The molecular basis for ER tubule formation

Author:
Jacob Peter Brady

Supervisor:
Prof. Jason R. Schnell

A thesis submitted for the degree of Doctor of Philosophy

Department of Biochemistry

Green Templeton College

January 2015

"We can only see a short distance ahead,
but we can see plenty there that needs to be done."

Alan Turing

Abstract

The molecular basis for
ER tubule formation

by Jacob Peter Brady

Integral membrane proteins of the DP1 and reticulon families are responsible for maintaining the high membrane curvature required for both smooth ER tubules and the edges of ER sheets. Mutations in these proteins lead to motor neurone diseases such as hereditary spastic paraplegia. Reticulon/DP1 proteins contain Reticulon Homology Domains (RHD) that have unusually long (≈30 aa) hydrophobic segments and are proposed to adopt intramembrane helical hairpins that stabilise membrane curvature. I have uncovered the secondary structure and dynamics of the DP1 protein Yop1p and identified a C-terminal conserved amphipathic helix that on its own interacts strongly with negatively charged membranes and is necessary for membrane tubule formation. Analyses of DP1 and reticulon family members indicate that most, if not all, contain C-terminal sequences capable of forming amphipathic helices. Together, these results indicate that amphipathic helices play a previously unrecognised role in RHD membrane curvature stabilisation. This work paves the way towards full structure determination of Yop1p by solution state NMR and marks the first high structural resolution study on an RHD protein.

Acknowledgements

Thanks to Jason Schnell for encouragement, supervision and teaching me good NMR data processing habits.

Jolyon Claridge for his supreme wisdom and mastery of all things NMR. You also encouraged and then fuelled my python addiction.

Jose you were a great companion during all those late nights in the lab in the early days.

Felipe and Nader for being great lab mates!

Antonio De Riso and Errin Johnson for teaching me so much about electron microscopy

Peter Smith for his work on the Yop1p amphipathic helix

Nick Soffe and Christina Redfield for help and training on Omega spectrometers.

Thanks to Vlado Gelev at Frankenstein Biotech (FBReagents) for making d_{27}-LMPG which made many of my experiments possible.

Meeting Hari Arthanari at Harvard Medical School was an enlightening experience. He sparked my interest in and encouraged me to implement NUS.

Sven Hyberts for Poisson gap sampling and writing hmsIST, thus making NUS accessible to mere mortals.

Kate, I'll miss our tea breaks. You are great and I hope I get the pleasure of doing science with you again one day.

Narin, Emma and Marie-Laure - you are inspirational people.

Kurt Drickamer and Maureen Taylor, without whose excellent advice and support I probably wouldn't be here in the first place!

Laura Gosset for love, support and keeping balance in my life.

My Mum and Dad for supporting and nurturing my interest in science!

Contents

List of Figures

List of Tables

Abbreviations

AcN	**Ac**eto**ni**trile
APH	**A**mphi**p**athic **h**elix
APS	**A**mmonium **p**er **s**ulphate
ATP	**A**denosine-5'- **t**ri**p**hosphate
BAR	**B**in-**A**mphiphysin-Rvs
BMRB	**B**iological **m**agnetic **r**esonance **b**ank
β**ME**	β-**m**ercapto**e**thanol
CAA	**C**as**a**mino **a**cids
CD	**C**ircular **d**ichroism
cmc	**C**ritical **m**icelle **c**oncentration
CNBr	**C**yanogen **b**romide
CSA	**C**hemical **s**hift **a**nisotropy
CV	**C**olumn **v**olume
DD	**D**ipole-**d**ipole
DDM	**D**o**d**ecyl maltoside
DFT	**D**iscrete **F**ourier **t**ransform
D(C_6)HPC	1,2-**D**i**h**exanoyl-*sn*-glycero-3-**p**hospho**c**holine
D(C_7)HPC	1,2-**D**i**h**eptanoyl-*sn*-glycero-3-**p**hospho**c**holine
DM	n-**D**ecyl-α-D-**m**altopyranoside

DOPC	1,2-**Dio**leoyl-*sn*-glycero-3-**p**hospho**c**holine
DOPE	1,2-**Dio**leoyl-*sn*-glycero-3-**p**hospho**e**thanolamine
DP1	**D**eleted in **p**olyposis 1
DPC	**D**odecyl**p**hos**c**holine
DPPC	1,2-**Dip**almitoyl-*sn*-glycero-3-**p**hospho**c**holine
DSS	4,4-**D**imethyl-4-**s**ilapentane-1-**s**ulfonic acid
DTT	**Dit**hio**t**hreitol
EM	**E**lectron **m**icroscopy
EPL	***E.*** *coli* **p**olar **l**ipids
ER	**E**ndoplasmic **r**eticulum
smER	**Sm**ooth **ER**
rER	**R**ough **ER**
FA	**F**ormic **a**cid
FFT	**F**ast **F**ourier **t**ransform
FID	**F**ree **i**nduction **d**ecay
FPLC	**F**ast **p**rotein **l**iquid **c**hromatography
GTP	**G**uanosine-5'- **t**ri**p**hosphate
HFIP	1,1,1,3,3,3-**H**exa**f**luoro-2-**p**ropanol
HMQC	**H**eteronuclear **m**ultiple **q**uantum **c**oherence
HPLC	**H**igh **p**erformance **l**iquid **c**hromatography
HSQC	**H**eteronuclear **s**ingle **q**uantum **c**oherence
IMAC	**I**mmobilised **m**etal **a**ffinity **c**hromatography
IPA	propan-2-ol
IST	**I**terative **s**oft **t**hresholding
LB	**L**uria **B**roth
LDAO	N,N-dimethyldodecylamine-N-oxide
LMPC	1-myristoyl-2-hydroxy-*sn*-glycero-3-phosphocholine
LMPG	1-myristoyl-2-hydroxy-*sn*-glycero-3-phospho-(1'-rac-glycerol) (sodium salt)
LPPC	1-palmitoyl-2-hydroxy-*sn*-glycero-3-phosphocholine
MS	**M**ass **s**pectrometry
MTSL	1-oxyl-2,2,5,5-tetramethyl-η^3-pyrroline-3-methyl-methanethiosulfonate
dMTSL	1-acetyl-2,2,5,5-tetramethyl-η^3-pyrroline-3-methyl-methanethiosulfonate
MWCO	**M**olecular **w**eight **c**ut **o**ff

NMR	**N**uclear **m**agnetic **r**esonance
NOE	**N**uclear **O**verhauser **e**ffect
NOESY	**N**uclear **O**verhauser **e**ffect **s**pectroscopy
NUS	**N**on-**u**niform **s**ampling
OD_{600}	**O**ptical **d**ensity at 600 nm
PA	**P**hosph**a**tidic acid
PAGE	**P**oly **a**crylamide **g**el **e**lectrophoresis
PC	**P**hosphatidyl**c**holine
PE	**P**hosphatidyl**e**thanolamine
PI	**P**hosphatidyl**i**nositol
PRE	**P**aramagnetic **r**elaxation **e**nhancement
PS	**P**hosphatidyl**s**erine
RDC	**R**esidual **d**ipolar **c**oupling
REEP	**R**eceptor **e**xpression **e**nhancing **p**roteins
RHD	**R**eticulon **h**omology **d**omain
RMSD	**R**oot **m**ean **s**quare **d**eviation
Rtn	**R**et**i**cul**n**
SDS	**S**odium **d**odecyl **s**ulphate
TCEP	**T**ris(2-**c**arboxy**e**thyl)**p**hosphine
TEM	**T**ransmission **e**lectron **m**icroscopy
TFA	**T**ri **f**luoro**a**cetic acid
TFE	**T**ri **f**luoro**e**thanol
TM	**T**rans**m**embrane
TRACT	**T**ROSY for **r**ot**a**tional **c**orrelation **t**imes
TROSY	**T**ransverse **r**elaxation **o**ptimised **s**pectroscop**y**
US	**U**niform **s**ampling
YPL	**Y**east **p**olar **l**ipids

To my Mum, Dad and Laura

1

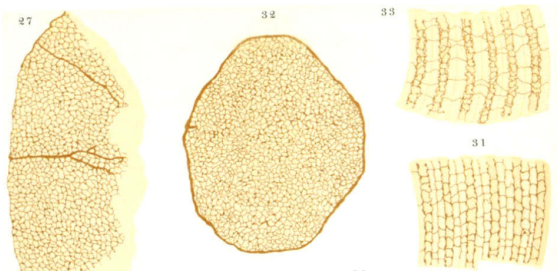
Introduction

"...there remains no dearth of authoritative investigators either in past epochs or in recent times who oppose this doctrine, affirming that the contractile part does not have a fibrillar form in the living elements, and that filaments isolable from muscle fibres with chemical reagents are artificial products."

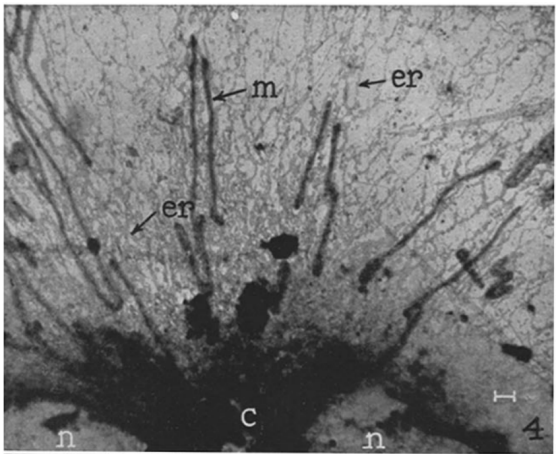
Dr. Emilio Veratti - discoverer of the ER (1902)

1.1 Organelles

The discovery of intracellular organelles is rooted in the desire to understand the basis of movement, one of the most important abilities inherent to all animals. Organelles are generally membrane-bound structures within cells that enable the compartmentalisation and fine control of cellular processes such as oxidative phosphorylation, post-translational modification and signalling. Some of the first successful attempts to peer beneath the plasma membrane and elucidate these intracellular compartments were made by Robert Hooke (Hooke, 1665) and Antony Van Leuwenhook (Leuwenhook, 1680) with their novel microscopic techniques. Hooke first defined the concept of "the cell" and the nucleus was first observed and described by Leuwenhook. It was not until 1833 however that Robert Brown, a Scottish botanist, gave the nucleus its name. Following



(A)



(B)

FIGURE 1.1: (A) Drawings of sarcoplasmic reticulum from various crustaceans by Veratti (1902) (Veratti, 1961). 27, transverse section of a limb muscle fibre from *Carcinus maenas* (Littoral crab). 31, longitudinal section from claw muscle fibre from *Astacus astacus* (European crayfish). 32 and 33, transverse and longitudinal sections of a limb muscle fibre from *Hydrophilus piceus* (Great Silver Water Beetle), respectively. (B) One of the first electron micrographs to unambiguously identify the ER (er) (Porter, 1953). Tubules can be seen emanating away from the nucleus (n) with more dense sheet-like structures towards the center. The images were collected from chick macrophage cells, hence the multiple nuclei. c = centrosome, m = mitochondria and scale bar represents 1 μm.

this, many other organelles were discovered such as mitochondria (Altmann, 1890) and the Golgi apparatus (Golgi, 1898) but conspicuously absent was the ER until the mid 20^{th} century.

In 1902 Emilio Veratti (Veratti, 1961)* (Figure 1.1A), a student of Camillo Golgi, first described the ER in a paper that was largely overlooked by the scientific community for the next 50 years. Most of Veratti's contemporaries suspected that the *"elegant reticulum of rounded or polygonal meshes"* he described were simply artefacts from Golgi's silver staining protocol *a.k.a* the *"black reaction"*. Their scepticism was not entirely without warrant; Golgi's black reaction was notoriously difficult to perform reliably and reproducibly. Veratti himself even commented on the difficulties of the staining protocol - *"...one arrives quickly at a stage in which the muscle fibers treated with the black reaction stain uniformly brown or black so that one can no longer distinguish any details of internal structure."* (Veratti, 1961).

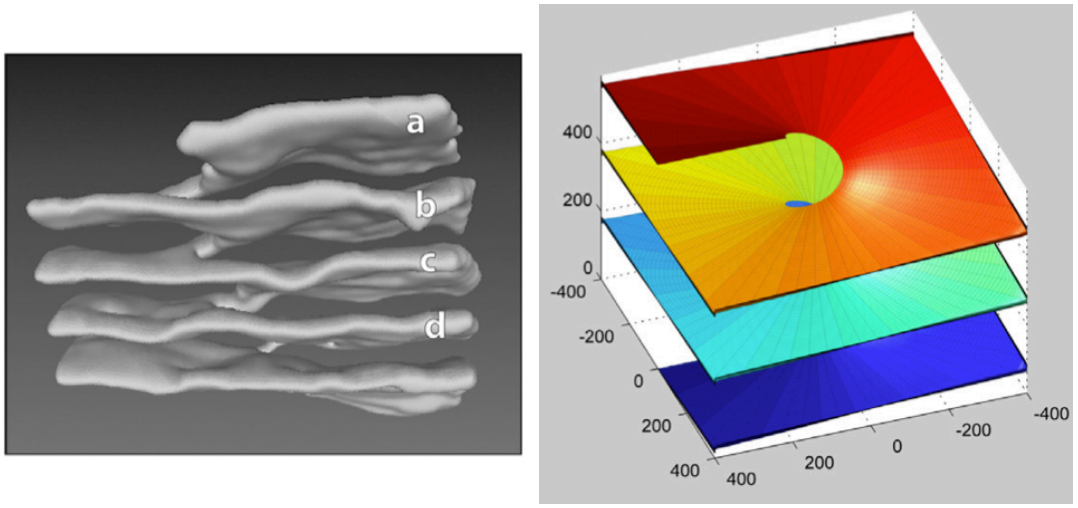
It was not until the pioneering electron microscopy studies of a Canadian cell biologist Keith Porter (Porter, 1953) and later in collaboration with George Palade (Porter and Palade, 1957) that the ER was given its name and accepted as an organelle by the greater scientific community. In Porter's seminal paper (Porter, 1953)† he described *"...a faintly stained, reticular pattern to be designated the endoplasmic reticulum"* and later alluded to the ER as *"...a lace-like pattern of fine strands"* (Figure 1.1B). These papers marked the beginning of a new and exciting era of research into the ER.

Now scientists the world over endeavour to dissect and understand the ER's complex structure and function. The steadily increasing rate of ER-related publications reflects the unfathomable complexity and importance of this organelle (Schuldiner and Schwappach, 2013) and ER dysfunction has been implicated in many human diseases, especially neurological disorders (Blackstone, 2012, Roussel et al., 2013).

1.2 ER structure and function

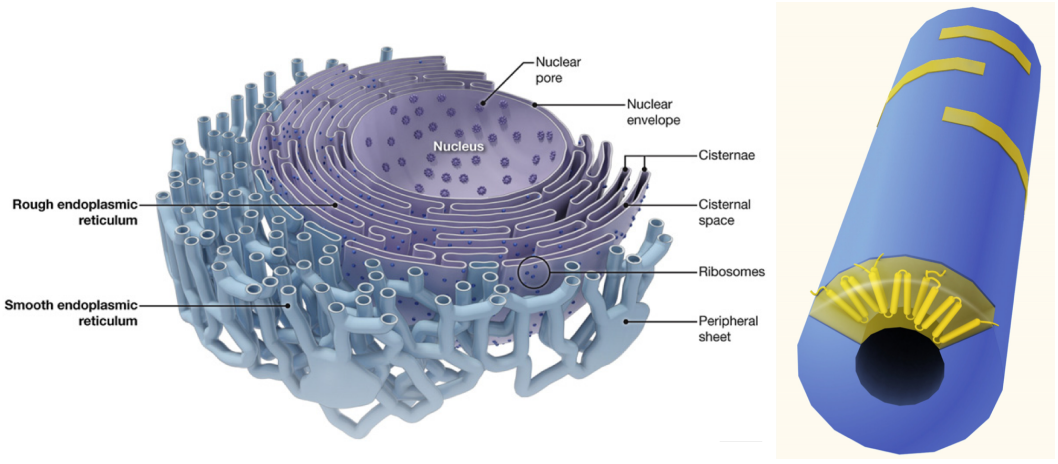
The ER is a vast organelle. In the simplest terms, it is categorised as being either "rough" (rER) or "smooth" (smER). The rER extends out from the nucleus as a continuation of the outer nuclear membrane to form reticular ribosome studded sheets which stack on top of one another in a helicoidal arrangement (Terasaki et al., 2013) (Figure 1.2). Ribosomal coverage becomes more sparse towards the periphery of the cell as the smER sheets transform into a dynamic polygonal network of narrow tubules (Figure 1.3) connected

*This is the first English translation of Veratti's work from 1902 which had been largely forgotten.
†Porter did not reference Veratti in this paper.



(A) (B)

FIGURE 1.2: Current model for sheet ER adapted from Terasaki et al. (2013). The sheets are connected in a helicoidal arrangement much like a multi-storey car park. This arrangement minimises the energetic cost (elastic energy) of having such highly curved membranes since the total (J) curvature tends to 0 with the negative curvature due to the helicoid and positive curvature at the edges of the sheets. (A) is a 3D reconstruction from cryo-EM sections of adult mouse tissue, while (B) is a computational model of a single helicoidal connection between ER sheets.



(A) (B)

FIGURE 1.3: (A) Schematic of ER structure from Goyal and Blackstone (2013). The branched tubular ER can be seen towards the periphery of the cell while the rough cysternal ER surrounds the nucleus. It should be noted that the depiction of sheet ER here is inconsistent with the model presented in Figure 1.2. (B) Schematic of an ER tubule from Hu et al. (2011) highlights the degree of membrane curvature required for tubulation. The yellow structures summarise the prevailing model for how this curvature is achieved by membrane protein scaffolding and wedging.

by 3-way junctions. The dynamic structure of the ER is highlighted in recent work from Francis Barr's lab (Gerondopoulos et al., 2014) where ER tubules can be seen growing, retracting and fusing with others on a timescale of seconds.

During a cell's life cycle the composition of the ER can change dramatically. For example, during transition from interphase to mitosis the smER undergoes a drastic transformation from a primarily tubular network into sheet structures (Wang et al., 2013). ER structure is also cell type dependent, with secretory pancreatic acinar cells containing primarily sheet ER while cells specialised in metabolism such as hepatocytes are replete with tubular ER (Park and Blackstone, 2010). This raises many interesting questions, how does ER morphology relate to its functions? How is ER morphology dynamically controlled and regulated? What happens when the structure of the ER is perturbed and why does this lead to disease? These questions are broadly unanswered but they reflect the fertility and importance of this area of research.

1.3 Membrane curvature

The majority of intracellular organelles are surrounded by lipid bilayers that are sculpted into characteristic shapes. For this to happen there is almost invariably a requirement for membrane curvature. For example the tubular ER, the edges of ER sheets, nuclear pores, mitochondrial cristae and the Golgi apparatus all require high levels of membrane curvature. There are also requirements on membrane curvature generation in signalling pathways such as the ESCRT pathway, COPI and II mediated trafficking and endocytosis.

1.3.1 Types of curvature

There are three main types of curvature (summarised in Table 1.1). The principle curvatures are described by Eqs. (1.1) and (1.2) where R_1 and R_2 are the radii of two intercepting planes of curvature (Figure 1.4). These planes are arcs across the region of maximum and minimum curvature for a given curved surface. J is the total curvature (Eq. (1.3)) and K is the Gaussian curvature (Eq. (1.4)). A cylindrically shaped membrane has *zero* Gaussian curvature and *positive* total curvature. Saddle shaped membrane surfaces have *zero* total curvature but *negative* Gaussian curvature. Spherical membrane surfaces have both *positive* Gaussian and *positive* total curvature. The main types of membrane curvature along with the organelles and membrane structures in which they are found are summarised in Table 1.1. In biological systems *positive* membrane curvature is conventionally defined as bulging towards an outside medium.

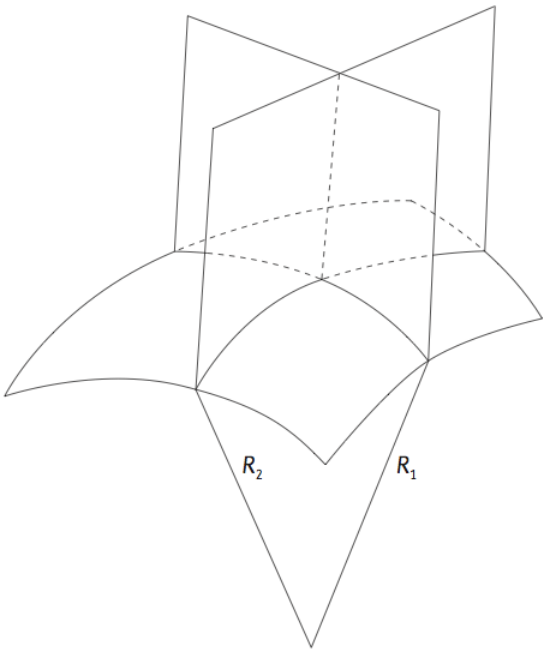


FIGURE 1.4: Schematic showing principle curvatures from Zimmerberg and Kozlov (2006). R_1 and R_2 are the radii of the two arcs of minimum and maximum curvature from which the principle curvatures are derived.

However, the sign of the curvature is somewhat arbitrary.

$$Principle\ curvature\ (c_1) = \frac{1}{R_1} \tag{1.1}$$

$$Principle\ curvature\ (c_2) = \frac{1}{R_2} \tag{1.2}$$

$$Total\ curvature\ (J) = c_1 + c_2 \tag{1.3}$$

$$Gaussian\ curvature\ (K) = c_1 \cdot c_2 \tag{1.4}$$

TABLE 1.1: Types of membrane curvature

Type	c_1	c_2	J	K	Example
Flat	0	0	0	0	sheet ER
Cylindrical	+	0	+	0	tubular ER/Golgi
Saddle	+	-	0	-	sheet ER/ vesicle scission points
Spherical	+	+	+	+	vesicles

1.3.2 Lipids involved in membrane curvature

Biological membranes have complex compositions and the shape, chemical properties and distributions of lipid species can influence membrane curvature. The wild-type *S. cerevisiae* "lipidome" comprises at least 162 different lipid species with a head-group composition of 20.3% Phosphatidylinositol (PI), 14.9% phosphatidylethanolamine (PE), 14.3% phosphatidylcholine (PC), 8.3% phosphatidic acid (PA) and 1.8% phosphatidylserine (PS) (Ejsing et al., 2009). This diversity coupled with organelle specific

lipid compositions and often asymmetric distributions between leaflets makes understanding the behaviour of biological membranes a formidable task (see van Meer and de Kroon (2011), Van Meer et al. (2008) for reviews).

