uL23 and uL29 through its 54 kD subunit (SRP54 in eukaryotes and Ffh in bacteria) near the exit of the tunnel, in doing so placing the M domain in a position to capture the exiting nascent chain. The M domain is made up of four central helices, ordered around a central hydrophobic core which houses the hydrophobic stretch of the emerging nascent chain (Keenan et al. 1998; Janda et al. 2010; Hainzl et al. 2011)

With a high level of diversity in length, shape and sequence within signal sequences (Zheng & Gierasch 1996), early studies showed that peptides containing a hydrophobic domain with a high helical propensity were the most likely candidates for SRP recognition (von Heijne 1985; Hatsuzawa et al. 1997). Recent work, however, suggests that although important, hydrophobicity alone may not be enough to determine an SRP interaction or indicate if a peptide is targeted by the SRP pathway. In fact, work carried out by Huber et al (2005) calculated the hydrophobicity of signal sequences required for SRP-dependant and SRP-independent targeting pathways were similar and found no way of determining through hydrophobicity alone whether an SRP interaction was capable. Evidence also suggested that sequences containing hydrophobicity above a certain threshold had an adverse effect on SRP binding, suggesting the possibility of additional factors involved in SRP targeting (Huber et al. 2005). Indeed, other features have been postulated to be required to enhance SRP-nascent chain interaction such as helical propensity, the presence of secondary structure (Tu et al. 2000), the presence of basic residues at the N-terminus of a signal sequence (Peterson et al. 2010), as well as additional factors that have not yet been determined.

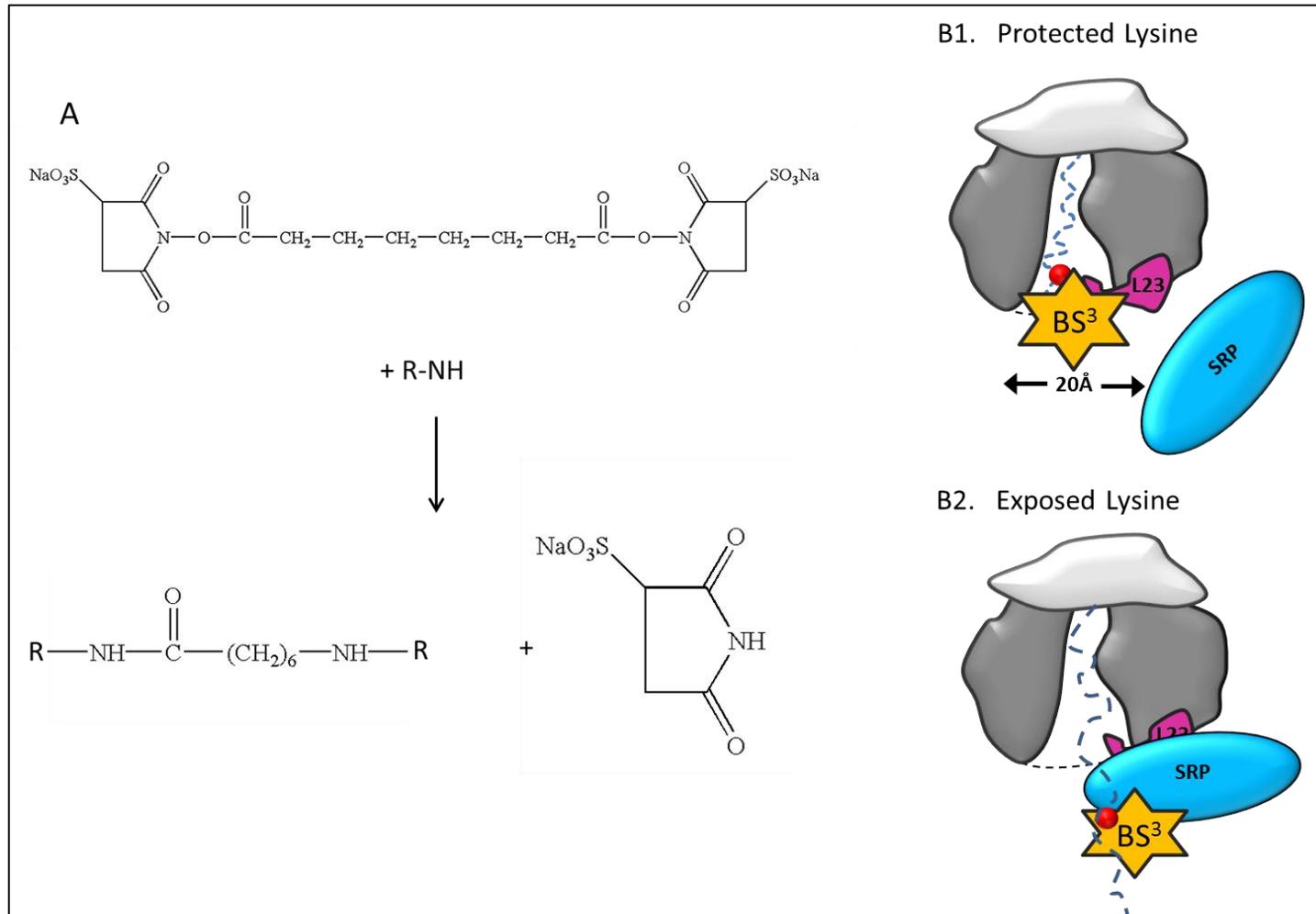


Figure 4.2 Cross-linking assay to measure the level of interaction with components of the SRP targeting pathway. (A) Chemical cross-linking BS³ molecule reacts with the amino group of the single lysine residue in the nascent chain and the free amino groups in SRP targeting components. The succinimidyl groups at either end of the BS³ molecule are cleaved and two proteins are crosslinked by the 6 carbon fatty acid chain (Adapted from Sarngardharan et al 2003). (B1) Schematic diagram representing the protection offered to the single lysine (red circle) in the nascent chain from a molecule of BS³ by the ribosome tunnel and (B2) the cross-linking interaction between BS³, SRP and the lysine when exposed from the ribosome exit

Alongside hydrophobicity, a feature that has often been remarked upon as enhancing the affinity for a SRP-nascent chain interaction is the presence of an α -helix within the signal sequence. Crystal structures of the M domain within SRP54 show the presence of a hydrophobic groove thought to accommodate mainly α -helical signal sequences (Keenan et al. 1998). With a number of studies now showing the formation of α -helices occurring within hydrophobic stretches that make up TM domains of IMPs, whilst in the ribosome tunnel (Kosolapov et al. 2004; Lu & Deutsch 2005; Robinson et al. 2012), it is surely a possibility that this feature could increase the selectivity for the SRP-dependant pathway (Tu et al. 2000; Berndt et al. 2009).

Although the final destination for GPR35 is the PM, not much work has been carried out to understand its targeting to the ER membrane. To assess if the GPR35 nascent chain is capable of using the SRP-dependant pathway and if so what features within the nascent chain are required, the method of chemical cross-linking will be utilized. Chemical cross-linking, as well as photo cross-linking have been used by a number of groups to show interactions between the nascent chain and components required to initiate or carry out SRP mediated targeting (Houben et al. 2005; Robinson et al. 2012; Nilsson et al. 2014). In particular, I will be following a similar chemical cross-linking method to that carried out by Robinson et al (2012), attempting to determine if an interaction can take place between Ffh/SRP54 protein (prokaryotic/eukaryotic SRP) and the ribosomal protein uL23. The chemical bis(sulfosuccinimidyl)suberate (BS^3) will be used to investigate the proximity between the nascent chain and selected proteins. The assay itself relies on two ester-reactive sulfo-NHS groups at either end of the BS^3 molecule reacting with proteins containing primary amines. When proteins are in close enough proximity, a bond is formed between the succinimdy groups of BS^3 and free amines of lysine residues within the nascent chain (Figure 4.2A). BS^3 has a diameter of $\sim 20\text{\AA}$, making it incapable of entering far into the ribosome tunnel. Therefore, as explained for the pegylation assay, the tunnel offers protection to the nascent chain from BS^3 (Figure 4.1 B1). Only once the lysine residues within the nascent chain reach the distal regions of the tunnel can BS^3 generate cross-links with proteins in close enough proximity (Figure 4.1 B2) (Robinson et al. 2012).

Therefore, the overall aim of this chapter will be to identify, via cross-linking assays, if interactions occur between components of the SRP mediated pathway and the translating GPR35 nascent chain. Specifically, we aim to investigate the importance of features such as secondary structure and hydrophobicity within the 1st TM domain and how they impact of SRP binding.

4.2.1 Is the GPR35 nascent chain capable of interacting with SRP?

The folding profile of GPR35 was assessed by pegylation assays in chapter 3 and showed the nascent chain to exist in a linear conformation as it made its way through the ribosome tunnel. As this is the case, experiments were carried out to investigate how the lack of secondary structure within the first TM domain of GPR35 impacted on targeting the ER membrane. As seen with many membrane proteins, the SRP pathway is the major route of targeting to the membrane and we would hypothesize that it would be the most likely method of targeting GPR35. To investigate if this is the case and whether the lack of secondary structure within the first TM domain has an effect, cross-linking assays were set up to analyse the interactions between the nascent chain and components of the SRP pathway.

A construct capable of generating cross-links with BS³ was required to carry out this cross-linking assay. GPR35 contains a single lysine residue 5aa upstream of the 1st TM domain in GPR35 (K20- marker lysine) (Figure 4.3A). This lysine provided a free amino group within the nascent chain for the formation of cross-links between itself, BS³, and other proteins. K20 is the only lysine present in the N-terminal sequence of GPR35, therefore it will also be capable of indicating the length of the nascent chain required to span the ribosome tunnel before cross-linking takes place. Therefore, as in chapter 3, intermediate lengths of the nascent chain used in the cross-linking assay will be named from the peptidyl-transferase centre (PTC) to the marker lysine (K20).

Initially, cross-linking experiments were set up in the prokaryotic S-30 *in vitro* transcription/translation expression system, as antibodies for the detection of SRP targeting components uL23 and Ffh (bacterial SRP54) were available (western blots displaying the specificity of respective antibodies are displayed in Appendix 4: Figure 1A and B). As the bacterial and eukaryotic SRP pathways share common ancestry we were confident that we could reliably show an interaction between GPR35 and the components of a prokaryotic system. Radiolabelled RNCs were once again synthesized from linear DNA lacking a termination codon and producing stable nascent chains of predetermined length. The RNC intermediates ranged from 25-65aa (from the PTC to K20), increasing at intervals of 10aa..

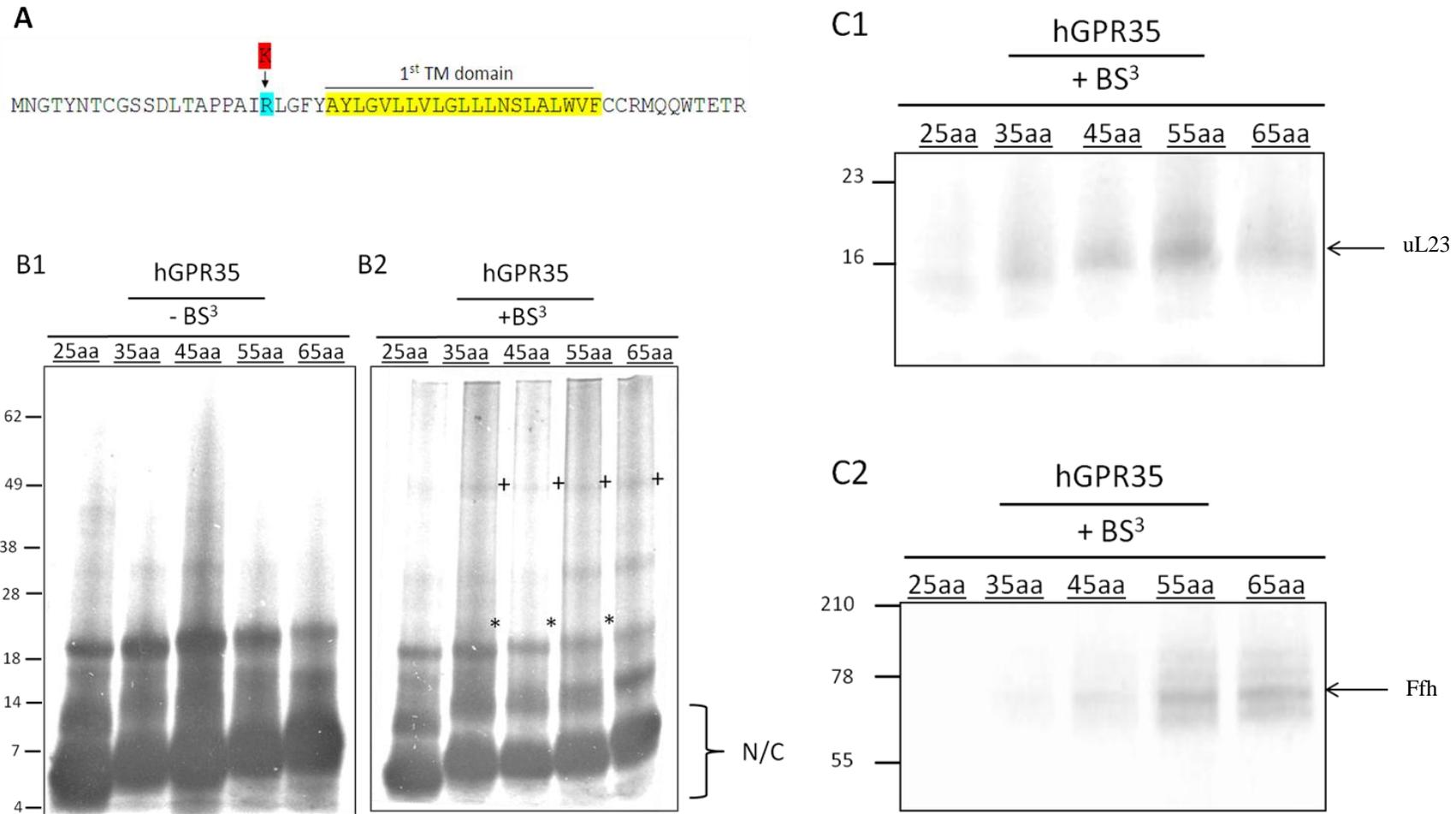


Figure 4.3 Cross-linking assay of GPR35 to uL23 and Ffh in the prokaryotic S-30 transcription/translation system. (A) Schematic diagram of the GPR35 gene highlighting the substitution mutation made by SDM to place a single lysine (marker cysteine) at position 20 for cross-linking to BS3. (B) Ribosomes displaying radiolabelled N-terminal intermediates between 25-65aa of GPR35 at intervals of 10 amino acids can be seen before (B1) and after (B2) treatment with BS3 when resolved on a 10% tricine gel. Crosslinks displaying the correct molecular weight for GPR35 translation products, GPR35-uL23 and GPR35-Ffh complexes have been marked N/C, * and + respectively. (C) Immunoprecipitation of cross-linking assay of GPR35 to components of the SRP pathway (C1) Products shown in the cross-linking assay of GPR35 were subjected to immunoprecipitation with anti-uL23 serum. (C2) Products shown in the cross-linking assay of GPR35 were subjected to immunoprecipitation with anti-Ffh serum.

This range allowed us to explore the interactions taking place both during and after TM1 has made its way through the ribosome tunnel. The results of the GPR35 cross-linking assay can be seen in Figure 4.3B1 and 4.3B2. In the absence of the cross-linker molecule BS³, the translated nascent peptides can be seen to steadily increase as they range from 25-65aa (Figure 4.3 B1). Also present are weak background bands believed to result from endogenous DNA/RNA in S-30 extract, which can be detected in the absence of input DNA. In the presence of BS³ the translated nascent peptides can once again be detected, with the first evidence of cross-linking apparent when the marker lysine is 35aa from the PTC, suggesting the K20 is extremely close to the exit of the ribosome at this point (Figure 4.3B2). This band can be immunoprecipitated with an antibody raised to the uL23 protein and is the first sign of cross-linking to a protein involved in the targeting of GPR35 (Figure 4.3C1). The same band representing cross-links between the nascent chain and uL23, also exists at intermediates of length of 45aa and 55aa, suggesting the nascent chain may reside in the distal regions of the tunnel for an extended period of time. Also between the lengths of 45-55aa, higher cross-link appears at ~55 kDa and can again be detected in the 65aa sample (Figure 4.3 B). These larger cross-links can be immunoprecipitated by the antibodies raised to the Ffh protein (Figure 4.3C2). The results seem to show a sequential interaction for the nascent chain passing from the uL23 to Ffh protein, with cross-links weakening in uL23 and strengthening in Ffh after ~55aa in a 'handover'-like manner of the nascent chain.

To confirm the results of the GPR35 cross-linking assay in the S-30 expression system, the same assay was carried out with Bacterioopsin (BO) as a prokaryotic control protein. The first TM domain of BO was seen to have a similar folding profile to that of GRP35 and therefore we would hypothesize that the cross-linking between the nascent chain, uL23 and Ffh proteins would follow a similar pattern. Unlike GPR35, BO has a signal peptide which we would expect to be the targeting signal to SRP (Figure 4.4A). As this was the case, the single lysine required at the N-terminus for cross-linking with BS³ was placed C-terminally to the signal peptide at residue 20 (K20-marker lysine). As well as this, two lysine residues (K43 and 59) found at the C-terminus of TM1 were mutated to arginines to ensure that no further cross-links could form during the synthesis of later intermediates. As hypothesized, in the presence of BS³, bands representing the nascent chain cross-linked to components of the SRP pathway appeared in an identical pattern to those in GPR35 (Figure 4.4 B1 and B2). Confirmed once again by immunoprecipitation, cross-links between the nascent chain and uL23 could be detected when the marker lysine was between 35 and 55aa from the PTC (Figure 4.4 C1). As with GPR35, immunoprecipitations showed cross-links after 55aa

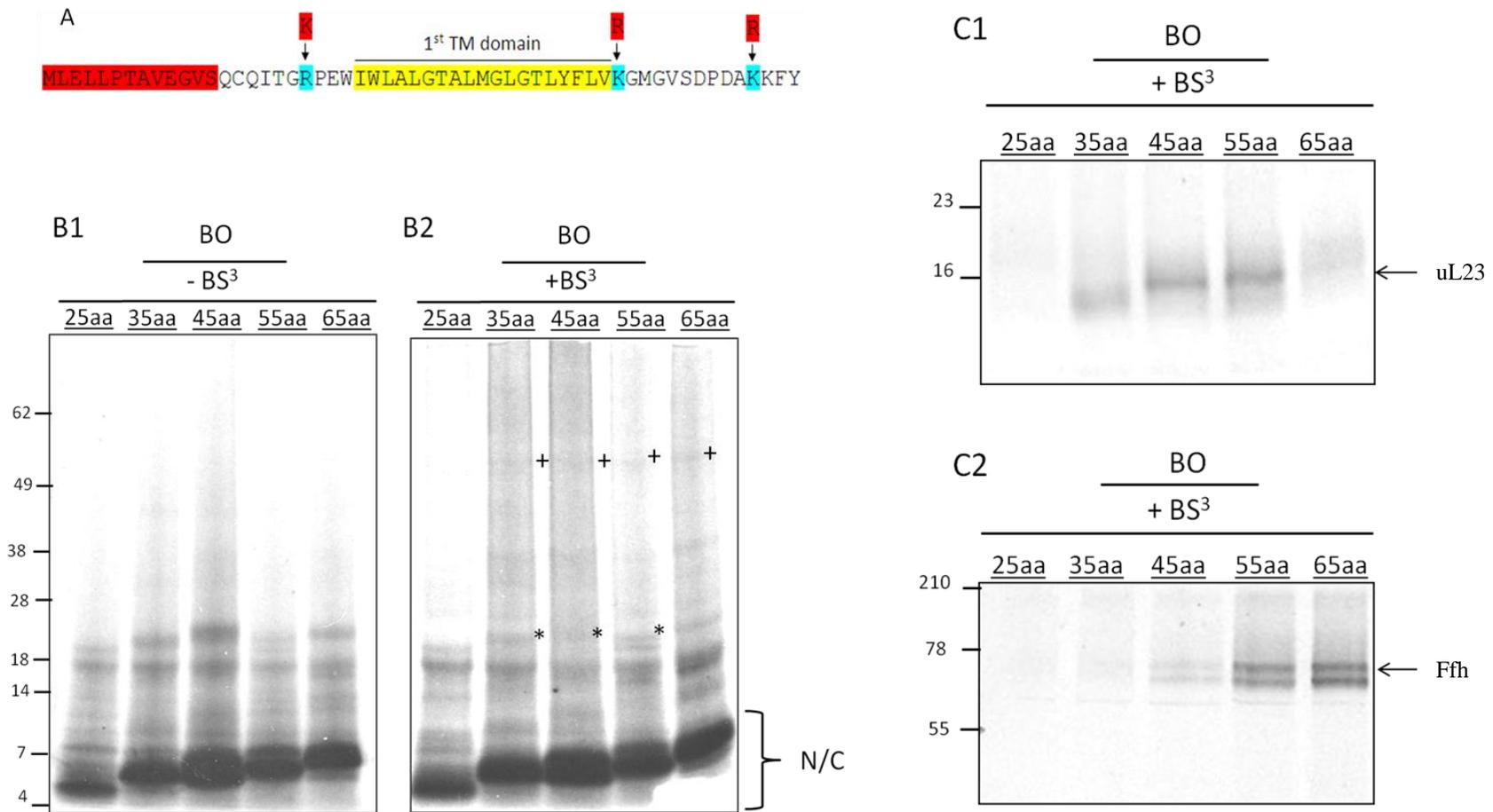


Figure 4.4 Cross-linking assay of BO to uL23 and Ffh in the prokaryotic S-30 transcription/translation system. (A) Schematic diagram of the BO gene highlighting the substitution mutations made by SDM to place a single lysine (marker cysteine) at position 20 for cross-linking to BS3 and removal of two native lysines downstream of the 1st TM domain. The TM domain and signal-peptide of BO are highlighted (yellow and red respectively). (B) Ribosomes displaying radiolabelled N-terminal intermediates between 25-65 amino acids of BO at intervals of 10 amino acids can be seen before (B1) and after (B2) treatment with BS3 when resolved on a 10% tricine gel. Crosslinks displaying the correct molecular weight for GPR35-uL23 and GPR35-Ffh complexes have been marked * and + respectively. (C) Immunoprecipitation of cross-linking assay of BO to components of the SRP pathway (C1) Products shown in the cross-linking assay of BO were subjected to immunoprecipitation with anti-uL23 serum. (C2) Products shown in the cross-linking assay of BO were subjected to immunoprecipitation with anti-Ffh serum.

began to weaken between the nascent chain and uL23, at the same time as an interaction with Ffh was seen to strengthen (Figure 4.4 C2). This again suggests a sequential interaction between the nascent chain of BO and uL23 as it nears the distal regions of the tunnel, followed by an interaction with Ffh as the nascent chain begins to exit the ribosome. The similarity between the results in GPR35 and BO suggest that the GPCR is fully capable of interacting with components of the prokaryotic SRP pathway, indicating that the first TM domain of GPR35 acts as a signal anchor for membrane targeting.

The similarity in the cross-linking patterns shown by GPR35 and BO suggests a high level of homology may exist between the key mechanisms that enable prokaryotic and eukaryotic SRP targeting. In particular, the M domain of Ffh and SRP54, the region essential for signal sequence binding, has been shown through crystallographic studies to be exceptionally well conserved (Keenan et al 1998, Janda et al 2010, Hainzl et al 2011). With this being the case, we hypothesized that similar results would be obtained when the same cross-linking assay is carried out in the eukaryotic Rabbit Reticulocyte Lysate (RRL) expression system (see section 2.5.4). The assay once again relied on intermediates containing a lysine at residue 20 and ranged in length from 25-65aa from the PTC. In the absence of BS³, translation product representative of each intermediate could be detected. Also present in each sample at ~16 kDa was a double band representative of the protein heam (found in all RRL samples isolated by centrifugation through a sucrose cushion) and an unknown band at ~28 kDa (Figure 4.5A1). In the presence of BS³, cross-links between the nascent chain at lengths 55 and 65aa could be detected with a protein of ~50 kDa (Figure 4.5A2). These bands that could be seen in the cross-linked samples and were immunoprecipitated with an antibody raised to SRP54 (Appendix: Figure 1C), indicating an interaction is taking place between the nascent chain and eukaryotic SRP (Figure 4.5 B).

The results from cross-linking assays in both prokaryotic and eukaryotic expression systems show that the nascent chain interacts with the SRP pathway. Results from the assay in the prokaryotic system show the nascent chain cross-linking with uL23, a protein found in the most distal regions of the ribosome tunnel, whilst the marker lysine is 35aa from the PTC. This would suggest the nascent chain is in a linear conformation when the first cross-links were observed and is in agreement with previous results seen in chapter 3. Interactions between the nascent chain and SRP proteins (Ffh and SRP54) occurred later, beginning when the K20 was 45aa from the PTC. A previous study, carried out by Robinson et al (2012), used a similar cross-linking assay with a membrane protein capable of forming secondary structure in the ribosome tunnel. In this study, IPs were also carried

out for the proteins uL23 and Ffh, showing interactions between the two proteins and the nascent chain to occur when the marker lysine was ~ 48 and 59aa from the PTC respectively. This highlights the difference in length required for an extended and compacted nascent chain to interact with SRP, but also indicates that GPR35 is able to interact with SRP regardless of secondary structure formation within the tunnel. The extended times in which the nascent chain interacts with both uL23 and SRP may indicate the first folding events occurring within the 1st TM domain of GPR35.

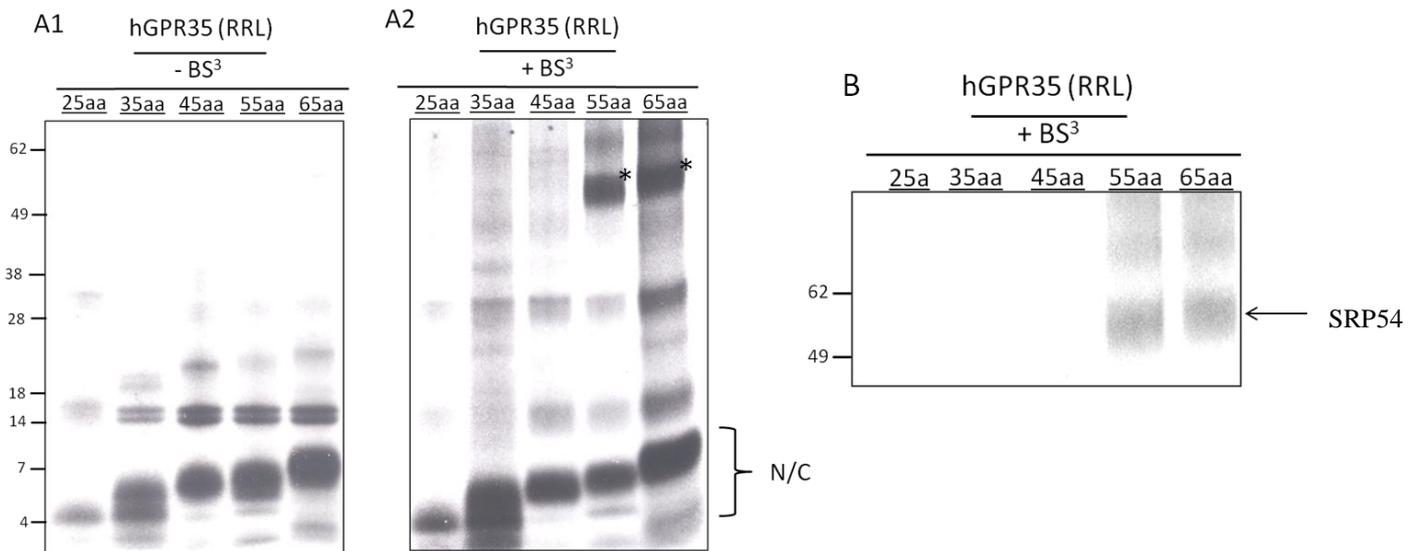


Figure 4.5 Cross-linking assay of GPR35 to SRP54 in rabbit reticulocyte lysate system. Ribosomes displaying radiolabelled N-terminal intermediates between 25-65aa of GPR35 at intervals of 10 amino acids can be seen before (A1) and after (A2) treatment with BS3 when resolved on a 10-20% SDS-PAGE gel. Cross-links displaying the correct molecular weight for GPR35-SRP54 complexes have been marked with a *. (B) Products shown in the cross-linking assay of GPR35 were subjected to immunoprecipitation with anti-SRP54 serum.

4.2.2. Does hydrophobicity drive the interaction between SRP and the first TM of GPR35?

The cross-linking experiments carried out in section 4.2.1 showed that SRP is able to interact with the nascent chain of GPR35, despite the lack of secondary structure within the first TM domain (the predicted signal anchor). Therefore, we set out to determine what drives the interaction between the nascent peptide and the SRP54 protein. The interaction between SRP and its targets is believed to be influenced by a number of factors, one of which was believed to be the helicity of the nascent chain. As we have shown, the nascent chain of GPR35 does not seem to compact whilst inside the ribosome tunnel, yet is still capable of interacting with SRP. Therefore, we seek to investigate another factor previously presented as being a key feature in signal anchors and signal sequences (Keenan et al. 1998), hydrophobicity, and its effect on the interaction between GPR35 and SRP.

To investigate exactly how critical the need for a hydrophobic first TM domain within GPR35 is for targeting via the SRP pathway, we set out to reduce the hydrophobicity of TM1 and analyse the effects this had on SRP cross-linking to the nascent chain. Firstly, we had to generate a construct which we believed could no longer act as a signal anchor domain. As signal anchors are generally found to be integral TM domains of the mature protein, the ‘Dense Alignment Surface’ (DAS) - TM filter server was used to assess the changes required to decrease the hydrophobicity in TM1 (Cserző et al. 1997). The algorithm used in the DAS-TM filter is designed to identify TM domains based on their hydrophobicity and provides a high precision hydrophobic profile of individual domains. Using this tool, the WT GPR35 sequence containing the first 85 residues (first and second TM domain included) was entered and a DAS profile score was obtained (Figure 4.5). The WT GPR35 protein can be seen to contain an extremely hydrophobic first TM domain that well exceeds the 2.2 ‘strict cut-off’ DAS profile score recommended for a TM segment. Therefore, to lower the hydrophobicity of the TM domain and make it unrecognisable as a signal anchor for SRP, we aim to generate a peptide whose hydrophobicity falls below the ‘strict cut-off’ rate.

Lowering the hydrophobicity of TM1 sufficiently and generating a domain unrecognisable as a TM sequence required changes to several of the most hydrophobic residues. Using the

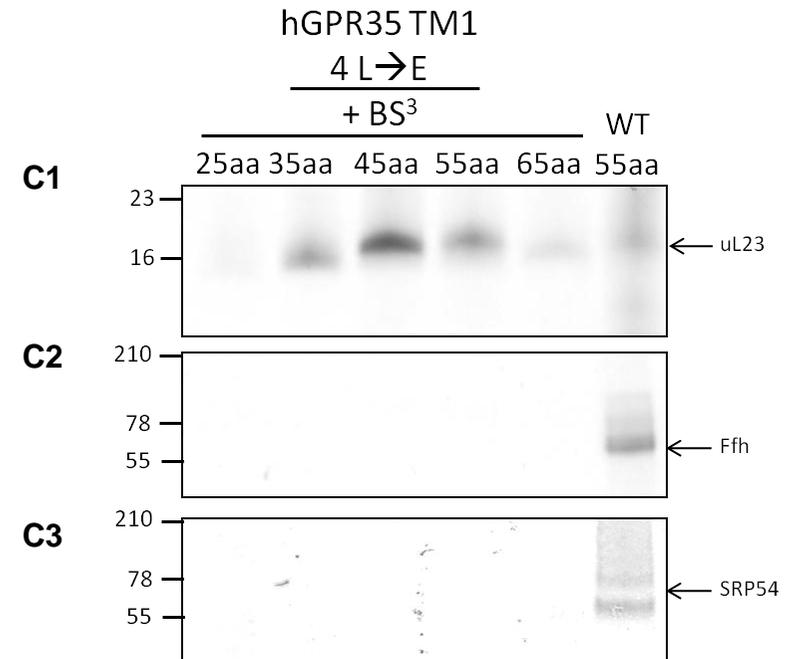
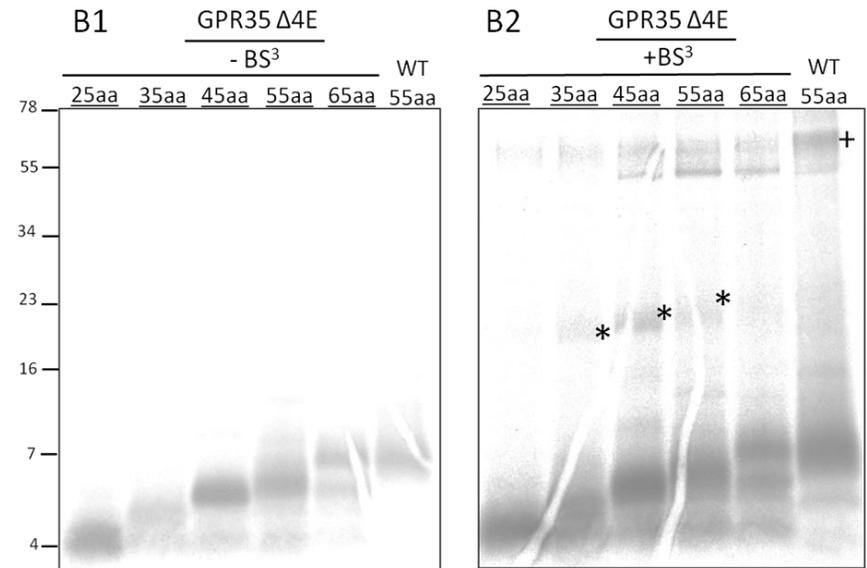
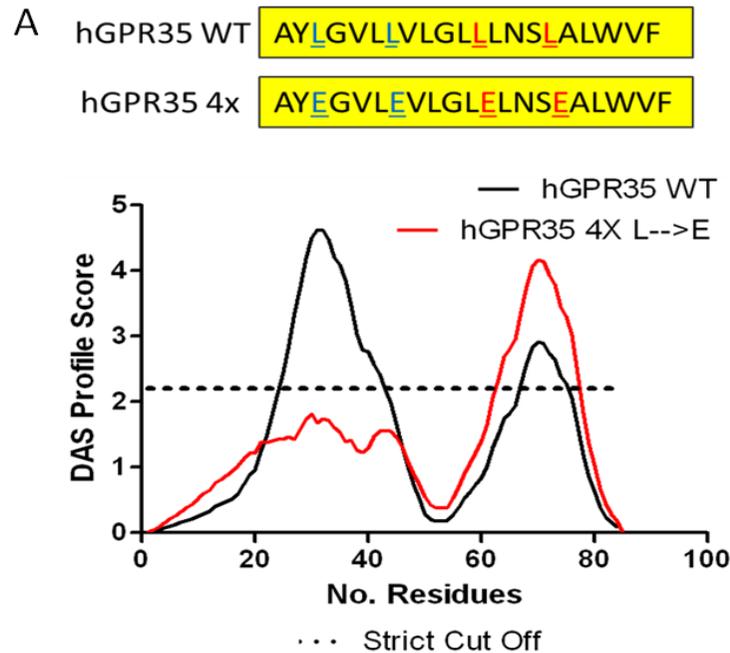


Figure 4.6 Cross-linking assay of GPR35 Δ 4E mutant to uL23 and Ffh in the prokaryotic S-30 transcription/translation system. (A) Site-directed mutagenesis of hydrophobic leucine residues at position 27,31,36 and 40 to glutamic acids with GPR35 reduce the hydrophobicity of the 1st TM domain below the strict cut off point set out by the Dense Alignment Surface (DAS) method indicating the presence of a TM domain (Cserzö et al. 1997). Ribosomes displaying radiolabelled N-terminal intermediates between 25-65aa of TM1 of the Δ 4E mutant at intervals of 10 amino acids can be seen before (B1) and after (B2) treatment with BS3 when resolved on a 10% tricine gel. Crosslinks displaying the correct molecular weight for GPR35 Δ 4E - uL23 and GPR35 Δ 4E - Ffh complexes have been marked * and + respectively. (C) Immunoprecipitation of cross-linking assay of Δ 4E to components of the SRP pathway (C1) Products shown in the cross-linking assay of Δ 4E were subjected to immunoprecipitation with anti-uL23 serum. (C2) Products shown in the cross-linking assay of Δ 4E were subjected to immunoprecipitation with anti-Ffh serum. (C3) Products shown in the cross-linking assay of Δ 4E were subjected to immunoprecipitation with anti-SRP54 serum.

DAS-TM filter, we could postulate the effect substitutions in a number of regions within the TM domain would have on its overall hydrophobicity. Several methods of lowering the hydrophobicity below the 'strict cut-off' point were attempted, with the results often leading to the substitution of over half the residues within the WT TM segment. These consisted of polar residues grouped in threes throughout TM1 or polar residues sporadically placed throughout the entire TM domain, both of which required a high number of changes to have an effect. Finally, it was decided that substituting four Leucine residues for Glutamic Acids (at positions 27, 31, 36 and 40) to create the construct GPR35 Δ 4E, would minimize the number of mutations required to maximize the effect on the overall hydrophobicity, leaving the majority of the TM sequence intact (Figure 4.6A).

The GPR35 Δ 4E construct was used to carry out cross-linking assays to SRP in both prokaryotic and eukaryotic *in vitro* expression systems. The single lysine required for BS³ cross-linking was again located at residue 20 and intermediate lengths ranging from 25-65aa to the PTC were once again generated. Initial assays were carried out in the S-30 expression system, enabling us to use the uL23 protein as a cross-linking control. As uL23 is found in the distal regions of the tunnel, we would expect the nascent peptide to be in close enough proximity for cross-linking to occur regardless of an interaction with TM1. This was indeed the case, as the reactions containing the 35-55aa Δ 4E nascent peptides showed a similar pattern of cross-linking to that seen in the WT GPR35 intermediates, when immunoprecipitated with an antibody raised to uL23 (Figure 4.6 C1). In cross-linked samples containing the 45-65aa Δ 4E intermediates, the ~55 kDa band representative of Ffh that could be detected in the WT nascent peptides, no-longer existed in the mutated constructs suggesting the interaction with SRP was no-longer occurring (Figure 4.6C2). The assay was also carried out using the eukaryotic RRL expression system and cross-links detected by immunoprecipitation with an antibody raised to SRP54. This assay confirmed the effect a loss of hydrophobicity within TM1 had on the cross-linking of GPR35 with SRP, as the cross-links were completely abolished in the 45-65aa samples (Figure 4.6).

The results obtained from the cross-linking assay using the GPR35 Δ 4E construct have clearly indicated the importance of hydrophobicity within the first TM domain of GPR35. Substitution of the four hydrophobic leucine residues to hydrophilic glutamic acids, lowering the overall hydrophobicity within the segment has abolished the interaction between the nascent chain and SRP. The strategically placed mutations, disguising first TM domain's role as a signal anchor, highlight its importance in the SRP targeting process,

as well as providing insights into how the SRP molecule relies on hydrophobicity within a nascent chain to target a peptide to the membrane.

4.2.3 Does SRP bind preferentially to a region of the first TM domain in GPR35?

Cross-linking assays to investigate SRP interaction with the GPR35 first TM domain has shown us that hydrophobicity over secondary structure formation is the critical driving force for the interaction between the SRP and the nascent chain. Mutagenesis to reduce the hydrophobicity in the entire 1st TM domain was enough to abolish the interaction between SRP and the signal anchor of GPR35. Indeed, similar studies in signal peptides are in agreement with results displayed in section 4.2.2 (Janda et al. 2010; Ataide et al. 2011; Nilsson et al. 2015). This being said, the hydrophobic h-domain of a signal peptide spans ~10aa and is often much smaller than the signal anchor domain of an integral membrane protein. Coupling this with studies that predict the hydrophobic groove of the M-domain, in both eukaryotic and prokaryotic SRP, is capable of encapsulating a sequence of ~10aa in length (Janda et al 2010, Keenan et al 1998), we suggest that it is possible that only a portion of the first TM domain may be required for SRP binding and postulate that the N-terminal region would be likely to be more critical as it will interact first.

To test if a region within the signal anchor domain of GPR35 did indeed play a more essential role in the binding of SRP, two further constructs were generated containing leucine to glutamic acid substitutions in either the N- or C-terminus of the 1st TM domain. The construct GRP35 Δ NT contained two leucine to glutamic acid mutations at residues 27 and 31 (Figure 4.7B), whilst the GPR35 Δ CT construct contained two leucine to glutamic acid mutations at residues 36 and 40 (Figure 4.7C). The two constructs once again enabled the analysis of the interaction taking place between the nascent chain and SRP by cross-linking. The 55aa intermediate (K20 is 55aa from the PTC) was selected as the length of nascent chain to carry out this cross-linking assay, as it has been shown to interact with uL23, Ffh and SRP54.

Expression of the Δ NT and Δ CT intermediates, along with the WT and Δ 4E intermediates, and cross-linking with BS³ provided the opportunity to compare the different binding affinities of each intermediate length to SRP. As before, the uL23 protein was used as a cross-linking control in the S-30 expression system and was shown to interact in the same manner with all four intermediates following an immunoprecipitation (Figure 4.8B and C1). However, differences can be noticed between the four intermediates when cross-

linked with both eukaryotic (Figure 4.8A and B) and prokaryotic SRP (Figure 4.8B). Immunoprecipitations were carried out using antibodies raised to either Ffh or SRP54, and were used to detect the efficiency at which each intermediate bound SRP. In the prokaryotic S-30 system, both the Δ NT and Δ CT intermediates form a much weaker interaction with Ffh (Figure 4.8B) and were quantified to bind with lower than 50% affinity than WT GPR35 (Figure 4.8 C2). Although this is the case, both intermediates are capable of rescuing the loss of Ffh binding seen with the Δ 4E intermediate. When the same experiment was carried out in the eukaryotic RRL system, a different effect could be seen when comparing the four intermediates. The WT and Δ 4E intermediates produced near identical results, with the four leucine to glutamic acid mutations generating a complete loss in SRP54 cross-linking to GPR35 intermediates. However, a noticeable difference was detected between the binding of SRP54 to the Δ NT and Δ CT intermediates. The Δ NT intermediate can be seen interact with SRP54 in a manner that is more representative of the Δ 4E intermediate, barely rescuing the SRP interaction at all (Figure 4.8B). The cross-linking efficiency between the nascent chain and SRP54 is reduced to approximately 20% (Figure 4.8 C3), a much greater effect than the one that occurred between the same construct and Ffh. The opposite effect occurs between the Δ CT intermediate and SRP54, with SRP binding approximately 75% of that seen with the WT construct (Figure 4.8B and C3). The results show a major difference between prokaryotic and eukaryotic SRP cross-linking, possibly alluding to a different mechanism of recognition or simply a higher level of complexity behind SRP binding in eukaryotes.

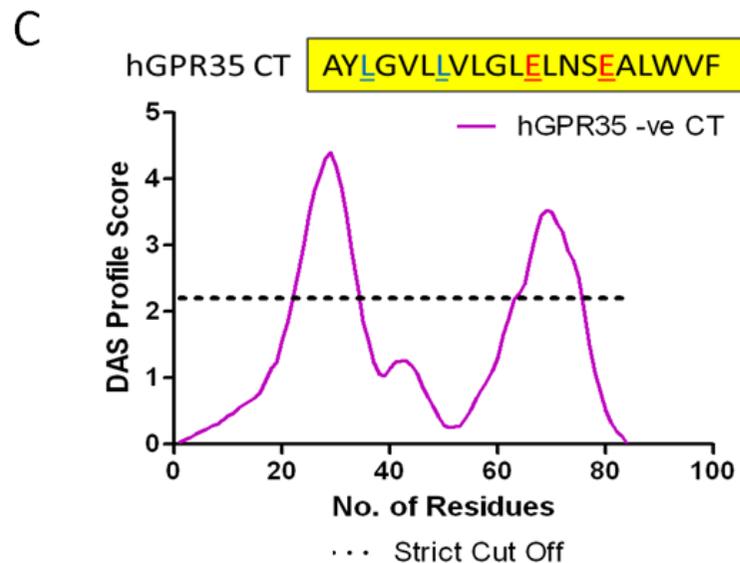
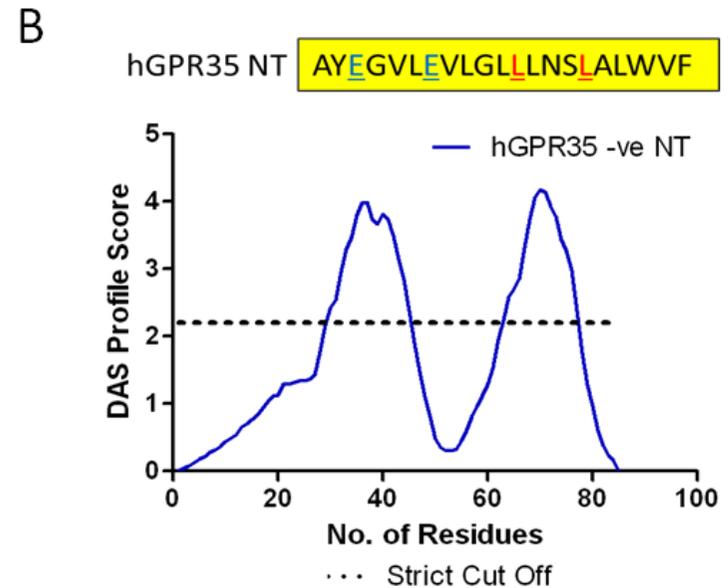
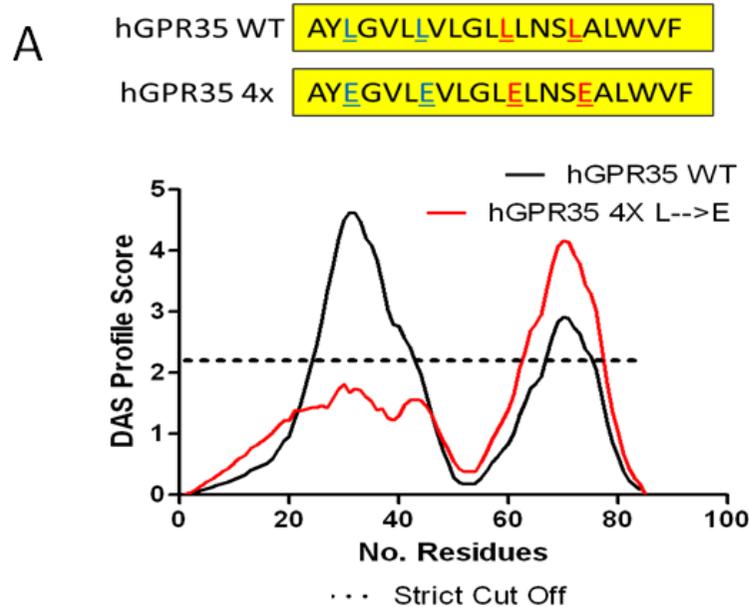


Figure 4.7 Negatively charged amino acids reduce the hydrophobicity in TM domain 1. (A) Site-directed mutagenesis of hydrophobic leucine residues at position 27,31,36 and 40 to glutamic acids with GPR35 reduce the hydrophobicity of the 1st TM domain below the strict cut off point set out by the Dense Alignment Surface (DAS) method indicating the presence of a TM domain (Cerzo et al 1994). (B) Two mutations substituting leucine 27 and 31 to glutamic acid residues provided a construct with a reduction in hydrophobicity within the N-terminal region of TM1. (C) Two mutations substituting leucine 36 and 40 to glutamic acid residues provided a construct with a reduction in hydrophobicity within the C-terminal region of TM1.

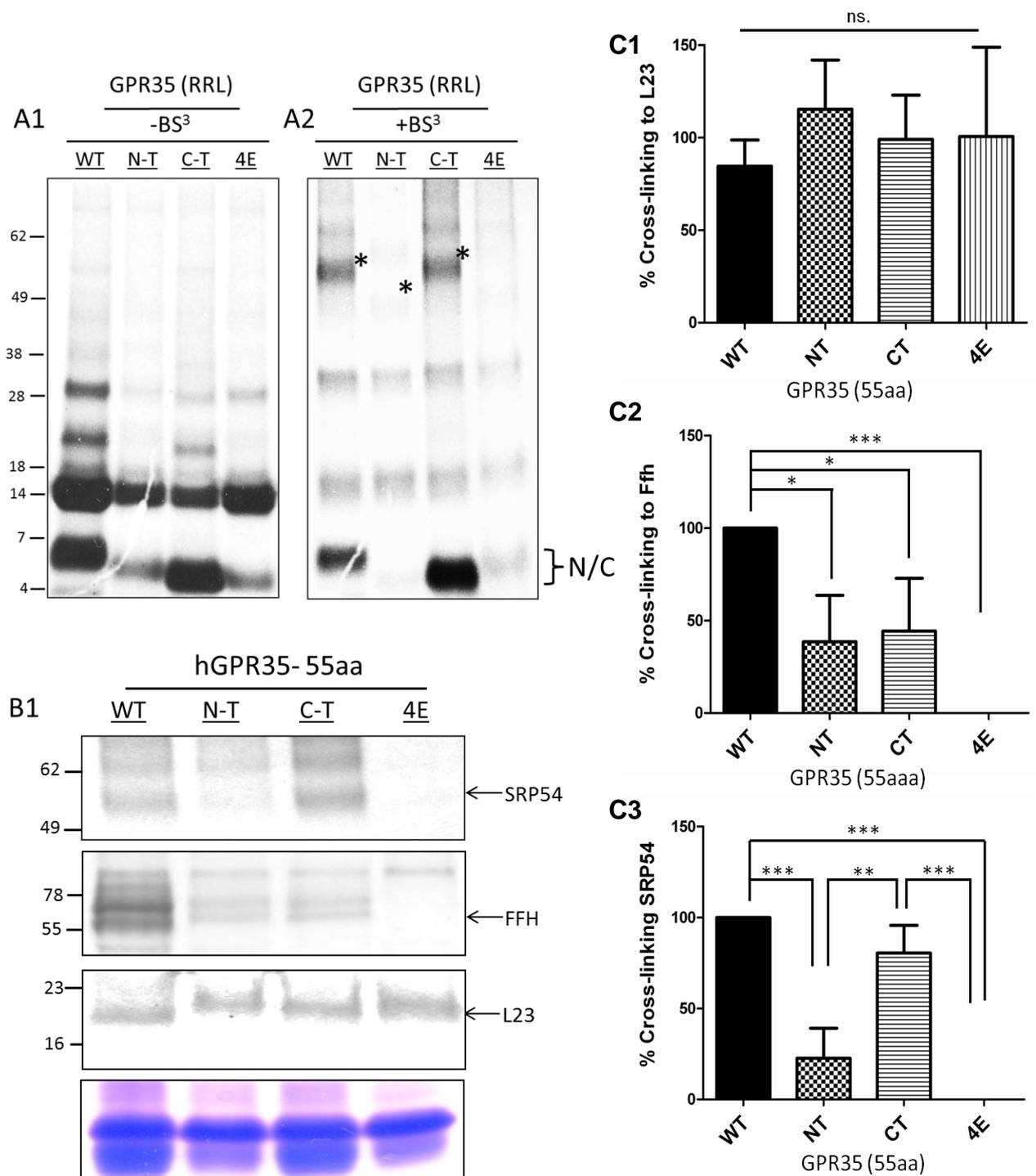


Figure 4.8 Cross-linking assay of GPR35 hydrophobic mutants to SRP54 in rabbit reticulocyte lysate system. Ribosomes displaying radiolabelled N-terminal intermediates of GPR35 hydrophobic mutants at (as described in Figure 4.8) at lengths of 55 amino acids can be seen before (A1) and after (A2) treatment with BS3 when resolved on a 10-20% SDS-PAGE gel. The uncross-linked nascent chain is marked N/C and cross-links displaying the correct molecular weight for GPR35-SRP54 complexes have been marked with a *. (B) Immunoprecipitation of cross-linking assay to components of the SRP pathway (B1) Products shown in the cross-linking assay were subjected to immunoprecipitation with anti-L23 serum, anti-Ffh serum or anti-SRP54 serum. A representative comassie blot acts as a loading control and indicates how IP gels are normalised for subsequent quantification (C) Bar graphs show the average % cross-linking to L23 (C1), Ffh (C2) or SRP54 (C3) for individual intermediates IP results were quantified using Image J software. All % cross-linking values were calculated using [IP product/comassie loading control] and adjusted for background. Average percentage cross-linking is calculated from an n of 3. Error bars indicate standard deviation and groups were compared using one-way ANOVA with Tukey's *post hoc* comparisons. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$.

4.3 Discussion

The targeting of integral membrane proteins (IMPs) to their destinations within the membrane in both eukaryotes and prokaryotes has been extensively studied. The SRP pathway is conserved in both and has been shown to be the mode of targeting to ensure the majority of membrane proteins find their correct location. The SRP pathway relies on a number of key interactions occurring between the nascent chain and the SRP molecule to enable the targeting process to proceed efficiently. SRP recognizes a signal sequence within the nascent chain's N-terminal region to carry out targeting effectively. Signal sequences, however, have a wide range of diversity with little conservation in their sequence. This being said, they do have various features such as charge, hydrophobicity and structure all known to play key roles in targeting of membrane proteins. In this chapter, we investigate the properties of a signal anchor sequence within GPR35 and highlight the key relationships occurring between the nascent chain and components of the SRP pathway.

In vitro cross-linking data confirmed what had previously been suspected for GPR35; like other GPCRs it was capable of interacting with SRP (Friedlander & Blobel 1985; Audigier et al. 1987) and therefore likely to be targeted through the SRP pathway. Intermediate lengths of GPR35 highlighted the sequential events believed to be taking place to trigger the SRP targeting pathway. As described in a number of studies (Ban 2000; Woolhead et al. 2004; Robinson et al. 2012), the nascent chain is believed to be exposed to and possibly interact with a number of sensory ribosomal proteins, uL23 and uL22, as it makes its way through the ribosome tunnel. In prokaryotes, uL23 has been shown to interact with the nascent chain to initiate the SRP targeting pathway. Cross-linking assays of GPR35 intermediates, with the reactive lysine at residue 20, confirm prior structural and biochemical data that positions the uL23 protein in the distal regions of the tunnel. The 25aa intermediate of GPR35, which was previously shown to exist as an extended nascent chain, fails to cross-link with uL23. At this point the reactive lysine residue should be positioned $\sim 30\text{\AA}$ from the exit site and suggests it is still buried too far inside the ribosome to contact BS³ and uL23 simultaneously. As the length progressed to 35aa, the first cross-links between uL23 and the nascent chain appeared, placing the signal anchor domain 30aa from the PTC and close to the bottom of the ribosome. This would suggest a nascent chain length of between 25aa and 30aa from the PTC would be required for the GPR35 signal anchor to come into contact with uL23 and potentially trigger the start of the SRP pathway. At this point, the nascent chain is still believed to be an extended peptide and indeed

similar cross-linking studies can confirm this (Houben et al. 2005; Robinson et al. 2012). Previously, experiments using a TM segment proven to compact within the ribosome tunnel required 48aa to traverse the distance between the PTC and the uL23 protein (Robinson et al. 2012), whilst a non-compacted bacterial membrane protein required ~27aa to reach uL23 (Houben et al. 2005). Therefore, it would appear that the first interactions between GPR35 and components of the SRP pathway occur with the nascent chain in an extended conformation and do not rely on prior structural formation within the ribosome.

The sensing of a nascent chain making its way through the ribosome is an essential process in initiating the binding of SRP to the ribosome. Affinity assays highlight the differences between translating and non-translating ribosomes and have confirmed the detection of the nascent chain by ribosomal proteins is vital for SRP docking (Flanagan et al. 2003). In prokaryotic ribosomes, conformational changes are believed to occur in the globular domain of uL23, driven by a finger loop domain that protrudes into the distal regions of the exit tunnel, therefore increasing the affinity for SRP binding (Gu 2003; Ullers et al. 2003; Bornemann et al. 2008). Bornemann et al. (2008) showed that by mutating the sensory finger loop region of uL23 in the prokaryotic ribosome, a subset of ribosomes could be generated that could no longer recognize cargo for the SRP pathway and failed to recruit SRP to the exit tunnel. The failure of SRP to bind to the ribosome leaves the first TM domain exposed as it exits the tunnel, likely to aggregate as it meets the cytosol. Therefore, events taking place as the nascent chain contacts uL23 and moves towards binding SRP must be co-ordinated, something that can be observed during the cross-linking assay. As the intermediate lengths increase from 55aa, cross-links to uL23 begin to weaken. At the same time, both prokaryotic and eukaryotic SRP begin to show the initial signs of cross-linking to the nascent chain. This seems to represent the nascent chain moving from its position of contact with uL23 at the lower regions of the tunnel, to taking up its position within the hydrophobic groove of the M-domain where TM1 will be protected from the cytosol. Again, when similar studies were carried out using a compacted TM domain in the nascent chain, cross-linking to SRP was not witnessed until the marker lysine was ~63aa from the PTC (Robinson et al. 2012). Once more, this shows a considerable difference in length upon interacting with SRP, suggesting that the GPR35 nascent chain remains extended. This data presents another example of how the nascent chain is sensed by uL23, which then prepares the ribosome environment for the binding of SRP and the capture of the nascent chain.