Many lipid species have intrinsic membrane curving properties. For example lysophopholipids (i.e. phospholipids containing only one acyl chain) induce positive curvature on the lipid leaflet they occupy while 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (DOPE) and diacylglycerol induce negative curvature (Zimmerberg and Kozlov, 2006). However, the relevance of these effects may be limited to particular cases where the given lipid species can be significantly enriched within a confined area. Examples include multivesicular endosomes, which are thought to be induced by high local concentrations of lysobisphosphatidic acid (LBPA) (Matsuo et al., 2004). It was also thought for some time that Endophilin-1 and other Bin-Amphiphysin-Rvs (BAR) domain proteins effected negative membrane curvature through acyltransferase activity by converting lysophosphatidic to phosphatidic acids (Schmidt et al., 1999, Weigert et al., 1999) though this was later debunked as an artefact of purification (Gallop et al., 2005).

1.3.3 Proteins involved in membrane curvature

The complex structures of organelles cannot be generated by lipids alone. A diverse set of proteins are required to regulate, generate and stabilise the membrane structures. These proteins can be mechanistically categorised as follows; however, some proteins utilise combinations of the following mechanisms (see Table 1.2 for summary).

1.3.3.1 Amphipathic helices

Amphipathic helices may be inserted into the lipid headgroup region thereby displacing lipids on one leaflet to induce curvature. The insertion depth is of great importance. For instance, shallow insertion will induce positive curvature while deeper insertion beneath the lipid headgroups can induce negative curvature (Campelo et al., 2008). Proteins that rely on this mechanism include Epsin (Ford et al., 2002), N-BAR domains (Gallop et al., 2006), Atlastins (Liu et al., 2012) and viral proteins such as the M2 channel from influenza (Roberts et al., 2013). Amphipathic helices often work in concert with scaffolding proteins by destabilising the leaflet to which they are bound and facilitating membrane curvature (see Section 1.3.3.4).

1.3.3.2 Loop insertion

Proteins involved in synaptic vesicle signalling such as the synaptotagmin C2 domain induce high levels of membrane curvature through insertion of hydrophobic loops (distal in primary structure) causing displacement of lipids from one leaflet (Martens et al., 2007).

1.3.3.3 Membrane coats

Membrane coats are complex assemblies of membrane-associated and accessory proteins that mostly sculpt membranes into vesicles. Examples include COPI and II coated vesicles and clathrins. A feature of these systems is the combinations of cargo adapter proteins, cargo receptor proteins and coat proteins that enable binding of specific protein substrates and facilitate complex cell sorting mechanisms (Guo et al., 2014).

1.3.3.4 Scaffolds

The most well studied scaffolding proteins involved in membrane curvature are the BAR domains. These proteins are conserved from yeast to humans and have a diverse set of functions (Frost et al., 2009). BAR domains generally form dimers of curved architecture that bind to membrane surfaces, either generating or stabilising membrane tubulation. Positive curvature is induced in the case of N-BAR and F-BARs and negative in the case of I-BAR domains. N-BAR domains are so called due to their N-terminal amphipathic helices. These helices destabilise the membrane reducing the energetic barrier for membrane curvature (Mim et al., 2012). Another amphipathic helix containing BAR domain is amphiphysin, which is found in the T-tubules of striated muscle cells (Lee et al., 2002). A key feature of many scaffolding proteins is oligomerisation, a process that enables the formation of large ordered assemblies on membrane surfaces resulting in curvature of the target membrane. Two great examples are the cryo-EM studies of Frost et al. (2008) and Mim et al. (2012) where helical arrangements of oligomerised N- and F-BAR domains, respectively, were visualised coating tubular membranes.

1.3.3.5 ATP/GTPases

Since membrane deformation invokes an energetic cost it is not surprising that proteins containing ATPase and GTPase activities play an important role. An example of a GTPase class are the Dynamin proteins which are involved in vesicle budding during Clathrin mediated endocytosis, mitochondrial fission and fusion and even resistance

to viral infections (Heymann and Hinshaw, 2009). The endosomal sorting complex required for transport (ESCRT) pathway is an example of a scaffolding mechanism which requires the action of an AAA-ATPase, Vsp4, for vesicle budding (Henne et al., 2012). For example, the ESCRT-III protein produces a helical scaffold inducing membrane curvature at the neck of a budding vesicle. The oligomeric Vsp-4 complex is then thought to mediate ATP dependent membrane fission (Henne et al., 2013).

1.3.3.6 Membrane embedded

A final and more mysterious class of membrane deforming proteins are the "membrane embedded". Examples include Caveolins, Reticulons (Rtn) and Receptor expression enhancing proteins (REEPs). A unifying feature of these proteins is the fact that their TM domains are predicted to be too short to fully traverse the bilayer but also too long to form canonical TM helices (Figure 1.5). Prediction algorithms also generally fail to predict clear boundaries compared with more typical transmembrane proteins (Figure 1.5, bottom). However, it is worth pointing out that the algorithms do predict clear TM regions upon insertion of short linker sequences between the suspected break points on the lumenal side of the TMs. This suggests that the algorithms may lack sensitivity to membrane proteins that contain very tight turns between TMs. The use of averaging window functions could contribute to this problem.

Of all the known classes of membrane curvature sensing, stabilising and generating proteins (categorised in Table 1.2), a structure from the Reticulon/REEP/Caveolin classes of transmembrane proteins remains conspicuously absent. Caveolins are beginning to be studied on a structural level by NMR (Lee and Glover, 2012, Rieth et al., 2012, Rui et al., 2014). Nevertheless, in-depth structural studies on the TM domains of Reticulon/REEP proteins are completely lacking.

The sub-cellular localisations of proteins involved in membrane curvature is nicely summarised by Shen et al. (2012) in Figure 1.6. This figure also highlights the global importance of membrane curvature in cellular structure and function.

Table 1.2: Membrane curving proteins

Category	Examples
Amphipathic helices	BAR domains
Loop insertions	Synaptotagmin C2-domain
Membrane coats	COPI/II, Clathrins
Membrane embedded	Reticulons, REEPs, Caveolins
Scaffolds	ESCRT, BAR
ATP/GTPases	Vsp-4, Dynamins

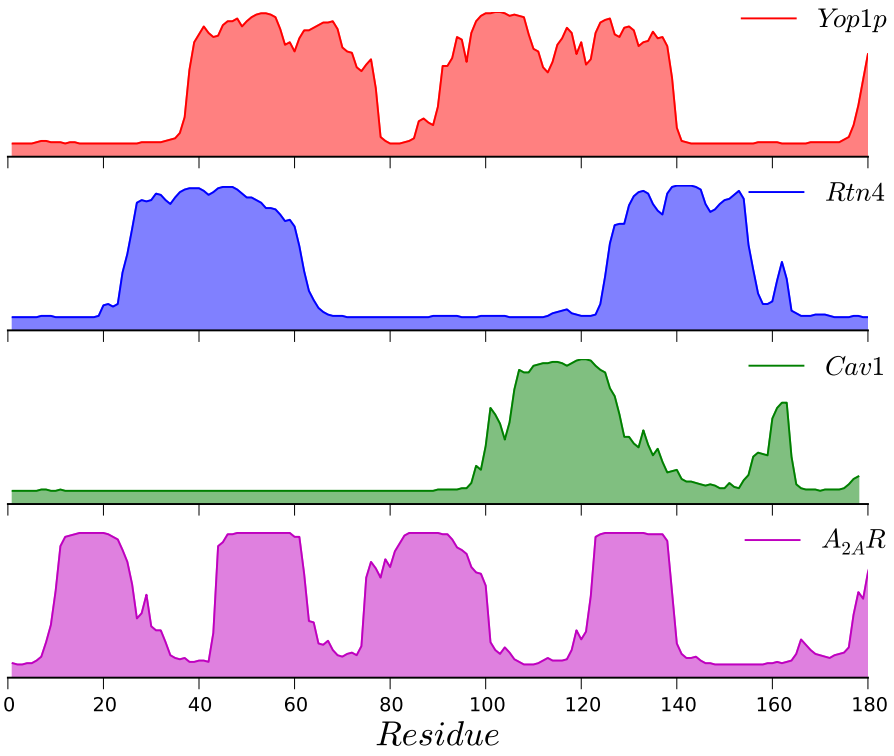


FIGURE 1.5: Yop1p (Uniprot:Q12402), Reticulon4 (Rtn4, Isoform A, Uniprot: Q9NQC3), Caveolin1 (Cav1, Uniprot:Q03135) and the Adenosine A_{2A} receptor ($A_{2A}R$, Uniprot:P29274) sequences were run through the Octopus TM prediction server (Viklund and Elofsson, 2008) and residues 1 to 180 were plotted to ease comparison. $A_{2A}R$ is included as an example of a canonical TM protein. Figure inspired by "Le petit prince".

A note about nomenclature and phylogeny

The receptor expression enhancing proteins (REEPs) or Deleted in Polyposis 1 (DP1) family of proteins refers to a family of 6 proteins in humans REEP1-6.

Yop1 is the *S. cerevisiae* functional homolog of REEP5 (also known as DP1) in humans. Reticulons 1 and 2 (Rtn1 and Rtn2) are functional homologs of the human reticulons (Rtn1-4). However, Rtn proteins in higher eukaryotes contain large C-terminal domains and many splice isoforms.

The sequence identity between *S. cerevisiae* Yop1 and Reticulon 1 (Rtn1) is 5.4% which essentially categorises the proteins as unrelated by primary sequence. The unifying features of these proteins are their unusual hydrophobic regions ($\approx$ 30 amino acids) known as reticulon homology domains (RHDs) and their propensity to localise exclusively to regions of high membrane curvature in the ER.

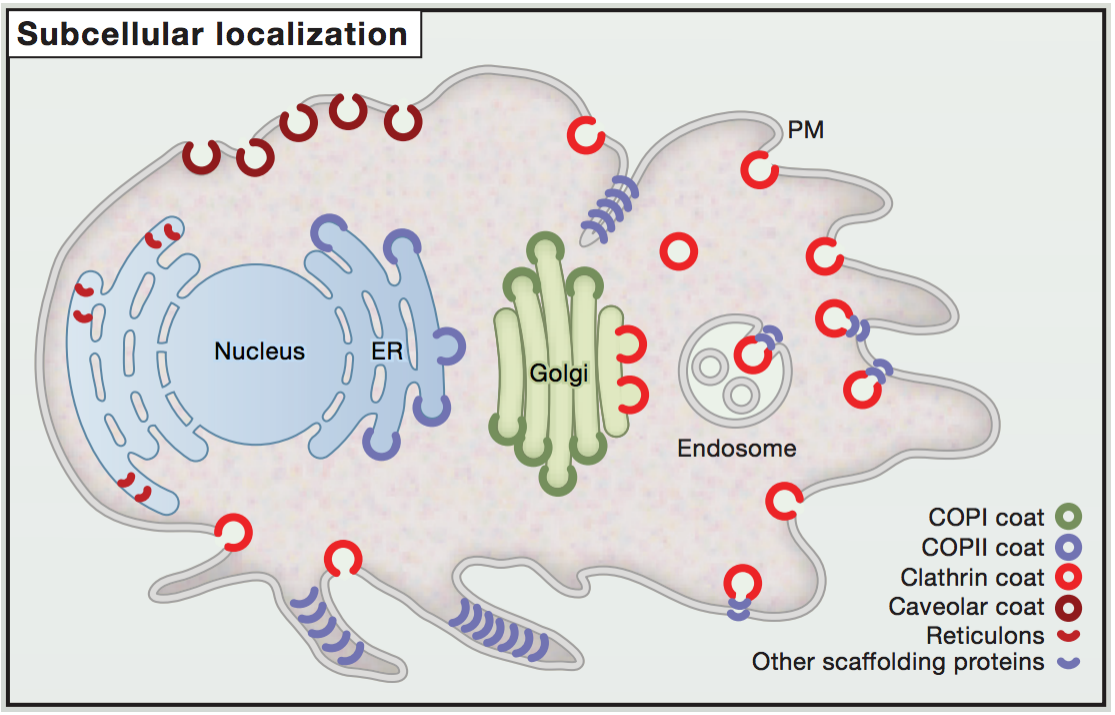


FIGURE 1.6: Subcellular localisation of membrane curvature sensing and generating proteins. "Other scaffolding proteins" include dynamins, ESCRT proteins (e.g. CHMP3), BAR domains and "loop inserting" synaptotagmins (e.g. Synaptotagmin 1 C2-domain). "Reticulons" includes DP1/REEP family proteins such as Yop1p. Figure taken from Shen et al. (2012)

1.4 The REEPs and the Reticulons

The DP1/REEP and reticulon classes of integral membrane proteins are responsible for stabilising the high membrane curvature found in ER tubules and the edges of sheets (Hu et al., 2008, Shibata et al., 2008, Voeltz et al., 2006). Consistent with their role in ER morphology, similar proteins have not been found in prokaryotes (Oertle and Schwab, 2003). There are six human DP1/REEP proteins originally identified as accessory proteins facilitating the expression of odorant receptors and thus termed receptor expression-enhancing proteins (REEPs) (Saito et al., 2004). The ability of REEPs to facilitate receptor trafficking has recently been related back to their ER shaping activities (Björk et al., 2013). The importance of REEPs in ER morphology is highlighted also by their implication in human diseases that are associated with neurons having long axons and requiring an extended tubular ER (Blackstone, 2012, Shibata et al., 2009). For example, mutations in the transmembrane domain of REEP1, which is primarily expressed in neurons (Hurt et al., 2014), lead to pure forms of hereditary spastic paraplegia (HSP) (Zhao et al., 2001, Züchner et al., 2006). The reticulon family in humans comprises four members, Rtn1-4 (Oertle and Schwab, 2003). Of particular interest is Rtn4, as it is primarily responsible for generating or stabilising the tubular ER in mammalian

cells (Voeltz et al., 2006). Like the REEPs, Rtn4 is critical in neuronal cell processes as it inhibits spontaneous neurite outgrowth (GrandPré et al., 2000) and restricts neural plasticity (Kempf and Schwab, 2013), and has been implicated in several neuronal diseases, including schizophrenia and motor neuron disease (Wang et al., 2012, Willi and Schwab, 2013, Yang and Strittmatter, 2007). The DP1 and reticulon proteins have in common a region containing two unusually long hydrophobic segments (≈35 amino acids in length each, see Figure 1.5 and Figure 1.8 for alignment of RHD region.) known as a Reticulon Homology Domain (RHD) (Oertle and Schwab, 2003). In humans, REEPs 1-6 are similar overall, but the N-terminus of the RHDs of REEPs 1-4 is truncated (Figure 1.7). REEPs 1-4 also possess C-terminal tubulin binding domains providing a link between ER morphology and the cytoskeleton (Park et al., 2010). The RHDs are proposed to form hydrophobic hairpins too short to fully traverse the membrane leading to greater displacement of lipids from the outer leaflet than the inner. By this mechanism a combination of hydrophobic wedging and oligomeric scaffolding could lead to the stabilisation of the tubular ER membrane. However, there is currently no structural information available for these intramembrane domains and the exact mechanism of curvature generation and stabilisation remains unknown. The mechanism for exclusive localisation of RHD proteins to highly curved membranes also remains unclear. Yet, the transmembrane domains have been implicated in both membrane localisation and curvature stabilisation (Tolley et al., 2010, Voeltz et al., 2006, Zurek et al., 2011).

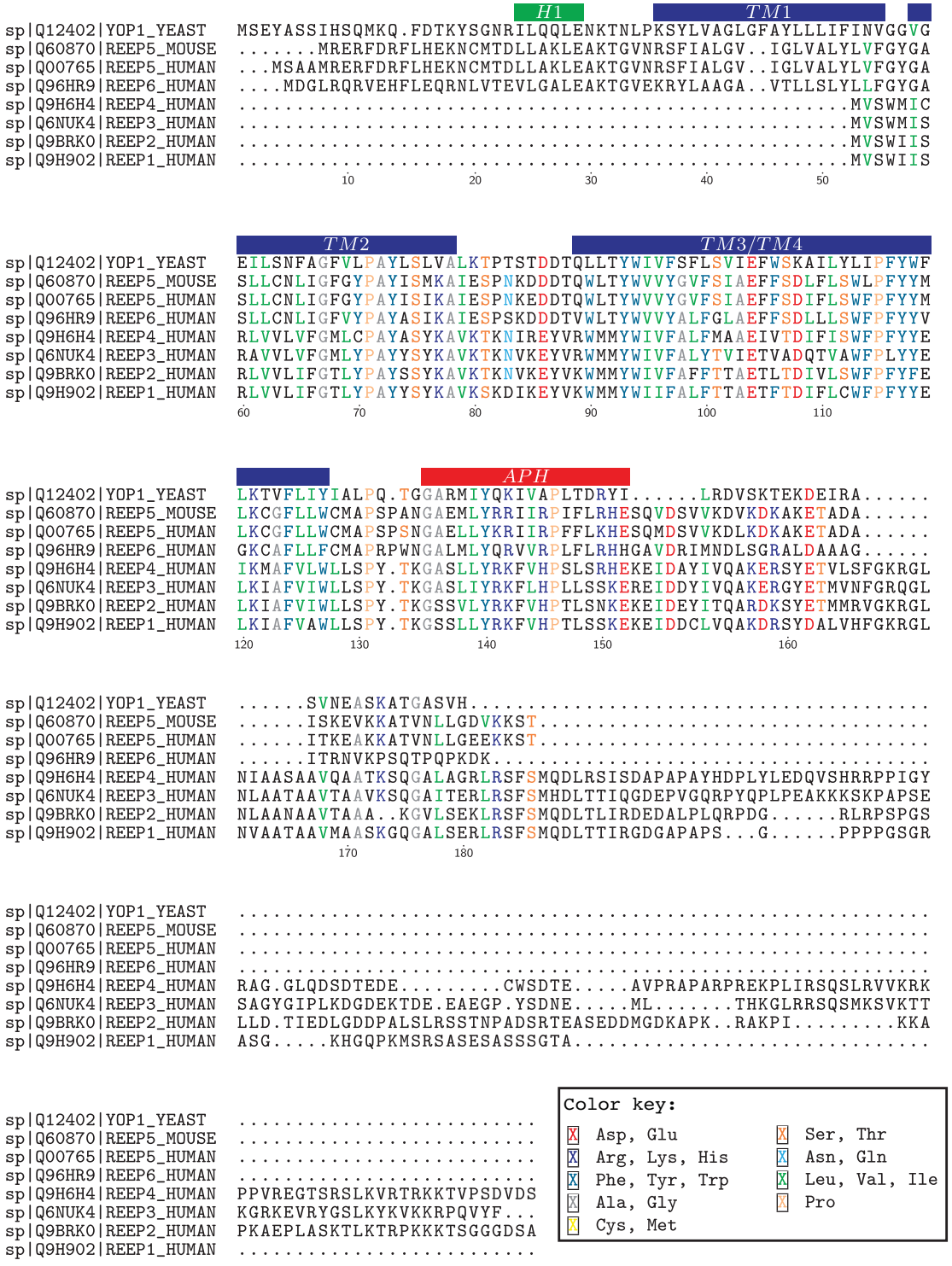


FIGURE 1.7: Sequence alignment of Yop1p with the human REEPs and mouse REEP5. Helical regions indicated above the sequence are derived from NMR chemical shift data using TALOS-N (Shen and Bax, 2013b). Figure generated with TEXshade (Beitz, 2000)

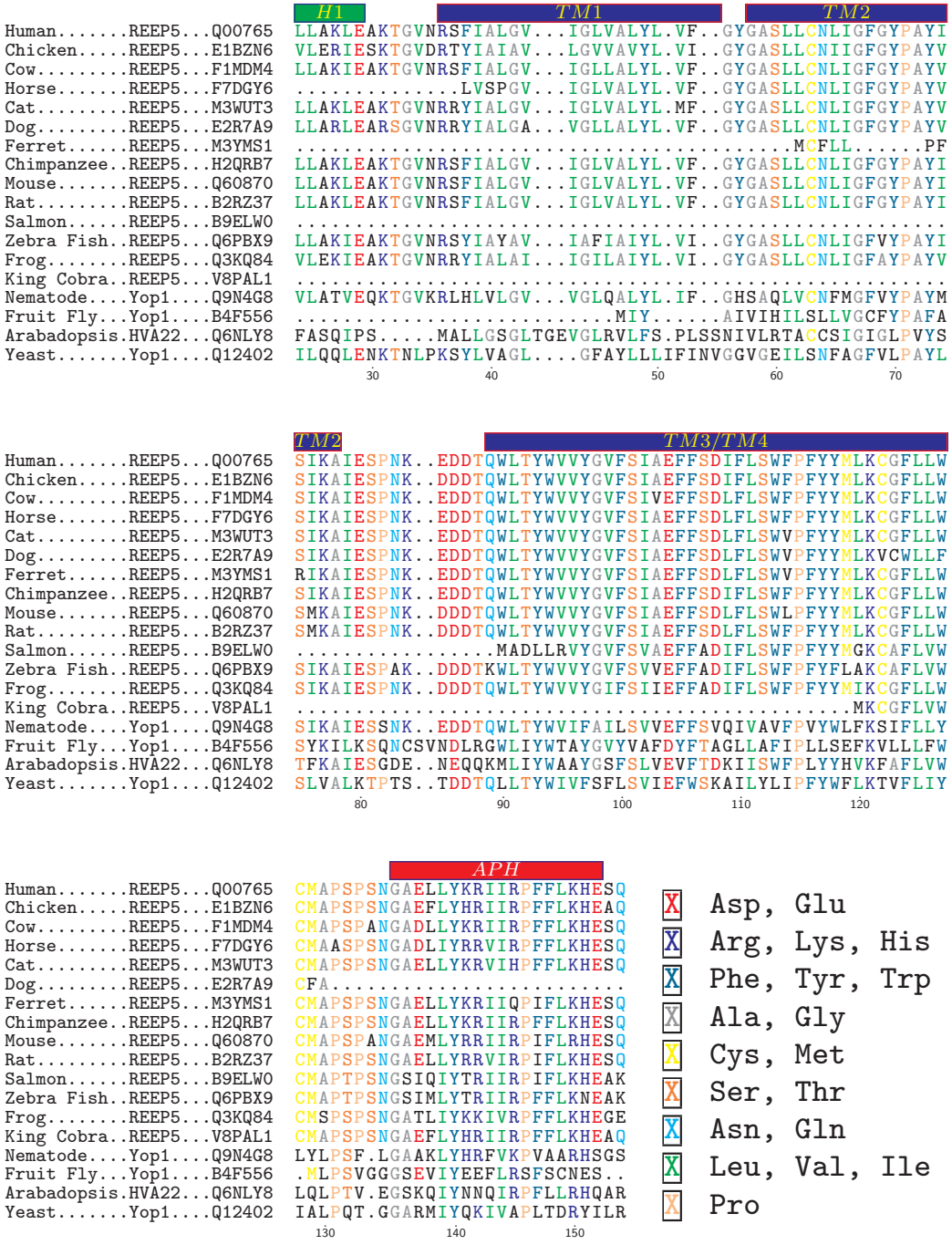


FIGURE 1.8: Sequence alignment of Yop1p with its functional homologs from a range of species. The Uniprot accession codes are also listed. Helical regions indicated above the sequence are derived from NMR chemical shift data using TALOS-N (Shen and Bax, 2013b). Figure generated with TEXshade (Beitz, 2000)

1.4.1 Plants

Plants have a vastly more complex system of proteins governing ER morphology. For example, *Arabidopsis thaliana* has 21 reticulon homologs (Tolley et al., 2010) and at least 11 DP1 homologs, HVA-22 like proteins A-K (Chen et al., 2002, UniProt Consortium, 2014). The closest homolog to Yop1p in *Arabidopsis thaliana* appears to be HVA-22 like protein K (Uniprot ID:Q9S7V4) which aligns well with Yop1p and has 4 putative TM helices (see Figure 1.9). The other HVA-22 like proteins bear more similarity to REEPs 1-4 with their truncated first TM regions (Figure 1.7). HVA-22-like protein A has been shown to localise to the tubular ER (Chen et al., 2011b) but its exact role has not yet been established. Another feature of the HVA-like proteins is an extended proline rich C-terminus much like REEPs 1-4 which could suggest other functions such as tethering to the cytoskeleton (Park et al., 2010). The HVA proteins were first discovered in *Hordeum vulgare* (Barley) as an abscisic acid (a plant hormone involved in response to stresses such as drought and salinity) and stress inducible gene (Shen et al., 1993).

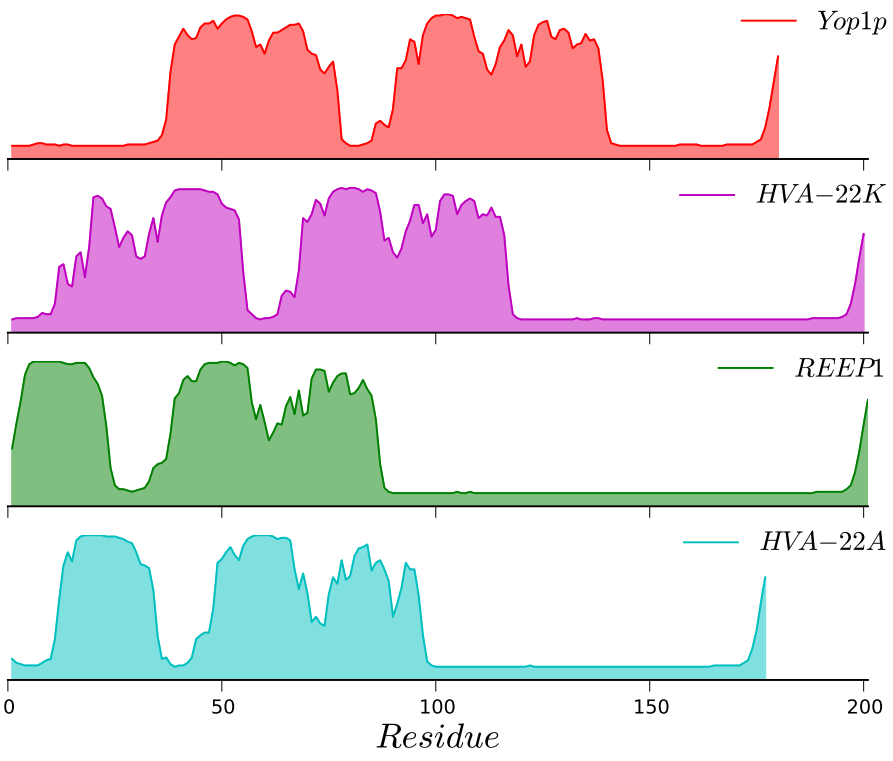


FIGURE 1.9: Comparison of Octopus TM prediction for HVA-22 like proteins A and K with Yop1p and human REEP1 (Viklund and Elofsson, 2008). HVA-22 like protein K (HVA-22K) bears most resemblance to Yop1p/REEP5 while HVA-22 like protein A (HVA-22A) bears most resemblance to REEP1.

1.4.2 Reticulons and viruses

There has been growing interest in the role of the ER in viral infections both in animal and plant viruses. In plants the tubular ER can form a continuous network between cells by passing through channels known as plasmodesmata. These contain narrow membrane tubules, known as desmotubules and are the primary route by which plant viruses spread between cells (Wu et al., 2011). Animal viruses generally hijack the endo/exocytosis machinery; however, reticulon proteins have also been implicated.

Brome mosaic virus (BMV) from the alpha-virus like superfamily (Including Semliki Forest and Sindbis viruses) has been shown to require RHD containing proteins in order to efficiently replicate (Diaz et al., 2010). While Enterovirus 71, a Picornavirus causing Polio like symptoms in children, has also been shown to require binding to the RHD of Reticulon 3 for efficient replication (Tang et al., 2007). However, another recent study showed that Reticulon 3 actually hinders replication of Hepatitis C virus (HCV) (Wu et al., 2014), suggesting the picture may be more complicated. These studies suggest that the RHD proteins may have a future role to play in the treatment of viral infections.

1.5 Overview of thesis

The overarching aim of this thesis is to develop a molecular understanding of the mechanisms involved in generation and stabilisation of high membrane curvature in the ER.

I have chosen to study Yop1p as a model system for a number of reasons. (1) Yop1p is one of the best characterised of the RHD containing proteins (Hu et al., 2008, Voeltz et al., 2006). (2) It is a member of the DP1 family and has the highest functional and sequence conservation with human REEP5 (Park et al., 2010)(also see Figure 1.7). (3) Yop1p can be purified from yeast membranes and reconstituted into lipids to generate tubules of ≈17nm diameter (Hu et al., 2008).

To investigate possible mechanisms of membrane deformation by RHDs, I have expressed, purified and determined the secondary structure and dynamics of Yop1p in lipid and detergent containing micelles. My results indicate that an N-terminal transmembrane domain contains two helices of 20-21 amino acids in length separated by a glycine-rich break. By contrast, the C-terminal transmembrane domain consists of 39 helical residues with two possible helical break points. Furthermore, an amphipathic helix (APH) was identified that is C-terminal to the transmembrane domains in a region previously shown to be important for function (Hu et al., 2008). I show here that the APH on its own interacts with membranes and that deletion of the APH abolishes membrane tubule formation in vitro. The amphipathic nature of the helix as well as its position C-terminal to the last transmembrane domain are highly conserved across the DP1 and reticulon families, and thus may play a role in membrane curvature stabilisation by all family members.

This structural data should be used to inform future experiments for further interrogation of the molecular mechanisms of membrane curvature in the ER.

2

Materials and Methods

"Hearken; Behold, there went out a sower to sow: And it came to pass, as he sowed, some fell by the way side, and the birds of the air came and devoured it up. And some fell on stony ground, where it had not much earth; and immediately it sprang up, because it had no depth of earth: But when the sun was up, it was scorched; and because it had no root, it withered away. And some fell among thorns, the thorns grew up, and choked it, and it yielded no fruit. And other fell on good ground, did yield fruit that sprang up and increased; and brought forth, some thirty, and some sixty, and some an hundred. He said unto them, He that hath ears to hear, let him hear.

Mark *Chapter 4, Verses 3-9*
-King James Bible

2.1 Cloning and mutagenesis

A DNA construct of Yop1 containing no ACA codons (ACA-less) (GeneArt, Invitrogen) was cloned into the pColdI vector (Takara) containing an N-terminal 6x-His tag and Factor Xa cleavage site using *NdeI* and *HindIII* restriction sites. The M138I mutant of Yop1p was made using QuikChange (Stratagene) and all subsequent mutants were produced using the site directed mutagenesis kit from GeneArt (Invitrogen) according to the manufacturer's protocol. A table of cloned and sequence verified constructs can be found in Appendix C.1.

2.2 Protein expression

Protein expression was carried out according to the SPP protocol (Suzuki et al., 2007). BL21 (DE3) chemically competent E. coli were transformed with pMazF using 25 μg mL^{-1} chloramphenicol and M9 minimal media agar supplemented with casamino acids (M9-CAA). Successful transformants were then made chemically competent using the Inoue method (Sambrook and Russell, 2001) and stored at -80 °C. Expression constructs were transformed into pMazF competent cells using 25 μg mL^{-1} chloramphenicol and 100 μg mL^{-1} ampicillin. The resulting colonies were used to inoculate 1 L of M9-CAA which was then incubated at 37 °C overnight at 280 rpm. The overnight was then pelleted at 4000 xg (25 °C) for 15 minutes and resuspended in 10 L of M9 (without CAA) for a starting OD_{600} ~0.2. The cells were incubated at 37 °C until OD_{600} ~0.6-0.7 at which point the cells were cold-shocked by shaking the culture flasks in water containing crushed ice for 5-10 minutes and then incubated at 15 °C for 45 minutes at 280 rpm. The cells were then washed by centrifugation and resuspension in M9-salts (in ^2H_2O if deuterium labelling). Finally, cells were suspended in a 20 fold condensed volume of isotopically labelled M9 media containing 5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated for up to 2 days at 15 °C. Expression levels were monitored by SDS-PAGE. It is also possible to follow the above protocol using C43 cells or pLemo (Wagner et al., 2008) and omitting the use of the MazF plasmid. In this case, 10 fold condensations gave the best protein yields.

2.3 SDS-PAGE

To monitor protein expression in a reproducible manner, samples were collected according to Eq. (2.1) where x is the OD_{600} of a sample of *E. coli* cells. Samples were then

centrifuged at 17,000 xg on a desktop centrifuge, the supernatant discarded and the pellet resuspended in 40 μL 8 M Urea. 30 μL loading dye (Table 2.1) was added to each sample and 10 μL of each was loaded onto a 12% NuPage gel (Invitrogen) and run at 200 V for 35 mins in 1x MES buffer (Invitrogen).

$$volume(\mu L) = \frac{0.6}{x} * 250 \tag{2.1}$$

TABLE 2.1: Recipe for SDS-PAGE loading dye

Reagent	**Volume**
NuPage LDS sample buffer	2 mL
0.5 M SDS	0.5 mL
1 M DTT	0.5 mL
Total	3 mL

2.4 Isolation of membrane fractions

Pelleted cells were resuspended 100 mL lysis buffer (50 mM Tris,150 mM NaCl; pH 8) and lysed using a high pressure cell disrupter and 1 mg DNAse I (Sigma-Aldrich). Lysates were centrifuged at 41,000 rpm (Beckman 45Ti rotor) for 1.5 hours at 4 °C to isolate membranes. Membranes were solubilised in 6 M guanidine-HCl with Triton X-100 (0.5%) and incubated overnight at room temperature. Prior to immobilised metal affinity chromatography (IMAC) purification, membranes were centrifuged at 41,000 rpm (45Ti rotor) for 30 minutes to remove insoluble material.

2.5 Immobilised metal affinity chromatography

Ni-Sepharose resin was equilibrated with 6 M guanidine-HCl (pH 8) and incubated with solubilised membranes for 1hr at room temperature. The flow through was passed through the column 3 more times. The column was washed with 20 column volumes (CV) of 6 M guanidine HCl (pH 8) containing 10 mM imidazole and eluted with 10 CV of 400 mM imidazole. The eluate was then dialysed into 4 L of water overnight at room temperature using Spectra/Por Dialysis Membrane (Molecular weight cut off (MWCO): 3500Da).

2.6 Cyanogen bromide cleavage

Precipitated protein was transferred to 50 mL tubes and centrifuged for 30 minutes at 1500 xg. Pellets were solubilised in 2 ml 70% formic acid (FA) with 0.5 g cyanogen bromide (CNBr). The reaction proceeded wrapped in tinfoil under a nitrogen stream for 2 hours at room temperature. Reactions were then transferred into 3500 MWCO dialysis cassettes (Thermo Scientific) and dialysed against 4 L of de-ionised water (dH_2O) for 1 hr. Samples were subsequently frozen and then lyophilised overnight.

2.7 High performance liquid chromatography

Lyophilised protein powder was dissolved in 1 mL 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) at 42 °C for 30 minutes before adding 0.5 mL FA and 1 mL of dH_2O. The mixture was then filtered and loaded onto a C3 reverse-phase high performance liquid chromatography (HPLC) column (Zorbax). The column was washed with 60% buffer B (37% acetonitrile (AcN), 58% propan-2-ol (IPA), 5% dH_2O, 0.1% trifluoroacetic acid (TFA)), 40% buffer A (95% dH_2O, 5% AcN, 0.1% TFA) and the protein was eluted with an increasing gradient of buffer B. Eluted protein samples were diluted 3 fold with dH_2O, frozen and lyophilised. Protein concentration was quantified in 50% AcN using absorbance at 280 nm in a quartz cuvette (theoretical ϵ_{280} = 29,910 $M^{-1}cm^{-1}$ for Yop1p(24-165)). Samples were then aliquoted, frozen and lyophilised.

2.8 Thin filming and reconstitution

Lyophilised protein powder was dissolved into 1 mL HFIP and added directly to the powdered detergent/lipid of choice. Using a nitrogen stream the mixture was made into a thin film against the walls of a glass test tube and lyophilised for at least 1 hr to remove residual solvent. NMR samples were then reconstituted into 350 μL of 20 mM potassium phosphate buffer (pH 6.5) with 5% D_2O and 400 μM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) and transferred to a Shigemi tube (Shigemi Inc.). EM samples were prepared by gel filtration of NMR samples into 0.01% 1-myristoyl-2-hydroxy-sn-glycero-3-phospho-(1'-rac-glycerol)(LMPG) using a Superdex 200 10/300 GL column (0.5 mL min^{-1}). The peak fractions were combined and concentrated to a final volume of 400 μL with a pre-rinsed 30,000 MWCO spin concentrator (Sartorius). 20% (w/v) *E. coli* polar lipids (Avanti Polar Lipids) solubilized in 1% decyl maltoside (DM) were then added at a 1:1 mass ratio and dialysed against 500 mL reconstitution

buffer overnight at 4 °C. Samples were dialysed for a further 4 hrs against fresh buffer before addition of Biobeads SM2 (Bio-Rad). Two Biobeads were added every hour for 4 hrs while the sample was incubated on a blood wheel at room temperature.

2.9 Negative stain electron microscopy

Negative staining was done using uranyl acetate. 3 μL of reconstituted protein sample (0.01 mg mL^{-1}) was placed onto a glow discharged carbon layered copper grid and incubated to 10 s before blotting with filter paper. The sample was then fixed by adding a drop of uranyl acetate, blotting and then immediately adding another drop of uranyl acetate and waiting 10-15 s. Excess uranyl acetate was blotted and the grids were dried. Images were collected on a TEM Philips CM120 (150 keV) using magnification of 45,000.

2.10 Expression and purification of amphipathic helix constructs

Amphipathic helix constructs were cloned into a pMMHb plasmid containing an N-terminal trpLE fusion and a methionine between the start of the construct and the trpLE fusion. Constructs were transformed into BL21 pLysS cells using Luria Broth (LB) agar plates with 50 μg ml^{-1} kanamycin selection and incubated overnight at 37°C. Colonies were used to start 100 mL overnights in LB which were then used to inoculate 4 L of LB. These steps were carried out at 37°C in a shaking incubator at 280 rpm and 50 μg ml^{-1} kanamycin. Once the cells reached an OD_{600} of 0.6 the cells were centrifuged at 4000 xg for 15 mins and resuspended in M9 salts and centrifuged again at 4000 xg for 15 mins and resuspended in a 4 fold decreased volume of M9 salts lacking NH_4Cl and Glucose. The cells were then incubated at 37°C for 20 mins before addition of $^{15}NH_4Cl$ and ^{12}C-glucose and then for a further 20 mins before induction with 1 mM IPTG. Expression was tested according to Section 2.3. Cells were lysed according to Section 2.4 however, instead of ultracentrifugation, inclusion bodies were harvested by centrifugation at 27,270 xg for 15 mins. The supernatants were then discarded and the inclusion bodies were resuspended in lysis buffer and pelleted again at 27,270 xg for 15 mins. The washed pellets were then resuspended in 40 mL 6 M guanidine-HCl (pH 8) and purified according to Section 2.5. All following purification steps were carried out according to Sections 2.6 and 2.7. Purified and lyophilyzed peptides were then dissolved in 20 mM potassium phosphate buffer (pH 6.5) buffer with or without detergents or lipids.

2.11 Crosslinking

25 mM glutaraldehyde (Sigma) was added to protein samples in a volume of 10 μL. All crosslinking reactions were carried out at room temperature and allowed to proceed for 30 minutes before quenching with 1 M Tris buffer (pH 7). Reactions were then analysed by NuPAGE Bis-Tris gels run in MES buffer (Invitrogen).

TABLE 2.2: 15 μL of crosslinked protein is loaded on to an SDS-PAGE and run at 200 V for 40 mins in 1x MES buffer (Invitrogen)

Reagent	Concentration	Volume (μL)
Protein	400 μM	10
Glutaraldehyde	500 mM	0.5
Incubate for 30 mins at room temperature		
Tris (pH 8)	1 M	1.5
Urea	8 M	8
Loading dye	-	10

2.12 Circular dichroism

Liposomes with average diameters of 200, 100 and 50 nm were generated by sonication and extrusion through polycarbonate membranes (Avanti Polar Lipids) from 10 mg mL^{-1} stocks in 20 mM KH_2PO_4 pH 6.5. Circular dichroism (CD) was carried out at 20 °C temperature using a Jasco J-815 spectrometer with 0.1 cm path length, 0.2 nm between each point and 10 scans per data set. Liposomes were mixed with protein at a 1:1 (w/w) ratio unless otherwise stated.

2.13 Bioinformatics

All bioinformatic analyses were performed using scripts written in the Python programming language (Rossum et al., 1989). Extensive use was made of the BioPython (Cock et al., 2009), numpy (Oliphant, 2006) and matplotlib (Hunter, 2007) packages.

2.13.1 Transmembrane helix break propensity

The transmembrane helix break propensity is calculated according to an experimentally defined scale (Monné et al., 1999). The average break propensity score over a specified window across a sequence or alignment of sequences was calculated.

2.13.2 Hydrophobic moment analysis

An alignment of sequences was generated using BLAST (Altschul et al., 1990) or Clustal Omega (Sievers et al., 2011) and filtered so that only sequences with no gaps relative to the query sequence window were analysed. If the start position of the secondary structure element was known (e.g. G135 for the APH from Yop1p) all sequences longer than 14 amino acids following the defined start position where analysed. Average hydrophobicity and hydrophobic moment (Terwilliger, 1982) (see Section 6.2) were calculated over an adjustable sliding window for every sequence in the alignment. The maximum value for hydrophobic moment was then plotted against its average hydrophobicity for each sequence. Helical wheels were plotted using HELIQUEST (Gautier et al., 2008) and my own helical wheel plotting program using MatPlotLib (Hunter, 2007).

2.14 Nuclear magnetic resonance spectroscopy

2.14.1 Detergent screening

For detergent screening and initial characterisation of sample conditions the SOFAST-HMQC pulse sequence was used (Schanda et al., 2005). This enables the rapid acquisition of high signal to noise spectra of proteins at low concentrations for screening.

2.14.2 Assignments

^{15}N-edited, ^{13}C-edited aliphatic and ^{13}C-edited aromatic NOESY spectra were collected at 950 MHz on an homebuilt Omega spectrometer (GE Healthcare) and triple resonance experiments, HNCO, HN(CA)CO, HNCA, HN(CO)CA, CBCA(CO)NH and HNCACB were collected at 600 MHz with a Bruker Avance II spectrometer with TXI cyroprobe. All spectra were processed using NMRPipe (Delaglio et al., 1995) and assignments were made using CARA (Keller, 2004). For spectra collected with non-uniform sampling (NUS), sinusoidally weighted Poisson gap schedules were generated using the Wagner group website (http://gwagner.med.harvard.edu/intranet/istHMSv2/gensched_new.html) and the data were processed using iterative soft thresholding (hmsIST) (Hyberts et al., 2012).

2.14.3 Relaxation

^{1}H-^{15}N heteronuclear NOE (hetNOE), ^{15}N longitudinal (R_1) and transverse (R_2) relaxation data were collected at 950 MHz and 600 MHz. For R_1 at 600 (950) MHz delay times of 4, 280 and 420ms (101, 302, 503, 705 and 1208 ms) were sampled. For R_2 delays were set to 0, 34 and 68 ms (5, 25, 40 and 55 ms). Data were fitted to decaying exponentials using the Follow Intensity Changes Module of Analysis 2.4 (Vranken et al., 2005) with Eqs. (2.2) and (2.3), where M_z^0 and $M_z(\tau)$ is the net magnetisation at 0 time and a given delay time τ. Errors were estimated using the covariance method (Press et al., 1992) and overlapped peaks were excluded from the analysis. Estimates for the rotational correlation, τ_c, were calculated using Eq. (2.4) which is derived from equation 8 in Kay et al. (1989). This assumes isotropic tumbling of the molecule and only takes $J(0)$ and $J(\omega_N)$ into account (i.e. ignores $J(\omega_H)$). These are strong assumptions, but without a 3D structure from which to derive inertia tensors and the relatively large errors inherent in the measurements for membrane proteins, these assumptions should provide reasonable estimates of the correlation times.

$$M_z(\tau) = M_z^0[1 - e^{(-R_1\tau)}] \tag{2.2}$$

$$M_{xy}(\tau) = M_{xy}^0 e^{(-2R_2\tau)} \tag{2.3}$$

$$\tau_c \approx \frac{1}{4\pi\omega_N}\sqrt{6\frac{R2}{R1} - 7} \tag{2.4}$$

Estimates of τ_c for distinct secondary structure regions were calculated by averaging the per-residue effective τ_c values for those regions. Data points above or below a standard deviation from the mean were excluded and the reported error in τ_c is the standard deviation.