SRP cross-linking to the GPR35 nascent chain begins at 45aa to 65aa from the PTC. This is likely to represent the initial interaction with an extended GPR35 nascent chain, leading to the capture of the nascent chain by the hydrophobic groove in the M-domain of SRP. Crystallographic studies have shown that as the nascent chain emerges from the ribosome, the hydrophobic groove of SRP encapsulates the signal sequence, protecting it from the cytosolic environment (Keenan et al. 1998). The length of the interaction between the nascent chain and SRP may be an indication that the signal anchor of GPR35 begins to form secondary structure upon binding SRP. Although we can only speculate, compaction on binding of the signal sequence to SRP has been previously documented to occur within TM proteins and signal peptides (Houben et al. 2005; Halic et al. 2006a). Structural data provide dimensions of the hydrophobic groove supporting the hypothesis that compaction may be required to enable SRP to protect the nascent chain from the cytosol (Batey et al. 2000). Although this may be the case, it seems unlikely from the results seen here that the structure of the GPR35 nascent chain drives the SRP interaction. Therefore, further experiments were carried out to explore what properties of the nascent chain are critical for a SRP interaction.

Hydrophobicity is a feature that is common within all signal sequences, whether they are integral signal anchors or cleavable signal peptides (Keenan & Freymann 2001). All signal sequences contain a hydrophobic stretch of residues which many cite as being essential in driving the SRP binding. To test if this is also the case for GPR35, the hydrophobic core of the 1st TM domain was altered to reduce its levels of hydrophobicity and analyse the effect it had on SRP binding. Four leucine residues, made up of hydrophobic hydrocarbon side chains, were substituted to glutamic acid residues containing carboxylic acid side groups, to produce a construct of reduced hydrophobicity; GPR35 Δ 4E (Figure 4.6A). Translation intermediates from Δ 4E were used to carry out similar cross-linking experiments as described previously. Cross-links to uL23 were detected in a similar pattern to those seen with the WT intermediates (Figure 4.6B2 and C). The proximity of the nascent chain to uL23 as it passes through the ribosome tunnel and the diameter of the BS³ cross-linker makes it impossible to know if a loss in hydrophobicity with the 1st TM domain would have impacted on any possible interaction taking place between the nascent chain and uL23.

The interaction between the nascent chain and SRP was, however, impacted upon by the loss of hydrophobicity. No interactions between the nascent chain and SRP could be detected at any intermediate length between 25aa and 65aa (Figure 4.6B2 and C). These

results were in agreement with previous studies carried out in signal peptides, suggesting the hydrophobic core (h-domain) was essential for SRP targeting (Janda et al. 2010; Ataide et al. 2011; Nilsson et al. 2015). In a similar experiment, Nilsson et al (2015) showed by deleting 3 or more hydrophobic leucine residues within a signal peptide was enough to drive SRP binding efficiencies below 25% of that seen in the WT protein. It was also enough to prevent targeting and translocation of the nascent chains into ER membrane representative rough microsomes (Nilsson et al. 2015). This suggests that hydrophobicity within a signal sequence is critical for the recognition, binding and also the targeting of a nascent chain by SRP, as well generating problems with translocation.

The exact reason as to why the loss in hydrophobicity within the signal anchor of GPR35 leads to an inefficient SRP binding could have been due to two reasons: the nascent chain may not be recognised as a membrane protein 1) by the ribosome or 2) by SRP. Nascent chain recognition by proteins of the ribosome exit tunnel has been shown to be an essential step in the SRP pathway (Woolhead et al. 2004; Berndt et al. 2009; Robinson et al. 2012). Altering the hydrophobicity of a TM domain however, may impact on how it is treated within the tunnel. Evidence suggests that detection of hydrophobic TM domain by the sensory proteins in the ribosome drives an increased affinity for SRP docking at the exit tunnel (Berndt et al. 2009). If this is the case, loss of hydrophobicity within GPR35 TM1 may result in a loss of recruitment of SRP to the ribosome. Alternatively, cross-linking and affinity assays have suggested that all translating ribosomes have a similarly high affinity to SRP and the presence of any nascent chain in the tunnel, creates a conformational change that enables SRP binding (Flanagan et al. 2003; Bornemann et al. 2008). This data would therefore suggest that SRP has a scanning mechanism that detects a hydrophobic segment as it exits the ribosome tunnel, rather than initial detection occurring within the tunnel itself.

To provide a greater insight as to whether SRP possesses a scanning mechanism for detecting the hydrophobic nascent chain as it exits the ribosome, two further hydrophobic mutant constructs were generated (4.7 B and C). Cross-linking assays showed both mutants, by restoring hydrophobicity to either half of the first TM domain, were capable of rescuing the SRP-NC interaction to some degree (Figure 4.8). However, differences between the eukaryotic and prokaryotic SRP binding to the Δ NT and Δ CT intermediates may suggest differences in the SRP recognition mechanism. In prokaryotes, both mutants showed binding efficiencies below 50% of the WT TM1, although the Δ CT intermediate showed a slightly greater cross-linking efficiency to Ffh than that of the Δ NT intermediate.

However, the difference between the SRP cross-linking to the individual intermediates in the eukaryotic system indicates a greater complexity in the recognition of a signal sequence (Figure 4.8). The Δ CT intermediate on this occasion cross-linked with ~80% efficiency to the WT, whereas the Δ NT intermediate represented a nascent chain much more like the Δ 4E intermediate. This suggests that the SRP molecule may have a more localised mode of scanning for stretches of hydrophobicity within the nascent chain, whilst it exits the ribosome. Eukaryotic SRP, due to its more complex structure, may not be capable of interacting with the hydrophobic region of the Δ NT intermediate in the same way as Ffh. Alternatively, it could allude to a different form of interaction between the ribosome and the nascent chain, in which it no longer recognises the Δ NT intermediate as SRP-dependant peptide. Nonetheless, these results provide increasing evidence that the signal anchor of GPR35 relies heavily on hydrophobicity to direct SRP targeting. Therefore, by restoring some hydrophobicity within TM1 of GPR35, an interaction between the nascent chain and the hydrophobic groove of SRP is rescued. How the alterations in hydrophobicity impact on the downstream effects of SRP targeting and translocation remain unknown. Nilsson et al (2015) carried out experiments using a hydrophobically-altered signal peptide and showed targeting and translocation to be significantly inhibited (Nilsson et al. 2015).

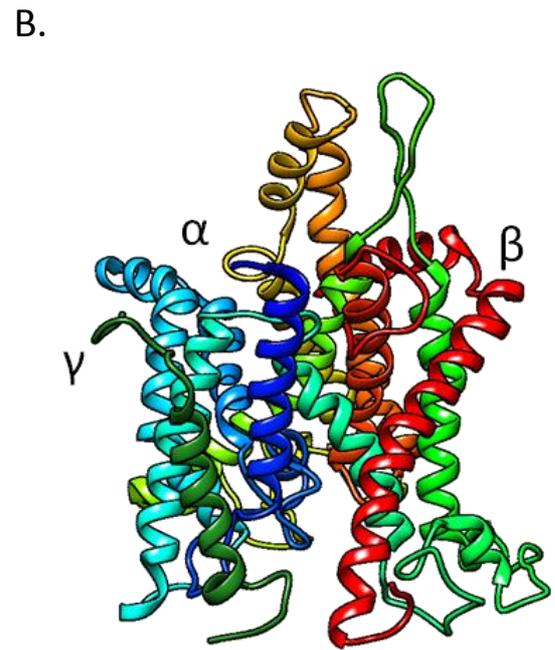
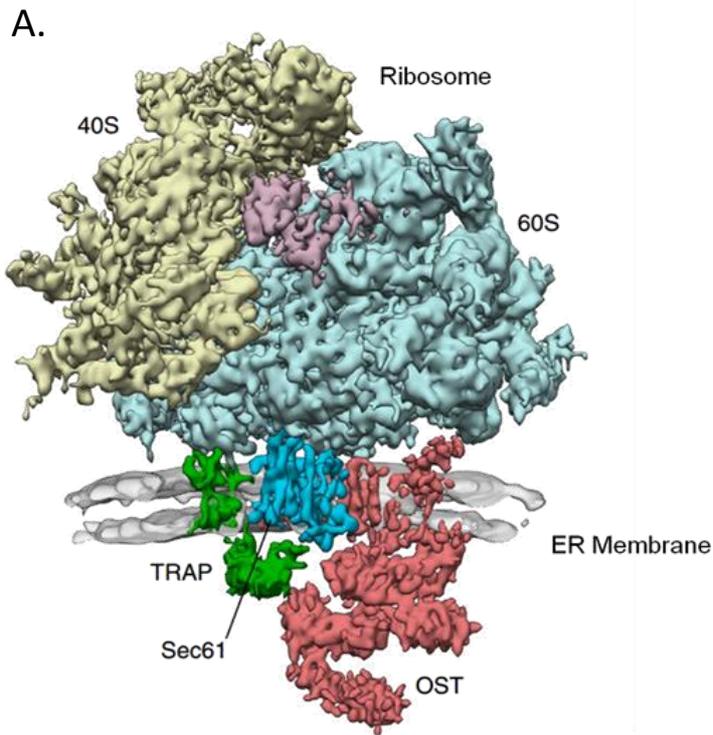
In summary, the results from this chapter have provided evidence for the targeting of GPR35 to occur in a SRP-dependant manner. Like many other membrane proteins, GPR35 can be shown to interact with SRP and other recognized components of the SRP pathway. Cross-linking to uL23 and SRP have indicated that the GPR35 nascent chain is likely to encounter the SRP pathway in a linear conformation and may at the point of SRP interaction begin to form a compacted helix. As this is the case, we are confident that the conformation of the nascent chain does not drive the SRP-NC interaction. Through the generation of mutants altering the hydrophobicity of the GPR35 signal anchor, it seems that hydrophobicity within TM1 is key for an SRP interaction and targeting to the ER membrane.

5: Targeting and integration of GPR35 to the Endoplasmic Reticulum

5.1 Introduction

The final steps in the biogenesis of a membrane protein begin as it becomes co-translationally inserted and integrated into the lipid bilayer. It is at this point that the TM domains of IMPs are moved into the membrane, their topology determined and their final structure achieved (Alder & Johnson 2004; Shao & Hegde 2011). In eukaryotes, this event occurs most often at the Sec61 translocon and is coupled to the co-translational targeting pathway as discussed in Chapter 4. A sequence of co-ordinated events between proteins synthesis and nascent chain integration allows for the safe passage of the hydrophobic TM segments into the lipid bilayer, preventing exposure to the hydrophilic cytosol (Halic et al. 2006a).

Targeting of most eukaryotic proteins (both soluble and membrane proteins) is centred on the movement of the peptides through the Sec61 translocon. The translocon resides in the ER membrane and is a heterotrimer, made up of α , β and γ -subunits (Figure 5.1B and C). High resolution structures of the Sec61 translocon have enabled us to deduce that the translocation of proteins across the membrane occurs through a narrow pore within the complex. This pore is believed to be made entirely of Sec61 α -subunits and is gated by a subdomain, known as the ‘plug’ (Berg et al. 2004). Interactions between the RNC and the translocon are believed to initiate the opening of the channel in preparation for translocation. As the ribosome engages the Sec61 translocon, a structural change occurs, switching the channel from a ‘closed’ to an ‘open’ conformation (Berg et al. 2004; Tsukazaki et al. 2008). Photo-crosslinking experiments have shown that the movement of TM helices 2b/3 and 7/8, within the α -subunit pore, open the Sec61 translocon to enable lateral movement of the nascent chain into the ER membrane (Plath et al. 1998). Originally, elongation of the nascent chain through the Sec61 channel was believed to aid the switch to the ‘open’ conformation by displacing the α subdomain ‘plug’ (Cannon et al. 2005), but recent



C.

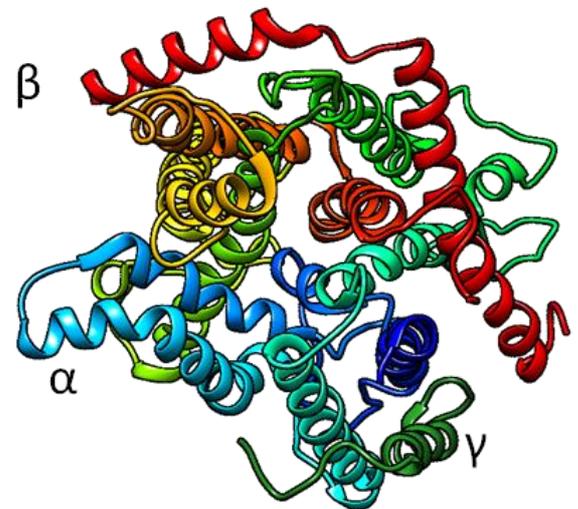


Figure 5.1 Sec61 translocon located in the mammalian ER membrane. (A) CryoEM structure of the translocation complex associated with the ER membrane. 40S (yellow), 60S (light blue), elongation factors (purple) Sec61 translocon (blue) TRAP (green) and OST (red). (Adapted from Pfeffer et al 2014) (B) Cryo-EM structure of mammalian Sec61 translocon (lateral view) (C) Cryo-EM structure of mammalian Sec61 translocon from the cytosol. PDB (4CG7).

cryo-EM data has deemed this may not be necessary, as the Sec61 channel has been captured in the 'open' state without the presence of a nascent chain (Pfeffer et al. 2015). Although a single Sec61 complex has been shown to permit insertion of small or single spanning membrane proteins into the ER, many suggest this is not the case for more complex membrane proteins. In fact, in prokaryotic systems, both structural and biochemical data show that oligomers of the SecY translocon may be necessary for the insertion of multi-domain membrane proteins, something that is also believed to be likely in eukaryotic systems (Mitra et al. 2005; Ménétret et al. 2005; Osborne & Rapoport 2007). Not only is it likely that Sec61 oligomers are involved in translocating polytopic membrane proteins, but it has been shown that a broader translocon complex exists. This is made up of membrane chaperones (such as BiP and Calnexin), accessory factors (such TRAM and TRAP), and enzymes (such as the oligosaccharyl transferase complex (OST)) (Figure 5.1A).

Insertion of the nascent chain begins as the RNC complex docks on the cytosolic surface of the Sec61 channel. Upon binding, the translation process, which is temporally stalled by SRP, is re-engaged and the nascent chain makes its way into the channel of the translocon. Photo-crosslinking studies have shown the Sec61 channel to be capable of housing the nascent chain, separating it from the lipid bilayer until such a time as it is to be released laterally into the ER membrane (McCormick et al. 2003; Sadlish et al. 2005). This lateral movement through TM domains 2/3 and 7/8 of the Sec61 α , which is known as the 'lateral gate' of the translocon, is most commonly used by hydrophobic TM segments and seems to have been designed as a method of separating the aqueous translocon pore from the hydrophobic bilayer. This lateral movement of membrane proteins has been well studied using both single spanning and polytopic membrane proteins. The movement of single spanning proteins has been proven to be relatively straightforward, something that cannot be said for multi domain membrane proteins (Booth & High 2004; Higy et al. 2004). Increased size and number of domains, coupled with obtaining the correct orientation and correct secondary structure before integration into the membrane, instantly makes a polytopic protein more challenging for the translocon to process.

The integration of IMPs into the ER presents a number of problems. IMPs contain multiple domains with vastly different properties i.e. hydrophobic TM domains, charged signal sequences and hydrophilic internal and external loops, making integration into the membrane more challenging. As previously stated, the final secondary structure of membrane proteins must be obtained before integration can begin, therefore TM domains in many IMPs must obtain their correct structure and orientation before translocation comes to an end. As discussed in previous chapters, large TM segments have been shown to be capable of taking up secondary structure prior to entering the translocon, either whilst in the ribosome tunnel (Woolhead et al. 2004; Lu & Deutsch 2005b) or upon interacting with SRP, when emerging from the ribosome (Robinson et al. 2012). This is not known to be the case for all TM proteins and hence it is plausible that the translocon pore, in some cases, may aid domain folding. Secondary structure within the Sec61 translocon is also poorly understood, but there is evidence that the environment provided by the pore could enable TM domains to sample multiple conformations. Cross-linking experiments show stabilizing interactions with the Sec61 α -subunits of the translocon (McCormick et al. 2003), as well as interactions with accessory proteins such as TRAM (Heinrich et al. 2000; Sadlish et al. 2005) which may impact on folding within the translocon. The preference towards a helical structure by the Sec translocon has been seen in both eukaryotic and prokaryotic systems. Indeed, experiments using proton motive force (PMF) calculations showed that the Sec61 channel, due to spatial confinement and surface properties, greatly favoured a helical over an extended nascent chain (White & Von Heijne 2005). In the SecY channel, a 27 residue, hydrophobic peptide was used to show that the translocon could provide an environment capable of generating an α -helical conformation within a protein. The structure of the channel, resulting in the confinement of the peptide, was believed to be key in driving the peptide from an extended to helical conformation (Ulmschneider et al. 2014). As integration of membrane proteins relies on the correct formation of secondary structure with TM domains, the translocon could play a significant role in ensuring that this is in place.

Once the orientation and secondary structure of TM domains are in place within the translocon, integration into the ER bilayer can begin. Early models of protein integration proposed a sequential mechanism, in which each TM domain moved into the membrane as it emerges from the ribosome (Blobel 1980). This is the simplest of the proposed models, relying on the intrinsic hydrophobic sequence of the individual TM segments to drive the insertion process (Alder & Johnson 2004; Sadlish et al. 2005; Ismail 2006). However, in the case of most polytopic membrane proteins, a number of TM segments must interact before they can be stably inserted into the ER. One possible mode of integration relies on the packing and assembly of the TM domains within the translocon before release (Lecomte et al. 2003). A second relies on the full translocation of a segment into the ER lumen before being ‘pulled’ back into the membrane upon the insertion of subsequent domains (Lu et al. 2000). Finally, in one case it could be seen that a first TM domain can enter the ER membrane alone, remaining close to the translocon and upon the presence of a second TM domain, return to aid its entry into the bilayer (Heinrich & Rapoport 2003).

In this chapter, the issues of TM domain folding and mode of integration of GPR35 will be accessed. Experiments will be carried out to investigate if GPR35 is capable of being successfully translocated and integrated into the ER membrane (dog pancreas microsomes (DPMs) used as representative ER membranes). Experiments taking advantage of the protection offered by the ER membrane, using a protease assay, will assess if translocation and insertion into the membrane was successful. Proteinase K, a digestive enzyme with broad specificity, was used to detect the integration of full length GPR35 into the DPM membrane by digestion assays targeting protein components exposed to the cytosolic environment. Further experiments to determine orientation and secondary structure within the TM domains of GPR35 during translocation will take advantage of the enzymatic complex found on the luminal side of the ER membrane, the OST complex (Figure 5.2 A). This complex transfers a 14-sugar oligosaccharide dolichol within the ER membrane to an asparagine within the nascent protein at a NXT site. This process, N-linked glycosylation, changes the MW of the nascent chain enabling the detection of this post-translocational modification by gel-shift. The position of the OST complex, relative to the translocon, (requiring approximately 15aa to be translocated into the ER before glycosylation can occur (Whitley et al. 1996)) is known and when coupled with the known distance required

to traverse the ribosome tunnel and translocon we can accurately access if a nascent chain is compacted or extended within the translocon (Figure 5.2 A). The final experiments to be carried out in this chapter will be to assess how GPR35 is integrated into the ER membrane. Chemical cross-linking assays, using the cysteine specific cross-linker bismaleimido-hexane (BMH, spacer arm length: 13Å), which can enter the Sec61 translocon channel, were carried out to determine the point at which integration of TM domains 1 and 2 took place (Figure 5.2 B). Through the assessment of interactions with Sec61 α -subunits in the channel pore with the nascent peptide, we can propose the mechanism by which the N-terminus of GPR35 integrates into the ER membrane.

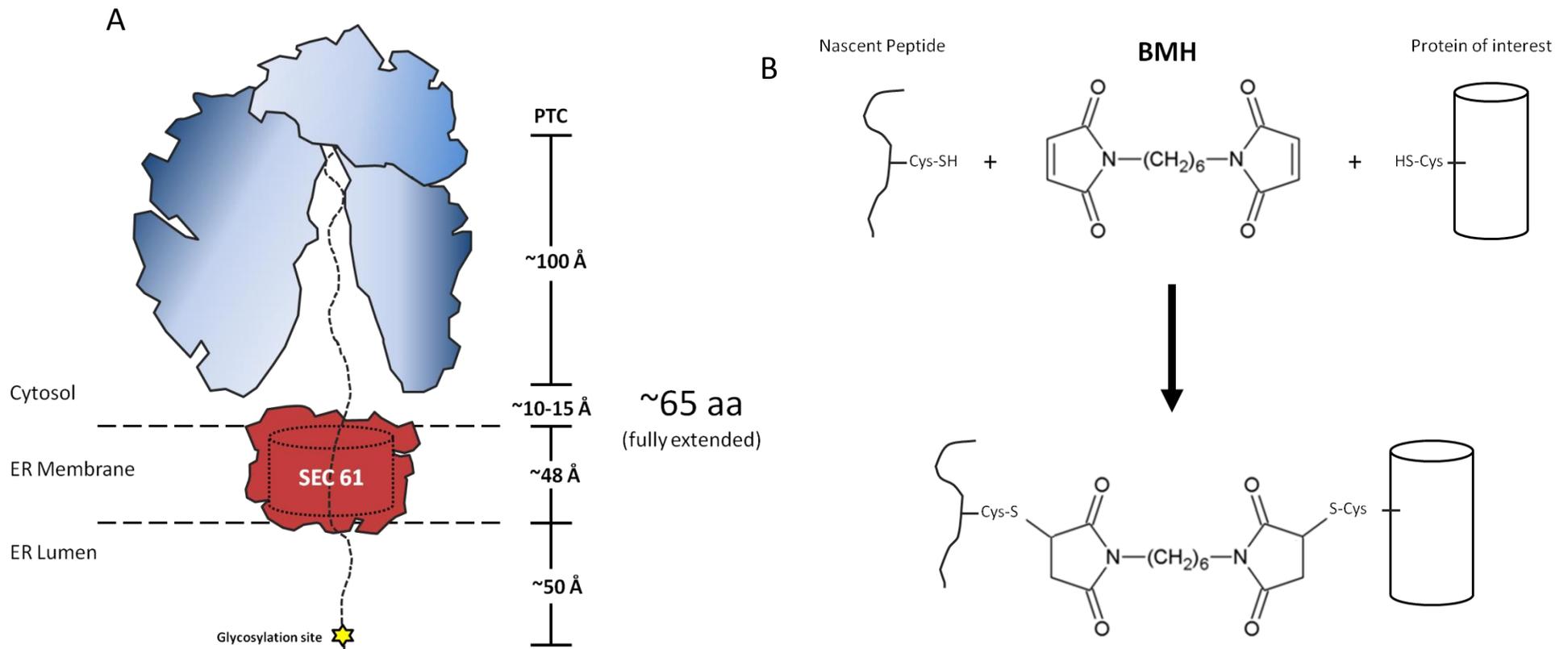


Figure 5.2 Experimental design of glycosylation and chemical cross-linking assays (A) Schematic diagram displaying the distance required to bridge the gap between the P-site of the ribosome and the OST site for glycosylation. (B) Chemical cross-linking molecule BMH reacts with the SH group of the single cysteine residue in the nascent chain and the free SH groups in components of the Sec61 translocon. The α - β double bond undergoes nucleophilic addition by the nascent peptide and protein thiols yielding a stable thioether linkage (Zucca & Sanjust 2014).

5.2 Results

5.2.1 Proteinase K digestion to determine successful targeting and integration of GPR35 to DPMs

Cross-linking experiments in chapter 4 have determined that GPR35 is most likely to be targeted to the ER membrane by the SRP pathway. As this is the case, experiments were carried out to investigate if GPR35 could be successfully targeted and integrated to the representative ER membranes, DPMs. Due to the interaction detected between GPR35 and SRP54, we expect translocation into the ER membrane to occur through the Sec61 translocon, as seen in the majority of other polytopic membrane proteins. To ensure successful integration, digestion and glycosylation experiments were set up to confirm GPR35 is in its correct orientation.

The initial experiment to test the success of insertion of GPR35 into DPMs was carried out using proteinase K (PK) and relied on the protection from digestion given by the membrane bilayer upon successful insertion. For the digestion and glycosylation experiments, radiolabelled GPR35 nascent chains were generated using the RRL *in vitro* translation system, to which DPMs were added to provide an ER membrane component and the end point for GPR35 insertion. On this occasion, full-length GPR35 RNCs were generated. The N-terminus of GPR35 contains a N-glycosylation site at residue Asn-2; therefore both insertion and correct orientation of the GPCR should be detectable upon isolating the membrane fraction.

The results of the digestion assay can be seen in Figure 5.3A. In the absence of PK, the translated full length GPR35 can be detected at approximately 30kD, in the presence of DPMs (Lane 1). Full length GPR35, however, showed no sign of a higher molecular weight (MW) band indicative of glycosylation occurring upon translocation. This suggested the GPCR N-terminus was not interacting with the OST complex, hence we could not determine if the correct orientation of GPR35 had been obtained. However, insertion of full length GPR35 was confirmed upon the addition

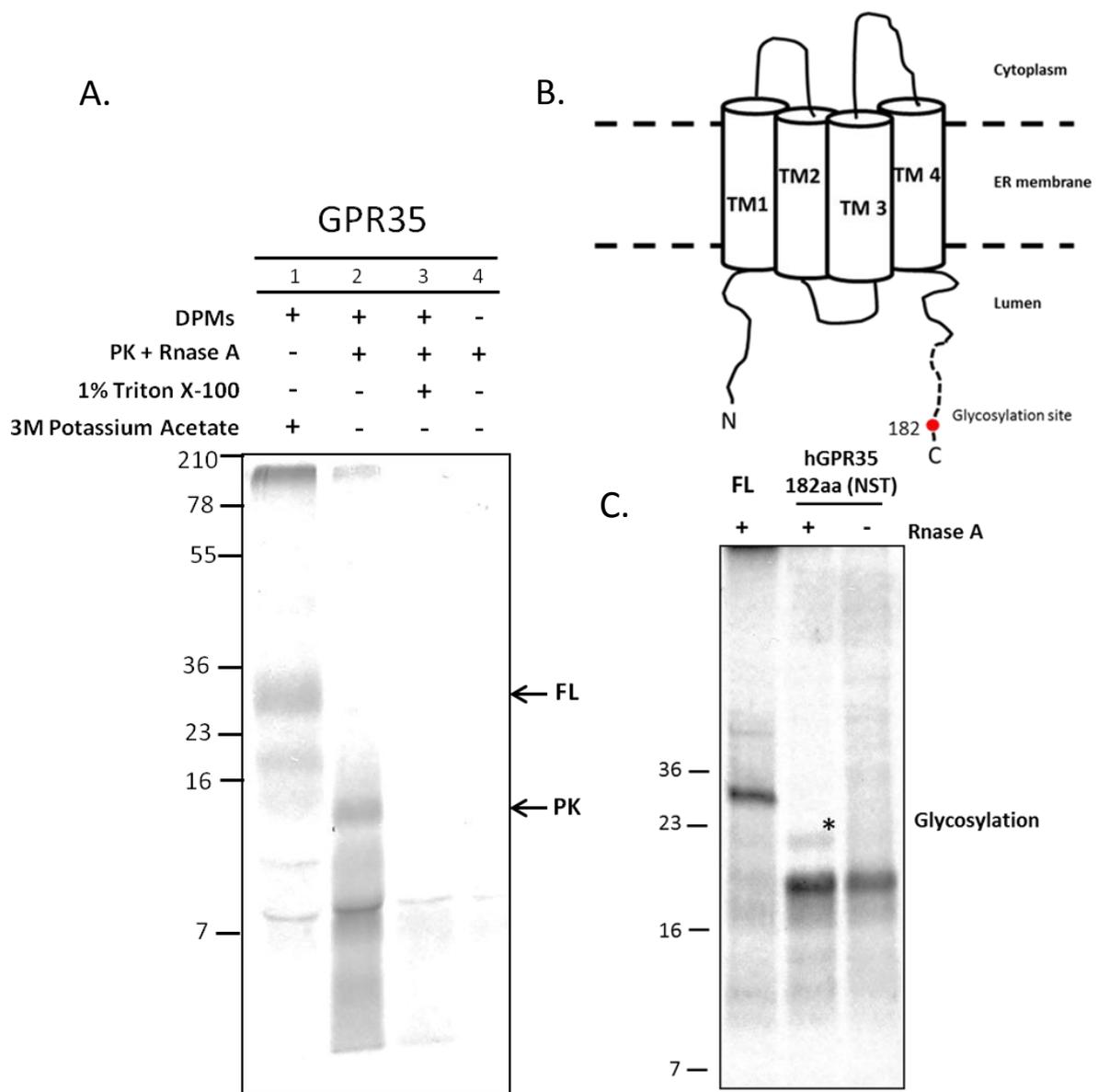


Figure 5.3 Assay to assess insertion and orientation of GPR35 into DPMs. (A) Radiolabelled full-length GPR35 was targeted to DPMs that were added to the RRL translation mix. Digestion assays using proteinase K assessed insertion of GPR35 into DPMs. (B) Schematic to show the expected orientation of the 182aa intermediate designed for glycosylation, when inserted into the ER membrane. (C) Selective release and glycosylation of the 182aa intermediate by RNaseA, shows GPR35 is orientated correctly in DPMs. * denotes glycosylation of C-terminus upon release from the PTC.

of PK (Lane 2) at which point the 30 kD band was no longer seen to exist and was replaced by bands at ~7 kD. These bands represent fragments of the GPR35 that have been protected from PK degradation by the lipid bilayer of the DPMs. Insertion of GPR35 into the membrane was further confirmed upon disrupting the permeability of the bilayer by a detergent (Lane 3). The bands present in lanes 1 and 2 could no longer be detected suggesting disruption of the membrane allows for the degradation of GPR35. Finally, a control experiment in the absence of DPMs was carried out to ensure GPR35 could not be isolated unless it was associated with the membrane (Lane 4). This resulted in no full-length GPR35, indicating that the appearance of any intermediates previously was due to its association with the DPM bilayer. The results therefore indicate that GPR35 can be successfully targeted and inserted into DPMs; however, the glycosylation of site Asn-2 is highly inefficient, therefore the orientation of the full-length protein remains unconfirmed.