HetNOEs were calculated according to Eq. (2.5) and errors were calculated according to Eq. (2.6). I_{ref} is the reference intensity and I_{sat} is the intensity after saturation. The noise values for the spectra (noiseSat and noiseRef) were measured in Analysis 2.4.

$$Intensity\ ratio(I_{ratio}) = \frac{I_{sat}}{I_{ref}} \tag{2.5}$$

$$error = I_{ratio} \cdot \sqrt{\left(\frac{noiseSat}{I_{sat}}\right)^2 + \left(\frac{noiseRef}{I_{ref}}\right)^2} \tag{2.6}$$

2.14.4 q-titrations

$$q = \frac{[lipid]}{[detergent]} \tag{2.7}$$

q-titrations were performed by reconstituting protein into high q LMPG:DPPC bicelles (e.g. q = 0.5) using the thin filming method described in Section 2.8. The q value was decreased by removing the protein sample from the Shigemi tube and mixing with a pre-weighed quantity of lyophilised detergent. Compound chemical shift perturbations were calculated using the method described by Mulder et al. (1999) where the ^{15}N chemical shift difference is scaled by a factor of 6.5 (R_{scale}) reflecting the ratio of variances of ^{15}N and ^{1}H chemical shifts for the 20 common amino acids across the BMRB.

$$\Delta\delta_{comp} = \sqrt{\Delta\delta_{HN}^2 + (\Delta\delta_N / R_{scale})^2} \tag{2.8}$$

Significant chemical shift perturbations were defined as those with magnitudes greater than 2 corrected standard deviations (σ) above the mean (Schumann et al., 2007). This corrected σ was calculated using a Python script (Appendix A.2).

2.14.5 Paramagnetic relaxation enhancements

Cysteine mutants were produced using site directed mutagenesis (GeneArt, Invitrogen) and then purified following the same protocols as for wild-type proteins with the exception that 5 mM β-mercaptoethanol (βME)* was added to all buffers during purification. It is noteworthy that CNBr cleavage was less efficient for the cysteine mutants, which could be due to side reactions due to the presence of a cysteine residue.

HPLC purified protein was thin filmed according to Section 2.8 and then reconstituted into 20 mM KH_2PO_4, pH 6.5 containing 5 mM dithiothreitol (DTT) and incubated at room temperature for 1 hr. Buffer exchange was then performed using a 5 mL HiTrap G25 superfine column (GE) running at 5 mL min^{-1} with 20 mM KH_2PO_4; 0.5 mM DTT, 0.05 % LMPG. 5 mM 1-oxyl-2,2,5,5-tetramethyl-η^3-pyrroline-3-methyl-methanethiosulfonate (MTSL) (25 μL of 200 mM stock dissolved in AcN.) was then added to the sample and incubated at room temperature for ≈20 hrs. The sample was loaded onto a Superdex 200 analytical column running at 0.5 mL min^{-1} with 20 mM KH_2PO_4, 0.5 mM DTT, 0.05 % LMPG at pH 6.5. The sample was then concentrated using a 30 000 MWCO Vivaspin 2 (Sartorius).

Myself and Dr. Jolyon Claridge implemented a pulse sequence for measurement of amide proton R_2 relaxation rates (Donaldson et al., 2001) on the Omega spectrometers. The

*DTT was found to be incompatible with IMAC despite manufacturers' guidelines.

same pulse sequence was also implemented by Prof. Jason Schnell on a 600 MHz Bruker spectrometer. The benefit of this experiment is that relaxation enhancements (Γ_2) are directly measured from ^{1}H R_2 values (Eq. (2.9)). This parameter is less sensitive to internal motions and cross-relaxation pathways. The effects of protein concentration changes upon addition of ascorbate (or equivalent reducing agent) are also mitigated (Donaldson et al., 2001, Iwahara et al., 2007).

$$\Gamma_2 = R_{2,para} - R_{2,dia} \tag{2.9}$$

However, due to the fast T_2 relaxation properties of the membrane proteins (and generally proteins >30 kDa) it is difficult to adequately sample the decay curve to extract reliable Γ_2 values. A potential solution to this problem was presented by Iwahara et al. (2007). They demonstrated accurate calculation of Γ_2 values from only two time point measurements for each para- and diamagnetic state according to Eq. (2.10), where T_a, T_b, I_{para} and I_{dia} are the first and second time points and the para- and diamagnetic intensities, respectively. However, the accuracy of this method relies strongly on the two time points chosen (Iwahara et al., 2007). A TROSY implementation of this experiment could also improve signal to noise and resolution for larger proteins; however, this has not been attempted to my knowledge.

$$\Gamma_2 = \frac{1}{T_b - T_a} ln \frac{I_{dia}(T_b) I_{para}(T_a)}{I_{dia}(T_a) I_{para}(T_b)} \tag{2.10}$$

Another important point is that repetition delays need to be long enough so that differences in magnetisation recovery between the para- and diamagnetic states are negligible. The use of diamagnetic forms of MTSL like 1-acetyl-2,2,5,5-tetramethyl-η^3-pyrroline-3-methyl-methanethiosulfonate (dMTSL) as demonstrated by Liang et al. (2006) can provide a more reliable control than reduction of the spin label with ascorbate or DTT since varying degrees of reduction and effects of altering the solvent conditions may reduce the reliability of intensity measurements.

Measurement of PRE effects on methyl protons could also be an avenue worth pursuing, especially in light of recent work from the Kay lab (Rosenzweig et al., 2013) where PRE restraints were calculated for protein complexes greater than 100kDa and up to 600 kDa. However, the interpretation of this data relied heavily on knowledge gleaned from a crystal structure of a homologous protein.

Distance restraints were gleaned from Eq. (2.11) using the τ_c and Γ_2 values.

$$r = \left(\frac{\beta}{\Gamma_2} \left(4\tau_c + \frac{3\tau_c}{1 + \omega_H^2 \tau_c^2} \right) \right)^{\frac{1}{6}} \tag{2.11}$$

Where $\beta = 1.23e^{-44}\ m^6\ s^{-2}$ and $\omega_H = 2\pi \cdot 600e^6\ rad\ s^{-1}$. See Figure 5.11

2.14.6 Residual dipolar couplings

Residual dipolar couplings (RDCs) were collected by soaking protein samples into 4.5% polyacrylamide gels from a 6 mm casting tube. The reagents in Table 2.3 were added to a 1.5 mL tube in the order they are listed and carefully mixed by pipetting up and down before transfer to the casting chamber. It is important to work quickly and tap out any bubbles; a P200 pipette set to 141 μL for transferring the gel mixture into the casting chamber gives more fine control than a P1000.

TABLE 2.3: Recipe for polyacrylamide gel for RDCs

Compound	**Volume (μL)**
dH_2O	2x 115.75
40% Acrylamide (37.5 : 1)	31.5
2% APS	17
TEMED	0.5
Total	280.5

After casting gels were transferred to a 50 mL tube containing deionised water and allowed to soak overnight with at least three changes of water. Gels were then dried out for 1 hour on each end on a piece of parafilm inverting regularly to avoid bulging at either end. Once the gel becomes sticky to the touch with a pipette tip it can be transferred into a tube containing the protein sample and left to soak for at least 48 hours. Once soaked the gel is again dried on a piece of parafilm until it is narrow enough to fit into the 6 mm casting chamber. Residual protein solution is used for lubrication and the gel is then carefully extruded into an NMR tube. See Section 7.5.

2.14.7 Estimation of τ_c from ^{15}N^{1}H-TRACT experiments

$^{15}N^1H-$TRACT spectra (Lee et al., 2006) were processed using NMRPipe (Delaglio et al., 1995) and each 1D spectrum was integrated between 8.2 and 8.8 ppm using Nmrglue (Helmus and Jaroniec, 2013). The integrals for the α and β states were then fit to a decaying exponential function (2.12) using a python script where Δ_i is the delay time in ms and σ_c is the integral over the selected spectral region. The values for R_α

and R_β (in Hz) were then used to calculate τ_c using the relation $2\eta_{xy} = R_\beta - R_\alpha$ where η_{xy} is the transverse cross-correlated relaxation rate.

$$\sigma(c) \propto e^{(-R_{\alpha/\beta}\Delta_i)} \tag{2.12}$$

The equations reported by Lee et al. (2006) were then rearranged in the following way to produce a cubic expression (Eq. (2.20)) which I solved with the numpy.roots() algorithm using $c1.6\omega^2, -2\eta_{xy}\omega_N^2, c2.8$ and $-2\eta_{xy}$ as roots a, b, c and d respectively with the real solution being equal to τ_c in seconds (see Appendix A.1).

$$2\eta_{xy} = 2\rho\delta_N(4J(0) + 3J(\omega_N))(3cos^2\theta - 1) \tag{2.13}$$

$$c = 2\rho\delta_N(3cos^2\theta - 1) \tag{2.14}$$

$$2\eta_{xy} = c(4J(0) + 3J(\omega_N))(3cos^2\theta - 1) \tag{2.15}$$

$$2\eta_{xy} = c4J(0) + c3J(\omega_N) \tag{2.16}$$

$$2\eta_{xy} = c1.6\tau_c{}^{\dagger} + \frac{c1.2\tau_c}{1 + (\tau_c\omega_N)^2} \tag{2.17}$$

$$2\eta_{xy} = \frac{c1.6\tau_c(1 + (\tau_c\omega_N)^2) + c1.2\tau_c}{1 + (\tau_c\omega_N)^2} \tag{2.18}$$

$$2\eta_{xy}(1 + (\tau_c\omega_N)^2) = c1.6\tau_c(1 + (\tau_c\omega_N)^2) + c1.2\tau_c \tag{2.19}$$

$$0 = c1.6\omega^2\tau_c^3 - 2\eta_{xy}\omega_N^2\tau_c^2 + c2.8\tau_c - 2\eta_{xy} \tag{2.20}$$

[†] 4J(0) simplifies to $1.6\tau_c$ since $\omega_N = 0$

3

Non-Uniform Sampling

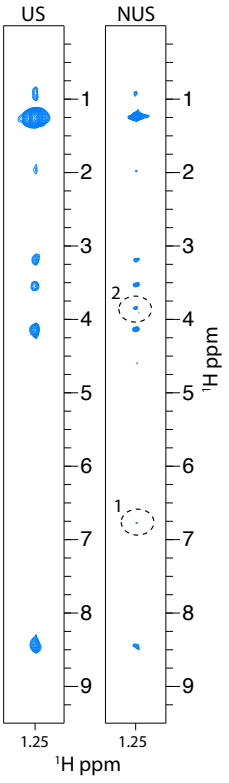


FIGURE 3.1: ^{1}H-^{1}H strips from ^{13}C-edited NOESYs for Ala 145 from Yop1_ΔC2 (400 μM). Uniformly sampled (US, left) spectrum was acquired using 64 (64) points in indirect proton (carbon). The 25% non-uniform sampled (NUS, right) spectrum was acquired with 300 (128) points in indirect proton (carbon) using Poisson gap sampling and hmsIST for reconstruction. Both spectra took ∼3 days to acquire at 950 MHz using 8 scans per point. The dashed circles highlight two peaks that are not observed in the US spectrum but can be seen clearly in the NUS spectrum due to the increased sensitivity (1) and resolution (2).

3.1 What is NUS?

Traditionally NMR data are collected using linear incrementation of chemical shift evolution times in the indirect dimensions. The resulting time domain data are then fast Fourier transformed to produce the frequency domain data.

In non-uniform sampling (NUS), only a fraction of the indirect data points are sampled resulting in a non-uniform spread of points across the interferograms. This incompletely sampled data set is then reconstructed to reproduce the full spectrum.

3.2 Why use NUS?

If you collect 64 points in ^{15}N at 600 MHz with a sweep width of 30 ppm you have a digital resolution of ≈30 Hz per point. If you collect the same experiment at 950 MHz then you get ≈48 Hz per point. To match the digital resolution at 600 MHz it is necessary collect ≈100 indirect points. Therefore, to take full advantage of the greater resolving power at high fields it is necessary to sample additional points in the indirect dimensions.

Collection of NUS experiments in a time equivalent manner relative to the uniformly sampled (US) experiment has been shown to increase the signal to noise ratio (SNR), resolution and enhance the ability to observe weak peaks that would otherwise be lost using traditional sampling methods and FFT (Hyberts et al., 2013). The benefits can be seen clearly in Figure 3.1 where the time-equivalent NUS ^{13}C-edited NOESY spectrum of Yop1_ΔC2 has superior resolution and also the appearance of peaks that were lost in the US spectrum.

However, it is important to note that the efficacy of NUS relies heavily on the reconstruction algorithms and sampling schedules used.

3.3 Sampling schedules

Defining a sampling schedule that gives optimal signal to noise, resolution and fidelity to the US spectrum per unit time is a difficult problem. For example, if you want to sample 256 out of 1024 indirect points on a Nyquist grid then there are more than 10^{248} unique combinations of points to choose from (number of combinations = $\frac{n!}{(n-r)!}$, where n = number of points and r is the number of samples)! If you use a random number generator (such as UNIX drand48) to select the points then the sampling schedules will

be strongly biased by the starting seed values. This in turn results in highly variable schedules and hence spectral qualities.

According to the Wagner group (Hyberts et al., 2014) optimal sampling schedules have a maximum evolution time of 1.2 T_2 and large gaps between sampling points should be minimised, especially at long and short acquisition times (Hyberts et al., 2010). To maximise the benefits of NUS the spectra should also be collected in a time equivalent manner where the reduction in sampling time is compensated by increased scans (Hyberts et al., 2013).

One way of assessing the effectiveness of a given sampling schedule is to calculate $L^2 = \sqrt{(f_i^{ref} - f_i^{rec})^2}$, where f^{ref} is the US spectrum and f^{rec} is the reconstructed NUS spectrum. The lower the value for L^2 the better the sampling schedule is at recapitulating the US spectrum (Hyberts et al., 2012). These methods require collection of both US and NUS spectra for comparison which is time consuming. Thus, a key goal in the NUS field has been creating algorithms that enable reliable generation of optimal sampling schedules.

In my experience, the most significant breakthroughs in recent times in this regard have come from the Wagner group with their introduction of sinusoidal weighted Poisson gap sampling (Hyberts et al., 2010) and hmsIST (Hyberts et al., 2012). These advances significantly reduced variations in spectral quality due to sampling schedules and have enabled user friendly processing of NUS data which, most importantly, can be performed on a laptop on the order of minutes.

Recently Aoto et al. (2014) developed a method for selecting Poisson gap sampling schedules that most faithfully reproduce peak intensities and positions relative to a US spectrum. This work has exciting applications in enabling accurate relaxation data collection on more challenging systems.

Table 3.1: List of sampling methods

Sampling types
Uniform random
Exponentially weighted random sampling
5*exponential sampling
US/NUS Sampling
Sinusoidal weighted Poisson gap distribution

Table 3.1 lists the main NUS sampling types available and below is a brief introduction to each of the types.

3.3.1 Uniform random

Uses UNIX random number generators and the sampling schedule is strongly biased by the initial seed values.

3.3.2 Exponentially weighted random sampling

An exponentially decaying function is applied over the probability function which results in more points being sampled at short delay times where the signals are most intense and less sampling as t_1 and t_2 increases.

3.3.3 5*exponential sampling

The exponential is 5 times over-weighted towards the beginning of the schedule. This results in few truncation artefacts due to gaps but it reduces the resolution of the resulting spectrum since you sample fewer points further out.

3.3.4 US/NUS Sampling

$1/8^{th}$ of the initial points are sampled uniformly and then the subsequent $7/8^{th}$s are sampled using uniform random sampling. This performs better than the above schedules giving rise to fewer artefacts such as "dips" or "trenches".

3.3.5 Sinusoidal weighted Poisson gap distribution

Large gaps in the sampling schedule are very detrimental to the fidelity of a reconstructed spectrum, especially if they occur in the early or late points. Sampling also has to be sufficiently random to satisfy the Nyquist theorem (Hyberts et al., 2010). For example if gap sizes between points in the sparse sampling grid is non-random then aliasing artefacts may arise since multiple frequencies can be fitted to the data.

In a Poisson distribution, the average gap size should be 1/(sparsity -1) where sparsity is the fraction of points collected. Therefore, the average gap size for a 25% sampled spectrum should be 3 (Figure 3.2). Poisson gap sampling gives the lowest L^2 values relative to other sampling methods (listed in above and in Table 3.1) and is least dependent on the initial seed values (Hyberts et al., 2010). The sinusoidal weighting on the distribution of gaps ensures that there are fewer gaps towards the beginning of the schedule and takes into account the fact that the measured signals are decaying.

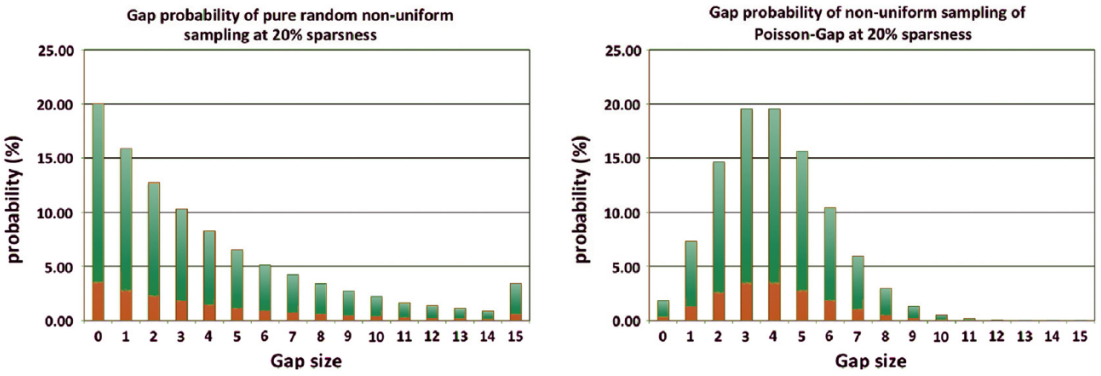


FIGURE 3.2: Comparison of gap distributions for pure random sampling (left) and Poisson gap sampling (right) taken from Hyberts et al. (2014). It can be seen that the Poisson gap distribution avoids large gaps compared with the random sampling schedule. The bar at 15 is for gap sizes of 15 or more.

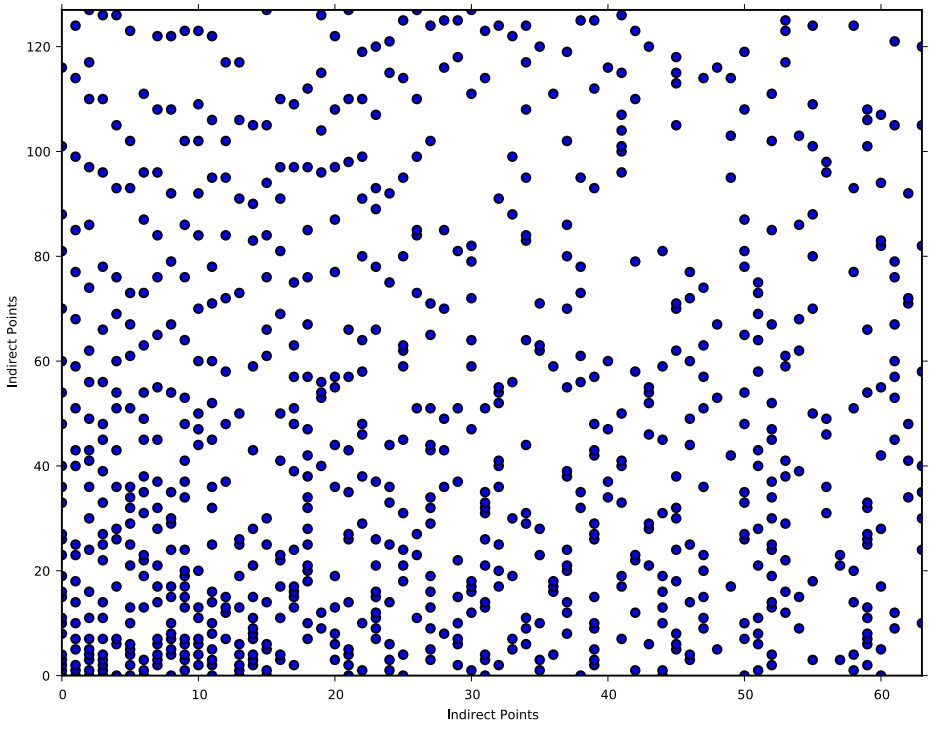


FIGURE 3.3: Example of a 25% sinusoidally weighted Poisson gap sampling schedule (Hyberts et al., 2010). Note that points near the beginning of the schedule are more densely sampled.

3.4 Reconstruction algorithms

The concept of the algorithm dates back to at least the 7^{th} century AD. The idea is to solve a complex mathematical problem by finding a simple set of instructions that can be executed step by step and iteratively. Ideally each iteration of the instruction set (algorithm) brings one closer to the best approximation of the answer. For a given problem there will be many different ways to approximate the solution and hence many different possible algorithms. However, the efficiencies of these algorithms can vary; some algorithms will rapidly produce a solution whereas others may take longer but produce a more accurate approximation. Conversely, some may take longer to produce a less accurate approximation of the solution. Reconstruction of NUS data is a prime example of algorithm development and the battle between accuracy and efficiency. Until very recently, computational resources have been insufficient to enable widespread use of NUS; even on a powerful computer (for the day), reconstruction could take as long as ∼2 days (Tugarinov et al., 2005b). There are now many algorithms available for reconstruction of NUS data and with increased computational resources they are becoming generally accessible with reconstruction times ranging from hours to minutes. Below is a brief introduction to some of the more prominent reconstruction algorithms. It remains to be seen which algorithm and accompanying software will progress into mainstream usage amongst the wider NMR community.

3.4.1 Iterative soft thresholding (IST)

NUS time domain data with all non-sampled data points set to 0 are FFTed. Signals above a threshold are removed and stored in a separate file. The now truncated frequency domain data are inverse FFTed and the grid points that were not initially sampled are again set to zero. This process is repeated using a lower threshold each time until the noise is reached and there are no residual signals (Hyberts et al., 2012). The result is reconstructed frequency domain data with few artefacts.

3.4.2 Iterative re-weighted least squares (IRLS)

IRLS works by minimising the $l_p - norm$ penalty function of a spectrum (Kazimierczuk and Orekhov, 2011). It performs similarly to IST but requires more computational time. I have found this method to be reasonably successful using the qMDD software (Orekhov and Jaravine, 2011); however, hmsIST is preferable. Both IST and IRLS are examples of compressed sensing algorithms.

3.4.3 Forward maximum entropy (FM) reconstruction

All non-sampled points are set to 0 and the resulting FIDs are FFTed to give a starting function $Q(\mathrm{f})$ which can be the sum of the absolute values of the frequency domain points, or the Shannon Entropy. The values of the non-sampled points are then perturbed and re-FFTed to give a set of new $Q(\mathrm{f})$s, $Q(\mathrm{f}+\delta_i)$. Polak-Riviere conjugate gradient minimisation is used and only the non-sampled data are perturbed. In layman's terms, it finds the best spectrum that is consistent with the NUS data.

3.4.4 Multi-dimensional decomposition (MDD)

MDD uses parallel processing of spectral regions (Orekhov and Jaravine, 2011). This approach is very computationally expensive and requires much more time than IST or IRLS. In my hands, the results are generally inferior to IST since only diagonal peaks and some of the more intense crosspeaks are reconstructed, weak signals are mostly lost.

An advantage of MDD is that it can be used to reconstruct multiple high dimensionality datasets ($\geq$4D) in order to improve the resultant spectra. This approach uses prior knowledge from the shared dimensions and is called coupled MDD (Co-MDD). Co-MDD was used successfully to reconstruct NUS 4D methyl-methyl and amide-methyl NOESYs along with a 2D ^{13}C-HMQC collected on VDAC-1, a β-barrel membrane protein with a tumbling time 70 ns (Hiller et al., 2009).

3.5 Choosing a reconstruction algorithm

In my experience, the algorithm that most efficiently and faithfully reconstructs NUS spectra is hmsIST (Hyberts et al., 2012) used in conjunction with Poisson gap sampling methods (Hyberts et al., 2010). Compared with other algorithms such as IRLS and MDD, hmsIST is generally able to reconstruct weak signals even in high dynamic range datasets. Another advantage is that hmsIST relies primarily on the discrete Fourier transform (DFT) for reconstruction so that the computational time required to reconstruct spectra increases linearly with the size of a dataset as opposed to other algorithms (see below) (Hyberts et al., 2014).

IST is implemented in $n \cdot \log(n)$ time,
Other algorithms such as IRLS and MDD are implemented in n^2 time,
Where n is the number of calculations.

This can mean the difference between hours and minutes when it comes to the time taken for reconstruction of large multidimensional data sets.

3.6 Limitations of NUS

In general the advantages of NUS far outweigh its limitations. However, NUS should be used with care since under-sampling or choosing incrementation periods that are too long or short can have detrimental effects on the resultant spectra. Also, experiments must usually be modified to enable acquisition of NUS data. Most modern spectrometers now have updated pulse sequence libraries and collection of NUS data is becoming routine. Some of the pitfalls and my recommendations for NUS data collection are discussed below.

3.6.1 Choosing an optimal sampling density

High dynamic range in peak intensities for some biological samples, in particular membrane proteins where overlap is also common, can make reconstruction problematic. In general the sampling density should be proportional to the number of signals expected for each transient. Therefore it is generally recommended to sample more densely for NOESY experiments than triple resonance experiments, especially since the number of peaks per transient will vary. According to Hyberts et al. (2014) a rule of thumb for maximum sparsity is 20% in each dimension which would mean 4% and 0.8% of 3D and 4D datasets, respectively. However, if peak overlap and high dynamic range between peaks is a concern, sampling densities $> 20\%$ per dimension should be used.

$$\text{Sampling density for each dimension } (\%) = \sqrt[n]{\frac{x}{100}} \times 100$$

Where n is the number of NUS dimensions and x is the overall sampling density.

From my experience, 40-50% sampling of each NUS dimension results in faithful reconstruction of all peaks including those overlapped and weak peaks.

For samples with unfavourable relaxation properties (such as membrane proteins) it may not be possible to increase the resolution markedly; however, by reducing the sampling density one may be able to improve the SNR by collecting more scans. It is important

TABLE 3.2: Tried and tested sampling densities. These sampling densities collected using Poisson gap sampling resulted in faithfully reconstructed spectra with few artefacts, improved S/N and resolution. In the case of the NOESY experiments the digital resolution in both dimensions is up to double that of the traditionally sampled experiments collected in Oxford. All spectra were reconstructed using 400 iterations of the hmsIST software (Hyberts et al., 2012) in conjunction with NMRPipe (Delaglio et al., 1995). The triple resonance experiments were collected at 600 MHz all others were collected at 950 MHz.

Sample	Experiment	Labelling	t1 points	t2 points	Sampling Density
Yop1_ΔC2	^{15}N-edited NOESY	^{2}H^{15}N	128	256	4-10%
Yop1_ΔC2	^{15}N-edited NOESY	^{15}N	128	196	20-25%
Yop1_ΔC2	^{13}C-edited NOESY	^{13}C^{15}N	128	300	20-25%
Yop1_ΔC2	HNCA	^{13}C^{15}N	42	128	4-10%
Yop1_ΔC2	HNCO	^{13}C^{15}N	42	150	4-10%
Yop1_ΔC2	HSQC-type	^{2}H^{15}N	192-256	-	25-50%

to reiterate that it is most effective to use NUS in a time equivalent manner. For example, if an experiment takes 3 days with traditional sampling, take 3 days for the NUS experiment. By increasing the number of scans and or resolution the resulting spectra will be markedly improved. Some tried and tested NUS experiments are listed in Table 3.2.

3.7 NUS on the Omega spectrometers

To fully utilise the resolving power of the high field magnets in Oxford (namely 950 and 750 MHz) I set out to implement NUS methods on these spectrometers. The following section documents the implementation of NUS on the Omega spectrometers and instructions for future reference.

Below is an example of a sampling schedule that can be read by an Omega spectrometer:

```
/* Example NUS sampling schedule */
nuslist_name[0] = {
1,
2,
3,
4,
5,
6,
etc
};
```

The pulse sequence then needs to be modified in the following way:

```
int nuslist_name[4096];
$
#include "nuslist_name"
@
```

3.7.1 Two dimensional spectra

Most of the 2D pulse sequences can be modified in the following manner:

Complex

```
int t1inc   = nbc / 2;
real evt1   = t1de + (nuslist_name[t1inc%256] * t1dw );
```

For **Complex** indirect detection use **acqmode -f na 0 nb 0**, where nb is the number of complex points.

i.e. 2× The number of points in your nuslist

Echo-Antiecho

```
int t1inc   = ncc;
int stppi   = ( ncc %2 ) * 2;
real evt1   = t1de + (nuslist_name[t1inc%256]*t1dw);
```

For **Echo-Antiecho** experiments use **acqmode -f na 0 2 0 nc 0**, where nc is the number of real points.

i.e. The number of points in your nuslist

3.7.2 Three dimensional spectra

For experiments with Complex indirect dimensions use **acqmode -f na 0 2 0 nc 0**, where nc is **2x the number of points** in your sampling schedule.

```
int t1inc = ncc/2;

real evt1  = t1de+(increment1[t1inc%4096] * t1dw);
real evt2  = t2de+(increment2[t1inc%4096] * t2dw);

int f1sign = (nbc%2);        /* For F1 sign discrimination;  0101 */
int stppi1 = (nbc%2)*2; /* 0  2  0  2  */
int f2sign  = (ncc%2);
int stppi2  = ((ncc/2)%2)*2;
```

For experiments with Echo-Antiecho in the HSQC block:

Use **acqmode -f na 0 4 0 nc 0**, where nc is **the number of points** in your sampling schedule.

```
int t1inc = ncc;

real evt1  = t1de+(increment1[t1inc%4096] * t1dw);
real evt2  = t2de+(increment2[t1inc%4096] * t2dw);

int f1sign = (nbc/2);       /* For F1 sign discrimination;  0011 */
int stppi2 = (ncc%2)*2; /* 0  2  0  2  */
```

It is necessary to recompile a pulse sequence each time a new sampling schedule is used.

```
sequel -l list -o seqname.out seqname.s;
tail list
```

3.8 Processing NUS data with hmsIST

Example processing scripts for a ^{15}N-NOESY collected at 950 MHz using non-uniform sampling with Poisson gap sampling and hmsIST. In some cases the data must be collected in chunks since sampling schedules exceeding ~3000 points cause the homebuilt Omega spectrometer console to crash. In these cases the ".bin" files are added together after acquisition using the Unix "dd" command, an example of which is shown below.

```
#!/bin/csh
rm -r fid
#(dd if=$1 bs=8192 count=9948;\ #Use dd to add multiple binary files together
# dd if=$2 bs=8192 count=9948)\ #Count = 4 * number of  indirect real points
#|bin2pipe -ge -neg\     #bs = 4 * block size (4 bytes per 32 bit float)
bin2pipe -in $1.bin -ge -neg\
-xN              2048  -yN                 4  -zN   length of nuslist   \
-xT              1024  -yT                 2  -zT   length of nuslist/2 \
-xMODE        Complex  -yMODE          real  -zMODE          real      \
-xSW            12500  -ySW         2994.012  -zSW           10000      \
-xOBS     949.7929580  -yOBS      96.2524520  -zOBS    949.7929580      \
-xCAR           4.618  -yCAR          116.502  -zCAR          4.618      \
-xLAB              HN  -yLAB              15N  -zLAB             1H      \
-ndim               3  -aq2D           States-TPPI                      \
| nmrPipe -fn MAC -macro $NMRTXT/bin_ranceY.M -noRd -noWr   \  # Echo-Antiecho
| pipe2xyz -x -out ./fid/test%03d.fid -verb -ov -to 0
```

The raw data are then Fourier transformed to obtain the incompletely sampled interferograms for reconstruction with hmsIST. Separate scripts are marked by a line of "#" symbols.

```
#!/bin/csh -f
rm -rf yzx # clean up
rm -rf yzx_ist # clean up
mkdir yzx
mkdir yzx_ist

xyz2pipe -in fid/test%03d.fid -x                    \
| nmrPipe  -fn SOL                                  \
| nmrPipe  -fn SP  -off 0.5 -end 0.98 -pow 1 -c 1   \
| nmrPipe  -fn ZF  -zf 2 -auto                      \
| nmrPipe  -fn FT  -verb                            \
| nmrPipe  -fn PS  -p0 52 -p1 0.0 -d                \
| nmrPipe  -fn EXT -x1 7ppm -xn 9ppm -sw            \
| pipe2xyz -ov -out yzx/test%03d.nus -z
#################################
#Parallel is used to run the hmsIST processes on multiple cores
parallel -j 100% './ist.csh {} > /dev/null; echo {}' ::: yzx/test*.nus
#################################
#!/bin/csh
set F = $1
set in = $F:t
set out = $F:t:r.phf
echo $in $out
hmsIST -dim 2 -incr 1 -autoN 1 -user 1 \
    -itr 100  -verb 0 -ref 0 -vlist ./nuslist1 \
    < ./yzx/${in} >! ./yzx_ist/${out}
#################################
#!/bin/csh
xyz2pipe -in yzx_ist/test%03d.phf | phf2pipe -user 1 -xproj xz.ft1 -yproj yz.ft1↩
     | pipe2xyz -out rec/test%03d.ft1
#################################
#!/bin/csh -f
# 3D States-Mode HN-Detected Processing.
xyz2pipe   -in rec/test%03d.ft1 -y              \
| nmrPipe  -fn SP -off 0.5 -end 0.98 -pow 1     \
| nmrPipe  -fn ZF -zf 2 -auto                   \
| nmrPipe  -fn FT -auto -verb                   \
| nmrPipe  -fn PS -p0 -90.0 -p1 180 -di         \
| nmrPipe  -fn TP                               \
#| nmrPipe  -fn LP -f -x1 1 -xn 128 -pred 64    \
| nmrPipe  -fn SP -off 0.5 -end 0.98 -pow 1     \
| nmrPipe  -fn ZF -zf 2-auto                    \
| nmrPipe  -fn FT -neg -verb                    \
| nmrPipe  -fn PS -p0 -90.0  -p1 180.0 -di      \
| nmrPipe  -fn TP                               \
#| nmrPipe  -fn ZTP                             \
> rec/data.pipe
pipe2xyz -in rec/data.pipe -out rec/test%03d.ft3 -z
```

4

Bacterially Expressed Yop1p Oligomerises and Forms Membrane Tubules *in vitro*

4.1 Expression and purification of Yop1p

A prerequisite to the study of protein structure by NMR is the production of samples enriched with NMR observable nuclei such as ^{15}N and ^{13}C. In more challenging cases deuteration and selective labelling may also be required. In order for this to be economically viable and in some cases possible it is generally necessary to use bacterial expression systems.

I chose to express protein in *E. coli* (BL21 and C43) cells using a slightly modified form of the Single Protein Production (SPP) system (Suzuki et al., 2007) (Section 2.2). The main benefit of the SPP system is that the 10 to 20 fold condensation at the point of induction greatly reduces the cost of labelling enabling the production of milligram amounts of labelled membrane protein from relatively little labelled media. Importantly also, aside from a 6x His tag, the protein was able to be expressed without the need for additional expression tags (such as trpLE or MBP) thus simplifying the purification strategy.

Initially I expressed full length Yop1 (FL-Yop1p) (Figure 4.1) and purified it by detergent extraction (Table 4.1). However, this approach did not give promising results and NMR spectra were not obtainable from these samples. An overview of the purification strategies is shown in Table 4.1.

TABLE 4.1: Preliminary purifications of FL-Yop1p gave pure samples but no NMR signal likely due to aggregation or polydispersity

Extraction	Exchanged	Pure	Spectra	Crosslinking
LDAO (1%)	-	yes	poor	-
DDM (1%)	-	yes	no signal	aggregated
DDM (1%)	0.5 % FC-10	yes	no signal	aggregated

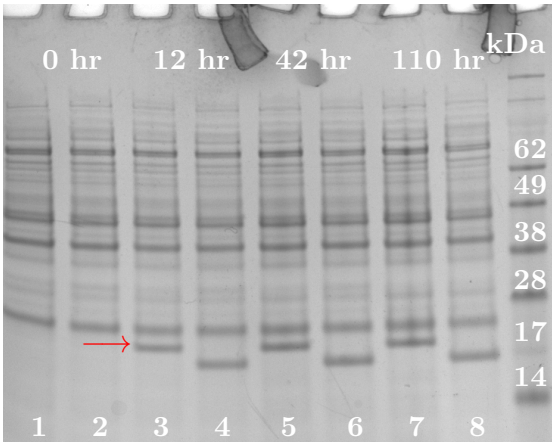


FIGURE 4.1: Time course of expression trial of FL-Yop1p. Samples were taken at intervals 0, 12, 42 and 110 hr and loaded according to Section 2.3. The odd numbered wells contain samples of FL-Yop1p. The evenly numbered wells contain samples of Yip3p a Rab-GDI displacement factor ortholog for which expression and purification trials were performed but no further progress was made. The position of the FL-Yop1p expression band is marked by a red arrow.

Initial extractions from the membrane fraction did not give promising results (see Figure 4.1 and Table 4.2) since detergent extractions were inefficient and good NMR spectra were unobtainable.

TABLE 4.2: Initial detergent extractions of membrane fractions for FL-Yop1p were unsuccessful. Harsher extraction in 6M GuHCl was successful. Samples were categorised as solubilised by visibility of the original expression band (Figure 4.1) and absence of visible aggregation in the sample wells of an SDS-PAGE.

Extraction	Solubilised
DDM (4%)	no
DM (4%)	no
DPC (2%)	partial
LDAO (4%)	no
GuHCl (6M)	yes
GuHCl (6M)+Triton(X100)	yes

I decided to try expressing a truncated construct of Yop1p comprising residues 24-165 (Yop1_ΔC2) since it was previously shown to be more efficient at tubule formation than FL-Yop1p (Hu et al., 2008). Being 4 kDa smaller this construct had the additional potential advantage of being more tractable for study by NMR.

His-tagged Yop1p was previously shown to be non-functional (Hu et al., 2008). Therefore an efficient cleavage method for the His-tag was required. CNBr cleavage was chosen since it has very high efficiencies and is compatible with the non-native purification strategy. Also, preliminary attempts to cleave the tag using Factor Xa were less than 50% efficient. Since CNBr cleaves peptide bonds C-terminal to methionine, it was necessary to remove M138. I chose to mutate to isoleucine as it is the consensus amino acid for that position (see Figure 4.2). For the rest of this thesis I will refer to this conservative mutant as Yop1_ΔC2.

Native purifications of this construct resulted in poor spectra and low protein yields. I next attempted purification using 6 M GuHCl and IMAC (see Sections 2.4 and 2.5) followed by HPLC and refolding by thin filming lyophilised protein with organic solvent in the presence of detergents and lipids (see Sections 2.7 and 2.8). A benefit of this approach is the ease with which one can screen conditions since lyophilised protein can be aliquoted and mixed with detergents and lipids of choice in a relatively rapid manner. Table 4.3 displays some of the sample conditions tested with the resulting spectra shown in Figure 4.3. To test these samples for oligomeric behaviour I used glutaraldehyde crosslinking experiments followed by SDS-PAGE analysis (see Section 2.11). These results are also summarised in Table 4.3.

Finally, reconstitution of the protein into LMPG micelles resulted in relatively homogeneous samples that gave high quality NMR spectra and stable protein. LMPG has

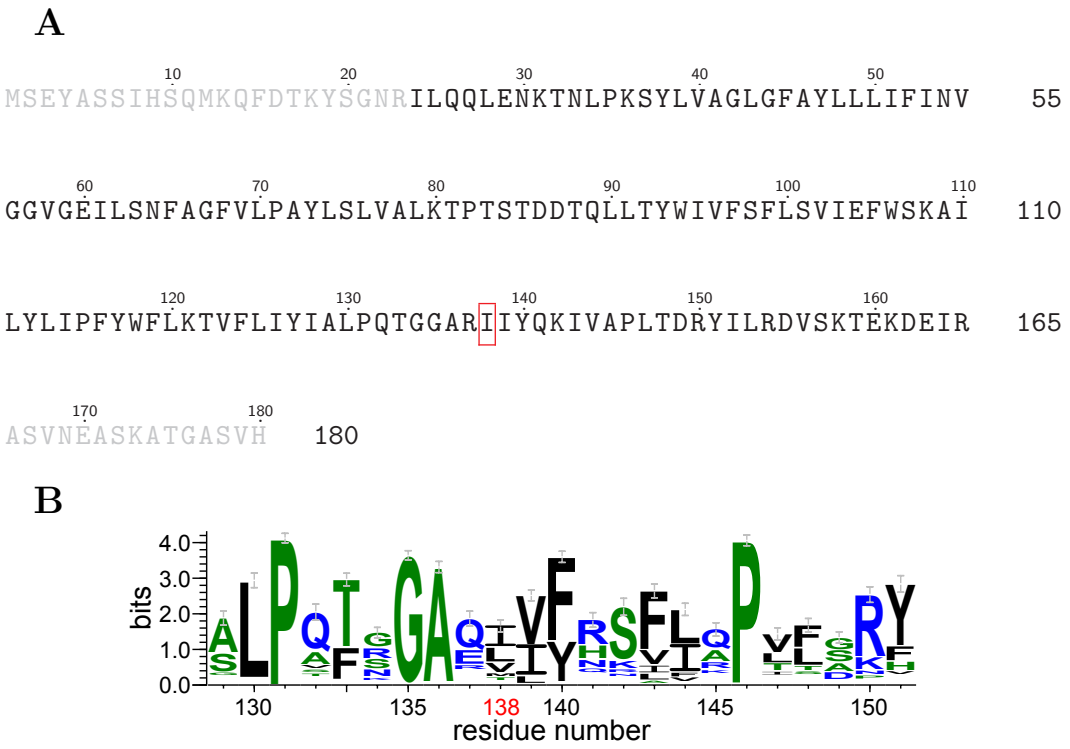


FIGURE 4.2: (A) Yop1p construct used for structural studies (Yop1_ΔC2). The black region comprising residues 24-165 was expressed. Residue 138 is boxed in red and was mutated from Met to Ile to facilitate CNBr cleavage of the His-tag from the construct. (B) Sequence Logo for mutated region demonstrates that Ile is the consensus amino acid for that position (36% I, 26% L, 18% V, 6% M, 14% other). The sequence logo was generated using the first 250 sequences from a BLAST alignment (Altschul et al., 1990) which were then re-aligned using Clustal Omega (Sievers et al., 2011) and fed into the WebLogo3 server (Crooks et al., 2004).

TABLE 4.3: Initial conditions explored for the Yop1_ΔC2 construct. Samples were prepared by thin filming from HFIP and Predominant species was determined by chemical crosslinking with 25 mM Glutaraldehyde for 30 mins at room temperature. Spectra were collected using the SOFAST-HMQC pulse sequence (Schanda et al., 2005). Conditions thought to form bicelles are starred (*).