To investigate whether the orientation of GPR35 after integrating into the membrane bilayer was correct, a further glycosylation assay was designed. On this occasion an engineered glycosylation site was placed on the C-terminal end of an 182aa intermediate (Figure 5.3 B). This glycosylation site was incorporated into the reverse primer, which was designed to be specific for extracellular loop 2 (positioned on the luminal side of the ER) of the mature GPR35. The reverse primer extended extracellular loop 2 by 18aa, with the glycosylation site situated 16aa downstream from TM4 and theoretically in range of the OST complex (Figure 5.3 B). The reverse primer was designed without a stop codon, generating a ribosome-bound nascent chain that could only be fully translocated by the addition of RNaseA and EDTA to the translation mix. This selective release from the ribosome enables the comparison between the glycosylated and unglycosylated ribosome-bound peptides, confirming if indeed GPR35 is inserted in the correct orientation.

The results of this assay show convincingly that, GPR35 is capable of inserting into DPMs with the correct orientation (Figure 5.3 C). Upon translation and isolation of the membrane integrated 182aa intermediate, a single band could be detected at ~16 kDa representing an unglycosylated ribosome bound peptide. Upon addition of RNaseA and EDTA to the translation mix, breaking the peptidyl-tRNA bond and removing the ribosomal subunits, the C-terminus is released and translocation of the

engineered glycosylation site occurs across the DPM membrane. Translocation of the C-terminal end produced 2 bands when analysed by gel electrophoresis, one representative of the 182aa intermediate and a second higher band representative of the glycosylated product. The appearance of the higher product suggests that the glycosylation site in loop 2 of GPR35 was capable of interacting with the OST complex on the luminal side of the DPM membrane, therefore confirming GPR35 was in the correct orientation.

Experiments set up to investigate the targeting and insertion of GPR35 into the membrane of DPMs shows that the protein is capable of efficiently integrating and adopting the correct orientation within the membrane. Protease experiments show that GPR35 is targeted to DPMs in the translation mix and offered protection from Proteinase K by insertion into the membrane. Insertion of GPR35 in the correct orientation was confirmed by the glycosylation of the NST site placed specifically in extracellular loop 2. Translocation of the C-terminal end of this 182aa intermediate across the membrane enables the interaction of the peptide with the OST complex, a complex which will become a useful tool in the future experiments to assess structure and movement of GPR35 within the DPM membrane.

5.2.3. Does a loss of hydrophobicity in the 1st TM domain result in a loss of insertion into the translocon?

Upon showing GPR35 could be successfully inserted into the ER membrane, an experiment was set up to assess what impact disrupting the hydrophobicity of a signal anchor domain would have on insertion. As seen in Chapter 4.2.3, the hydrophobicity of the 1st TM domain in GPR35 was critical for driving the interaction between the nascent chain and SRP. Reducing the hydrophobicity at either terminus of the TM domain or removing hydrophobicity entirely, so it was no longer recognised as a TM domain, adversely affected its interaction with SRP. As GPCRs are thought to be targeted to the ER membrane almost exclusively via the SRP pathway, experiments were set up to investigate how the loss of hydrophobicity would impact on the targeting and subsequent integration of GPR35 intermediates into the ER membrane.

The four constructs designed for cross-linking experiments in Chapter 4.2.3 will be used (GPR35 WT, Δ NT, Δ CT and Δ 4E). The constructs were once again expressed in the RRL *in vitro* translation system and targeted to DPMs. Each intermediate was a total length of 75aa and would contain the native glycosylation site at residue 2. A new batch of DPMs was used in this experiment and provided a much more efficient rate of glycosylation at Asn-2 than had been previously seen in Section 5.2.1. Successful targeting and translocation would enable the individual intermediates to interact with the OST complex and become glycosylated, hence producing an intermediate with a higher MW.

The effect of a loss of hydrophobicity in the 1st TM domain was clear to see in the resulting translocation reactions (Figure 5.4). In the absence of DPMs, one band representative of the translation band exists (Figure 5.4 A). Upon the addition of DPMs to the RRL translation reaction, the GPR35 WT intermediate, as expected, becomes targeted and successfully inserted. Approximately half the translation product shows a shift in MW from ~7 to ~14 kD, representative of glycosylation of at the N-terminal glycosylation site, within the nascent peptide (Figure 5.4 B and D). To ensure the higher MW band was indeed due to the translation product becoming

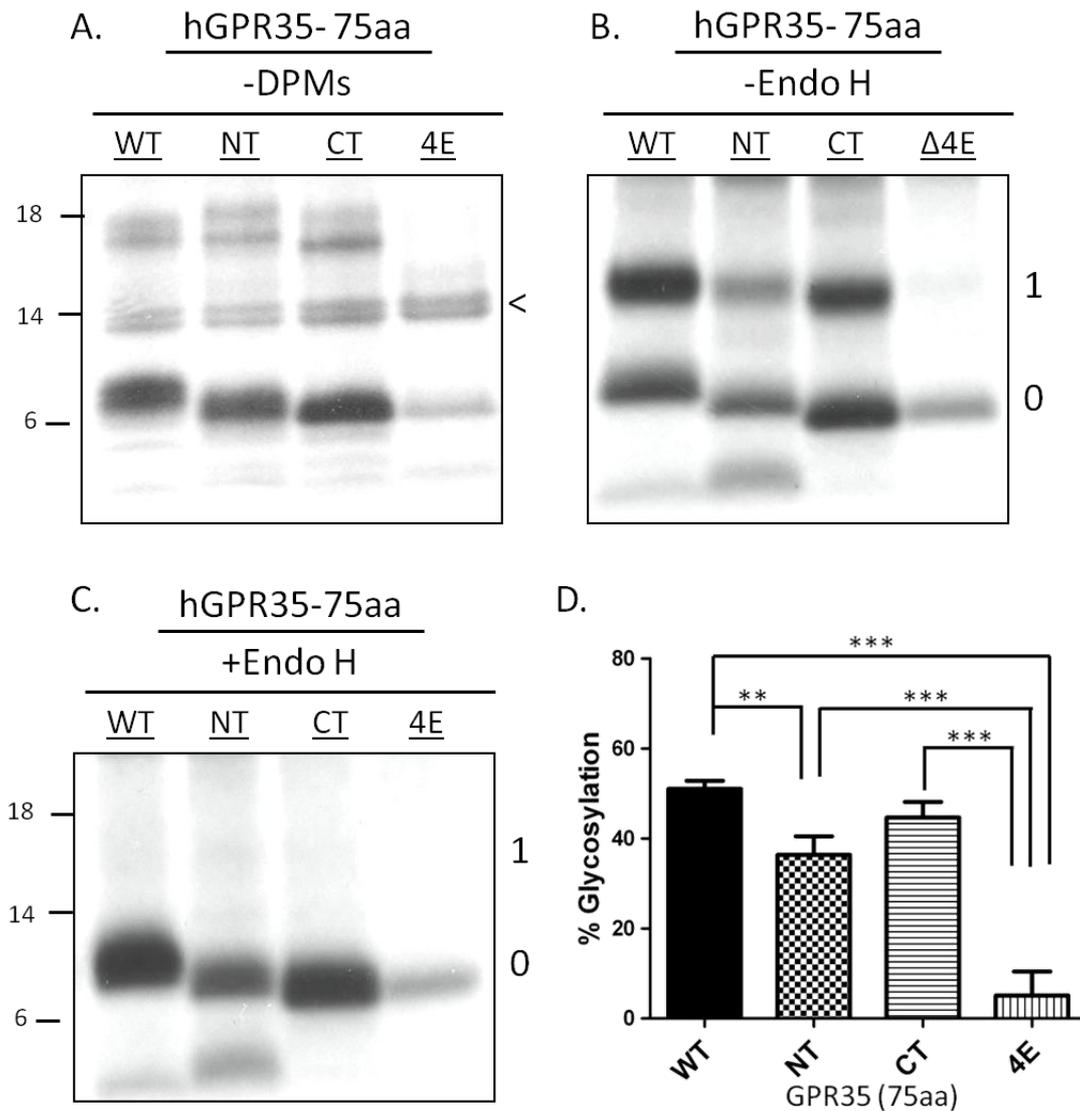


Figure 5.4 Glycosylation assay to assess translocation of GPR35 TM1 hydrophobic mutants. Ribosomes displaying radiolabelled N-terminal intermediates of GPR35 hydrophobic mutants (as described in Figure 4.8) at lengths of 75 amino acids before (A) and after insertion into DPMs (B and C). The addition of endoH to the DPMs results in a loss of the glycosylated translation product (C). Translation product and glycosylated translation product are indicated by 1 and 2 respectively, with the presence of the background heme band marked with a <. Products were resolved on a 10-20% SDS-PAGE gel. (D) Bar graph showing the average % of glycosylated translation product in the presence of DPMs. Glycosylation results were quantified using Image J software. All % glycosylation values were calculated using [glycosylated product/total translation product] and adjusted for background. Average percentage of glycosylation is calculated from an n of 3. Error bars indicate standard deviation and groups were compared using one-way ANOVA with Tukey's *post hoc* comparisons. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$.

glycosylated, the DPMs were incubated with the enzyme Endo H, which is capable of removing sugar group generated from a glycosylation event. In the presence of Endo H, the higher MW bands associated with glycosylation were successfully removed (Figure 5.4 C). GPR35 Δ NT and Δ CT intermediates were also successfully inserted into the ER membrane (Figure 5B and D). However, both intermediates showed a reduction in translocated product in comparison to the WT GPR35 intermediate, which may have been representative of a loss of interaction between the nascent chain and SRP, as seen in the cross-linking results (Figure 4.8 B). Generally, the effect on translocation in the Δ NT intermediate was significantly greater when compared to the WT intermediate, than what was seen in the Δ CT intermediate (Figure 5.4 D). Finally, the Δ 4E intermediate shows a much lower level of translation product isolated within the DPMs, with extremely low levels of the higher MW product representing glycosylation being detected when compared to each of the other intermediates (Figure 5.4 B and D). This would suggest that the N-terminus of the Δ 4E intermediate was unsuccessfully translocated across the ER membrane as a result of a loss in hydrophobicity leading to poor SRP targeting or poor insertion into the DPM membrane.

The results of this experiment highlight the importance of the hydrophobic signal as a requirement for targeting and translocation of the nascent chain into the ER membrane. As with the cross-linking assay in Chapter 4.2.3, the absence of hydrophobicity in the N-Terminus and C-Terminus reduced the levels of translation product that was successfully inserted into the ER membrane. The Δ CT intermediate is less affected, than the Δ CT intermediate, with WT like levels becoming successfully targeted. However, the complete reduction in hydrophobicity in the Δ 4E intermediate resulted in a complete loss translocated product, possibly due to the SRP no longer recognising TM1 as a signal anchor and hence generating a loss in targeting to the translocon machinery.

5.2.2 Assessing the secondary structure of GPR35 transmembrane domain 1 during translocation.

Throughout this study, one of the primary aims was to assess the structure and folding at the N-terminus of a model GPCR as it undergoes biogenesis. Through pegylation and chemical cross-linking experiments in chapter 3 and 4 respectively, we have shown as GPR35 makes its way through the ribosome tunnel and first makes contact with SRP while the nascent chain exists in an extended conformation. Upon interaction with SRP and throughout the process of targeting, it remains unclear as to whether the helical secondary structure of TM1 begins to form. To investigate if secondary structure exists prior to or takes place during insertion, we initially set out to use a photo-crosslinking assay set up by McCormick et al (2003). This assay relied on three successive probe sites being placed within the centre of the first TM domain. To detect whether or not a helical TM domain was present within the Sec61 channel, the symmetry of photoadducts was to be measured. A symmetrical pattern in cross-linking suggests the presence of an extended nascent chain; however asymmetrical cross-linking patterns would allude to the presence of an α -helical TM domain (McCormick et al. 2003). However, upon attempted incorporation of the fluorescent probes into the GPR35 intermediates only truncated peptides, due to the existence of an amber stop codon, could be detected, suggesting a failure of the probes to incorporate. Therefore, a change in experimental design was in order and we decided to utilize the known spanning distance between the PTC and the OST complex, with glycosylation as a marker of distance.

To use glycosylation as a marker of distance, we can take advantage of the previous work carried out by several groups who have mapped the distance of a number of key components that make up the RNC-translocon complex. Firstly, the ribosome tunnel has been shown experimentally to be $\sim 100\text{\AA}$ in length, requiring a total of ~ 30 residues in an extended conformation to traverse this distance (Lu & Deutsch 2005b). Secondly, the length of the Sec61 translocon channel has mapped by cryo-EM experiments and shown to $\sim 70\text{\AA}$ (Pfeffer et al. 2015), suggesting that ~ 20 aa in an extended conformation would be capable of covering this distance. Thirdly, the point

at which a peptide becomes glycosylated at the OST complex is believed to require ~15aa of additionally translocated protein to have passed through the Sec61 translocon and into the ER lumen. Finally, a controversial space between the cytoplasmic face of the translocon and the bottom of the ribosome upon formation of the RNC-translocon complex is believed to exist, spanning ~15Å in distance (Patterson et al. 2015). In total the RNC must traverse a distance of ~200Å to become glycosylated, requiring ~65 residues of an extended peptide (Whitley et al. 1996). Therefore, whether or not the nascent chain contains secondary structure while spanning the Sec61 complex can be calculated by the length at which an intermediate of GPR35 first becomes glycosylated.

As glycosylation of the N-terminus in full-length GPR35 could not be recognised in the initial experiments of Chapter 5, it was decided that another site further from the start codon should be inserted to enable an increased likelihood that glycosylation would occur. As the wild type GPR35 already contained an Asn (N) and Tyr (T) at residues 5 and 6 respectively, a Ser (S) was inserted between the two residues by site-directed mutagenesis (see section 2.4.6.1), generating an artificial glycosylation (NST) site. As in previous experiments, reverse primers without a stop codon were designed to generate ribosome bound intermediates of various lengths. The length of each intermediate was calculated from the PTC to the first residue in the artificial glycosylation sequence (N6). Expression of the intermediates was carried out in the RRL *in vitro* translation system containing DPMs, with lengths ranging from 65-100aa.

The results of the glycosylation experiment to assess the folding profile of GPR35 during translocation can be seen in Figure 5.3 A and B. In the absence of DPMs from the RRL translation system, a band correlating to the size of each unglycosylated intermediate could be detected. Also present was the background haem band (~16 kD) seen previously arise upon isolation of RNCs without DPMs present (Figure 5.5 Lane 1). Upon the addition of DPMs, the appearance of two bands of increased molecular weight (MW) could be detected along with the unglycosylated intermediate (Figure 5.5 Lane 2). The two bands of increased MW were believed to be the result of a single or double glycosylation event taking place at the N-terminus. A change in source of DPMs was believed to be a possible cause for glycosylation now occurring at the first

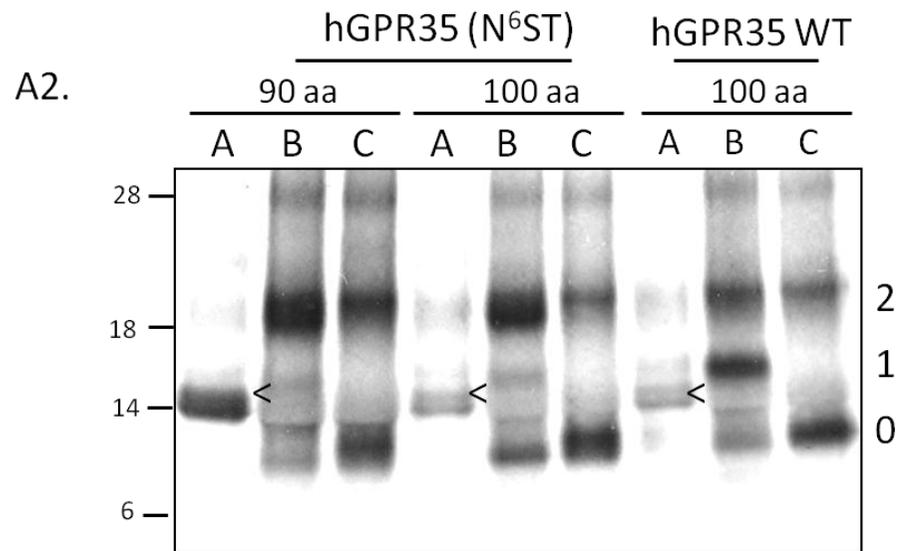
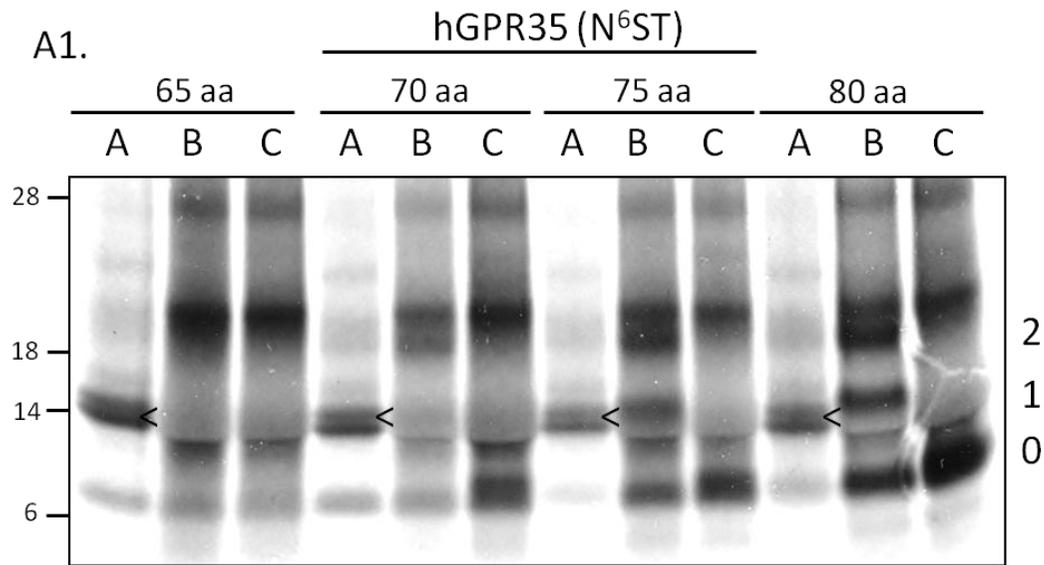


Figure 5.5. Analysis GPR35 glycosylation in DPMs to detect the formation of secondary structure in the Sec61 translocon. Intermediate lengths ranging from 65-100aa of GPR35 were expressed in an in vitro RRL translation reaction. Reactions were split into three, with one third being incubated without DPMs (Lane A), another being incubated in the presence of DPMs (Lane B) and the the final third being incubated in the presence of DPMs and Endo H (Lane C). Lane A contains the unglycosylated translation product (0). Also present in Lane A is a background band at ~14 kD, which is representative of heme, indicated by a <. Lane B represents glycosylation occurring in the translated product. The translated product (0) has two glycosylation sites present. In the presence of DPMs a single (1) or double (2) glycosylation event can be detected by a gel-shift of the translation product by ~ 5kDa. Lane C represents EndoH treated DPMs, removing glycans (1 and 2) and returning the glycosylated product to the initial MW of the translated product (0).

site (which could not be seen in original glycosylation assays with FL GPR35). The 65aa intermediate shows no sign of glycosylation at either glycosylation site, suggesting that it is long enough at this point to interact with the OST complex. Although the first glycosylation site within this intermediate is 70aa from the PTC and theoretically should be capable of low levels of glycosylation, there is none to be detected. This could indicate that this site in particular does not become glycosylated efficiently or may not be recognised by the OST due to its close proximity to the start codon, as suggested by Bano-Polo et al (2011). In the 70aa intermediate however, signs of weak glycosylation can be detected at both Asn-2 and Asn-6, suggesting that ~70 residues are required for glycosylation to occur. Glycosylation then appears to occur at increased levels in the 75aa intermediate, suggesting it is at an optimum length for interaction with the OST; this level of glycosylation which was maintained throughout the longer peptide lengths. To ensure the two higher MW bands were indeed as a result of glycosylation, the isolated DPMs were treated with the enzyme endoglycosidase H (endoH) which is capable of removing glycans from the native protein (Figure 5.5 Lane 3). Upon addition of the enzyme, endoH, the existence of the two higher MW bands were no longer present, hence suggesting that they were indeed a result of glycosylation in the native intermediate.

Glycosylation as a marker for the presence of secondary structure in translocating intermediates of GPR35 was capable of determining that the N-terminus and specifically the first TM domain forms a compacted structure in the Sec61 translocon. The lack of glycosylation products at 65aa from the PTC and the presence of glycosylated intermediates at 70aa, and increasingly at 75aa, indicate the presence of a helix-like domain during translocation. At this point we are unable to confirm whether it is due to an interaction with SRP or the translocation events that begins folding of TM1. In either case, it is an essential event in the biogenesis of GPR35, preparing it for integration.

5.2.4. Analysis of GPR35 integration into the ER membrane by site-specific cross-linking.

The integration of TM domains into the ER membrane is a critical step in the biogenesis of polytopic membrane proteins. To address this point, with GPR35, we have shown it can be successfully translocated and integrated into the membrane of DPMs; however, the mechanism of integration remains unknown. Integration of polytopic membrane proteins tends to be complex and in many cases differs between proteins, hence making it difficult to hypothesize a model of integration. However, as GPCRs are one of the largest and most widely studied classes of eukaryotic membrane proteins, a model of their integration may be of significant interest. Therefore, the aim of the following set of experiments is to study the mode and timing of integration of the first 2 TM domains of GPR35, enabling us to establish if integration is a co-ordinated event between domains.

In this study we examine the movement of the first 2 TM domains of GPR35 using a site-specific cross-linking assay that relies on the reagent bismaleimido-hexane (BMH); a sulfhydryl-to-sulfhydryl cross-linker. A single cysteine residue was inserted into the 1st or 2nd TM domain of GPR35 and was capable of being cross-linked to multiple cysteines in the Sec61 α domains of the translocon. To make the cross-linking site-specific, multi-site mutagenesis was carried out to remove other native cysteines within GPR35. Various lengths of truncated mRNA lacking a stop codon were expressed in the RRL *in vitro* translation system generating stable RNCs, which provided intermediates that could be trapped in the process of integration. Intermediates between 75 and 155aa were generated to assess the environment surrounding the nascent chain during integration and hence elucidate the timing of movement of individual TM domains.

To assess the mode of integration used by TM1 of GPR35, a construct containing a single cysteine in TM1 (GPR35-TM1) was used. The results of the TM1 cross-linking with BMH can be seen in Figure 5.6A and B. In the absence of the cross-linker, translated peptides can be detected in their unglycosylated and glycosylated states. The N-glycosylation, at the site found at the N-terminus of GPR35, indicates a

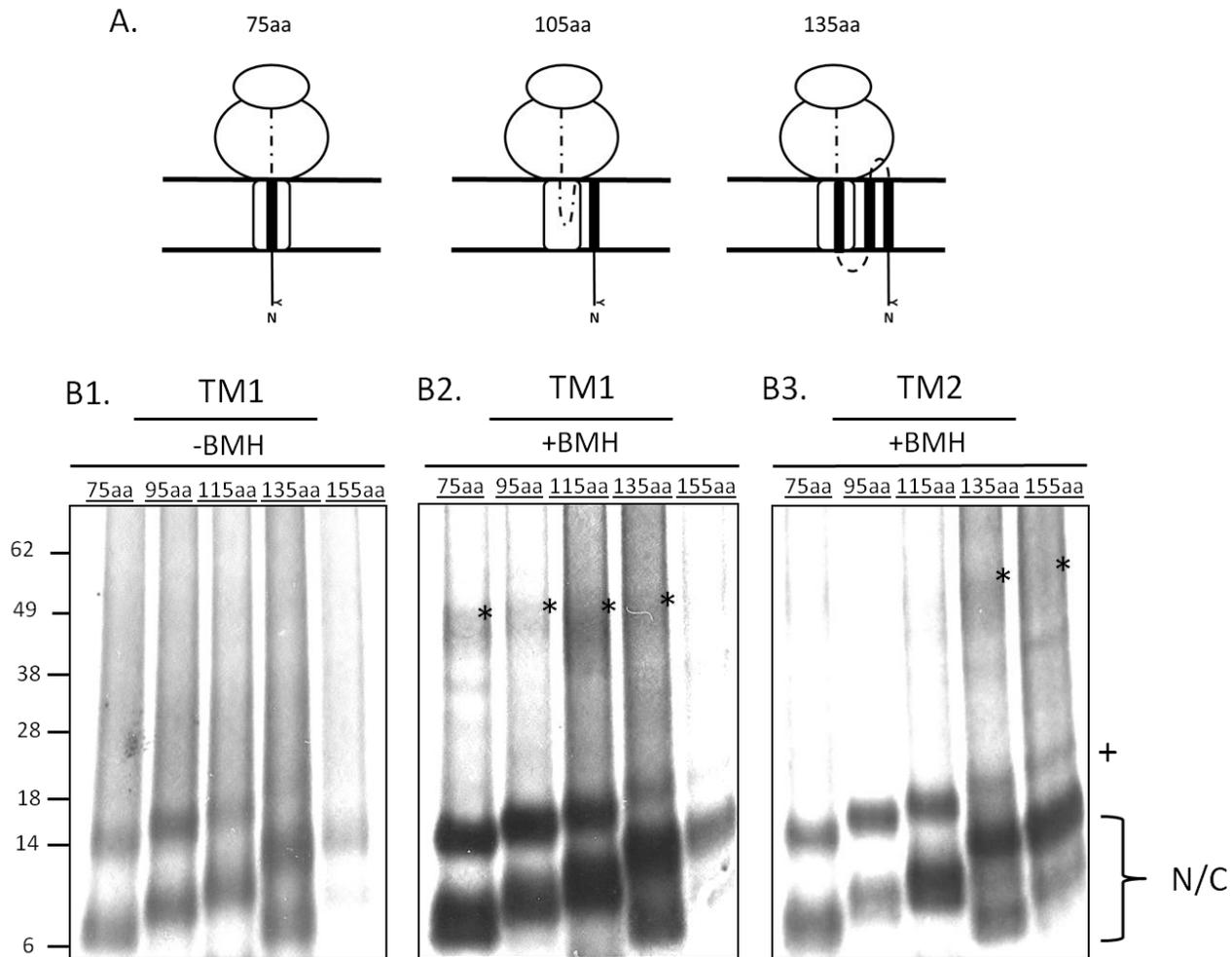


Figure 5.6. Integration GPR35 TM domain 1 into the DPM membrane. (A) Schematic to show the expected position of TM1 and TM2 of GPR35 at different intermediate lengths in the DPM membrane. (B) Ribosomes displaying radiolabelled N-terminal intermediates between 75-155aa of GPR35 at intervals of 20 amino acids can be seen before (B1) and after (B2 and B3) treatment with BMH when resolved on a 10-20% Bis-Tris gel. Crosslinks displaying the correct molecular weight for GPR35-Sec61 α complexes have been marked *. Glycosylated nascent chain is indicated with a + and is consistently reduced in intermediates 135 and 155aa.

successful integration of the GPR35 intermediates into the DPM membrane (Figure 5.6B1). In the presence of BMH, the translated peptides can once again be detected, as well as strong higher MW band at ~45 kD, at intermediate lengths ranging from 75-115aa. This higher band can also be detected weakly in the 135aa intermediate. These bands correspond to a possible cross-linking reaction taking place between the nascent chain and components of the Sec61 channel. (Figure 5.6 B2).