Solubilising Agent	Concentration (μM)	[Protein] (μM)	Predominant species	Spectra
DDM	100	8	Dimer	Poor
DDM : SDS	100 : 20	8	Dimer/Monomer	Poor
DMPC : DHPC*	50 : 125 (q = 0.4)	26	Hexamer	Poor
DMPC : DHPC*	50 : 200 (q = 0.25)	26	Dimer	Poor
DMPC : DHPC : SDS*	50 : 200 : 20	26	Dimer	Poor
LMPG	100	27	Monomer	Good
LMPG : DOPC*	100 : 10 (q=0.1)	27	Monomer	Good
LMPG : DMPC*	100 : 50 (q=0.5)	27	Monomer/Ladder	Poor

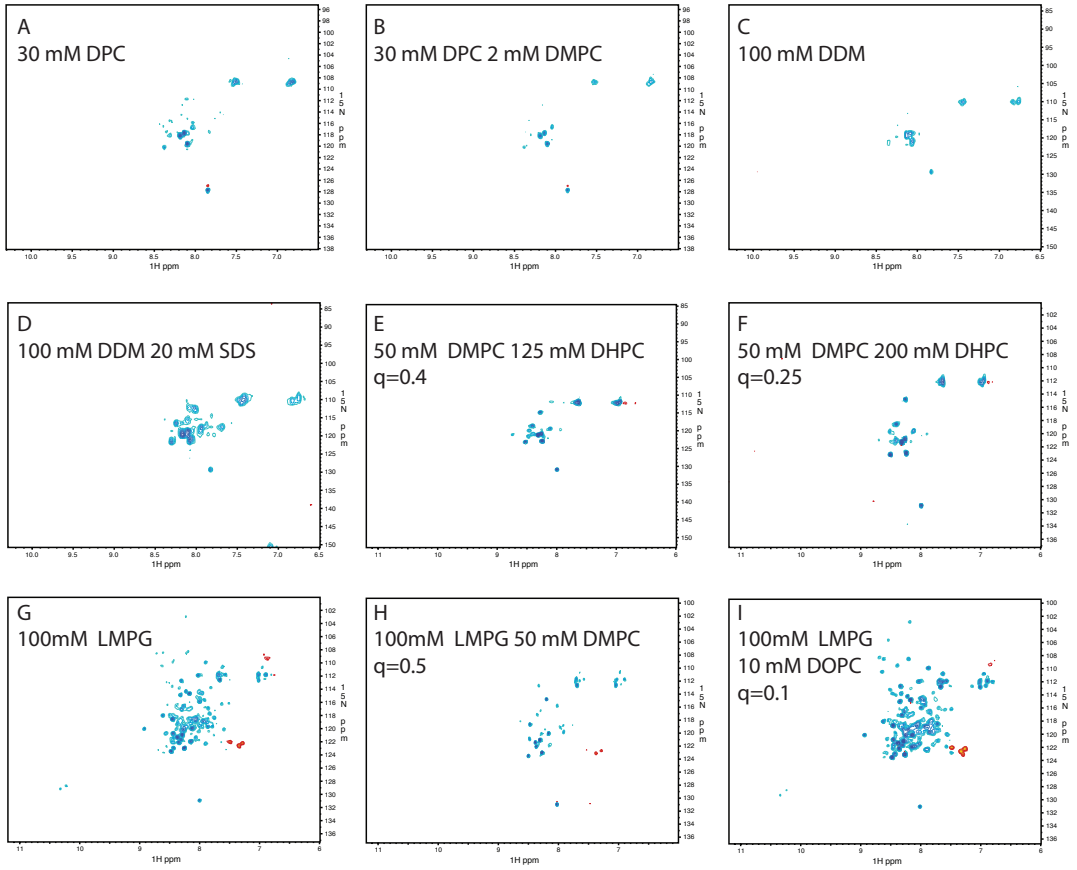
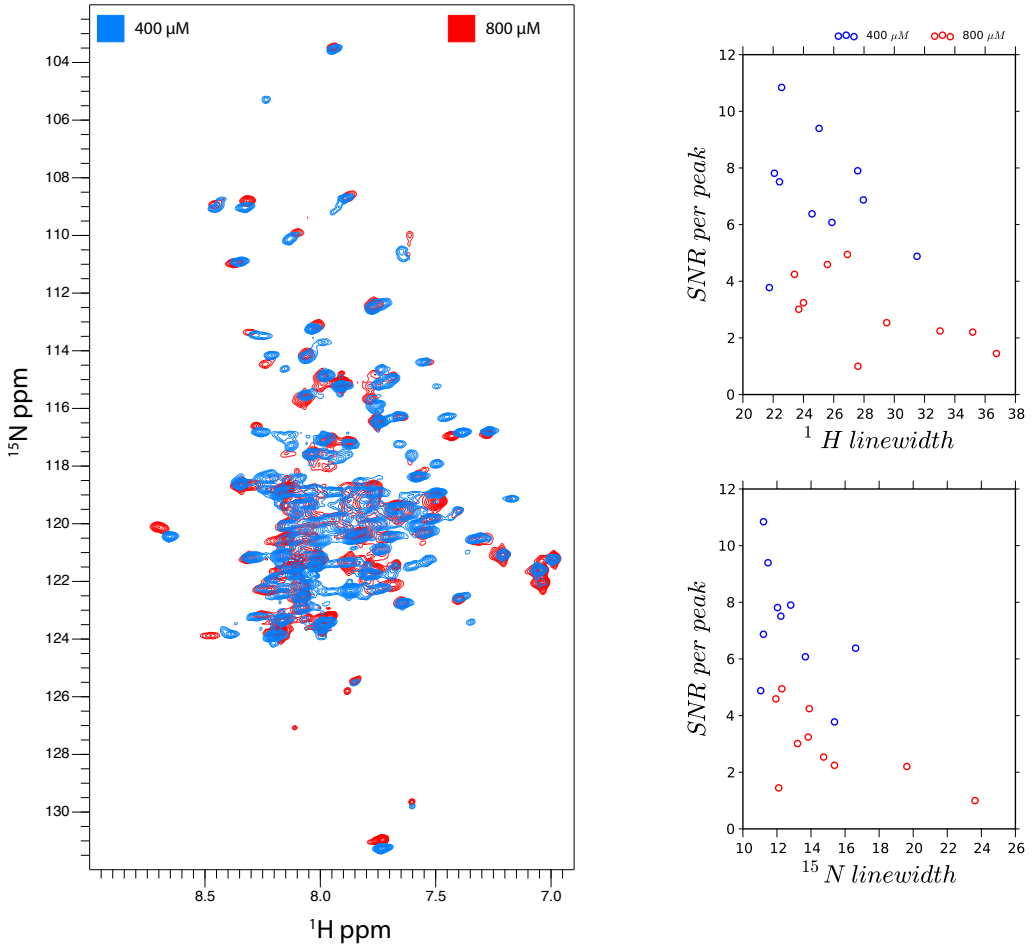


FIGURE 4.3: Sample spectra from detergent screening using the SOFAST-HMQC pulse sequence (Schanda et al., 2005). A-B, 950 MHz 37 °C. C-I, 600 MHz 37 °C. DDM, DPC and DHPC:DMPC bicelles gave broad linewidths and too few peaks. LMPG was shown to give narrower line widths and approximately the right number of peaks assuming some overlap.

been previously successful for NMR studies of membrane proteins (Barrett et al., 2012, Klammt et al., 2012, Tian et al., 2007) in particular Klammt et al. (2012) extensively screened detergents for solubilising human membrane proteins and found LMPG to be one of the most successful detergents. However, use of LMPG is not without controversies. Shenkarev et al. (2010) showed that while providing good spectra and sample longevity, LMPG and LPPC were unable to maintain activity of the voltage domain from a K+ KvAP channel under investigation. Similarly, Poget and Girvin (2007) showed that LMPG was unable to support the multidrug resistance transporter, Smr, in a functional state. With these considerations in mind I set out to characterise the functional state of Yop1_ΔC2 after reconstitution into LMPG micelles.

It is worth noting that I chose to collect NMR data for assignments at concentrations of 400 μM or below. Initially I produced an 800 μM U^{13}C-^{15}N-^{2}H sample which upon



(A) (B)

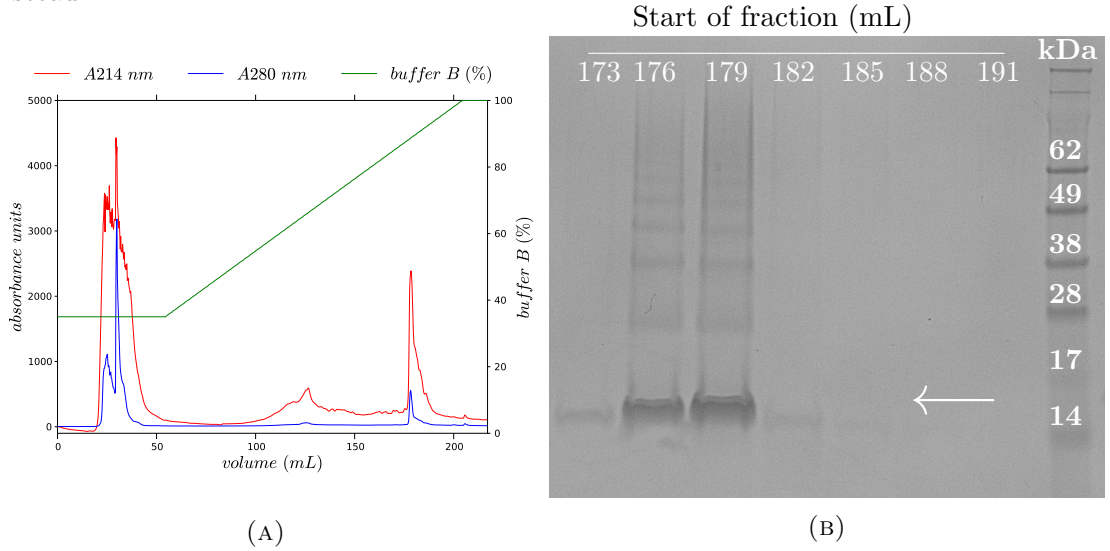
FIGURE 4.4: (A) Overlay of TROSY-HSQC spectra collected on Yop1_ΔC2 at 800 μM (red) and 400 μM (blue) at 500 MHz (42 °C, pH 6.5). 512 and 128 complex points were collected for the ^{1}H and ^{15}N dimensions respectively with 16 scans for each. The direct (indirect) sweep width was 6009.615 Hz (1824.152 Hz) and the spectra were processed identically. The 800 μM (300 mM LMPG) protein sample was diluted to 400 μM (150 mM LMPG) using 20 mM KH_2PO_4 (pH 6.5) to keep the protein to detergent ratio the same. (B) 10 non-overlapping peaks present in both spectra were selected and the signal to noise ratio (SNR) was calculated based on peak height and the noise figure from Analysis 2.4. The peak heights for the 800 μM sample were divided by 2 to correct for the concentration difference. An average SNR increase of 2.7 (upon dilution to 400 μM) was calculated using $(1/N)\sum_k (S_{ik}/N_i)/(S_{jk}/N_j)$ where N is the number of peaks, S_i and S_j are the k^{th} peaks heights at 400 and 800 μM respectively and N_i and N_j are the noise floors for each experiment. Equation was taken from (Yang and Kay, 1999b).

dilution to 400 μM gave a 2.7 fold improvement in overall signal to noise (SNR) and decreased linewidths. Figure 4.4 shows an overlay of the two conditions and a comparison of 10 non-overlapped peaks that appear in both spectra (at the same chemical shifts). This reduction in signal to noise at higher concentrations could reflect the oligomeric behaviour of Yop1_ΔC2.

4.2 Reconstitution

Since Yop1_ΔC2 was purified using "non-native" techniques it was necessary to establish a protocol for reconstitution of the protein into detergent and lipid systems. Thin filming has been shown to be a robust approach to α-helical membrane protein reconstitution and refolding (Pielak et al., 2011, Schnell and Chou, 2008). This is also a convenient method for thorough screening of lipid and detergent conditions (see Section 2.8). I tested multiple solvents including FA, chloroform:methanol, TFA, trifluoroethanol (TFE) and HFIP with the most effective being HFIP. So far in the lab the most robust membrane protein solubilising agent has proven to be HFIP since it does not appreciably affect the downstream pH of samples and is highly volatile and amenable to thin filming. It is also a very good solubilising agent for a wide range of lipids and detergents.

One pitfall of this methodology comes during production of cysteine mutants. TCEP is extremely insoluble in most organic solvents so care must be taken to use DTT or βME instead.



(A) (B)

FIGURE 4.5: (A) HPLC trace of M138I construct following CNBr cleavage. Yop1_ΔC2 elutes when the concentration of buffer B exceeds 90% (see Section 2.7). (B) 100 μL of each 3 mL peak fraction was lyophilised and the powder was dissolved into 20 μL 8 M Urea and 15 μL of LDS Loading Dye. 10 μL was loaded per well and run at 200 V for 35 mins. The monomer band for Yop1_ΔC2 is labelled with a white arrow. It is noteworthy that bands corresponding to Yop1_ΔC2 oligomers can be observed. The volume (mL) at which each fraction starts is shown at the top of the wells. Molecular weight marker = SeeBlue Plus 2 (Invitrogen)

Figure 4.6 shows a general strategy for expression, purification and study of small α-helical membrane proteins by NMR. This strategy and variations thereof have been used successfully in the lab to produce functional membrane protein samples for NMR studies (Claridge et al., 2013, Ortega-Roldan et al., 2013).

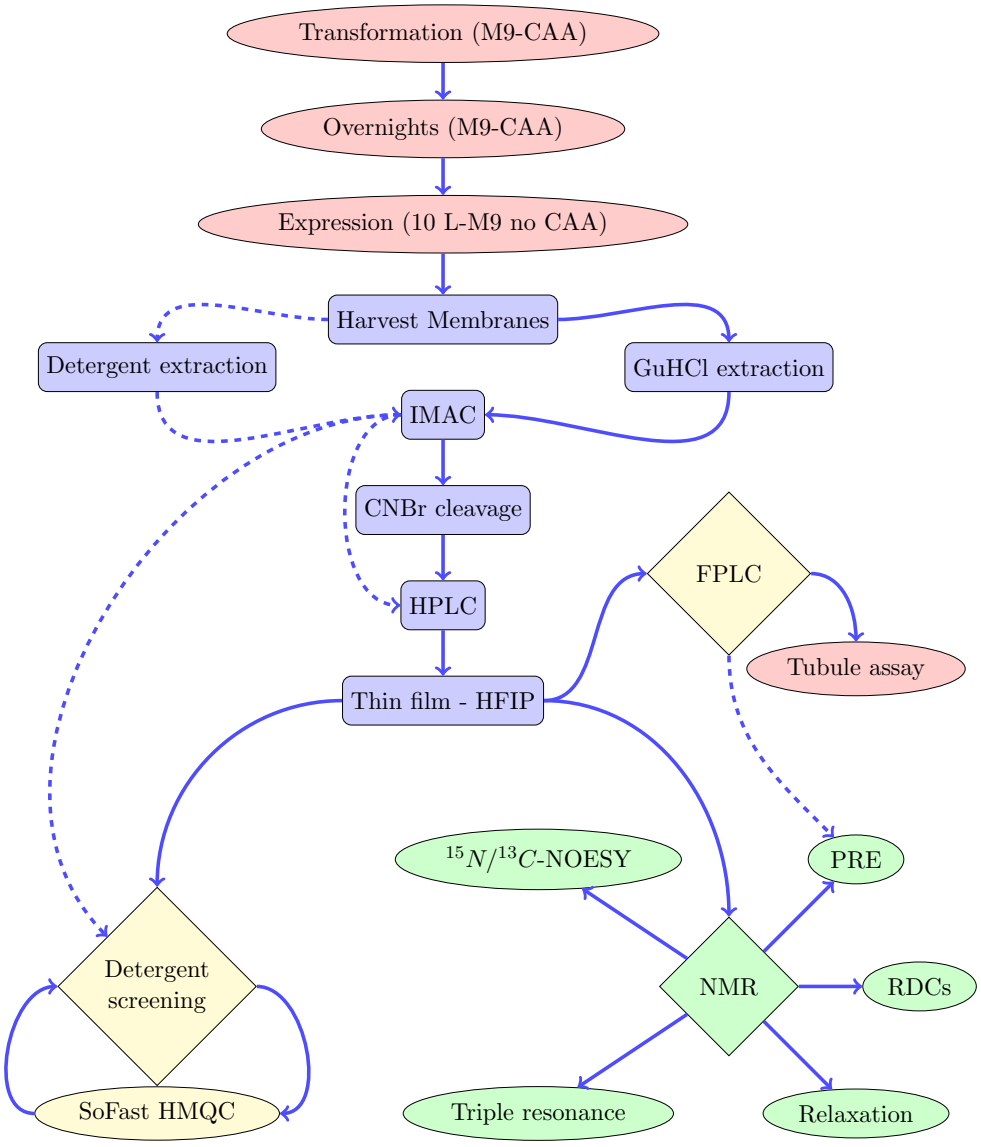


FIGURE 4.6: General strategy for NMR studies of Yop1. 6-8 mg of pure protein can be obtained per 10 L growth. This strategy and slight variations thereof are likely to be applicable to most small helical membrane proteins. Dotted lines represent alternative strategies that deviate from the general protocols (Chapter 2). The dotted line between FPLC and PRE represents the fact that this step has yet to be optimised (see Section 2.14.5 and Figure 5.11).

4.3 Yop1_ΔC2 forms polydisperse oligomers

Oligomerisation of Yop1p has been demonstrated previously by crosslinking (Shibata et al., 2008), sucrose gradient ultracentrifugation (Hu et al., 2008), and FRAP (Shibata et al., 2008), all of which suggest that Yop1p forms oligomers containing ∼4-7 copies. Glutaraldehyde crosslinking was used to assess the oligomeric state of E. coli expressed Yop1_ΔC2, and indicated that Yop1_ΔC2 was primarily monomeric in 2% LMPG. By contrast, crosslinking at a lower LMPG concentration of 0.01% resulted in a distribution of oligomeric species up to at least trimers being observed with no apparent aggregation (Figure 4.7).

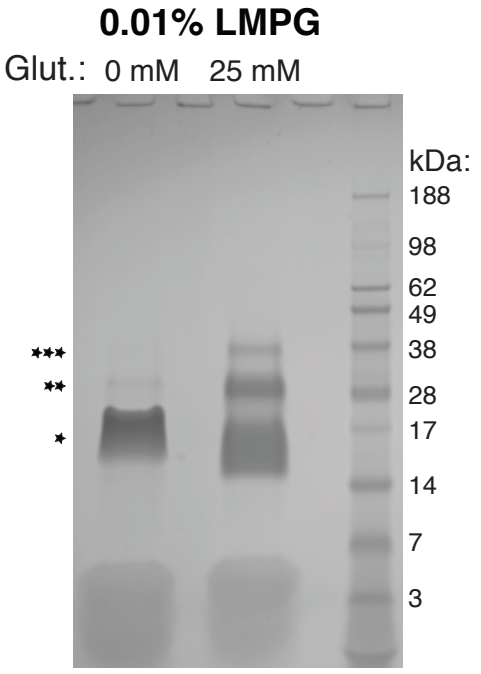
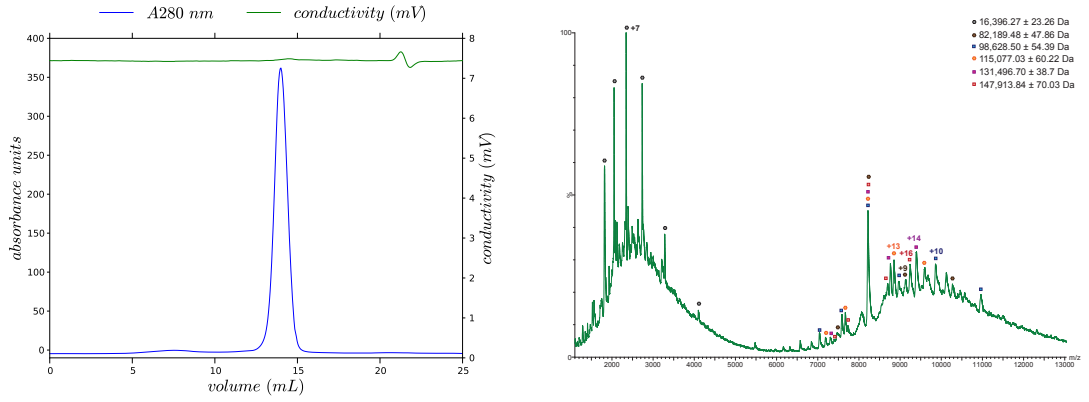


FIGURE 4.7: Chemical crosslinking of Yop1_ΔC2 in 0.01% LMPG. Samples were incubated at room temperature for 30 mins in the presence of the indicated glutaraldehyde concentrations and then dissolved with 8 M urea and SDS loading dye before running on NuPAGE in 1x MES buffer (see Section 2.11). Gels were stained with Coomassie Brilliant Blue.

To further characterise the oligomeric state of Yop1_ΔC2 I used native mass spectrometry (MS) (in collaboration with Eammon Reading and Dr. Arthur Laganowsky in Prof. Carol Robinson's lab). The protein was exchanged into 3 times the cmc of the non-ionic detergent C8E4 in ammonium acetate (pH 6.9) which gave a mono-disperse gel filtration trace similar to Figure 4.8A. The mass spectra collected on this sample show the presence of multiple oligomeric species ranging from 5mer to 9mers (Figure 4.8B). The data in Figure 4.8B were collected on the 10% ^{13}C U-^{15}N labelled sample from



(A) (B)

FIGURE 4.8: (A) Yop1_ΔC2 exchanged from 2% LMPG (20 mM KH_2PO_4 at pH 6.5) to 0.01% LMPG (50 mM Hepes, 150 mM KCl at pH 7) on a superdex 200 column with a flow rate of 0.5 mL min^{-1}. (B) The native mass spectrum of Yop1_ΔC2 in 3x cmc C8E4, sodium acetate pH 6.9. The masses of the peaks (top right) correspond to oligomers of Yop1_ΔC2 ranging from 5-mer to 9-mer.

Figure 5.18. In an attempt to probe the dynamics of any exchange processes between oligomeric states, I produced an unlabelled sample of Yop1_ΔC2. This was done so that upon mixing the labelled and unlabelled samples it might be possible to monitor exchange of subunits by time-resolved MS. This has been demonstrated previously for αB-crystallin (Baldwin et al., 2011); however, the mass difference was not large enough to allow these measurements with Yop1_ΔC2. It would be interesting to attempt this again using a truncated construct such as Yop1p(ΔAPH) (see Figure 6.4).

It could also be possible to tease out information on the shape and architecture of Yop1p oligomeric complexes by using ion-mobility mass spectrometry (Ruotolo et al., 2008). However, this is again complicated by the presence of multiple oligomeric species. The prospect of observing specific protein-lipid interactions by native MS is also highlighted in recent work by Laganowsky et al. (2014). This could potentially be applied to Yop1p in order to study any specific lipid interactions it might have.

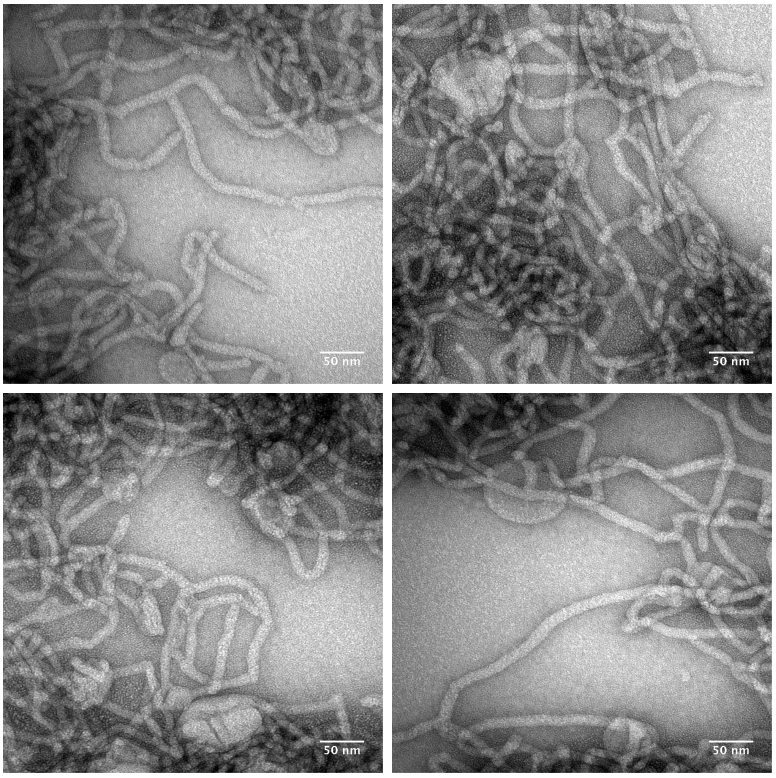


FIGURE 4.9: Electron micrographs of 2 mg ml^{-1} Yop1_ΔC2 reconstituted into *E. coli* polar lipids with a protein to lipid ratio of 1:1(w/w). Scale bar = 50 nm

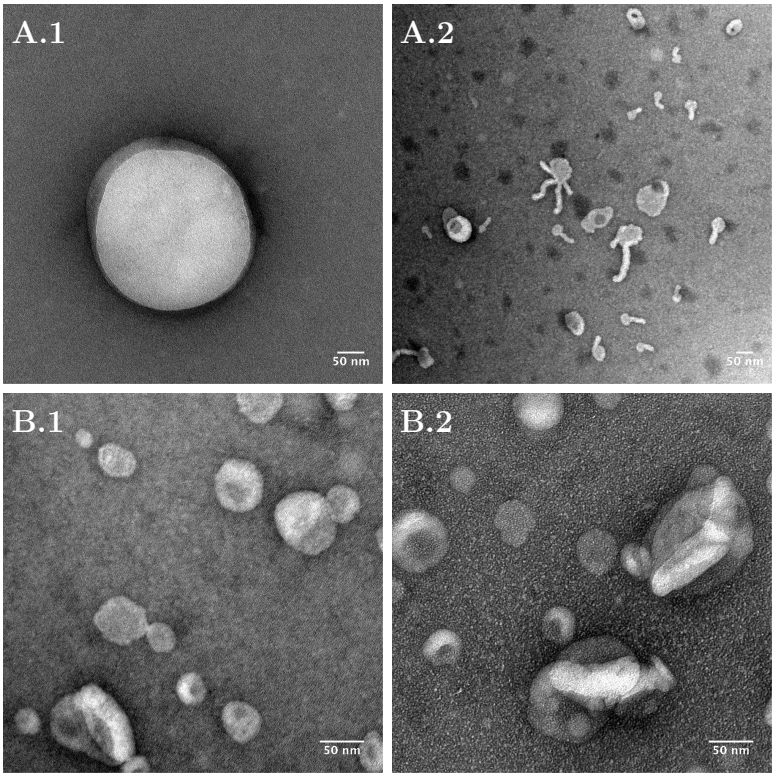


FIGURE 4.10: Control electron micrographs for *E. coli* polar lipids reconstituted without protein according to Section 2.8. A.1 and A.2, reconstitution from DM with 0.01% LMPG. B.1 and B.2, reconstitution from DM without LMPG. Scale bar = 50 nm.

4.4 Refolded Yop1_ΔC2 is functional, and forms membrane tubules *in vitro*

To confirm that the bacterially expressed and refolded Yop1_ΔC2 was competent to form tubules, the protein was reconstituted from 0.01% LMPG into *E. coli* polar lipids (EPL) solubilized in 1% DM using an adaptation of the method described by Hu et al. (2008)(see Section 2.8). The membrane tubules formed by Yop1_ΔC2 under these conditions had uniform diameters close to the ≈17 nm observed previously (Figure 4.9). Higher order oligomers up to at least 7-mers were observed after chemical crosslinking of Yop1_ΔC2 in membrane tubules (Figure 4.11). Although higher order oligomers were observed in tubules relative to that in low LMPG concentrations, the persistence of oligomerisation under both conditions suggests that the overall structure of Yop1_ΔC2 in micelles is related to that in membrane tubules.

The appearance of some short tubules in the control experiment starting with EPL dissolved in 1% DM and 0.01% LMPG (Figure 4.10, right) could be explained by the propensity of LMPG to induce spontaneous positive curvature (Zimmerberg and Kozlov, 2006). The considerably higher numbers of tubules in the presence of Yop1_ΔC2 coupled with the near absence of vesicles suggests that the result in Figure 4.9 is highly significant; therefore, I conclude that Yop1_ΔC2 purified from *E. coli* and reconstituted into membranes is functional.

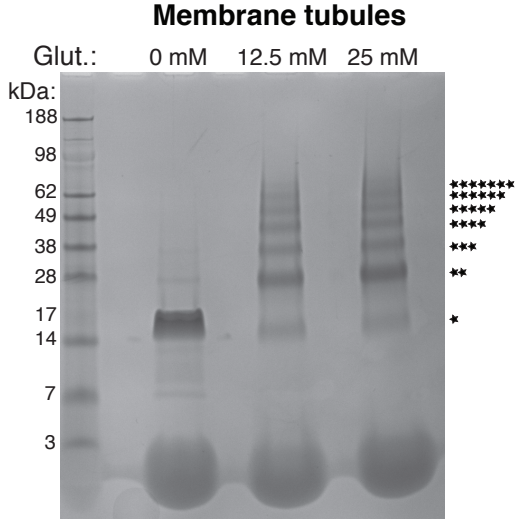


FIGURE 4.11: Chemical crosslinking of Yop1_ΔC2 after tubule formation by exchange into *E. coli* polar lipids followed by removal of LMPG. Samples were incubated at room temperature for 30 mins in the presence of the indicated glutaraldehyde concentrations and then dissolved with 8 M urea and SDS loading dye before running on NuPAGE in 1x MES buffer (see Section 2.11). Gels were stained with Coomassie Brilliant Blue.

5

Assignments, Secondary Structure and Dynamics of Yop1_ΔC2

5.1 Studying Yop1p by NMR

Since bacterially expressed Yop1_ΔC2 was shown to be functionally competent under NMR conditions, as evidenced by oligomeric behaviour and tubule formation (Figures 4.7 to 4.9 and 4.11), I decided to pursue structural studies of the protein.

Yop1p is predicted to have two peculiar TM regions, each ∼35 amino acids in length, which are proposed to be too short to span the membrane fully and hence form wedge-like hairpins in the bilayer (Voeltz et al., 2006). The first step towards validating or disproving these hypotheses on a structural level is to assign the protein's backbone resonances. The backbone chemical shifts are then used in combination with solvent accessibility and paramagnetic relaxation enhancements to determine the topology of Yop1p. The following chapter documents this work.

5.2 Backbone resonance assignment

The first and often most time consuming step towards characterising a protein's structure by NMR is the assignment of its backbone and side chain resonances*. This can be a particular challenge when dealing with proteins that lack large chemical shift dispersion such as membrane proteins and intrinsically disordered proteins (IDP). These problems can be overcome for IDPs by use of high dimensionality experiments (Kazimierczuk et al., 2013) but the inferior relaxation properties of membrane proteins will often preclude the use of such experiments.

Near complete backbone assignment (>95%) of Yop1_ΔC2 was achieved using the standard suite of triple resonance experiments on a combination of U-^{2}H,^{13}C,^{15}N and ^{13}C,^{15}N labelled samples collected at 500 and 600 MHz (Section 2.14). Initially, assignments were made on samples solubilised with 42 mM LMPG. The protein was then re-assigned in the LMPG:DPPC (40 mM: 20 mM, q=0.5) bicelle condition due to improved spectra quality (see Section 5.3.1). All assignments for Yop1_ΔC2 in bicelles were obtained utilising NUS with Poisson gap sampling and hmsIST for reconstruction (see Chapter 3).

Side chain assignments were achieved using ^{13}C and ^{15}N-edited NOESY experiments collected at 750 and 950 MHz.

*This step took me ∼1 year to complete!

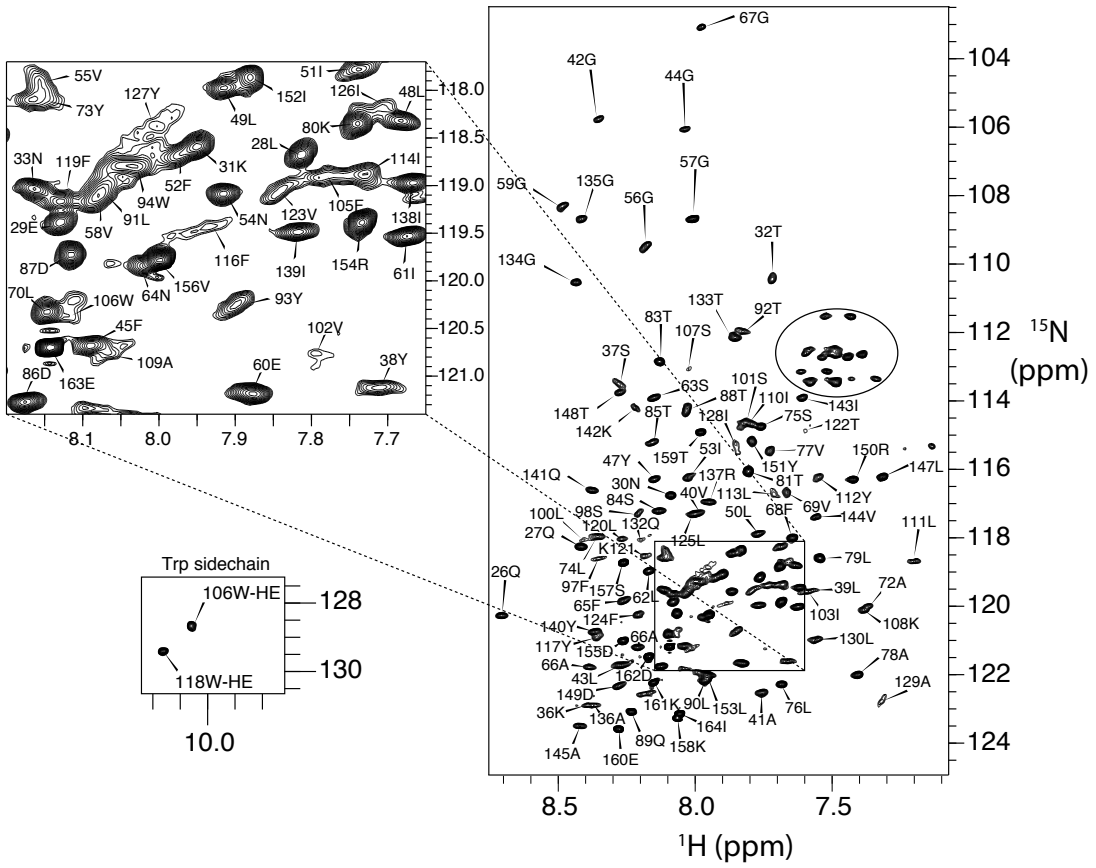


FIGURE 5.1: A ^{1}H-^{15}N TROSY-HSQC of 400 mM Yop1_ΔC2 in DPPC:LMPG mixed micelles (q = 0.5, 20mM DPPC: 40 mM LMPG) collected at 950 MHz (^{1}H) with residue assignments indicated. Partially refocused sidechain NH_2 crosspeaks from asparagine and glutamine amino acids are enclosed by an oval.

5.3 Chemical shift derived secondary structure

The backbone ϕ and ψ angles determined from chemical shifts (Shen and Bax, 2013a) indicated that Yop1_ΔC2 in mixed micelles contained at least five helical regions (Figure 5.2). A short N-terminal helix comprising residues 25-29 and possibly truncated in this construct was followed by three predominantly hydrophobic helices: 36-55 (TM1), 58-78 (TM2), and 89-127 (TM3/TM4). A fifth helix was found in residues 135-151 (APH; Figure 6.1). The N-terminal hydrophobic region consisted of two transmembrane helices, TM1 and TM2. TM1 and TM2 appeared to be canonical transmembrane helices comprising 20 and 21 amino acids, respectively, and separated by a "GGVG" containing sequence, which has been seen previously to form short loops connecting secondary structure elements (Hu et al., 2009, Khademi et al., 2004, Zeth et al., 2000). TM2 contains a conserved proline (P71), reducing the apparent helicity for residues F68 and G67, which are three and four positions N-terminal to the proline, respectively, and thus consistent with the presence of a proline-induced helical kink (Cordes et al., 2002). The N-terminal hydrophobic helical region (TM1 and TM2) was separated from the TM3/TM4 region by 10 residues (positions 79-88). The TM3/TM4 region consisted of a stretch of 39 helical and largely hydrophobic residues (89-127). Unlike the clear demarcation between the TM1 and TM2 helices, chemical shift-based predictions indicated the presence of two regions of decreased helicity in the TM3/TM4 region near W106 (sequence context EFWSK) and L113 (sequence context LYLIP) (Figures 1.7, 4.2 and 5.2). Similarly, a second algorithm for determining secondary structure from chemical shifts (Camilloni

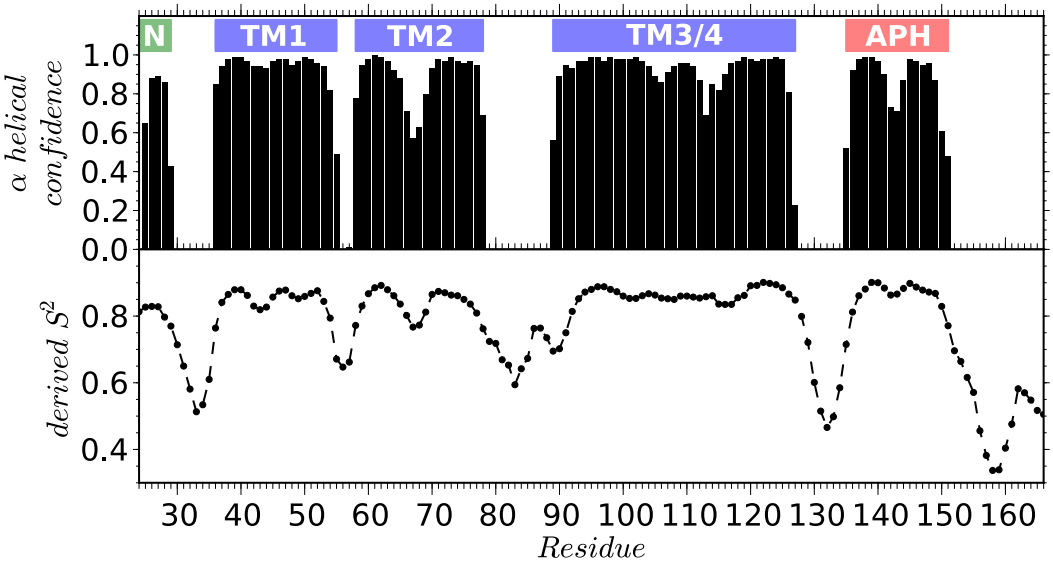


FIGURE 5.2: TALOS-N secondary structure prediction for Yop1_ΔC2 in LMPG:DPPC bicelles (q = 0.5)

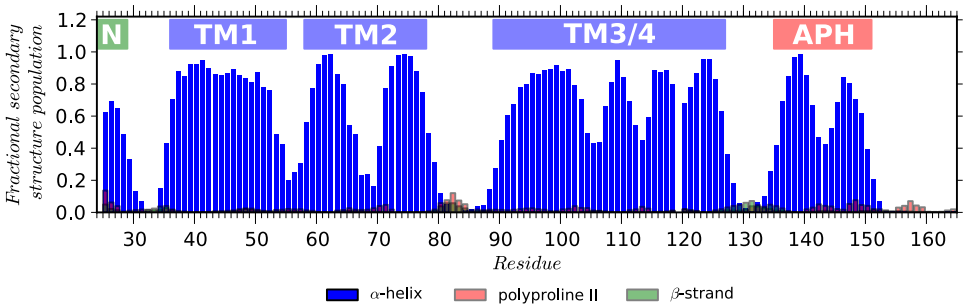


FIGURE 5.3: Chemical shift-based secondary structure populations calculated in $\Delta 2\delta$. Proportion of α-helix (blue), polyproline-II (red), and β-strand (green) as a function of residue number for Yop1_ΔC2 in DPPC/LMPG (q = 0.5) mixed micelles as determined from chemical shifts by the program $\Delta 2\delta$ (Camilloni et al., 2012).

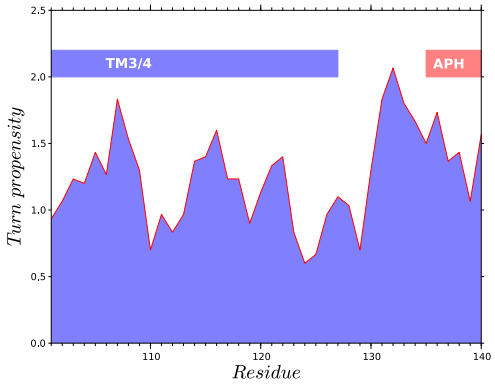


FIGURE 5.4: A three amino acid sliding window was used to calculate the average transmembrane helix break propensity (Monné et al., 1999) (see section 2.13.1). Residues 101 to 139 were analysed and the values are centred on the central residue of each triplet. Position 107 corresponding to WSK gave the highest break propensity for the TM3/4 region

et al., 2012), predicted decreased helical populations in residues 104-106 and 112-114 (Figure 5.3). Both regions also scored highly on an empirically-based transmembrane helix turn propensity scale (Monné et al., 1999) (Figure 5.4).

5.3.1 Micelle *vs* bicelle

Initially I reconstituted Yop1_ΔC2 into LMPG micelles but to further address the question of topology in a more "native" like environment I reconstituted Yop1_ΔC2 into bicelles containing 20 mM DPPC and 40 mM LMPG (q=0.5). DPPC was chosen since its acyl chain length of 16 is commonly found in yeast membranes (Ejsing et al., 2009). Figures 5.5 and 5.6 show the chemical shift perturbations between q=0.5 and q=0 (see Section 2.14.4 for method). Overall the chemical shift changes observed mostly fall below two corrected standard deviations from the mean (Schumann et al., 2007) which would be consistent with them being largely non-specific changes. It could then be argued that the structure of Yop1_ΔC2 is not greatly perturbed by the presence or absence of lipids. This is an important consideration for structure determination purposes since the protein behaves markedly smaller in micelles compared to bicelles Figure 5.16. The lack of significant perturbations in the APH would suggest that the helix is buried beneath the phosphate headgroups of the lipids since the most significant changes to chemical environment would be expected to arise from displacement of choline by glycerol headgroups of LMPG. However, another explanation could be that the APH has specific interactions with the negatively charged PG headgroups and may sequester the LMPG molecules. The APH region also displays similar tumbling times to the TM regions further suggesting that it is micelle associated (Figure 5.14 and Table 5.1).

Figure 5.7 shows build up of NOEs from the amide protons of Yop1_ΔC2 to the DPPC methyl and methylene groups demonstrating that the protein is in direct contact with the DPPC acyl chains. An increase in estimated τ_c (from 7.89 to 11.87 ns, Figure 5.15) is observed from the $^{15}N^{1}H$-TRACT experiment (Lee et al., 2006) which is also consistent with Yop1_ΔC2 being associated with bicelles. This corresponds to an increase in apparent molecular weight from ≈30-35 kDa in micelles to ≈50-55kDa in bicelles (Figure 5.16, see Section 5.5 for a detailed description of the TRACT data). The theoretical molecular weight of Yop1_ΔC2 is 16.2 kDa.

Comparison of the TALOS-N secondary structure prediction between Yop1_ΔC2 in micelles and bicelles reveals a slight overall increase in helical propensity. This is most striking in TM1, the kink at P71 in TM2 and in the APH (Figure 5.8). It is also noteworthy that the HSQC peak for G44, which in the LMPG condition is broadened, gives a much more homogeneous signal in bicelles (Figure 5.5, G44 is marked with a star). This result could reflect the improved membrane mimetic environment provided by the bicelles.

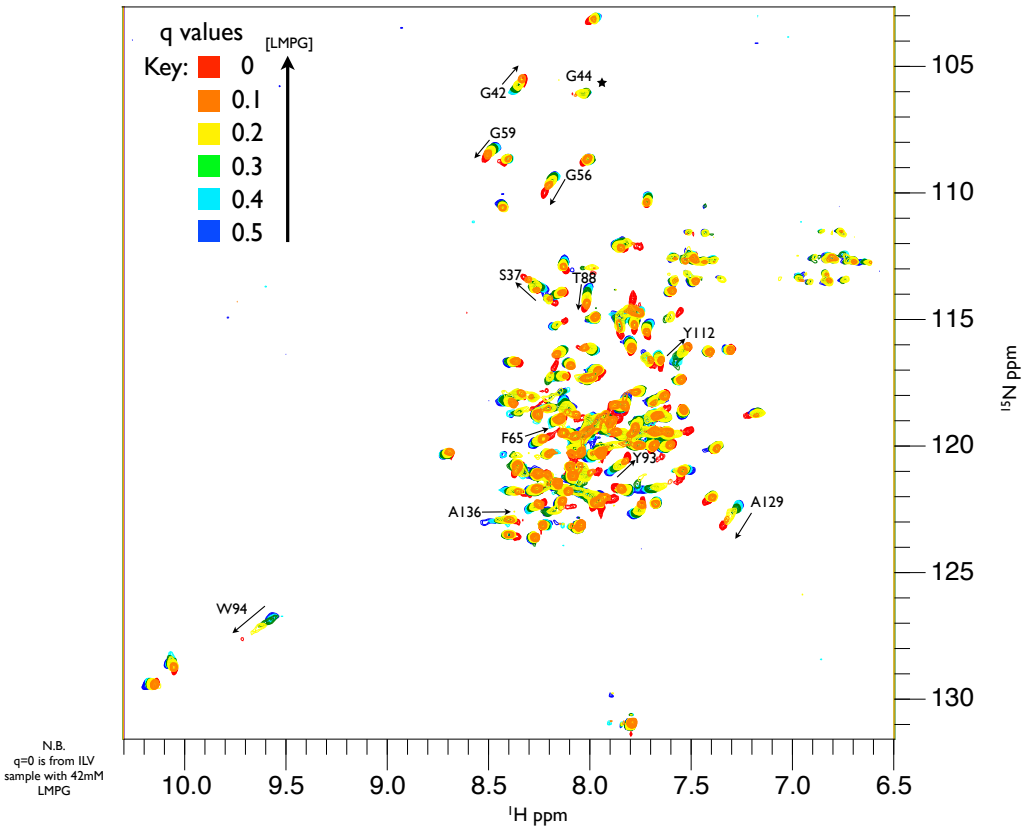


FIGURE 5.5: Spectral overlays showing chemical shift perturbations upon introduction of DPPC to form LMPG/DPPC mixed micelles. The spectra are coloured according to the ratio of DPPC to LMPG (q = [DPPC]/[LMPG]). Spectra were collected using TROSY-HSQC.