Immunoprecipitations with the antibody raised to the Sec61 α protein were attempted to determine if indeed this was an interaction between TM domain and the Sec61 channel. Unfortunately due to weak IP results and time mitigating factors, we were unable to confirm that the ~45 kD band was due to an interaction with Sec61 α .

However, a band of near identical size was seen in similar cross-linking assays carried out by Watson et al (2013) and was shown to be an interaction between the nascent chain and Sec61 α (Watson et al. 2013). A weakening of the band at ~45 kD after the 115aa intermediate suggests it is at this point that GPR35 TM1 is moved out of the Sec61 channel and possibly into the lipid bilayer of the ER membrane. Although there is a strong possibility that the 45 kD band is indeed Sec61 α , due to the timing and positioning of the nascent chain in the translocon, the protein TRAM also produces putative cross-links at this MW. The TRAM protein has been shown reside close to the translocon (Ismail 2006; Ismail et al. 2008), as well as cross-link to the nascent chain as it moves through the lateral gate (McCormick et al. 2003). However, it is unusual to see an interaction between TRAM and the nascent chain at such early lengths as 75aa from the PTC; hence I would suggest it is unlikely to be the cause of this higher MW band. The reduction in glycosylation in the longer intermediates may also suggest that the N-terminus, containing the N-glycosylation site, is being moved away from the OST which has been shown to be located near and interacting with the translocon (Pfeiffer et al. 2013).

The same experiment was carried out with TM2 (GPR35-TM2) to investigate the mode of integration of GPR35 TM domains; do they integrate sequentially or is it coordinated by the interaction of more than one domain? The GPR35-TM2 construct again contained a single cysteine residue in the middle of the 2nd TM domain. This ensured cross-linking was only capable when TM2 resided in the environment of the Sec61 translocon. The same intermediate lengths ranging from 75-155aa were used to investigate the timing of the TM2 insertion and whether TM1 was still present in the

translocon environment. Upon the addition of BMH, a higher MW band corresponding to that believed to be the cross-linked nascent chain and Sec61 α , was again detected but on this occasion at lengths 115-155aa (Figure 5.6 B3). The cross-links between Sec61 α and nascent chain at TM2 remain until 155aa, suggesting that TM2 is remains in the Sec61 channel as TM1 is integrated into the lipid bilayer. This suggests synthesis and insertion of TM2 may be required for the integration of TM1 into the ER membrane.

The results of this set of experiments provide the basis for a model of integration at the N-terminus of GPR35. The possible interaction between the nascent chain and Sec61 α would be consistent with the TM domains taking up a position within the channel of Sec61. A loss in the perceived cross-linking between Sec61 α subunits and the nascent chain could be seen as the partitioning to TM domains from the protein environment of the translocon, into the lipid bilayer of the ER. The loss in putative cross-linking between TM1 and the Sec61 coincides with appearance of cross-links in TM2; therefore suggesting a co-ordinated mode of integration occurs between the TM domains. Finally, glycosylation of the Asn-2 glycosylation site begins to become reduced after synthesis of the 115aa intermediate, coinciding with the lateral movement of TM1 from its position within the translocon. This would represent the N-terminus moving away from the OST site, as TM1 integration begins, hence agreeing with the described model of integration of the GPR35 N-terminus.

5.3. Discussion

The insertion and integration of TM domains into the ER membrane is an essential step in the biogenesis of polytopic membrane proteins that either reside in the ER or are trafficked along the secretory pathway. As this is the case, integration has been extensively studied using a variety of proteins, providing several different models. Prior to release from the translocon, membrane proteins are required to ensure their TM domains are in the correct orientation and have obtained the necessary secondary structure; once this has occurred integration can proceed. An increased number of TM domains in polytopic proteins adds to the complexity of integration, with several studies revealing a number of mechanisms for lateral release. As it was uncertain whether GPR35 had formed secondary structure prior to translocation and a model for its integration into the ER membrane was unknown, this chapter aimed to use a variety of techniques to shed light on what are key steps in the biogenesis of this GPCR.

Initial experiments were set up to confirm that full length GPR35 intermediates could be effectively targeted and integrated into the membrane of DPMs. Cross-linking results from chapter 4 suggested, as with many other integral membrane proteins, a SRP-dependant mode of targeting and therefore integration through the Sec61 translocon machinery. Indeed previous studies investigating the integration of GPCRs, specifically Opsin, have shown insertion into the ER membrane to be exclusively via the Sec61 translocon (Ismail 2006). Digestion assays using the broad specificity of proteinase K suggested, through partial protection of the full length GPR35, insertion and integration into the membrane of DPMs had taken place efficiently (Figure 5.2 A). The orientation of GPR35 at this point remained unknown as glycosylated intermediates had failed to be detected. This was believed to be due to a number of possibilities; the close proximity of the N-glycosylation site to the start codon, the quality of DPMs or, as discovered later, the suppressed rate in glycosylation of larger intermediates due to their movement away from the OST site at the Sec61 translocon. As the correct orientation for future experiments was essential, a second assay was set up to assess glycosylation at the C-terminus of an 182aa intermediate containing the first 4 TM domains. An engineered N-glycosylation site added to the extracellular loop 2 of GPR35 (found on the luminal face of the ER membrane) was

then used to assess the orientation of GPR35 in the membrane (Figure 5.3 B). Glycosylation of this intermediate, upon ribosome release, provided evidence that integration of GPR35 was occurring in the correct orientation (Figure 5.3 C); therefore further experiments investigating GPR35 translocation and integration could be carried out.

Successful targeting of polytopic membrane proteins to the ER membrane is essential for efficient integration. This process involves the presentation and transfer of the signal anchor from SRP to the Sec61 translocon in a highly co-ordinated series of events. Although the exact mechanism is poorly understood, the combination of translating ribosome, SRP and translocon is essential to begin the process of insertion. The recognition of a TM domain by the ribosome and the resulting structural changes is thought to prepare both the SRP and translocon for the incoming membrane protein (Liao et al. 1997; Pool 2009). The amino acid sequence and in particular the hydrophobicity of a TM domain has been proven to be essential for its recognition by SRP (see chapter 4). The amino acid sequence has also been proposed to be of high importance in directing the nascent chain to the translocon and deciding its orientation (Hessa et al. 2005; Nilsson et al. 2014). To test if altering the amino acid sequence in the first TM domain of GPR35 resulted in a loss of translocation, the hydrophobic mutants (GPR35 Δ NT, Δ CT and Δ 4E) designed for the cross-linking experiment in chapter 4.2.3 were used in insertion assays. The results followed similar patterns to that seen with the cross-linking to SRP, with N-glycosylation used as the marker for translocation. The Δ CT mutant was translocated with the highest efficiency out of the 3 mutants, followed by the Δ NT mutant (Figure 5.4). Both had suffered a loss in translocated product when compared to the WT intermediate, however, a loss of hydrophobicity at the N-terminus had a far greater effect than at the C-terminus. The Δ 4E intermediate unsurprisingly showed glycosylation levels no higher than background level, correlating exactly to the total loss of cross-linking with SRP witnessed in the previous chapter (Figure 5.4). The loss of efficiency in translocation correlates well with a similar study carried out to assess the impact of specific amino acid changes within the hydrophobic core of a signal sequence (Nilsson et al. 2014). In this study, deletions of Leu residues in the hydrophobic core result in a dramatic loss in both SRP targeting and insertion. This does raise the question as to whether a loss in SRP targeting has a direct effect on the interaction between the ribosome and

translocon, therefore preventing the nascent chain from being in a position to become translocated. However, biophysical studies have also shown that replacing polar residues with non-polar or charged residues can have an adverse effect on TM domain insertion, with position of certain amino acids within a TM domain crucial. Asn residues for example, have a far more detrimental effect on insertion when they are placed at the centre of a TM domain than when at either end (Hessa et al. 2005; Hessa et al. 2007). This has been suggested, along with cross-linking data, to upset the orientation of possible α -helices forming in the translocon and hence preventing specific interactions from occurring between the nascent chain and the translocon channel that enable integration (Hessa et al. 2005; McCormick et al. 2003). The formation of secondary structure within TM domains is essential before a peptide can leave the translocon and become integrated into the ER membrane. With GPR35 it had yet to be seen as to when α -helix formation began; therefore, the following experiments were set up to assess if it was occurring during translocation.

The secondary structure of TM domains whilst in the translocon is poorly understood. Much research in recent years has focused on the folding environment provided by the ribosome tunnel with secondary structure formation prior to insertion is a better understood process (Woolhead et al. 2004; Robinson et al. 2012). TM domains within the translocon are believed to be highly dynamic and studies show that they are capable of sampling a number of conformations (Goder & Spiess 2001). Secondary structure within membrane proteins is a requirement before integration into the ER can begin, something that was unseen in GPR35 prior to insertion into the translocon. Therefore, to test if secondary structure within TM1 of GPR35 was occurring during translocation, glycosylation assays were carried out. Previous work carried out by Whitley et al 1996, provides evidence that 65 amino acids in an extended conformation ($3.5\text{\AA}/\text{aa}$) is sufficient to bridge the gap between the P-site and the site of glycosylation at the OST complex. Intermediates of GPR35 witness glycosylation first occurring weakly when the construct is 70 residues in length (Figure 5.5), which is approximately a 5 amino acid increase on a fully extended nascent chain, and an intermediate of 75aa increases the level of glycosylation substantially. These results are in agreement data described by Whitley et al (1996) who suggest that, when an 18aa hydrophobic segment is believed to compact in the Sec61 translocon, 75aa are required to span the distance from the PST to the OST. With the first TM domain of

GPR35 being 20 residues in length, compaction within the translocon would delay glycosylation by ~10 residues. As the TM domains of GPCRs are known to form α -helices, it is likely that the first TM domain of GPR35 has compacted to form a helix in preparation for integration into the ER membrane. Indeed, cross-linking assays provide evidence of stabilizing forces within the translocon that could induce helix formation (McCormick et al. 2003). Biophysical assays have also suggested a helical structure will position the most hydrophobic residues towards the lipid bilayer and least hydrophobic towards the polar surface of the translocon channel (Hessa et al 05). Positioning of the α -helical TM domain in this manner, prepares the TM segment for integration into the lipid bilayer of the ER.

The integration process is one of the final steps in IMP biogenesis at the ER membrane. Integration has been studied in a number of model proteins, with various models of integration being suggested (Audigier et al. 1987; Friedlander & Blobel 1985; Ismail et al. 2008). TM domains upon integration into the ER membrane have previously been shown to partition laterally from the Sec61 translocon, through the lateral gate, into the phospholipid environment of the ER (Mothes et al. 1997; McCormick et al. 2003). Previous studies have used chemical and photo-crosslinking techniques to investigate the timing of TM domain integration, demonstrating the movement of nascent chain from the protein environment of the Sec61 channel to the lipid environment of the ER membrane (Meacock et al. 2002; McCormick et al. 2003; Ismail 2006; Ismail et al. 2008; Hou et al. 2012). Using a similar assay to those mentioned above, the movement of GPR35 domains 1 and 2 from the translocon to lipid bilayer was demonstrated by the loss of single cross-linking adducts to Sec61 α . In the case of TM1 of GPR35, we demonstrate by cross-linking with Sec61 α , that it resides in the vicinity translocon pore until the point at which the entire TM2 domain is inserted (~115aa). At this point, cross-links between TM1 and Sec61 α weaken at ~135aa and an interaction can no longer be detected at 155aa, suggesting a relocation of TM1 into the lipid bilayer (Figure 5.6B2). The efficiency of glycosylation at the N-terminal N-glycoylation site may also be a means of measuring the timing at which TM1 is laterally partitioned from the translocon (Figure 5.6). Glycosylation in intermediates 75-125aa is much stronger than that detected in either the 135aa or 155aa intermediate. This may be due to length of RNCs artificially holding the N-terminus of GPR35 in a position where it will be exposed to the OST for a longer period

of time. The 2nd TM domain begins to show cross-linking adducts to Sec61 α at the same time as the TM1 adducts begin to weaken (Figure 5.7A). This suggests that completion of insertion of TM2 into the translocon triggers the relocation of TM1 domain into the lipid bilayer. Previous studies have shown that integration of one TM domain is often dependant on the synthesis and insertion of the subsequent domains (Ismail et al 2006). Cross-links in TM2 remain beyond 135aa, suggesting TM2 remains in the vicinity of the channel and does not get portioned out along with TM (Figure 5.7A2). This could suggest a sequential model of insertion, but subsequent experiments would be required to determine this fully.

6. Final Discussion

6.1 Analysing secondary structure formation within the ribosome tunnel of TM 1 in a model GPCR.

A major aim of this thesis was to investigate the early folding events taking place in the N-terminal 1st TM domain of GPR35. As very little is known about the co-translational folding of these GPCRs, this work strived to shed light on how this major class of eukaryotic membrane protein folds during translation and whether like other IMPs they are likely to form structure in the ribosome tunnel. Using GPR35 as a model GPCR, assays were set up in both prokaryotic and eukaryotic *in vitro* translation systems to determine if the first TM domain was capable of folding, as well as investigating differences that may exist between the tunnel environment of prokaryotic and eukaryotic ribosomes.

Experiments investigating the folding profile of GPR35 were carried out in both prokaryotic and eukaryotic ribosomes (refer to Figures 3.3 and 3.5). Extensive structural studies have suggested a high level of conservation between different ribosome species (Ben-Shem et al. 2010; Jenner et al. 2012; Ban 2000), including that of the role played by the ribosome tunnel in nascent chain compaction (Woolhead et al. 2004; Houben et al. 2005; Bhushan et al. 2010). Following the outcome of studies in both GPR35 and Bacterioopsin, we too believe a high level of conservation exists in the ribosome tunnel, as the experiments discussed below provide nearly identical results in both prokaryotic and eukaryotic ribosomes.

A number of studies have highlighted the role played by the ribosome tunnel in aiding the formation of secondary structure within TM segments of a translating nascent chain (Woolhead et al. 2004; Lu & Deutsch 2005a; Robinson et al. 2012). In particular, several studies have provided evidence for preferred ‘folding zones’ within the tunnel where compaction of the nascent chain takes place. Woolhead et al (2004) were first to provide evidence through energy transfer experiments that the upper tunnel, near the PTC, is one such region where folding can exist. Following the synthesis of the final residues of the VSV-G TM domain in the 111p peptide, compaction of the nascent chain occurred. Photocrosslinking studies were capable of demonstrating a possible interaction between the nascent chain and ribosomal proteins, uL4 and uL22. As these interactions could not be seen with the extended peptide pPL, this suggested that the interaction was driven due to the existence of secondary structure in the upper regions of the tunnel. Studies carried out

by Lu and Deutsch (2005) using a polyalanine nascent chain, also confirmed the upper tunnel as a viable region for the formation of secondary structure and further suggested it was the strongest zone for supporting compaction. The upper tunnel has also been shown to compact TM segments of IMPs, with the C-terminus of TM domain 6 in Kv1.3, a voltage-gated potassium channel, one such example (Lu & Deutsch 2005b). From the work in this thesis, we established that neither GPR35 nor BO, a structurally similar prokaryotic protein, can compact in the upper region of the ribosome tunnel (refer to Figures 3.3-3.6). Based on the structure of the large ribosomal subunit, the exit tunnel is known to be $\sim 100\text{\AA}$ in length from the PTC to the ribosome exit point, both pegylation and cross-linking assays suggest that compaction of GPR35 in the upper tunnel is highly unlikely. Both nascent peptides can be pegylated at $\sim 30\text{aa}$ and cross-linked to uL23 as early as 25aa from the PTC (refer to Figure 4.3). At this point the entire TM1 domain would be synthesised and shows little evidence of secondary structure formation during the early stages of synthesis.

The subsequent regions of the ribosome tunnel have also been assessed for their ability to form compacted nascent chains. Approximately 30\AA from the PTC is a constriction point, which has been shown to be an area of the tunnel where compaction of a nascent chain cannot take place (Lu & Deutsch 2005a; Bhushan et al. 2010). However, it has been suggested that the uL4 and uL22 r-proteins may play a role in stabilizing a structure that had previously formed during the early stages of synthesis. As this is the case, it was highly unlikely that GPR35 or BO would begin to compact in this region of the tunnel. Pegylation and cross-linking data can confirm this, adding to the data already published suggesting the central region of the tunnel is not a viable 'folding zone' (refer to Chapter 4 and 5).

The lower region of the tunnel however, has been highlighted as a 'folding zone' for nascent chains, with both secondary and tertiary structure capable of forming whilst close to the exit point (Lu & Deutsch 2005a; Bhushan et al. 2010; Tu et al. 2014; O. B. Nilsson et al. 2015). This region of the tunnel has been the most frequently described 'folding zone' for secondary structure in TM domains of IMPs and the only region described for the formation of tertiary structure, such as β -hairpins or the complete formation of small proteins (Tu et al. 2014; Nilsson et al. 2015; Marino et al. 2016). Tu and Deutsch (2009) indicate that the lower region of the tunnel is critical for compaction of all six TM helices of the Kv1.3 potassium channel. However, their studies show that individual TM segments do not fold equivocally or in a sequential manner, whilst making their way through the tunnel. Nascent chain compaction within the distal regions of the tunnel has been described

to occur in a number of ways; firstly, some TM domains can begin folding upon entry of their N-terminus into the lower regions of the tunnel (Tu and Deutsch 2009; Bhushan et al 2010; Robinson et al 2012). Secondly, a TM domain can reach the exit port of the ribosome in an extended conformation before beginning to form secondary structure (Houben et al. 2005). Finally, at least one example exists of a TM domain having an extended N-terminus, whilst its C-terminus compacts upon reaching the final 20Å of the ribosome tunnel (Tu & Deutsch 2010). Our investigation into the folding profile of the first TM domain of GPR35, suggested that the nascent chain remains as an extended peptide until the point at which the 1st TM domain reaches the exit port of the ribosome. Pegylation of the N-terminal TM 1 in F_{0c}, previously shown to fold upon entering the lower regions of the ribosome tunnel (Robinson et al 2012), provided us with a control for what would be expected if the 1st TM of GPR35 were fold as it moved into the lower tunnel (refer to Figure 3.7). A delay in pegylation of ~10aa between F_{0c} and GPR35 intermediates, confirmed the first TM of GPR35 exists as an extended peptide as it passes into the lower regions of the tunnel. Subsequent experiments cross-linking GPR35 to uL23 suggested the 1st TM domain remains as an extended peptide as it reaches the ribosome exit (refer to Figure 4.3). At this point we suggest one of two scenarios could occur; 1) TM1 of GPR35 exits the ribosome as an extended peptide, where it interacts with SRP and begins the co-translational targeting process (Figure 6.1A) or 2) the N-terminus GPR35 begins to compact upon arrival at the exit port, possibly forming a helical nascent chain which binds to SRP (Figure 6.1 B).

The two scenarios above, suggest the 1st TM domain of GPR35 could exit the ribosome in two contrasting conformations, thus relying on two entirely different mechanisms of interacting with SRP. A number of studies have highlighted hydrophobicity and α -helical structure as two key features within the nascent chain for enabling an interaction with SRP. The first scenario would require SRP to bind an extended nascent chain as it leaves the exit tunnel, thus relying solely on intrinsic hydrophobicity within the 1st TM domain. A number of biochemical studies cite the importance of hydrophobicity within the nascent chain for recognition by SRP (Hessa et al., 2005, 2007; Nilsson et al., 2014). Indeed, the binding groove of SRP has been suggested to accommodate hydrophobic stretches of nascent chain, which normally correspond to signal sequences and signal anchor domains (Keenan et al. 1998). SRP binding at this point in time enables the safe passage of a hydrophobic stretch of amino acids from the ribosome to the target membrane, preventing aggregation upon exposure to the cytosol. Cross-linking results between GPR35 and Ffh or SRP54 begin as the marker lysine reaches 45aa from the PTC, suggesting the N-terminus of the 1st

TM domain is 40aa from to the PTC (refer to Figures 4.3 and 4.5). Therefore, given the results of previous experiments carried out in GPR35 and the dimensions of the exit tunnel, it would be expected that ~10aa of TM1 would be out with the tunnel and interacting with SRP. However, structural data analysing the interaction between a signal sequence and the M domain of SRP suggests 10 extended residues would be unlikely to fit in the hydrophobic groove (Janda et al. 2010).

To date, most structural data suggests that the binding groove of SRP, which interacts directly with the nascent chain as it exits the ribosome, is likely to house an α -helical peptide of approximately 10aa in length (Keenan et al. 1998; Halic, Blau, et al. 2006; Janda et al. 2010; Voorhees & Hegde 2015). Therefore, based on the structure of SRP, scenario 2 suggesting a compacted N-terminus in the 1st TM of GPR35 would be more favourable. Cross-linking data suggests GPR35 encounters both uL23 and SRP as an extended peptide (refer to Figures 4.3 and 4.5); however cross-links remain between the nascent chain and both proteins for extended periods of time. This would suggest that the nascent chain takes up multiple conformations whilst making its way through the tunnel. This model of TM domain folding in the ribosome tunnel is consistent with previous experiments by Houben et al (2005), in which the TM domain of the Lep protein reached the exit of the ribosome before compacting at the N-terminus. Cryo-EM structures of the SRP54 protein interacting with the signal anchor sequence of the yeast dipeptidyl aminopeptidase B protein, also showed a compacted N-terminal region interacting with SRP54, but no further secondary structure in the remainder of the TM domain (Janda et al. 2010). This suggests that only a short stretch of amino acids require α -helical structure for an interaction with SRP and would agree with this model of GPR35 folding in the ribosome tunnel.

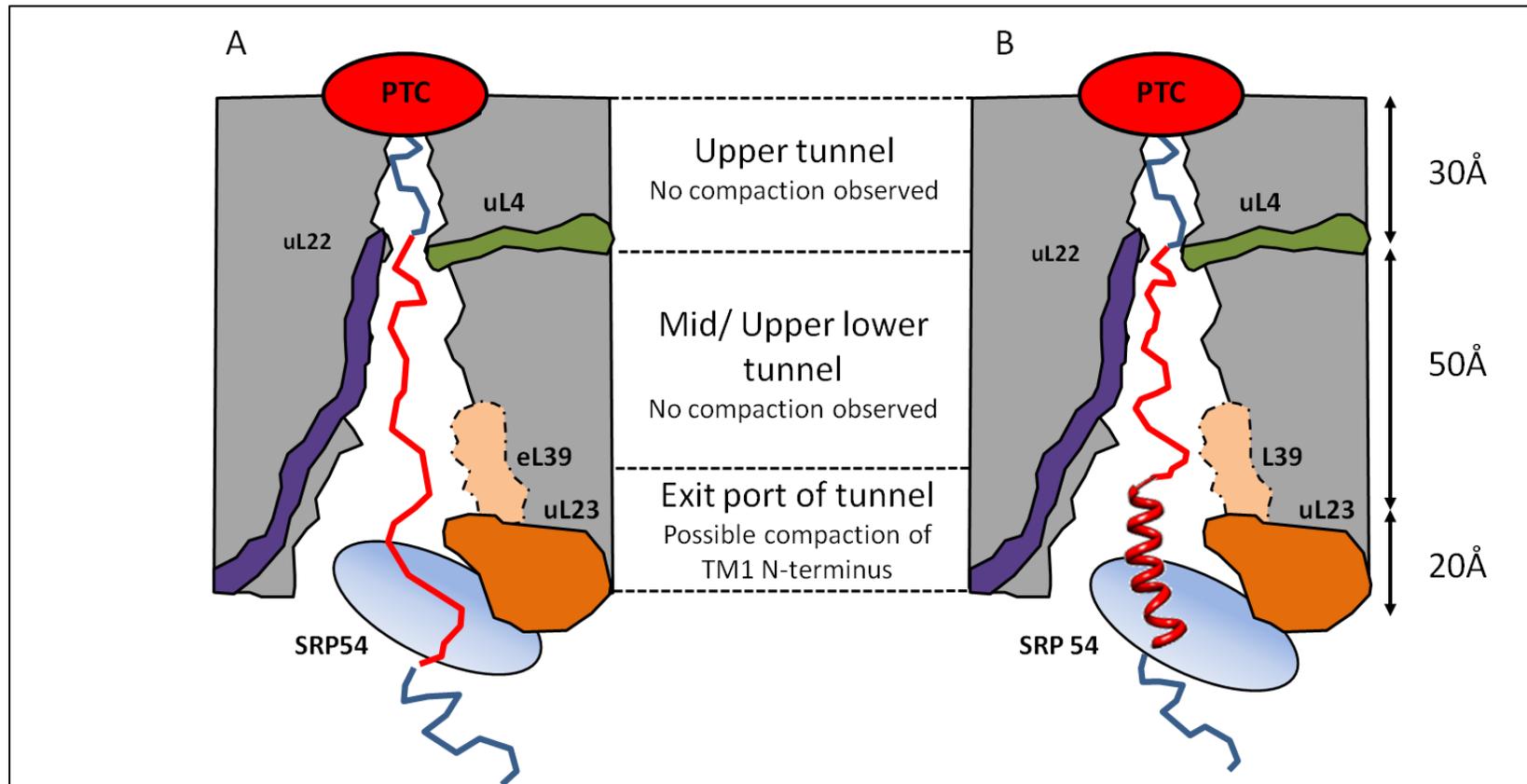


Figure 6.1: Schematic diagram of two possible scenarios describing the folding profile occurring in the 1st TM domain of GPR35 as it makes its way through the ribosome exit tunnel. (A) Scenario 1: suggesting an extended TM1 of GPR35 is capable if interacting with SRP (B) Scenario 2: upon reaching the distal regions of the ribosome tunnel, folding at the N-terminus of GPR35 TM1 occurs and an interaction with SRP takes place. The position of important r-proteins in the tunnel has been specified. The N-terminus or hydrophilic loop regions of the nascent chain are coloured blue, an extended TM1 is a solid red line and a compacted TM1 is represented by a red helix.

6.2 Co-translational targeting and TM domain biogenesis in GPR35.

Following on from the investigations into the early folding events within TM1 of GPR35, subsequent experiments were carried out to study the co-translational targeting and biogenesis of this model GPCR. As with all IMPs, correct targeting and biogenesis is essential for the production of a fully functional protein. Assays analysing GPR35-SRP interactions were set up both in prokaryotic and eukaryotic systems, enabling us to investigate similarities and differences that may arise between the different species of SRP.

Initially we set out to investigate if GPR35 followed the traditional route of IMP targeting to the ER, via the SRP-dependant pathway. Previous studies into the GPCR Opsin have indicated an SRP-dependant mechanism of ER targeting (Audigier et al. 1987; Laird & High 1997). Cross-linking of GPR35 intermediates with both prokaryotic and eukaryotic SRP have confirmed it to can also be targeted in an SRP-dependant manner. Indeed results in the prokaryotic system confirm previous studies (Houben et al. 2005; Bornemann et al. 2008; Robinson et al. 2012), suggesting nascent chains interacting with proteins of the ribosome tunnel are essential for the recruitment of SRP. In these assays we show that the TM domain of GPR35 is in the vicinity of uL23, before a subsequent interaction with SRP (Ffh) occurs (refer to Figure 4.3). This is consistent with previous studies suggesting that uL23 plays a role in recognition of the nascent chain and acts as a docking point for SRP. A prolonged interaction with uL23 may suggest, as earlier discussed, a conformational change within the nascent chain. An α -helical nascent chain has often been suggested as a prerequisite for SRP binding and previous studies have indicated that uL23 may play a key role generating a compacted peptide (Robinson et al 2012). An overlap in cross-linking of the nascent chain from uL23 to SRP highlights the sequential and co-ordinated manner in which the peptide is recognised by the ribosome before its capture by SRP.

The efficient capture of the nascent chain by SRP is required to protect hydrophobically sensitive segments of the nascent chain during the targeting process. For SRP to carry out this role effectively, it must be capable of distinguishing between regions such as hydrophilic loops and TM domains within the nascent chain as they leave the ribosome. Cross-linking of both prokaryotic and eukaryotic SRP to GPR35 intermediates confirm recognition of the 1st TM domain as the signal anchor sequence. Similarities in the cross-linking patterns of SRP to GPR35 in both systems suggest a conserved mechanism for the recognition and binding of the nascent chain (refer to Figures 4.3 and 4.5). Previous studies

investigating the interaction between SRP and signal sequences cited the importance of hydrophobicity within the nascent chain for its recognition by SRP (Nilsson et al. 2014).

As SRP-dependant targeting and translocation is highly reliant on a hydrophobic stretch of residues within the signal sequence/signal anchor of IMPs, alterations to the hydrophobicity of the 1st TM domain of GPR35 was carried out. Previous studies investigating the effect of lowering the hydrophobicity of h-region within a signal sequence found both SRP binding and nascent chain translocation to be reduced (Nilsson et al. 2014). By lowering the hydrophobicity of GPR35 TM1, as described in chapter 4, to the point at which it was no longer recognisable as a TM domain (Δ 4E), resulted in the complete loss of cross-linking to both Ffh and SRP54, as well as a loss in translocated product thus showing a similar trend to the studies in signal sequences (Nilsson et al. 2014). The lack of hydrophobicity within the nascent chain would most likely prevent SRP from recognising it as potential cargo, therefore successful binding and targeting is unlikely to occur (refer to Figures 4.6 and 4.8).

Reducing the hydrophobicity at either the N-terminus (Δ NT) or C-terminus (Δ CT) of the 1st TM domain had substantially different results in binding efficiencies with Ffh and SRP54. In bacterial SRP, the efficiency of cross-linking had decreased to below 50% in both cases in comparison to the WT GPR35. However, reducing the hydrophobicity in the N-terminus of GPR35 had a far more dramatic effect on cross-linking to SRP54 and therefore the subsequent translocation of the N-terminus across the ER membrane. The cross-linking efficiency in the Δ NT intermediate was reduced to ~25% of that of the WT, were as the Δ CT intermediate was only reduced to ~80% (refer to Figure 4.8). These differences in SRP cross-linking efficiencies between species could suggest additional layers of complexity in the recognition and binding process of SRP. Indeed structurally, eukaryotic SRP is known to be a far more complex molecule than bacterial Ffh (Pool 2005), therefore it may contain an added degree of complexity with regards to signal sequence recognition. Equally, the role of the ribosome in recognising a signal sequence for SRP-dependant targeting should not be forgotten. A large number of studies have suggested the importance of cross-talk between the ribosome and SRP as a signal sequence makes its way through the tunnel (Woolhead et al. 2004; Houben et al. 2005; Robinson et al. 2012), therefore it is plausible that a loss in hydrophobicity could impact upon the affinity with which SRP binds to the ribosome, hence reducing the efficiency at which SRP captures the nascent chain. Alternatively, the reduced hydrophobicity within the N-terminus of the nascent chain may encourage the binding of different eukaryotic protein

biogenesis factors that are not available in prokaryotes. Recent research has suggested that the exit site in both the prokaryotic and eukaryotic ribosomes becomes increasingly crowded as the nascent chain emerges; therefore it may be possible another targeting factor or chaperone competes to interact with the less hydrophobic Δ NT intermediate (Kramer et al 2009 rev). Although a number of the above reasons may be possible, it still remains unclear as to what properties, along with hydrophobicity drive the SRP interaction and how this is regulated in the absence of critical driving force such as hydrophobicity.

Following a successful co-translational targeting process, insertion and integration of IMPs take place at the ER membrane, as shown in Chapter 5. In GPCRs, Opsin has been the most frequently studied protein to investigate this process, with a model for its integration being produced in the last 10 years (Ismail 2006; Ismail et al. 2008). However, the biogenesis of polytopic membrane proteins is complex, with individual TM domains being required to complete the formation of their secondary structure during insertion, before being partitioned into the ER membrane. Previous studies have been capable of detecting the folding of peptides within the Sec61 translocon, by using N-linked glycosylation as a marker (Whitley et al. 1996). In the 1st TM of GPR35, the exact timing of when the full formation of secondary structure takes place has been uncertain. Leading on from previous studies regarding N-terminal compaction of GPR35, it is most likely that the 1st TM domain enters the translocon in a partially folded state. However, as the N-terminus of GPR35 enters the translocon, which is known to favour compacted or α -helical peptides (Hessa et al. 2005), it is likely that the nascent chain becomes nucleated and forms secondary structure throughout the entire 1st TM domain (refer to Figure 5.5).

Glycosylation results suggest TM1 of GPR35 is compacted in the translocon, with the timing of the glycosylation event agreeing with the model described by Whitley et al (1996). This is the first point at which the full formation of a compacted TM1 can be detected, most likely preparing the nascent TM for integration into the lipid bilayer.

The final stage in the biogenesis of a polytopic membrane protein is the integration process. Various modes of integration have been described in a number of IMPs and it is clear that much variability exists between individual TM domains. Previous models have shown that TM domains of polytopic membrane proteins can be co-translationally integrated sequentially, in pairs or in bulk (McCormick et al 2005; Ismail et al 2008; Hou et al 2010). As shown in chapter 5, site-specific chemical cross-linking assays detected a possible interaction between TM1 and TM2 of GPR35, and the Sec61 α subunit of the translocon channel (refer to Figure 5.6 B). These assays determined that GPR35 TM1

leaves the translocon channel, upon the insertion of the following TM domain, in what appears to be a sequential mode of integration. This correlates with results previously described for another GPCR, opsin. In this set of experiments Ismail et al (2006) used a similar cross-linking technique to suggest TM1-3 in Opsin, integrates into the lipid bilayer of the ER sequentially. Further experiments by the same group (2008) produced a complete model for the integration of Opsin protein, suggesting TM domains 5-7 integrate in bulk (Ismail et al. 2008). Time restrictions prevented further studies investigating the integration of subsequent GPR35 TM domains; however initial results suggest a similar pattern of integration to opsin could be possible. This would be increasingly interesting as it may suggest model of integration followed specifically by GPCRs.

6.3 Conclusion

The use of GPR35 as a model protein to study the early folding events and biogenesis of a GPCR highlights the complexity of processes such IMPs go through to take up their mature form. The results I have obtained have enabled me to produce what I believe could be a working model for folding and biogenesis of the GPR35 N-terminus (Figure 6.2). The N-terminus containing the first TM domain displays an extended structure until it reaches the distal regions of the ribosome tunnel. Upon interaction with UL23 and SRP a compaction of the nascent chain seems to occur. SRP binds to the 1st TM domain as it emerges from the ribosome and targets of GPR35 to the Sec61 translocon. GPR35 is successfully inserted into the ER translocon, with TM1 taking up a compacted (most likely helical) structure in preparation for integration into the lipid bilayer. Finally, partitioning of GPR35 TM1 into the lipid bilayer occurs due to the synthesis and arrival of TM2 into the translocon, producing a sequential mode of integration at the N-terminus of GPR3

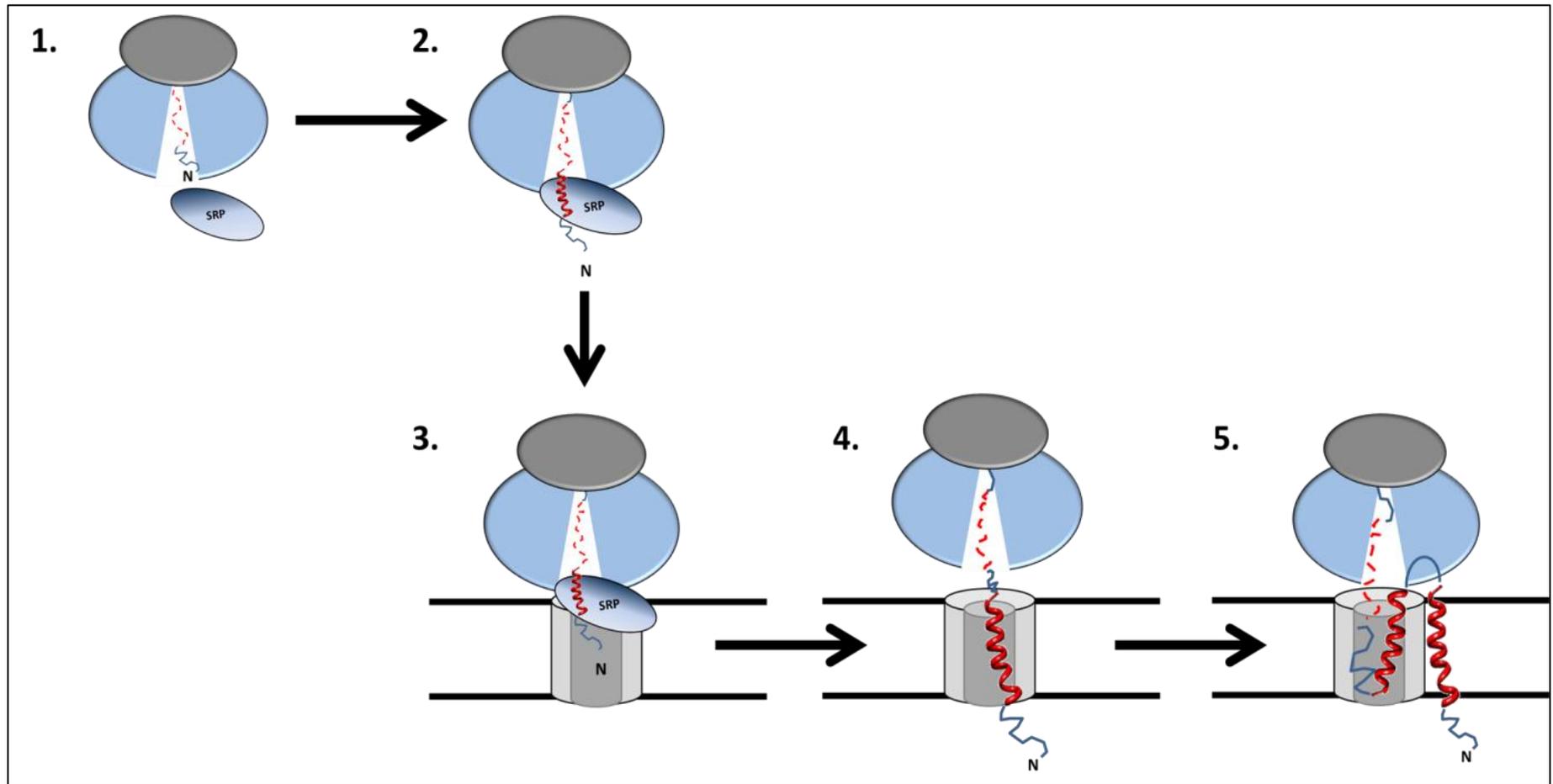


Figure 6.2: Schematic diagram of our suggested model of biogenesis at the N-terminus of GPR35. (1) As the nascent chain is synthesised it elongates through the tunnel in an extended conformation. (2) As the extended N-terminus reaches the exit point of the ribosome, an interaction with uL23 and SRP drives a conformational change. (3) Interaction between the compacted N-terminus of the GPR35 TM1 and SRP drives a co-translation targeting event. (4) The GPR35 TM1 is inserted into the Sec61 translocon, where it takes up its full α -helical conformation. (5) Synthesis and insertion of TM2 into Sec61 drives the lateral partitioning of TM1 from the translocon into the lipid bilayer.

6.4 Future directions

GPCRs are a major source of interest throughout scientific research, particularly in the field of pharmaceuticals, with the diversity of their function, despite the consistency of their structure, a major factor. The work in this thesis has attempted to increase the understanding of the early folding, targeting and biogenesis events in GPCRs, that directly impacts on the formation of their mature 7 TM domain structure. Following on from the research presented in this thesis, a number of directions could be taken to further investigate the early biogenetic processes of a GPCR and in particular GPR35.

Firstly, there is still much work to be done in investigating the folding profiles of individual GPR35 TM domains. Although we have speculated that TM1 may begin to fold in the distal regions of the ribosome and most certainly completes folding upon entry into the Sec61 translocon, the precise timings of such an event cannot be given. With many structural or bio-physical techniques available, such as crystallography, cryo-EM and FRET, the specific timing of TM domain folding could be investigated, as well as the specific interactions taking place between the nascent chain and the ribosome tunnel. Previously, high-resolution structures within the ribosome tunnel have been visualised by crystallography and cryo-EM (Bhushan et al. 2010; O. B. Nilsson et al. 2015), however none have been of the scale of an integral TM domain, which may be a considerable problem in itself. Obtaining clear structural data during the early stages of folding within the WT GPR35 could potentially allow us to investigate known mutations within TM domains that may generate disease phenotypes. This would enable us to discover how big a role the early folding events have in generating a mature GPCR structure.

Secondly, this thesis aimed to investigate the initial integration of the N-terminus of GPR35 from the translocon into the ER membrane. Previous studies investigating the integration of a number of IMPs, have suggested various models for the integration of TM domains (Laird & High 1997; Meacock et al. 2002; Ismail et al. 2008). We have suggested that N-terminus (TM1 followed by TM2) leaves the Sec61 translocon in sequential fashion as it is integrated into the ER membrane. The 1st TM domain relies of the insertion of TM2 in the translocon to aid its partitioning into the lipid bilayer. However, integration into the ER membrane has been shown to differ between TM

domains (Ismail et al. 2008); therefore it is feasible that the remaining TM domains do not follow the mode of integration seen at the N-terminus of GPR35. Producing a more thorough study on TM integration of GPR35 would enable a comparison with other integration models available, hence possibly moving a step closer to providing a known mode of integration for GPCRs as a family.

Finally, the investigation into the biogenesis of a mature GPR35 could be taken further by studying the effect of post-translational modifications whilst in the ER membrane. It has been previously shown that modifications such as glycosylation and disulphide bonds play a vital role in generating the structure of mature IMPS. In GPCRs, a number of studies highlighted the importance of the role played by two conserved cysteines in the extracellular loops (ECL) 1 and 2 (Hwa et al. 2001; Peeters et al. 2012). Investigations have shown that an impaired disulphide bonds between the two cysteines, impacts on the ability of mature receptors to generate their correct conformation in the ER and therefore leads to topological defects, poor trafficking to the plasma membrane and impaired ligand binding. Further experiments investigating the role of these two conserved mutations in GPR35 (work not shown in this thesis) have confirmed the importance of the conserved disulphide bond between ECL 1 and 2. *In vivo* studies showed a mutagenesis of the conserved Cys 89 and 162 to Ala, generated a GPR35 that could no longer be trafficked to the ER membrane. This could lead on to future work investigating if biogenesis at the ER membrane occurs, or if the aberration in the mature GPR35 generates an ER trapped substrate.

Appendix

Appendix 1: Table of constructs

Constructs for the generation of intermediates for use in the S30 *in vitro* transcription/translation system are in the pTrc99a plasmid.

Constructs for the generation of intermediates for use in the RRL translation system are in the pcDNA3.1 plasmid.

Appendix 1.1: Chapter 3 constructs

Construct name	Amino acid change	Original nucleotide sequence	Mutated nucleotide sequence
GPR35 C15	C8A	CTGTGTCCA	CTG GCT CCA
	A15C	ACCGCTCCC	AC CTGT CCC
GPR35Δ4I	L27I	TACTTGGGC	TAC ATT GGC
	L31I	CTGCTGGTG	CTG ATC GTG
	L36I	CTGCTGCTC	CTG ATC CCTC
	L40E	AGCCTGGCG	AGC ATC GCG

Table A1.1. GPR35 construct mutations for pegylation assays.

Construct name	Amino acid change	Original nucleotide sequence	Mutated nucleotide sequence
BO C15	A15C	CAGGCCAGA	CAG TGC AGA
F₀c C5	A5C	CTGAATATG	CTG TGT ATG

Table A1.2. BO and F₀c construct mutations for pegylation assays.

Appendix 1.2: Chapter 4 constructs

Construct name	Amino acid change	Original nucleotide sequence	Mutated nucleotide sequence
BO SRP	R20K	GGACGTCCG	GGAA AA CCG
	K43R	GTAAAGGT	GTT AG AGGT
	K59R	GCGAAAAAA	GCG AG AAA

Table A1.3. BO construct mutations for L23 and SRP cross-linking assays.

Construct name	Amino acid change	Original nucleotide sequence	Mutated nucleotide sequence
GPR35Δ4E	L27E	TACTTGGGC	TAC GAG GGC
	L31E	CTGCTGGTG	CTG GAG GTG
	L36E	CTGCTGCTC	CTG GAG CTC
	L40E	AGCCTGGCG	AGC GAG GCG

Table A1.4. GPR35 Δ NT, Δ CT and Δ 4E construct mutations for L23 and SRP cross-linking assays.

Appendix 1.2: Chapter 5 constructs

Construct name	Amino acid change	Original nucleotide sequence	Mutated nucleotide sequence
N-linked Glycosylation	NS ⁷ T	AACACC	AACT CC ACC
Cysteine BMH cross-linking	L37C (TM1 only)	CTGCTCAAC	CTGT G CAAC
	C46F, C47G	TTCTGCTGCCGC	TTCT TCGG CCGC
	R56C	ACCCGCATC	ACCT G CATC

Table A1.5. GPR35 construct mutations for N-linked glycosylation assays.

Appendix 2: Primers used for the generation of linear DNA intermediates.

Linear DNA for use in S30 *in vitro* transcription/translation system used the pTrc99a forward primer: 5' CTG AAA TGA GCT GTT GAC AAT TAA TCA TCC GG'3

Linear DNA for the generation of mRNA intermediates for use in RRL *in vitro* translation system used the pcDNA 3.1 forward primer: 5' GCA GAG CTC TCT GGC TAA CTA GAG AAC CCA C '3

Appendix 2.1: Chapter 3

Experiment	Primer Name	Primer Sequence (5'--->3')
Pegylation	GPR35 25aa rev	CAGGCTGTTGAGCAGCAGG
	GPR35 30aa rev	GAACACCCAGAGCGCCAGGC
	GPR35 35aa rev	CACTGCTGCATGCGGCAGC
	GPR35 40aa rev	ATGCGGGTCTCCGTCCAC
	GPR35 45aa rev	AGGTTGGTCATGTAGATGC
	GPR35 50aa rev	AGGTCGGCCACCGCCAGGTTGG

Table A2.1. Primers used in to generate GPR35 intermediates for pegylation assays.

Experiment	Primer Name	Primer Sequence (5'--->3')
Pegylation	BO 25aa rev	CATCATGAAGTACAGGGGTGC
	BO 30aa rev	CATCATCATACCTTTAACCAG
	BO 35aa rev	CATCATCGGATCCGAAACACC
	BO 40aa rev	CATCATGAATTTTTTCGCATC
	BO 45aa rev	CATCATGGTGGTGATAGCGTA G
	BO 50aa rev	CATCATGATAGCCGGCACCAC

Table A2.2. Primers used in to generate BO intermediates for pegylation assays.

Experiment	Primer Name	Primer Sequence (5'--->3')
Pegylation	F ₀ c 35aa rev	CGCTGCGCCTTCCAGGAATTTAC
	F0c 45aa rev	GCAGAGGAATCAGATCAG
	F0c 50aa rev	AGAACTGAGTACGCAGCAG
	F0c 55aa rev	GACCCATAACGATAAAGAAC
	F0c 70aa rev	GAACATCACGTACAGACCC

Table A2.3. Primers used in to generate F₀c intermediates for pegylation assays.

Appendix 2.2: Chapter 4

Experiment	Primer Name	Primer Sequence (5'--->3')
SRP cross-linking	GPR35 25aa rev	CATCATCACCCAGAGCGC
	GPR35 35aa rev	CATCATCTCCGTCCACTG
	GPR35 45aa rev	CATCATCACCGCCAGGTTGG
	GPR35 55aa rev	CATCATGGGCAAGGTGCAC
	GPR35 65aa rev	CATCATTGAGGTGTCTCG

Table A2.4. Primers used in to generate GPR35 intermediates for SRP cross-linking assays.

Experiment	Primer Name	Primer Sequence (5'--->3')
SRP cross-linking	BO 25aa rev	CATCATACCTCTAACCAG
	BO 35aa rev	CATCATTTTCTCGCATC
	BO 45aa rev	CATCATAGCCGGGCACCAG
	BO 55aa rev	CATCATCAGCATAGACAG
	BO 65aa rev	CATCATACCGAACGGTAC

Table A2.5. Primers used in to generate BO intermediates for SRP cross-linking assays.

Appendix 2.3: Chapter 5

Experiment	Primer Name	Primer Sequence (5'--->3')
PK Digestion	Full-length GPR35 rev	TTAGGCGAGGGTCACGCAC
Glycosylation	GPR35 182aa rev	AACAGCACCCGGCACAATTCA TACACTGAT ACTAGATCTGATAACAGCACCTTATGTCAA

Table A2.6. Primers used in to generate GPR35 intermediates for PK digestion and C-terminal N-linked glycosylation assays.

Experiment	Primer Name	Primer Sequence (5'--->3')
N-terminal Glycosylation	65aa NT Gly Rev	CGACAGCAGGCAGAGGTC
	70aa NT Gly Rev	GAAGGGCAAGGTGCACAG
	75aa NT Gly Rev	CAGGGAGTGCAGCACGAA
	80aa NT Gly Rev	GTCTGAGGTGTCTCGCAG
	90aa NT Gly Rev	GATGCCCTGGGAGAGCTG
	100aa NT Gly Rev	GCTGATGCTCATGTACC

Table A2.7. Primers used in to generate GPR35 intermediates for N-terminal N-linked glycosylation assays.

Experiment	Primer Name	Primer Sequence (5'--->3')
Integration	Cysteine CL 55aa rev	CATCATCGTCCACTGCTGCAT
	Cysteine CL 75aa (TM1) rev	CATCATCAAGGTGCCCAGCAG
	Cysteine CL 75aa (TM2) rev	CATCATCAAGGTGCACAGCAG
	Cysteine CL 95aa rev	CATCATTGGGAGAGCTGGCAC
	Cysteine CL 115aa rev	CATCATGTCCACGGCGATGGC
	Cysteine CL 135aa rev	CATCATAGCCTGCCTGGG
	Cysteine CL 155aa rev	CATCATCCCCAGGAGCCAGCG

Table A2.8. Primers used in to generate GPR35 intermediates for site specific cysteine cross-linking in TM1 and TM2 of GPR35

Appendix 3: Gene Sequences

GPR35

Translated Sequence

MNGTYNTCGSSDLTWPPAIKLGFYAYLGVLLVLGLLLNSLALWVFCCRMQQWTE
TRIYMTNLA^VADLCLLCTLPFVLHSLRDT^SDTPLCQLSQGIYLTNRYMSISLVTAIA
VDRYVAVRHPLRARGLRSPRQAAAVCAVLWVLVIGSLVARWLLGIQEGGFCFRS
TRHNFNSMAFPLLGFYLP LAVVVFCSLKVVTALAQRPPTDVGQAEATRKAARMV
WANLLVFVVCFLPLHVGLTVRLAVGWNACALLETIRRALYITSKLSDANCCLDIAI
CYYMAKEFQEASALAVAPSAKAHKSQDSL CVTLA

ATGAATGGCACCTACAACACCGCTGGCTCCAGCGACCTCACCTGTCCCCAGCGATCA
AGCTGGGCTTCTACGCCTACTTGGGCGTCTGCTGGTGGCTAGGCCTGCTGCTCAACAGC
CTGGCGCTCTGGGTGTTCTGCTGCCGCATGCAGCAGTGGACGGAGACCCGCATCTACA
TGACCAACCTGGCGGTGGCCGACCTCTGCCTGCTGTGCACCTTGCCCTTCGTGCTGCAC
TCCCTGCGAGACACCTCAGACACGCCGCTGTGCCAGCTCTCCAGGGCATCTACCTGA
CCAACAGGTACATGAGCATCAGCCTGGTCACGGCCATCGCCGTGGACCGCTATGTGGC
CGTGCGGCACCCGCTGCGTGCCCGCGGGCTGCGGTCCCCAGGCAGGCTGCGGCCGTG
TGC GCGTCTCTGGGTGCTGGTCATCGGCTCCCTGGTGGCTCGCTGGCTCCTGGGGAT
TCAGGAGGGCGGCTTCTGCTTCAGGAGCACCCGGCACAATTTCAACTCCATGCGGTT
CCGCTGCTGGGATTCTACCTGCCCTGGCCGTGGTGGTCTTCTGCTCCCTGAAGGTGGT
GACTGCCCTGGCCAGAGGCCACCCACCGACGTGGGGCAGGCAGAGGCCACCCGCAA
GGCTGCCCGCATGGTCTGGGCCAACCTCCTGGTGTTCGTGGTCTGCTTCCTGCCCTGC
ACGTGGGGCTGACAGTGCGCCTCGCAGTGGGCTGGAACGCCTGTGCCCTCCTGGAGAC
GATCCGTCGCGCCCTGTACATAACCAGCAAGCTCTCAGATGCCAACTGCTGCCTGGAC
GCCATCTGCTACTACTACATGGCCAAGGAGTTCAGGAGGCGTCTGCACTGGCCGTGG
CTCCCCGTGCTAAGGCCACAAAAGCCAGGACTCTCTGTGCGTGACCCTCGCCTAA

Bacterioopsin

Translated Sequence

MLELLPTAVEGVSQAQITGRPEWIWLALGTALMGLGTYFLVKGMGVSDPDAKK
FYAITTLVPAIAFTMYLSMMLLGYGLTMVPGGEQNPIYWARYADWLFTTPLLILLD
LALLVDADQGTILALVGADGIMIGTGLVGALTKVYSYRFVWWAISTAAMLYILYV
LFFGFTSKAESMRPEVASTFKVLRNVTVVLWSAYPVVWLIGSEGAGIVPLNIETLL
FMVLDVSAKVGFGILLRSRAIFGEAEPEPSAGDGAATSD

ATGTTGGAGTTATTGCCAACAGCAGTGGAGGGGGTATCGCAGTGCCAGATCACCGGAC
GTCCGGAATGGATCTGGCTAGCTCTGGGCACCGCTCTGATGGGTCTGGGCACCCTGTA
CTTCCTGGTTAAAGGTATGGGTGTTTCGGATCCGGATGCGAAAAAATTCTACGCTATC
ACCACCCTGGTGCCGGCTATCGCATTACCATGTACCTGTCTATGCTGCTGGGTTACGG
TCTGACCATGGTACCGTTCGGTGGTGAACAGAACCCGATCTACTGGGCCCGTTACGCT
GACTGGCTGTTACCAACCCCGCTGCTGCTGCTAGATCTGGCTCTGCTGGTTGACGCTGA
TCAGGGCACCATCCTGGCTCTGGTTGGCGCCGACGGTATCATGATCGGCACCGGCCTG
GTTGGCGCGCTGACCAAGGTTTACTCTTACCGTTTCGTTTGGTGGGCTATCTCTACTGC
AGGCATGCAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGAT
TAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGC
GGTGGTCCCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCCGTAGCGCCGATGGT
AGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACCTGCCAGGCATCAAATAAAACGAAA
GGCTCAGTCGAAAGACTGGGCCTTTCGTTTATCTGTTGTTTGTGCGGTGAACGCTCTCC
TGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACG

F₀C

Translated Sequence

MENLNMDLLYMAAAVMMGLAAIGAAIGIGILGGKFLEGAARQPDLIPLLRTOFFI
VMGLV DAIPMIAVGL GLYVMFAVA

ATGGAAAACCTGAATATGGATCTGCTGTACATGGCTGCCGCTGTGATGATGGGTCTGG
CGGCAATCGGTGCTGCGATCGGTATCGGCATCCTCGGGGGTAAATTCCTGGAAGGCGC
AGCGCGTCAACCTGATCTGATTCCTCTGCTGCGTACTCAGTTCTTTATCGTTATGGGTC
TGGTGGATGCTATCCCGATGATCGCTGTAGGTCTGGGTCTGTACGTGATGTTGCTGTC
CGTGA

Appendix 4: Western blot analysis of antibodies used in chapter 4.

Antibodies used to show the specific interaction of the GPR35 intermediates and components of the SRP targeting pathway in immunoprecipitation assays in chapter 4, can be shown to be specific to the proteins they are raised against using western blot analysis. Samples of S-30, WG and RRL extract were resolved on an SDS-PAGE gel, transferred to nitrocellulose membrane, before marking the target protein using a primary antibody (antibody raised to uL23, Ffh or SRP54) and a secondary antibody (antibody raised to HRP) to visualize by ECL.

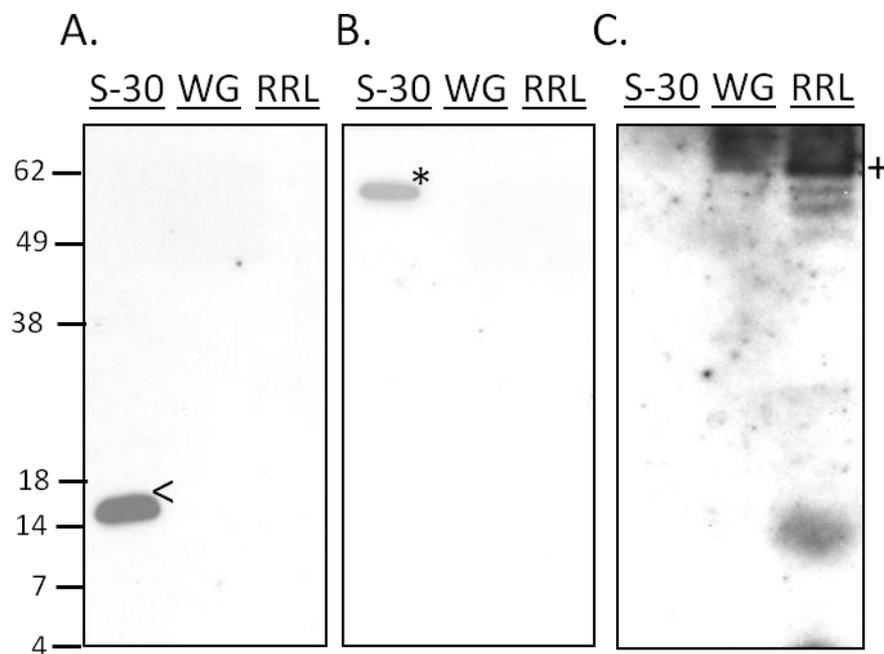


Figure 1. Western blot analysis of components of the SRP mediated pathway in the S-30, WG and RRL extract. The proteins uL23(<), Ffh (*) and SRP54 (+) have been detected using specific antibodies raised to the respective proteins.

Bibliography

- Akopian, D., Shen, K., Zhang, X., & Shan, S. (2013). Signal recognition particle: an essential protein-targeting machine. *Annual Review of Biochemistry*, 82, 693–721.
- Alder, N. N., & Johnson, A. E. (2004). Cotranslational membrane protein biogenesis at the endoplasmic reticulum. *Journal of Biological Chemistry*, 279(22), 22787–22790.
- Alder, N. N., Shen, Y., Brodsky, J. L., Hendershot, L. M., & Johnson, A. E. (2005). The molecular mechanisms underlying BiP-mediated gating of the Sec61 translocon of the endoplasmic reticulum. *Journal of Cell Biology*, 168(3), 389–399.
- Alkalaeva, E. Z., Pisarev, A. V., Frolova, L. Y., Kisselev, L. L., & Pestova, T. V. (2006). In Vitro Reconstitution of Eukaryotic Translation Reveals Cooperativity between Release Factors eRF1 and eRF3. *Cell*, 125(6), 1125–1136.
- Anger, A. M., Armache, J. P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D. N., & Beckmann, R. (2013). Structures of the human and Drosophila 80S ribosome. *Nature*, 497(7447),
- Ataide, S. F., Schmitz, N., Shen, K., Ke, A., Shan, S., Doudna, J. a, & Ban, N. (2011). The crystal structure of the signal recognition particle in complex with its receptor. *Science*, 331(6019), 881–
- Audigier, Y., Friedlander, M., & Blobel, G. (1987). Multiple topogenic sequences in bovine opsin. *Proceedings of the National Academy of Sciences of the United States of America*, 84(16), 5783–5787.
- Bacher, G., Lütcke, H., Jungnickel, B., Rapoport, T. a, & Dobberstein, B. (1996). Regulation by the ribosome of the GTPase of the signal-recognition particle during protein targeting. *Nature* 381(6579) 248-251.
- Ban, N. (2000). The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution. *Science*, 289(5481), 905–920.
- Ban, N., Beckmann, R., Cate, J. H. D., Dinman, J. D., Dragon, F., Ellis, S. R., Lafontaine, D. L. J., Lindahl, L., Liljas, A., Lipton, J. M., McAlear, M. A., Moore, P. B., Noller, H. F., Ortega, J., Panse, V. G., Ramakrishnan, V., Steitz, T. A., Tchorzewski, M., Tollervey, D., Warren, A. J., Williamson, J. R., Wilson, D., Yonath, A., & Yusupov, M. (2014). A new system for naming ribosomal proteins. *Current Opinion in Structural Biology*, 24(1), 165–169.
- Bañó-Polo, M., Baldin, F., Tamborero, S., Marti-Renom, M. A., & Mingarro, I. (2011). N-Glycosylation efficiency is determined by the distance to the C-terminus and the amino acid preceding an Asn-Ser-Thr sequon. *Protein Science*, 20(1), 179–186.
- Batey, R. T., Rambo, R. P., Lucast, L., Rha, B., & Doudna, J. A. (2000). Crystal Structure of the Ribonucleoprotein Core of the Signal Recognition Particle. *Science*, 287(5456), 1232–1239.
- Becker, T., Bhushan, S., Jarasch, A., Armache, J.-P., Funes, S., Jossinet, F., Gumbart, J., Mielke, T., Berninghausen, O., Schulten, K., Westhof, E., Gilmore, R., Mandon, E. C., & Beckmann, R. (2009). Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. *Science*, 326(5958), 1369–1373.
- Beltzer, J. P., Fiedler, K., Fuhrer, C., Geffen, I., Handschin, C., Wessels, H. P., & Spiess, M. (1991). Charged residues are major determinants of the transmembrane orientation of a signal-anchor sequence. *Journal of Biological Chemistry*, 266(2), 973–978.
- Ben-Shem, A., Jenner, L., Yusupova, G., & Yusupov, M. (2010). Crystal Structure of the Eukaryotic Ribosome. *Science*, 330(6008), 1203–1209.
- Berg, B. Van Den, Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C. & Rapoport T. A. (2004). X-ray structure of a protein-conducting channel. *Nature*, 427(6969), 36–44.
- Beringer, M., Bruell, C., Xiong, L., Pfister, P., Bieling, P., Katunin, V. I., Bottager, E. C., & Rodnina, M. V. (2005). Essential mechanisms in the catalysis of peptide bond formation on the ribosome. *Journal of Biological Chemistry*, 280(43), 36065–36072.
- Berndt, U., Oellerer, S., Zhang, Y., Johnson, A. E., & Rospert, S. (2009). A signal-anchor sequence stimulates signal recognition particle binding to ribosomes from inside the exit tunnel. *Proceedings of the National Academy of Sciences of the United States of America*, 106(5), 1398–1403.

- Bhushan, S., Gartmann, M., Halic, M., Armache, J.-P., Jarasch, A., Mielke, T., Berninghausen, O., Wilson, D., & Beckmann, R. (2010). alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nature Structural & Molecular Biology*, *17*(3), 313–317.
- Bibi, E. (2011). Early targeting events during membrane protein biogenesis in Escherichia coli. *Biochimica et Biophysica Acta - Biomembranes*, *1808*(3), 841–850.
- Blaber, M., Zhang, X. J., Lindstrom, J. D., Pepiot, S. D., Baase, W. A., & Matthews, B. W. (1994). Determination of alpha-helix propensity within the context of a folded protein. Sites 44 and 131 in bacteriophage T4 lysozyme. *Journal of Molecular Biology* *235* (2) 600-624
- Blanchard, S. C., Kim, H. D., Gonzalez, R. L., Puglisi, J. D., & Chu, S. (2004). tRNA dynamics on the ribosome during translation. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(35), 12893–12898.
- Blau, M., Mullapudi, S., Becker, T., Dudek, J., Zimmermann, R., Penczek, P. A., & Beckmann, R. (2005). ERj1p uses a universal ribosomal adaptor site to coordinate the 80S ribosome at the membrane. *Nature Structural & Molecular Biology*, *12*(11), 1015–6.
- Blobel, G. (1980). Intracellular protein topogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *77*(3), 1496–500.
- Booth, P. J., & High, S. (2004). Polytopic membrane protein folding and assembly in vitro and in vivo. *Molecular Membrane Biology*, *21*(3), 163–170.
- Bornemann, T., Jöckel, J., Rodnina, M. V., & Wintermeyer, W. (2008). Signal sequence-independent membrane targeting of ribosomes containing short nascent peptides within the exit tunnel. *Nature Structural & Molecular Biology*, *15*(5), 494–499.
- Bousset, L., Mary, C., Brooks, M. a, Bousset, L. U. C., Scherrer, A., Strub, K., & Cusack, S. (2014). Crystal structure of a signal recognition particle Alu domain in the elongation arrest conformation. *Rna*, *20*, 1–8.
- Braakman, I., & Bulleid, N. J. (2011). Protein folding and modification in the Mammalian endoplasmic reticulum. *Annual Review of Biochemistry*, *80*, 71–99.
- Bradshaw, N., Neher, S. B., Booth, D. S., & Walter, P. (2009). Signal sequences activate the catalytic switch of SRP RNA. *Science*, *323*(5910), 127–130.
- Breiman, A., Fioulaine, S., Meinnel, T., & Giglione, C. (2015). The intriguing realm of protein biogenesis: Facing the green co-translational protein maturation networks. *Biochimica et Biophysica Acta*, *1864*(5), 531–550.
- Burda, P., & Aebi, M. (1999). The dolichol pathway of N-linked glycosylation. *Biochimica et Biophysica Acta - General Subjects*, *1426*(2), 239–257.
- Buskiewicz, I. A., Jöckel, J., Rodnina, M. V., & Wintermeyer, W. (2009). Conformation of the signal recognition particle in ribosomal targeting complexes. *RNA* *15*(1) 44-54.
- Calloni, G., Chen, T., Schermann, S. M., Chang, H. C., Genevaux, P., Agostini, F., Tartaglia, G. G., Hayer-Hartl, M., & Hartl, F. U. (2012). DnaK functions as a central hub in the E. coli chaperone network. *Cell Reports*, *1*(3), 251–264.
- Cannon, K. S., Or, E., Clemons, W. M., Shibata, Y., & Rapoport, T. A. (2005). Disulfide bridge formation between SecY and a translocating polypeptide localizes the translocation pore to the center of SecY. *Journal of Cell Biology*, *169*(2), 219–225.
- Clemons Jr., W. M., Brodersen, D. E., McCutcheon, J. P., May, J. L., Carter, A. P., Morgan-Warren, R. J., Wimberly, B. T., & Ramakrishnan, V. (2001). Crystal structure of the 30 S ribosomal subunit from Thermus thermophilus: purification, crystallization and structure determination. *J Mol Biol*, *310*(4), 827–843.
- Clérico, E. M., Szymanska, A., & Gierasch, L. M. (2009). Exploring the interactions between signal sequences and E. coli SRP by two distinct and complementary crosslinking methods. *Biopolymers*, *92*(3), 201–211.
- Connell, S. R., Takemoto, C., Wilson, D. N., Wang, H., Murayama, K., Terada, T., Shirouzu, M., Rost, M., Schuler, M., Giesebrecht, J., Dabrowski, M., Mielke, T., Fucini, P., Yokoyama, S., & Spahn, C. M. T. (2007). Structural Basis for Interaction of the Ribosome with the Switch Regions of GTP-Bound Elongation Factors. *Molecular Cell*, *25*(5), 751–764.
- Conti, B. J., Devaraneni, P. K., Yang, Z., David, L. L., & Skach, W. R. (2015). Cotranslational Stabilization of Sec62/63 within the ER Sec61 Translocon Is Controlled by Distinct Substrate-Driven Translocation Events. *Molecular Cell*, *58*(2), 269–283.

- Conti, B. J., Elferich, J., Yang, Z., Shinde, U., & Skach, W. R. (2014). Cotranslational folding inhibits translocation from within the ribosome-Sec61 translocon complex. *Nature Structural & Molecular Biology*, *21*(3), 228–35.
- Cornish, P. V., Ermolenko, D. N., Noller, H. F., & Ha, T. (2008). Spontaneous intersubunit rotation in single ribosomes. *Molecular Cell*, *30*(5), 578–588.
- Cross, B. C. S., & High, S. (2009). Dissecting the physiological role of selective transmembrane-segment retention at the ER translocon. *Journal of Cell Science*, *122*(11), 1768–1777.
- Crowley, K. S., Liao, S., Worrell, V. E., Reinhart, G. D., & Johnson, A. E. (1994). Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore. *Cell*, *78*(3), 461–471.
- Cruz-Vera, L. R., Sachs, M. S., Squires, C. L., & Yanofsky, C. (2011). Nascent polypeptide sequences that influence ribosome function. *Current Opinion in Microbiology*, *14*(2), 160–166.
- Cserzö, M., Wallin, E., Simon, I., von Heijne, G., & Elofsson, A. (1997). Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Engineering*, *10*(6), 673–676.
- Dallas, A., & Noller, H. F. (2001). Interaction of translation initiation factor 3 with the 30S ribosomal subunit. *Molecular Cell*, *8*(4), 855–864.
- Denks, K., Vogt, A., Sachelaru, I., Petriman, N., Kudva, R., & Koch, H. (2014). The Sec translocon mediated protein transport in prokaryotes and eukaryotes. *Molecular Membrane Biology*, *31*(2-3), 58–84.
- Deshaies, R. J., Sanders, S. L., Feldheim, D. a., & Schekman, R. (1991). Assembly of yeast Sec proteins involved in translocation into the endoplasmic reticulum into a membrane-bound multisubunit complex. *Nature*, *349*(6312), 806–808.
- Dunkle, J. A., Wang, L., Feldman, M. B., Pulk, A., Chen, V. B., Kapral, G. J., Noeske, J., Richardson, J. S., Blanchard, S. C., & Cate, J. H. D. (2011). Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. *Science*, *332*(6032), 981–984.
- Egea, P. F., Stroud, R. M., & Walter, P. (2005). Targeting proteins to membranes: Structure of the signal recognition particle. *Current Opinion in Structural Biology*, *15*(2) 213-220
- Emr, S. D., Hanley-Way, S., & Silhavy, T. J. (1981). Suppressor mutations that restore export of a protein with a defective signal sequence. *Cell*, *23*(1), 79–88.
- Fang, H., & Green, N. (1994). Nonlethal sec71-1 and sec72-1 mutations eliminate proteins associated with the Sec63p-BiP complex from *S. cerevisiae*. *Molecular Biology of the Cell*, *5*(9), 933–942.
- Finke, K., Plath, K., Panzner, S., Prehn, S., Rapoport, T. A., Hartmann, E., & Sommer, T. (1996). A second trimeric complex containing homologs of the Sec61p complex functions in protein transport across the ER membrane of *S. cerevisiae*. *The EMBO Journal*, *15*(7), 1482–94.
- Flanagan, J. J., Chen, J. C., Miao, Y., Shao, Y., Lin, J., Bock, P. E., & Johnson, A. E. (2003). Signal recognition particle binds to ribosome-bound signal sequences with fluorescence-detected subnanomolar affinity that does not diminish as the nascent chain lengthens. *Journal of Biological Chemistry*, *278*(20), 18628–18637.
- Fons, R. D., Bogert, B. A., & Hegde, R. S. (2003). Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. *Journal of Cell Biology*, *160*(4), 529–539.
- Foster, W., Helm, A., Turnbull, I., Gulati, H., Yang, B., Verkman, A. S., & Skach, W. R. (2000). Identification of sequence determinants that direct different intracellular folding pathways for aquaporin-1 and aquaporin-4. *Journal of Biological Chemistry*, *275*(44), 34157–34165.
- Frank, J., & Agrawal, R. K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature*, *406*(6793), 318–322.
- Frank, J., Zhu, J., Penczek, P., Li, Y., Srivastava, S., Verschoor, A, Radermacher, M., Grassucci, R., Lata, R. K., & Agrawal, R. K. (1995). A model of protein synthesis based on cryo-electron microscopy of the *E. coli* ribosome. *Nature* *376*(6539) 441-444.
- Frauenfeld, J., Gumbart, J., Sluis, E. O. van der, Funes, S., Gartmann, M., Beatrix, B., Meilke, T., Berninghausen, O., Schulten, K., & Beckmann, R. (2011). Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. *Nature Structural & Molecular Biology*, *18*(5), 614–621.

- Freistroffer, D. V., Pavlov, M. Y., MacDougall, J., Buckingham, R. H., & Ehrenberg, M. (1997). Release factor RF3 in *E. coli* accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner. *EMBO Journal*, *16*(13), 4126–4133.
- Friedlander, M., & Blobel, G. (1985). Bovine opsin has more than one signal sequence. *Nature*, *318*(6044), 338–343.
- Gerhardy, S., Menet, A. M., Pena, C., Petkowski, J. J., & Panse, V. G. (2014). Assembly and nuclear export of pre-ribosomal particles in budding yeast. *Chromosoma*, *123*(4), 327–344.
- Goder, V., & Spiess, M. (2001). Topogenesis of membrane proteins: Determinants and dynamics. *FEBS Letters*, *504*(3), 87–93.
- Gogala, M., Becker, T., Beatrix, B., Armache, J.-P., Barrio-Garcia, C., Berninghausen, O., & Beckmann, R. (2014). Structures of the Sec61 complex engaged in nascent peptide translocation or membrane insertion. *Nature*, *506*(7486), 107–110.
- Gong, F., & Yanofsky, C. (2002). Instruction of translating ribosome by nascent peptide. *Science (New York, N.Y.)*, *297*(5588), 1864–1867.
- Görlich, D., Hartmann, E., Prehn, S., & Rapoport, T. A. (1992). A protein of the endoplasmic reticulum involved early in polypeptide translocation. *Nature*, *357*(6373), 47–52.
- Görlich, D., & Rapoport, T. A. (1993). Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* *75*(Mdc), 615–630.
- Greber, B. J., Bieri, P., Leibundgut, M., Leitner, A., Aebersold, R., Boehringer, D., & Ban, N. (2015). The complete structure of the 55S ribosome. *Nature*, *524*(7569), 303–308.
- Green, R., & Noller, H. F. (1997). Ribosomes and translation. *Annual Review of Biochemistry*, *66*, 679–716.
- Gross, J. D., Moerke, N. J., Von Der Haar, T., Lugovskoy, A. A., Sachs, A. B., McCarthy, J. E. G., & Wagner, G. (2003). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell*, *115*(6), 739–750.
- Grosshans, H., Deinert, K., Hurt, E., & Simos, G. (2001). Biogenesis of the signal recognition particle (SRP) involves import of SRP proteins into the nucleolus, assembly with the SRP-RNA, and Xpo1p-mediated export. *Journal of Cell Biology*, *153*(4), 745–761.
- Grudnik, P., Bange, G., & Sinning, I. (2009). Protein targeting by the signal recognition particle. *Biological Chemistry*, *390*(8), 775–782.
- Gu, S.-Q. (2003). The signal recognition particle binds to protein L23 at the peptide exit of the *Escherichia coli* ribosome. *RNA*, *9*(5), 566–573.
- Gundelfinger, E. D., Krause, E., Melli, M., & Dobberstein, B. (1983). The organization of the 7SL RNA in the signal recognition particle. *Nucleic Acids Research*, *11*(21), 7363–7374.
- Hainzl, T., Huang, S., Merilainen, G., Brannstrom, K., & Sauer-Eriksson, A. E. (2011). Structural basis of signal-sequence recognition by the signal recognition particle. *Nature Structural Molecular Biology*, *18*(3), 389–391.
- Halic, M., Becker, T., Pool, M. R., Spahn, C. M. T., Grassucci, R. a, Frank, J., & Beckmann, R. (2004). Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature*, *427*(6977), 808–814.
- Halic, M., Blau, M., Becker, T., Mielke, T., Pool, M. R., Wild, K., Sinning, I., & Beckmann, R. (2006a). Following the signal sequence from ribosomal tunnel exit to signal recognition particle. *Nature*, *444*(7118), 507–511.
- Halic, M., Gartmann, M., Schlenker, O., Mielke, T., Pool, M. R., Sinning, I., & Beckmann, R. (2006b). Signal recognition particle receptor exposes the ribosomal translocon binding site. *Science*, *312*(5774), 745–747.
- Hamman, B. D., Hendershot, L. M., & Johnson, A. E. (1998). BiP maintains the permeability barrier of the ER membrane by sealing the luminal end of the translocon pore before and early in translocation. *Cell*, *92*(6), 747–758.
- Harada, Y., Li, H., Li, H., & Lennarz, W. J. (2009). Oligosaccharyltransferase directly binds to ribosome at a location near the translocon-binding site. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(17), 6945–6949.
- Harley, C. A., Holt, J. A., Turner, R., & Tipper, D. J. (1998). Transmembrane protein insertion orientation in yeast depends on the charge difference across transmembrane segments, their total hydrophobicity, and its distribution. *Journal of Biological Chemistry*, *273*(38), 24963–24971.

- Hartmann, E., Görlich, D., Kostka, S., Otto, A., Kraft, R., Knespel, S., Burger, E., Rapoport, T. A., & Prehn, S. (1993). A tetrameric complex of membrane proteins in the endoplasmic reticulum. *European Journal of Biochemistry / FEBS*, *214*(2), 375–381.
- Hartmann, E., Rapoport, T. A., & Lodish, H. F. (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc Natl Acad Sci U S A*, *86*(15), 5786–5790.
- Hartz, D., Binkley, J., Hollingsworth, T., & Gold, L. (1990). Domains of initiator transfer-Rna and initiation codon crucial for initiator transfer-Rna selection by Escherichia-coli If3. *Genes & Development*, *4*(10), 1790–1800.
- Hatsuzawa, K., Tagaya, M., & Mizushima, S. (1997). The hydrophobic region of signal peptides is a determinant for SRP recognition and protein translocation across the ER membrane. *Journal of Biochemistry*, *121*, 270–277.
- Heinrich, S. U., Mothes, W., Brunner, J., & Rapoport, T. A. (2000). The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. *Cell*, *102*(2), 233–44.
- Heinrich, S. U., & Rapoport, T. A. (2003). Cooperation of transmembrane segments during the integration of a double-spanning protein into the ER membrane. *EMBO Journal*, *22*(14), 3654–3663.
- Helmers, J., Schmidt, D., Glavy, J. S., Blobel, G., & Schwartz, T. (2003). The β -subunit of the protein-conducting channel of the endoplasmic reticulum functions as the guanine nucleotide exchange factor for the β -subunit of the signal recognition particle receptor. *Journal of Biological Chemistry*, *278*(26), 23686–23690.
- Hermans, J., Anderson, A. G., & Yun, R. H. (1992). Differential helix propensity of small apolar side chains studied by molecular dynamics simulations. *Biochemistry*, *31*(24), 5646–5653.
- Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S. H., & von Heijne, G. (2005). Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature*, *433*(7024), 377–81.
- Hessa, T., Meindl-Beinker, N. M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., Nilsson, I., White, S. H., & von Heijne, G. (2007). Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature*, *450*(7172), 1026–1030.
- Higy, M., Junne, T., & Spiess, M. (2004). Topogenesis of membrane proteins at the endoplasmic reticulum. *Biochemistry*, *43*(40), 12716–12722.
- Hirashima, A., & Kaji, A. (1973). Role of elongation factor G and a protein factor on the release of ribosomes from messenger ribonucleic acid. *Journal Biological Chemistry*, *248*(21), 7580–7587.
- Hoffmann, A., Becker, A. H., Zachmann-Brand, B., Deuerling, E., Bukau, B., & Kramer, G. (2012). Concerted action of the ribosome and the associated chaperone trigger factor confines nascent polypeptide folding. *Molecular Cell*, *48*(1), 63–74.
- Hou, B., Lin, P. J., & Johnson, A. E. (2012). Membrane protein TM segments are retained at the translocon during integration until the nascent chain cues FRET-detected release into bulk lipid. *Molecular Cell*, *48*(3), 398–408.
- Houben, E. N. G., Zarivach, R., Oudega, B., & Lührink, J. (2005). Early encounters of a nascent membrane protein: Specificity and timing of contacts inside and outside the ribosome. *Journal of Cell Biology*, *170*(1), 27–35.
- Huber, D., Boyd, D., Xia, Y., Olma, M. H., Gerstein, M., & Beckwith, J. (2005). Use of thioredoxin as a reporter to identify a subset of Escherichia coli signal sequences that promote signal recognition particle-dependent translocation. *Journal of Bacteriology*, *187*(9), 2983–2991.
- Hung, D., Falcone, D., Lin, J., Andrews, D. W., & Johnson, A. E. (1996). The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell*, *85*(3), 369–378.
- Hwa, J., Klein-Seetharaman, J., & Khorana, H. G. (2001). Structure and function in rhodopsin: Mass spectrometric identification of the abnormal intradiscal disulfide bond in misfolded retinitis pigmentosa mutants. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(9), 4872–6.
- Ibba, M., Curnow, A. W., & Söll, D. (1997). Aminoacyl-tRNA synthesis: Divergent routes to a common goal. *Trends in Biochemical Sciences*, *22*(2), 39–42.
- Ismail, N. (2006). Active and passive displacement of transmembrane domains both occur during opsin biogenesis at the Sec61 translocon. *Journal of Cell Science*, *119*(13), 2826–2836.

- Ismail, N., Crawshaw, S. G., Cross, B. C. S., Haagsma, A. C., & High, S. (2008). Specific transmembrane segments are selectively delayed at the ER translocon during opsin biogenesis. *The Biochemical Journal*, *411*(3), 495–506.
- Ito, K., Uno, M., & Nakamura, Y. (2000). A tripeptide “anticodon” decipheres stop codons in messenger RNA. *Nature*, *403*(6770), 680–684.
- Jackson, R. J., Hellen, C. U. T., & Pestova, T. V. (2012). *Termination and post-termination events in eukaryotic translation. Advances in Protein Chemistry and Structural Biology* *1*(86). 45–93
- Marino, J., von Heijne, G., & Beckmann, R. (2016). Small domains fold inside ribosome exit tunnel. *Manuscript Submitted and under Publication*, *590*, 655–660.
- Jadhav, B., McKenna, M., Johnson, N., High, S., Sinning, I., & Pool, M. R. (2015). Mammalian SRP receptor switches the Sec61 translocase from Sec62 to SRP-dependent translocation. *Nature Communications*, *6*, 10133.
- Janda, C. Y., Li, J., Oubridge, C., Hernandez, H., Robinson, C. V., & Nagai, K. (2010). Recognition of a signal peptide by the signal recognition particle. *Nature*, *465*(7297), 507–510.
- Jenner, L., Melnikov, S., de Loubresse, N. G., Ben-Shem, A., Iskakova, M., Urzhumtsev, A., Meskauskas, A., Dinman, J., Yusupova, G., & Yusupov, M. (2012). Crystal structure of the 80S yeast ribosome. *Current Opinion in Structural Biology*, *22*(6), 759–767.
- Kaiser, C. M., Chang, H.-C., Agashe, V. R., Lakshminpathy, S. K., Etchells, S. A., Hayer-Hartl, M., Hartl, U., & Barral, J. M. (2006). Real-time observation of trigger factor function on translating ribosomes. *Nature*, *444*(7118), 455–460.
- Kalies, K. U., Rapoport, T. A., & Hartmann, E. (1998). The β subunit of the Sec61 complex facilitates cotranslational protein transport and interacts with the signal peptidase during translocation. *Journal of Cell Biology*, *141*(4), 887–894. 7
- Karamyshev, A. L., Patrick, A. E., Karamysheva, Z. N., Griesemer, D. S., Hudson, H., Tjon-Kon-Sang, S., Nilsson, I., Otto, H., Liu, Q., Rospert, S., von Heijne, G., Johnson, A. E., & Thomas, P. J. (2014). Inefficient SRP interaction with a nascent chain triggers a mRNA quality control pathway. *Cell*, *156*(1-2), 146–157.
- Kauko, A., Hedin, L. E., Thebaud, E., Cristobal, S., Elofsson, A., & von Heijne, G. (2010). Repositioning of Transmembrane α -Helices during Membrane Protein Folding. *Journal of Molecular Biology*, *397*(1), 190–201.
- Keenan, R., & Freymann, D. (2001). The signal recognition particle. *Annual Review of Biochemistry*, *70*, 755–75.
- Keenan, R. J., Freymann, D. M., Walter, P., & Stroud, R. M. (1998). Crystal structure of the signal sequence binding subunit of the signal recognition particle. *Cell*, *94*(2), 181–191.
- Kolupaeva, V. G., Breyne, S. de, Pestova, T. V., & Hellen, C. U. T. (2007). *In Vitro* Reconstitution and biochemical characterization of translation initiation by internal ribosomal entry. *Methods in Enzymology* *430* (7) 409-439
- Kosolapov, A., & Deutsch, C. (2003). Folding of the voltage-gated K⁺ channel T1 recognition domain. *Journal of Biological Chemistry*, *278*(6), 4305–4313.
- Kosolapov, A., & Deutsch, C. (2009). Tertiary interactions within the ribosomal exit tunnel. *Nature Structural & Molecular Biology*, *16*(4), 405–411.
- Kosolapov, A., Tu, L., Wang, J., & Deutsch, C. (2004). Structure acquisition of the T1 domain of Kv1.3 during biogenesis. *Neuron*, *44*(2), 295–307.
- Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, *44*(2), 283–292.
- Kramer, G., Kudlicki, W., McCarthy, D., Tsalkova, T., Simmons, D., & Hardesty, B. (1999). N-terminal and C-terminal modifications affect folding, release from the ribosomes and stability of in vitro synthesized proteins. *International Journal of Biochemistry and Cell Biology*, *31*(1), 231–241.
- Kramer, G., Ramachandiran, V., & Hardesty, B. (2001). Cotranslational folding--omnia mea mecum porto? *The International Journal of Biochemistry & Cell Biology*, *33*(6), 541–553.
- Kroczyńska, B., Evangelista, C. M., Samant, S. S., Elguindi, E. C., & Blond, S. Y. (2004). The SANT2 domain of the murine tumor cell DnaJ-like protein 1 human homologue interacts with α 1-antichymotrypsin and kinetically interferes with its serpin inhibitory activity. *Journal of Biological Chemistry*, *279*(12), 11432–11443.
- Laird, V., & High, S. (1997). Discrete cross-linking products identified during membrane protein biosynthesis. *Journal of Biological Chemistry*, *272*(3), 1983–1989.

- Lakkaraju, a. K. K., Mary, C., Scherrer, A., Johnson, A. E., & Strub, K. (2008). SRP Keeps Polypeptides Translocation-Competent by Slowing Translation to Match Limiting ER-Targeting Sites. *Cell*, *133*(3), 440–451.
- Lang, S., Benedix, J., Fedeles, S. V., Schorr, S., Schirra, C., Schauble, N., Jalal, C., Greiner, M., Haßdenteufel, S., Tatzelt, J., Kreutzer, B., Edelmann, L., Krause, E., Rettig, J., Somlo, S., Zimmermann, R., & Dudek, J. (2012). Different effects of Sec61, Sec62 and Sec63 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. *Journal of Cell Science*, *125*, 1958–1969. doi:10.1242/jcs.096727
- Laurberg, M., Asahara, H., Korostelev, A., Zhu, J., Trakhanov, S., & Noller, H. F. (2008). Structural basis for translation termination on the 70S ribosome. *Nature*, *454*(7206), 852–7.
- Laursen, B., & Sørensen, H. (2005). Initiation of protein synthesis in bacteria. *Microbiology and Molecular Biology Reviews*, *236*(3), 747–771.
- Li, L., Park, E., Ling, J., Ingram, J., Ploegh, H., & Rapoport, T. A. (2016). Crystal structure of a substrate-engaged SecY protein-translocation channel. *Nature*, *531*(7594), 395–399.
- Liao, S., Lin, J., Do, H., & Johnson, A. E. (1997). Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. *Cell*, *90*(1), 31–41.
- Lin, P. J., Jongasma, C. G., Pool, M. R., & Johnson, A. E. (2011). Polytopic membrane protein folding at L17 in the ribosome tunnel initiates cyclical changes at the translocon. *Journal of Cell Biology*, *195*(1), 55–70.
- Lockwood, A. H., Chakraborty, P. R., & Maitra, U. (1971). A complex between initiation factor IF2, guanosine triphosphate, and fMet-tRNA: an intermediate in initiation complex formation. *Proceedings of the National Academy of Sciences of the United States of America*, *68*(12), 3122–6.
- Lu, J., & Deutsch, C. (2005b). Secondary Structure Formation of a Transmembrane Segment in Kv Channels Secondary Structure Formation of a Transmembrane Segment in Kv Channels. *Proteins*, *44*, 8230–8243.
- Lu, J., & Deutsch, C. (2005a). Folding zones inside the ribosomal exit tunnel. *Nature Structural & Molecular Biology*, *12*(12), 1123–1129.
- Lu, J., & Deutsch, C. (2014). Regional Discrimination and Propagation of Local Rearrangements along the Ribosomal Exit Tunnel. *Journal of Molecular Biology*, *426*(24), 4061–4073.
- Lu, J., Kobertz, W. R., & Deutsch, C. (2007). Mapping the electrostatic potential within the ribosomal exit tunnel. *Journal of Molecular Biology*, *371*(5), 1378–1391.
- Lu, Y., Turnbull, I. R., Bragin, a, Carveth, K., Verkman, A S., & Skach, W. R. (2000). Reorientation of aquaporin-1 topology during maturation in the endoplasmic reticulum. *Molecular Biology of the Cell*, *11*(9), 2973–2985.
- Luirink, J. (2004). SRP-mediated protein targeting: structure and function revisited. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1694* (1-3), 17–35.
- Lycklama A Nijeholt, J. A., De Keyzer, J., Prabudiansyah, I., & Driessen, A. J. M. (2013). Characterization of the supporting role of SecE in protein translocation. *FEBS Letters*, *587*(18),
- MacKenzie, A. E., Caltabiano, G., Kent, T. C., Jenkins, L., McCallum, J. E., Hudson, B. D., Nicklin, S. A., Fawcett, L., Markwick, R., Charlton, S. J., & Milligan, G. (2014). The antiallergic mast cell stabilizers lodoxamide and bufrolin as the first high and equipotent agonists of human and rat GPR35. *Molecular Pharmacology*, *85*, 91–104.
- Mackenzie, A. E., & Milligan, G. (2015). The emerging pharmacology and function of GPR35 in the nervous system. *Neuropharmacology*. 1-11
- Malkin, L. I., & Rich, A. (1967). Partial resistance of nascent polypeptide chains to proteolytic digestion due to ribosomal shielding. *Journal of Molecular Biology*, *26*(2), 329–346. doi:10.1016/0022-2836(67)90301-4
- Mccormick, P. J., Miao, Y., Shao, Y., Lin, J., & Johnson, A. E. (2003). Cotranslational protein integration into the ER membrane is mediated by the binding of nascent chains to translocon proteins. *Molecular Cell*, *12*, 329–341.
- Meacock, S. L. Lecomte F. J. L. Cranshaw, S. G. High, S. (2002). Different Transmembrane Domains Associate with Distinct Endoplasmic Reticulum Components during Membrane Integration of a Polytopic Protein. *Molecular Biology of the Cell*, *13*(12), 4114–4129.

- Ménétré, J. F., Hegde, R. S., Heinrich, S. U., Chandramouli, P., Ludtke, S. J., Rapoport, T. A., & Akey, C. W. (2005). Architecture of the ribosome-channel complex derived from native membranes. *Journal of Molecular Biology*, 348(2), 445–457.
- Milligan, G. (2011). Orthologue selectivity and ligand bias: Translating the pharmacology of GPR35. *Trends in Pharmacological Sciences*, 32(5), 317–325.
- Mingarro, I., Nilsson, I., Whitley, P., & von Heijne, G. (2000). Different conformations of nascent polypeptides during translocation across the ER membrane. *BMC Cell Biology*, 1, 3.
- Mitchell, S. F., & Lorsch, J. R. (2008). Should I stay or should I go? Eukaryotic translation initiation factors 1 and 1A control start codon recognition. *Journal of Biological Chemistry*, 283(41), 27345–27349.
- Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C. L., Ban, N., & Frank, J. (2005). Structure of the E. coli protein-conducting channel bound to a translating ribosome. *Nature*, 438(7066), 318–324.
- Mothes, W., Heinrich, S. U., Graf, R., Nilsson, I., von Heijne, G., Brunner, J., & Rapoport, T. A. (1997). Molecular Mechanism of Membrane Protein Integration into the Endoplasmic Reticulum. *Cell*, 89(4), 523–533.
- Muller, L., de Escauriaza, M. D., Lajoie, P., Theis, M., Jung, M., Muller, A., Burgard, C., Greiner, M., Snapp, E. L., Dudek, J., & Zimmermann, R. (2010). Evolutionary gain of function for the ER membrane protein Sec62 from yeast to humans. *Molecular Biology of the Cell*, 21(5), 691–703.
- Nakatogawa, H., & Ito, K. (2002). The Ribosomal Exit Tunnel Functions as a Discriminating Gate. *Cell*, 108, 629–636.
- Nicchitta, C. V., & Blobel, G. (1993). Luminal proteins of the mammalian endoplasmic reticulum are required to complete protein translocation. *Cell*, 73(5), 989–998.
- Nilsson, I., Lara, P., Hessa, T., Johnson, A. E., von Heijne, G., & Karamyshev, A. L. (2015). The code for directing proteins for translocation across ER membrane: SRP cotranslationally recognizes specific features of a signal sequence. *Journal of Molecular Biology*, 427(6), 1191–1201.
- Nilsson, O. B., Hedman, R., Marino, J., Wickles, S., Bischoff, L., Johansson, M., ... von Heijne, G. (2015). Cotranslational Protein Folding inside the Ribosome Exit Tunnel. *Cell Reports*, 12(10), 1533-1540
- Nissen, P. (2000). The Structural Basis of Ribosome Activity in Peptide Bond Synthesis. *Science*, 289(5481), 920–930.
- Nyathi, Y., Wilkinson, B. M., & Pool, M. R. (2013). Co-translational targeting and translocation of proteins to the endoplasmic reticulum. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1833(11), 2392–2402.
- O'Brien, E. P., Hsu, S. T. D., Christodoulou, J., Vendruscolo, M., & Dobson, C. M. (2010). Transient tertiary structure formation within the ribosome exit port. *Journal of the American Chemical Society*, 132(47), 16928–16937.
- O'Dowd, B. F., Nguyen, T., Marchese, A., Cheng, R., Lynch, K. R., Heng, H. H., Kolakowski, L. F., George, S. R. (1998). Discovery of three novel G-protein-coupled receptor genes. *Genomics*, 47(2), 310–3.
- Ogle, J. M., Brodersen, D. E., Clemons, W. M., Tarry, M. J., Carter, a P., & Ramakrishnan, V. (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science*, 292(5518), 897–902.
- Öjemalm, K., Halling, K. K., Nilsson, I., & Von Heijne, G. (2012). Orientational preferences of neighboring helices can drive ER insertion of a marginally hydrophobic transmembrane helix. *Molecular Cell*, 45(4), 529–540.
- Okumura, S. I., Baba, H., Kumada, T., Nanmoku, K., Nakajima, H., Nakane, Y., Hioki, N., & Ikenaka, K. (2004). Cloning of a G-protein-coupled receptor that shows an activity to transform NIH3T3 cells and is expressed in gastric cancer cells. *Cancer Science*, 95(2), 131–135.
- Osborne, A. R., & Rapoport, T. A. (2007). Protein Translocation Is Mediated by Oligomers of the SecY Complex with One SecY Copy Forming the Channel. *Cell*, 129(1), 97–110.
- Oubridge, C., Kuglstatter, A., Jovine, L., & Nagai, K. (2002). Crystal structure of SRP19 in complex with the S domain of SRP RNA and its implication for the assembly of the signal recognition particle. *Molecular Cell*, 9(6), 1251–1261.

- Pace, C. N., & Scholtz, J. M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. *Biophysical Journal*, 75(1), 422–427.
- Passmore, L. A., Schmeing, T. M., Maag, D., Applefield, D. J., Acker, M. G., Algire, M., ... Ramakrishnan, V. (2007). The Eukaryotic Translation Initiation Factors eIF1 and eIF1A Induce an Open Conformation of the 40S Ribosome. *Molecular Cell*, 26(1), 41–50.
- Patterson, M. A., Bandyopadhyay, A., Devaraneni, P. K., Woodward, J., Rooney, L., Yang, Z., & Skach, W. R. (2015). The ribosome-Sec61 translocon complex forms a cytosolically restricted environment for early polytopic membrane protein folding. *Journal of Biological Chemistry*, 290(48), 28944–28952.
- Peeters, M. C., Wisse, L. E., Dinaj, a., Vroling, B., Vriend, G., & Ijzerman, A. P. (2012). The role of the second and third extracellular loops of the adenosine A1 receptor in activation and allosteric modulation. *Biochemical Pharmacology*, 84(1), 76–87.
- Pestova, T. V., Borukhov, S. I., & Hellen, C. U. (1998). Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature*, 394(6696), 854–859.
- Pestova, T. V., & Hellen, C. U. (2001). Preparation and activity of synthetic unmodified mammalian tRNAⁱ(Met) in initiation of translation in vitro. *Rna*, 7(10), 1496–1505.
- Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I., & Hellen, C. U. (2001). Molecular mechanisms of translation initiation in eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, 98(13), 7029–36.
- Pestova, T. V., Lomakin, I. B., Lee, J. H., Choi, S. K., Dever, T. E., & Hellen, C. U. (2000). The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature*, 403(6767), 332–335.
- Peterson, J. H., Woolhead, C. a., & Bernstein, H. D. (2010). The conformation of a nascent polypeptide inside the ribosome tunnel affects protein targeting and protein folding. *Molecular Microbiology*, 78(1), 203–217.
- Petrelli, D., La Teana, A., Garofalo, C., Spurio, R., Pon, C. L., & Gualerzi, C. O. (2001). Translation initiation factor IF3: Two domains, five functions, one mechanism? *EMBO Journal*, 20(16), 4560–4569.
- Pfeffer, S., Burbaum, L., Unverdorben, P., Pech, M., Chen, Y., Zimmermann, R., Beckmann, R., & Förster, F. (2015). Structure of the native Sec61 protein-conducting channel. *Nature Communications*, 6, 8403 1-7.
- Pfeffer, S., Dudek, J., Gogala, M., Schorr, S., Linxweiler, J., Lang, S., Becker, T., Beckmann, R., Zimmermann, R., & Förster, F. (2014). Structure of the mammalian oligosaccharyl-transferase complex in the native ER protein translocon. *Nature Communications*, 5, 3072.
- Pfeiffer, N. V., Dirndorfer, D., Lang, S., Resenberger, U. K., Restelli, L. M., Hemion, C., Miesbauer, M., Frank, S., Neutzer, Zimmermann, R., Winklhofer & Tatzelt, J. (2013). Structural features within the nascent chain regulate alternative targeting of secretory proteins to mitochondria. *The EMBO Journal*, 32(7), 1036–51.
- Pisarev, A. V., Skabkin, M. A., Pisareva, V. P., Skabkina, O. V., Rakotondrafara, A. M., Hentze, M. W., ... Pestova, T. V. (2010). The Role of ABCE1 in Eukaryotic Posttermination Ribosomal Recycling. *Molecular Cell*, 37(2), 196–210. 4
- Plath, K., Mothes, W., Wilkinson, B. M., Stirling, C. J., & Rapoport, T. A. (1998). Signal sequence recognition in posttranslational protein transport across the yeast ER membrane. *Cell*, 94(6), 795–807.
- Politz, J. C., Yarovoi, S., Kilroy, S. M., Gowda, K., Zwieb, C., & Pederson, T. (2000). Signal recognition particle components in the nucleolus. *Proceedings of the National Academy of Sciences of the United States of America*, 97(1), 55–60.
- Pool, M. R. (2005). Signal recognition particles in chloroplasts, bacteria, yeast and mammals. *Molecular Membrane Biology*, 22(1-2), 3–15.
- Pool, M. R. (2009). A trans-membrane segment inside the ribosome exit tunnel triggers RAMP4 recruitment to the Sec61p translocase. *Journal of Cell Biology*, 185(5), 889–902.
- Powers, T., & Walter, P. (1995). Reciprocal stimulation of GTP hydrolysis by two directly interacting GTPases. *Science*, 269(C), 1422–1424.
- Rapoport, T. A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature*, 450(7170), 663–9.
- Robinson, P. J., Findlay, J. E., & Woolhead, C. A. (2012). Compaction of a prokaryotic signal-anchor transmembrane domain begins within the ribosome tunnel and is stabilized by SRP during targeting. *Journal of Molecular Biology*, 423(4), 600–612.

- Rodnina, M. V., Pape, T., Fricke, R., Kuhn, L., & Wintermeyer, W. (1996). Initial binding of the elongation factor Tu-GTP-aminoacyl-tRNA complex preceding codon recognition on the ribosome. *Journal of Biological Chemistry*, 271(2), 646–652.
- Rodnina, M. V., Savelsbergh, A., Katunin, V. I., & Wintermeyer, W. (1997). Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature*, 385(6611), 37–41.
- Rorbach, J., Richter, R., Wessels, H. J., Wydro, M., Pekalski, M., Farhoud, M., Kuhl, I., Gaisne, M., Bonnefoy, N., Smeitink, J. A., Lightowers, R. N., & Chrzanoska-Lightowers, Z. M. A. (2008). The human mitochondrial ribosome recycling factor is essential for cell viability. *Nucleic Acids Research*, 36(18), 5787–5799.
- Rosendal, K. R., Sinning, I., & Wild, K. (2004). Crystallization of the crenarchaeal SRP core. *Acta Crystallographica Section D: Biological Crystallography*, 60(1), 140–143.
- Rothblatt, J. A., Deshaies, R. J., Sanders, S. L., Daum, G., & Schekman, R. (1989). Multiple genes are required for proper insertion of secretory proteins into the endoplasmic reticulum in yeast. *The Journal of Cell Biology*, 109(6), 2641–2652.
- Sadlish, H., Pitzonzo, D., Johnson, A. E., & Skach, W. R. (2005). Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein. *Nature Structural & Molecular Biology*, 12(10), 870–878.
- Sato, M., Hresko, R., & Mueckler, M. (1998). Testing the charge difference hypothesis for the assembly of a eucaryotic multispanning membrane protein. *Journal of Biological Chemistry*, 273(39), 25203–25208.
- Savelsbergh, A., Katunin, V. I., Mohr, D., Peske, F., Rodnina, M. V., & Wintermeyer, W. (2003). An Elongation Factor G-Induced Ribosome Rearrangement Precedes tRNA-mRNA Translocation. *Molecular Cell*, 11(6), 1517–1523.
- Schaffitzel, C., Oswald, M., Berger, I., Ishikawa, T., Abrahams, J. P., Koerten, H. K., Koning, R. I., & Ban, N. (2006). Structure of the E. coli signal recognition particle bound to a translating ribosome. *Nature*, 444(7118), 503–506.
- Schäuble, N., Lang, S., Jung, M., Cappel, S., Schorr, S., Ulucan, Ö., Linxweiler, J., Dudek, J., Blum, R., Helms, V., Paton, A. W., Paton, J., Cavalie, A., & Zimmermann, R. (2012). BiP-mediated closing of the Sec61 channel limits Ca²⁺ leakage from the ER. *The EMBO Journal*, 31(15), 3282–96.
- Schmeing, T. M., Voorhees, R.M., Kelley, A. C., Gao, Y., Murphy IV, F. V., Weir, J. R., & Ramakrishnan, V. (2010). The Crystal Structure of the Ribosome. *Risk Management*, 688(October), 688–695.
- Schmeing, T. M., & Ramakrishnan, V. (2009). What recent ribosome structures have revealed about the mechanism of translation. *Nature*, 461(7268), 1234–1242.
- Schütz, P., Bumann, M., Oberholzer, A. E., Bieniossek, C., Trachsel, H., Altmann, M., & Baumann, U. (2008). Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 105(28), 9564–9.
- Seidelt, B., Innis, C. A., Wilson, D. N., Gartmann, M., Armache, J.-P., Villa, E., ... Beckmann, R. (2009). Structural insight into nascent polypeptide chain-mediated translational stalling. *Science (New York, N.Y.)*, 326(5958), 1412–1415.
- Shan, S. O., Chandrasekar, S., & Walter, P. (2007). Conformational changes in the GTPase modules of the signal reception particle and its receptor drive initiation of protein translocation. *Journal of Cell Biology*, 178(4), 611–620.
- Shan, S., Stroud, R. M., & Walter, P. (2004). Mechanism of association and reciprocal activation of two GTPases. *PLoS Biology*, 2(10), 572–582.
- Shao, S., & Hegde, R. S. (2011). Membrane protein insertion at the endoplasmic reticulum. *Annual Review of Cell and Developmental Biology*, 27, 25–56.
- Shatkin, A. J. (1976). Capping of eucaryotic mRNAs. *Cell*, 9(4), 645–653.
- Sheets, M. D., & Wickens, M. (1989). Two phases in the addition of a poly (A) tail. *Genes & Development*, 1401–1412.
- Shen, K., Arslan, S., Akopian, D., Ha, T., & Shan, S. (2012). Activated GTPase movement on an RNA scaffold drives co-translational protein targeting. *Nature*, 492(7428), 271–5.
- Shibatani, T., David, L. L., McCormack, A. L., Frueh, K., & Skach, W. R. (2005). Proteomic analysis of mammalian oligosaccharyltransferase reveals multiple subcomplexes that contain Sec61, TRAP, and two potential new subunits. *Biochemistry*, 44(16), 5982–5992.
- Shine, J., & Dalgarno, L. (1974). The 3'-terminal sequence of Escherichia coli 16S ribosomal

- RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America*, 71(4), 1342–6.
- Sievers, A., Beringer, M., Rodnina, M. V., & Wolfenden, R. (2004). The ribosome as an entropy trap.. *Proceedings of the National Academy of Sciences of the United States of America* 101(21), 7897–901.
- Sommer, N., Junne, T., Kalies, K. U., Spiess, M., & Hartmann, E. (2013). TRAP assists membrane protein topogenesis at the mammalian ER membrane. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1833(12), 3104–3111.
- Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell*, 136(4), 731–745.
- Spahn, C. M. T., Beckmann, R., Eswar, N., Penczek, P. A., Sali, A., Blobel, G., & Frank, J. (2001). Structure of the 80S ribosome from *Saccharomyces cerevisiae* - tRNA-ribosome and subunit-subunit interactions. *Cell*, 107(3), 373–386.
- Thomas, Y., Bui, N., & Strub, K. (1997). A truncation in the 14 kDa protein of the signal recognition particle leads to tertiary structure changes in the RNA and abolishes the elongation arrest activity of the particle. *Nucleic Acids Research*, 25(10), 1920–1929.
- Tsukazaki, T., Mori, H., Fukai, S., Ishitani, R., Mori, T., Dohmae, N., Perederina, A., Sugita, Y., Vassilyev, D. G., Ito, K., & Nureki, O. (2008). Conformational transition of Sec machinery inferred from bacterial SecYE structures. *Nature*, 455(7215), 988–91.
- Tu, L., Khanna, P., & Deutsch, C. (2014). Transmembrane segments form tertiary hairpins in the folding vestibule of the ribosome. *Journal of Molecular Biology*, 426(1), 185–198.
- Tu, L. W., & Deutsch, C. (2010). A Folding Zone in the Ribosomal Exit Tunnel for Kv1.3 Helix Formation. *Journal of Molecular Biology*, 396(5), 1346–1360.
- Tu, L., Wang, J., Helm, A., Skach, W. R., & Deutsch, C. (2000). Transmembrane biogenesis of Kv1.3. *Biochemistry*, 39(4), 824–836.
- Tyedmers, J., Lerner, M., Wiedmann, M., Volkmer, J., & Zimmermann, R. (2003). Polypeptide-binding proteins mediate completion of co-translational protein translocation into the mammalian endoplasmic reticulum. *EMBO Reports*, 4(5), 505–10.
- Ullers, R. S., Houben, E. N. G., Raine, A., Ten Hagen-Jongman, C. M., Ehrenberg, M., Brunner, J., Oudega, B., & Lührink, J. (2003). Interplay of signal recognition particle and trigger factor at L23 near the nascent chain exit site on the *Escherichia coli* ribosome. *Journal of Cell Biology*, 161(4), 679–684.
- Ulmschneider, M. B., Ulmschneider, J. P., Schiller, N., Wallace, B. a., von Heijne, G., & White, S. H. (2014). Spontaneous transmembrane helix insertion thermodynamically mimics translocon-guided insertion. *Nature Communications*, 5, 4863.
- Van Der Laan, M., Bechduft, P., Kol, S., Nouwen, N., & Driessen, A. J. M. (2004). F1F0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. *Journal of Cell Biology*, 165(2), 213–222.
- Voigt, S., Jungnickel, B., Hartmann, E., & Rapoport, T. A. (1996). Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. *Journal of Cell Biology*, 134(1), 25–35.
- von Heijne, G. (1985). Signal sequences. *Journal of Molecular Biology*, 184(1), 99–105.
- von Heijne, G. (1990). The Signal peptide. *J Membr Biol*, 115, 195–201.
- von Heijne, G. (1992). Membrane protein structure prediction. *Journal of Molecular Biology*, 225(2), 487–494.
- Voorhees, R. M., & Hegde, R. S. (2015). Structures of the scanning and engaged states of the mammalian SRP-ribosome complex. *eLife*, 4, 1–21.
- Voss, N. R., Gerstein, M., Steitz, T. a., & Moore, P. B. (2006). The geometry of the ribosomal polypeptide exit tunnel. *Journal of Molecular Biology*, 360(4), 893–906.
- Wang, J., & Purisima, E. O. (1996). Analysis of thermodynamic determinants in helix propensities of nonpolar amino acids through a novel free energy calculation. *Journal of the American Chemical Society*, 118(5), 995–1001.
- Wang, S., Sakai, H., & Wiedmann, M. (1995). NAC covers ribosome-associated nascent chains thereby forming a protective environment for regions of nascent chains just emerging from the peptidyl transferase center. *Journal of Cell Biology*, 130(3), 519–528.
- Watson, H. R., Wunderley, L., Andreou, T., Warwicker, J., & High, S. (2013). Reorientation of the first signal-anchor sequence during potassium channel biogenesis at the Sec61 complex. *Biochemical Journal*, 456(2), 297–309.

- Weichenrieder, O., Stehlin, C., Kapp, U., Birse, D. E., Timmins, P. A., Strub, K., & Cusack, S. (2001). Hierarchical assembly of the Alu domain of the mammalian signal recognition particle. *RNA*, 7(5), 731–40.
- Weixlbaumer, A., Jin, H., Neubauer, C., Voorhees, R. M., Petry, S., Kelley, A. C., & Ramakrishnan, V. (2008). Insights into Translational Termination from the Structure of RF2 Bound to the Ribosome. *Science*, 322(5903), 953–956.
- White, S. H., & Von Heijne, G. (2005). Transmembrane helices before, during, and after insertion. *Current Opinion in Structural Biology*, 15(4), 378–386.
- Whitley, P., Nilsson, I., & Von Heijne, G. (1996). A nascent secretory protein may traverse the ribosome/endoplasmic reticulum translocase complex as an extended chain. *Journal of Biological Chemistry*, 271(11), 6241–6244.
- Wild, K., Halic, M., Sinning, I., & Beckmann, R. (2004). SRP meets the ribosome. *Nature Structural & Molecular Biology*, 11(11), 1049–1053.
- Wilkinson, B. M., Brownsword, J. K., Mousley, C. J., & Stirling, C. J. (2010). Sss1p is required to complete protein translocon activation. *Journal of Biological Chemistry*, 285(42), 32671–32677. d
- Wimberly, B. T., Brodersen, D. E., Clemons Jr, W. M., Morgan-Warren, R. J., Carter, A. P., Vornrhein, C., Hartsch, T., & Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. *Nature*, 407, 327–339.
- Woolhead, C. A., Johnson, A. E., & Bernstein, H. D. (2006). Translation arrest requires two-way communication between a nascent polypeptide and the ribosome. *Molecular Cell*, 22(5), 587–598.
- Woolhead, C. A., McCormick, P. J., & Johnson, A. E. (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.
- Yam, A. Y., Xia, Y., Lin, H.-T. J., Burlingame, A., Gerstein, M., & Frydman, J. (2008). Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. *Nature Structural & Molecular Biology*, 15(12), 1255–62.
- Yonath, A., Leonard, K. R., & Wittmann, H. G. (1987). A tunnel in the large ribosomal subunit revealed by three-dimensional image reconstruction. *Science*, 236(15), 813–816.
- Yosef, I., Bochkareva, E. S., & Bibi, E. (2010). Escherichia coli SRP, Its Protein Subunit Ffh, and the Ffh M Domain Are Able To Selectively Limit Membrane Protein Expression When Overexpressed. *mBio*, 1(2), e00020–10–e00020–16.
- Yusupova, G. Z., Yusupov, M. M., Cate, J. H. D., & Noller, H. F. (2001). The path of messenger RNA through the ribosome. *Cell*, 106(2), 233–241.
- Zavialov, A. V., Buckingham, R. H., & Ehrenberg, M. (2001). A posttermination ribosomal complex is the guanine nucleotide exchange factor for peptide release factor RF3. *Cell*, 107(1), 115–124.
- Zavialov, A. V., Haurlyliuk, V. V., & Ehrenberg, M. (2005). Guanine-nucleotide exchange on ribosome-bound elongation factor G initiates the translocation of tRNAs. *Journal of Biology*, 4(2), 9.
- Zhang, D., & Shan, S. O. (2012). Translation elongation regulates substrate selection by the signal recognition particle. *Journal of Biological Chemistry*, 287(10), 7652–7660.
- Zhang, X., Kung, S., & Shan, S. O. (2008). Demonstration of a Multistep Mechanism for Assembly of the SRP-SRP Receptor Complex: Implications for the Catalytic Role of SRP RNA. *Journal of Molecular Biology*, 381(3), 581–593.
- Zheng, N., & Gierasch, L. M. (1996). Signal sequences: The same yet different. *Cell*, 86(6), 849–852.
- Ziv, G., Haran, G., & Thirumalai, D. (2005). Ribosome exit tunnel can entropically stabilize alpha-helices. *Proceedings of the National Academy of Sciences of the United States of America*, 102(52), 18956–18961.
- Zopf, D., Bernstein, H. D., Johnson, A. E., & Walter, P. (1990). The methionine-rich domain of the 54 kd protein subunit of the signal recognition particle contains an RNA binding site and can be crosslinked to a signal sequence. *The EMBO Journal*, 9(13), 4511–7.
- Zucca, P., & Sanjust, E. (2014). Inorganic materials as supports for covalent enzyme immobilization: Methods and mechanisms. *Molecules*, 19(9), 14139–14194.