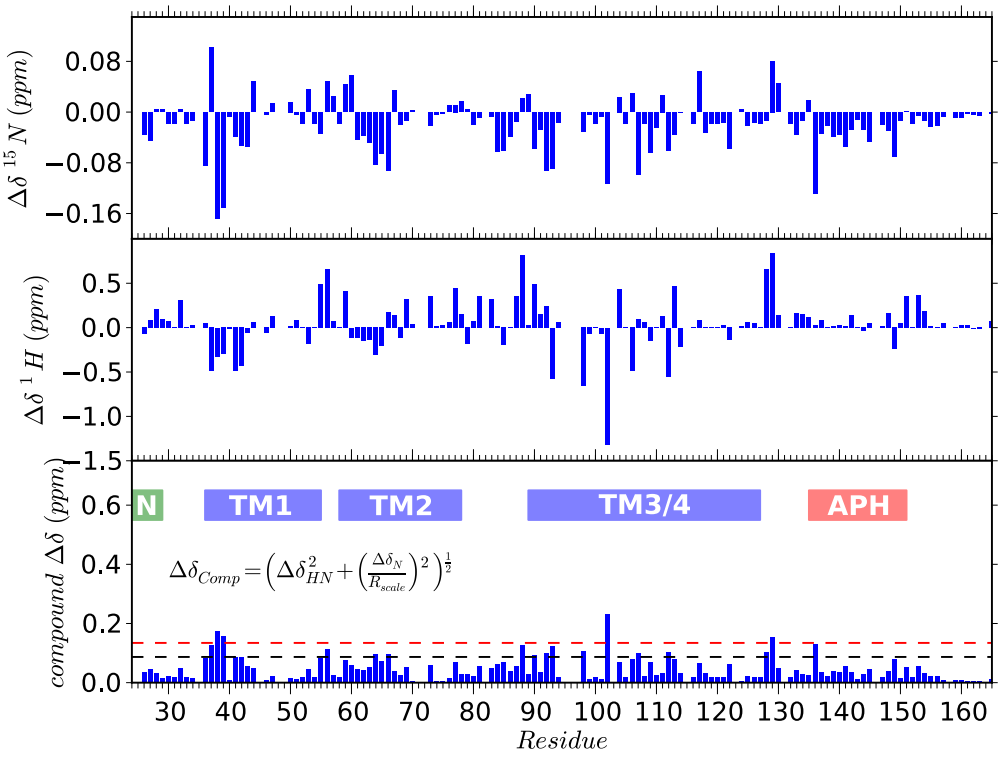


FIGURE 5.6: Amide chemical shift perturbations for Yop1_ΔC2 between a q = 0.5 sample and a q = 0 sample. The compound perturbations were calculated according to Mulder et al. (1999) where $R_{scale} = 6.5$ which, is the ratio of the variances for ^{15}N and $^{1}H_N$ chemical shifts for all amino acids across the BMRB. The red dashed line at 0.13 ppm represents 2 times the corrected standard deviation + mean while the black line at 0.09 ppm represents the standard deviation + mean. The corrected standard deviation was calculated by iteratively excluding data points above 3 standard deviations from the mean (Schumann et al., 2007). This demonstrates that overall chemical shift changes are largely non-specific and insignificant showing that the overall structure is probably not perturbed by the presence of lipids.

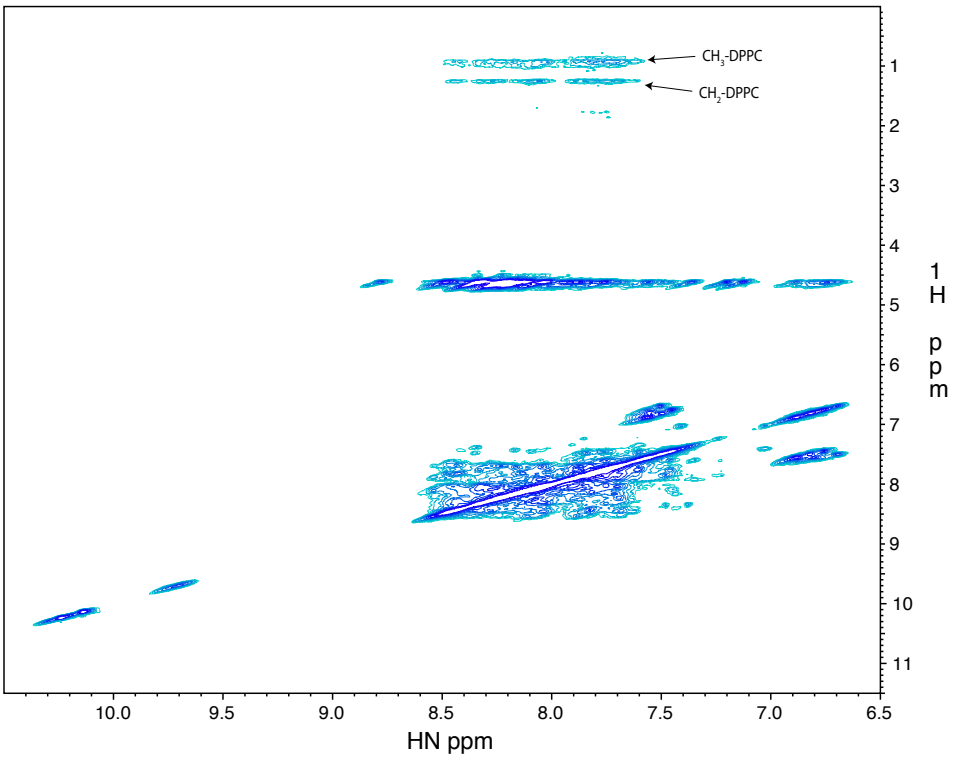


FIGURE 5.7: 2D ^{15}N- edited ^{1}H-^{1}H NOESY spectrum (300 ms mixing time) with strip of NOEs from the amide protons of Yop1_ΔC2 to the DPPC methyl and methylene groups indicated demonstrating direct contact of protein with lipid in mixed micelles. The sample contained Yop1_ΔC2 in DPPC/LMPG mixed micelles (q = 0.5) in which LMPG contained perdeuterated acyl chains. Signals from methyl and methylene protons of DPPC can be seen at 0.97 and 1.28 ppm, respectively.

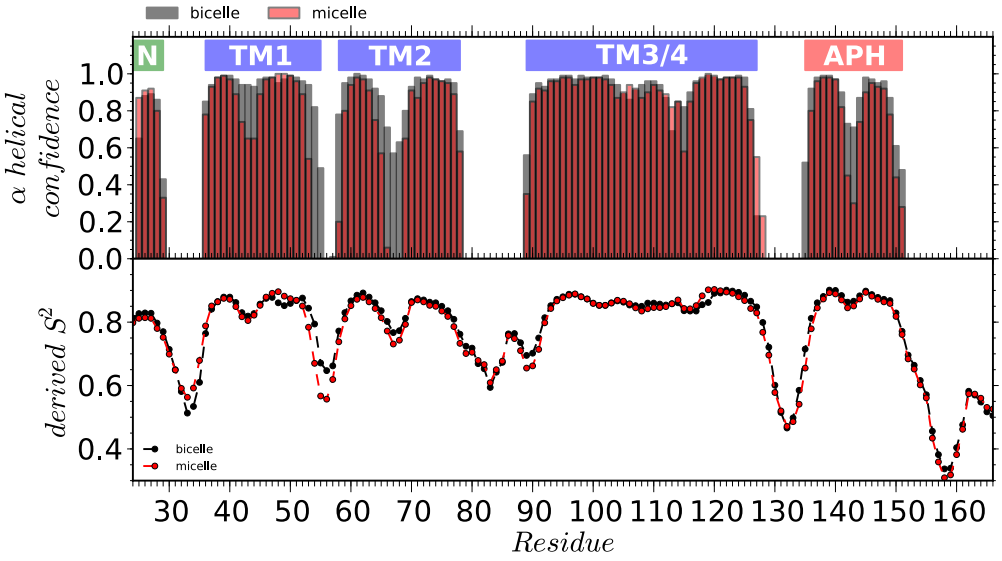


FIGURE 5.8: Comparison of the TALOS-N secondary structure prediction and the derived S^2 values for Yop1_ΔC2 in LMPG micelles and LMPG/DPPC (q = 0.5) mixed micelles.

5.4 Topology

In order to probe for regions of the protein that were solvent exposed, amide crosspeak intensities were measured in the absence and presence of the paramagnetic metal Mn^{2+} chelated to EDDA (Mn^{2+}EDDA). The Mn^{2+}EDDA complex is water soluble and excluded from the micellar hydrophobic core and polar headgroup region (Altenbach et al., 1994, Lau et al., 2008). As expected, the amide crosspeaks of the N- and C-terminal residues, the loop between TM1 and TM2 (comprising the GGVG motif), the flexible region connecting TM2 and TM3, and the region connecting the C-terminal end of TM4 and the APH were broadened indicating exposure to the water soluble Mn^{2+}EDDA (Figure 5.9). Although variation in protection was observed in the TM3/TM4 region, residues for which data could be obtained indicated that all were at least partially protected by either the mixed micelle or other regions of the protein. The hydropathy plot in Figure 5.9 also correlates with the Mn^{2+}EDDA data. The overall drop in solvent protection for TM2 relative to TM1 and TM3/4 is mirrored by more hydrophilic sequence characteristics in that region.

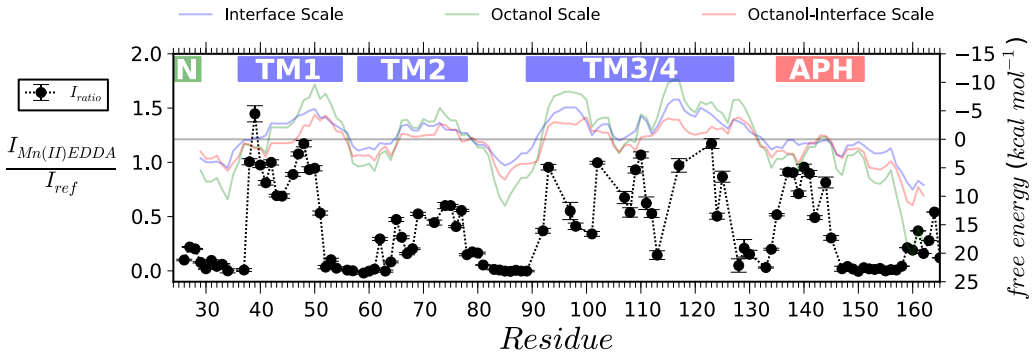


FIGURE 5.9: The ratio of crosspeak intensities after ($I_{Mn(II)EDDA}$) and before (I_{Ref}) addition of 500 μM Mn^{2+}EDDA to a 400 μM Yop1_ΔC2 LMPG:DPPC (40 mM: 20 mM, respectively, q=0.5) sample. The ratio is taken to be proportional to the Mn^{2+}-induced line-broadening, and therefore accessibility. In the background is the hydropathy plot centred on a 9 amino acid window over which the hydropathies are summed (Wimley and White, 1996). Residues for which data could not be reliably obtained due to crosspeak overlap in mixed micelles (particularly problematic for residues in the TM3/TM4 region) were excluded.

To further probe the topology of Yop1p I used the CLEANEX-PM experiment (Hwang et al., 1998) to measure amide proton exchange rates with water (Figure 5.10). The CLEANEX-PM data suggest that the loop between TM1 and TM2 (comprising the VGGVG motif) is not exchanging with water, while the loop between TM2 and 3/4 and the C-terminal residues are. Overall the CLEANEX-PM data correlate well with the predicted intrinsic exchange rates (K_{ex}) derived from SPHERE (Zhang, 1995). However,

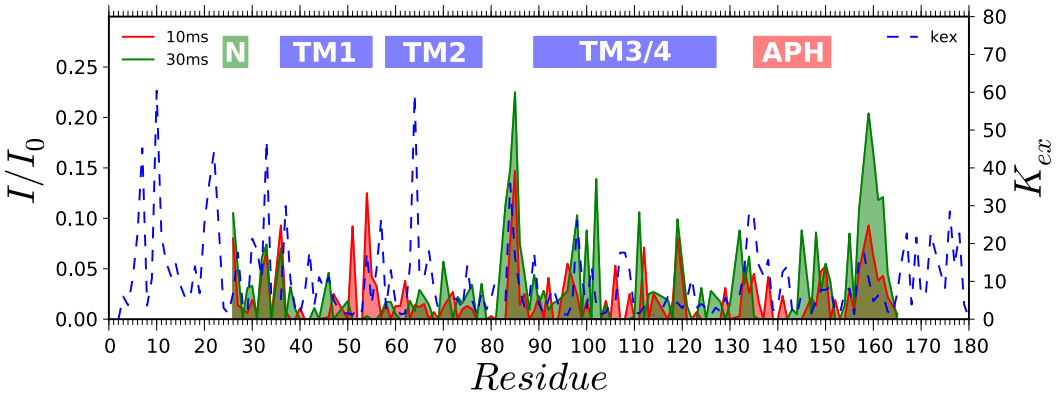


FIGURE 5.10: CLEANEX-PM data for Yop1_ΔC2 in q=0.3 bicelles (LMPG:DPPC). I/I_0 was calculated from a reference fHSQC (Mori et al., 1995) and the fHSQC incorporating the CLEANEX-PM spin-locking sequence during mixing periods of both 10 and 30 ms. The dashed blue line represents theoretically derived exchange rates with water (K_{ex}) from SPHERE (Zhang, 1995).

it could be argued that the data are not highly significant owing to the fact that more delay times would be required to extract K_{ex} parameters or observe trends (this is not possible when n=2!). The fact that the experiments were conducted at pH 6.5 would also lead to a reduction in any expected water exchange due slow amide proton exchange rates.

The lack of significant data to define a break point between TM3 and 4 could lead one to suggest that this region comprises one long transmembrane segment that would likely be highly tilted with respect to the membrane. However, previous work by Voeltz et al. (2006) has suggested, although not conclusively proved, that the C-terminal region is cytosolic and hence TM3 and 4 do form two transmembrane segments. My preliminary attempts to observe paramagnetic relaxation enhancements due to the presence of covalently bound MTSL free radicals (Figure 5.11, see Section 2.14.5 for methods) also strongly suggest that TM3 and 4 form full transmembrane spanning helices and that Yop1p comprises 3 cytosolic parts, the N-terminus, the 9 amino acid loop between TM2 and TM3 and the C-terminus (including the APH). The reasoning is as follows: Yop1_ΔC2 labelled with MTSL at residue 37 leads to relaxation enhancements at the N- and C-termini of the APH region (Figure 5.11,top). This is consistent with the APH being on the same side of the membrane as the N-terminus. The relaxation enhancements in the cytosolic loop, residues 79, 80 and 93 in particular also confirms the proposed topology. The fact that relaxation enhancements occur at both N- and C-termini of the APH could be indicative of dimeric behaviour. This would be consistent with presence of SDS resistant dimers seen in SDS-PAGE experiments (Figure 4.11). Although, based

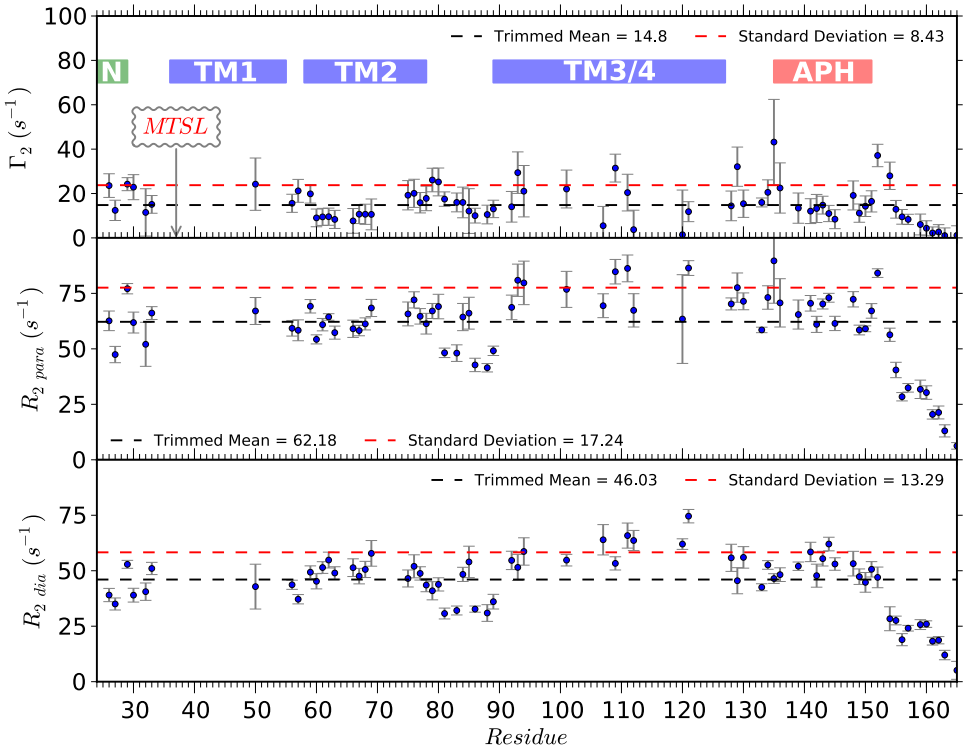


FIGURE 5.11: Paramagnetic relaxation enhancements for Yop1_ΔC2 (S37C) labelled with MTSL. The relaxation enhancement, Γ_2, is the difference between ^1H relaxation rates in the oxidised and reduced forms ($R_{2para} - R_{2dia}$). These values can be converted into approximate distances from the spin label according to Eq. (2.11). Trimmed means were calculated by excluding data outside the 10^{th} and 90^{th} percentiles.

on the crosslinking (Figure 4.7) and ^{15}N^{1}H-TRACT results (Figure 5.15) the dimer is mostly likely a minor population. There is also a slightly raised baseline for the $R_{2,para}$ data compared to the $R_{2,dia}$ which could reflect residual micelle associated MTSL that was not washed out during gel filtration. This could be due to non-specific binding of MTSL to either the protein or micelle and would be consistent with Kroncke and Columbus (2012) who demonstrate that the only effective way to remove MTSL from micelles is by binding the labelled protein to an IMAC column and washing extensively. This is currently not practical for Yop1_ΔC2 samples since the His-tag is removed. The non-specific binding of MTSL to detergent was controlled for in Klammt et al. (2012) where protein containing no cysteines was incubated with MTSL and then treated identically to the labelled proteins. The resulting peak intensities were used to account for any non-specific binding.

The topology presented here is consistent with those proposed previously in that the loop between the two hairpins would be exposed to the cytoplasm. However, the TMs defined by NMR are all around 20 aa in length and would likely be able to span the bilayer at least to the opposing leaflet's headgroup region. If one also takes into account the average width of the ER bilayer (37.5 Å from Mitra et al. (2004)) then it is possible to

speculate that the TMs of Yop1p are fairly conventional with their hydrophobic segments spanning the bilayer into the headgroup regions of each leaflet.

The δ2D secondary structure prediction (Figure 5.3) differs markedly from the TALOS-N (Figure 5.2) result for TMs 3 and 4 suggesting two likely break points at W105 and L113. The latter is likely due to P115; however, the reduction in helicity at W105 is preceded by slightly more PPII character which could suggest the start of a β turn (Camilloni et al., 2012). This would make TM3 very short at ≈15 residues, yet this is not unheard of since 3 out of 32 *S. cerevisiae* ER membrane proteins studied in Sharpe et al. (2010) had TM lengths of 15 aa. Increased break propensity was observed for the "WSKA" region compared with that of P115 using the transmembrane helix break propensity scale (Monné et al., 1999) (Figure 5.4) further implicating this as the break point between TM 3 and 4.

Yop1p contains two unusually long hydrophobic regions that have been proposed to form helical hairpins that differentially crowd one of the lipid monolayers and thereby stabilise membrane curvature (Voeltz et al., 2006). I have used a combination of backbone chemical shifts, solvent accessibility and dynamics (Section 5.5) to determine the possible topologies of Yop1p. Secondary structure determination from chemical shifts demonstrates that the N-terminal hydrophobic region contains two helices broken at a GGVG sequence. Thus TM1 and TM2 are 20 and 21 residues in length, respectively, and therefore long enough to fully traverse the hydrophobic portion of the ER membrane. The position of the Yop1p GGVG aligns well with the start of REEPs 1-4 (Figure 1.7) suggesting that REEPs 1-4 lack TM1. A helical kink was found at positions G67 and F68 in TM2 that is likely induced by P71. Interestingly, mutations related to disease have been mapped to the homologous region of REEP1, including P19R, P19L, A20E, and S23F (Beetz et al., 2008, Falk et al., 2014, Goizet et al., 2011) (Figure 6.15), all of which exhibit localization defects (Beetz et al., 2012, Falk et al., 2014). These REEP1 positions are conserved in Yop1p and lie within the C-terminal half of TM2 (Figure 1.7). TM2 is linked to the C-terminal hydrophobic region by a flexible linker. In the C- terminal transmembrane domain, secondary structure analyses identified two potential helical breaks, near W106 and L113. A single break at L113 would result in a conventional TM3 helix containing >20 residues but an unusually short TM4 of 13 residues, depending on the number of residues in the break. A single break near W106 would result in more conventional lengths for both TM3 and TM4 of 17-18 and 18-19 residues, respectively. Recent studies have suggested that transmembrane helices containing as few as 15 amino acids occur at only a 2-fold lower frequency than the most common length of 20 amino acids (Sharpe et al., 2010), although that result is based on single-pass membrane proteins. Additional constraints such as helix tilt and kinking imposed by other regions of the protein are likely to modulate the effective hydrophobic length

in polytopic proteins such as Yop1p. The possibility of two breaks in the C-terminal RHD, resulting in both TM3 and TM4 being relatively short (approximately 16 and 13 residues, respectively) and joined by a linker helix consisting of residues 110-114, cannot be excluded. As mentioned above, this linker helix would be highly hydrophobic (sequence ILYLI), and the lengths of the TM3 and TM4 helices could place this helix within the hydrophobic region of the luminal lipid monolayer. Membrane insertions within the hydrophobic region of the luminal monolayer could, theoretically, induce the necessary positive membrane curvature for the laterally uncoupled monolayers that are expected *in vivo* (Campelo et al., 2008). However, the hydrophobicity of the amino acids in positions 110-114 (sequence ILYLI) combined with the observation that helical kinks rather than breaks are often introduced in the helical turn preceding prolines (Cordes et al., 2002) suggests that these residues are more likely an extension of TM4. Furthermore, sequence conservation and HSP-associated genetic mutants more strongly implicate the residues around S107 and K108 for a structural or functional role. The position K108 is highly conserved as a charged residue with preference for glutamate and aspartate, respectively, and mutations at the positions homologous to S107 and K108 in REEP1 are known to cause HSP (Goizet et al., 2011, Schlang et al., 2008), although the molecular basis remains unknown (Falk et al., 2014). Thus, I postulate that the TM3/TM4 helical turn occurs near W106/S107/K108, and that Yop1p therefore contains four hydrophobic helices that are long enough to span fully the ER bilayer Figure 6.16. In contrast to the TM1/TM2 luminal loop, the TM3/TM4 luminal loop is unlikely to extend beyond the lipid headgroup region based on the observed solvent protection. In this model, TM4 is relatively short, and the kink observed at the conserved P115 would further decrease the membrane depth it could favourably accommodate. In addition, the fact that luminal loops connecting TM1 with TM2 and TM3 with TM4 are short and rigid may impart a "wedge-like" configuration to the protein fold, in a manner similar to that proposed previously (Shibata et al., 2009). However, the energetic costs of burying an exposed peptide bond in a helical turn (Roseman, 1988) may instead favour local rearrangements of the lipid bilayer. Indeed, effectively short transmembrane domains may play a role in RHD localization to regions of high membrane curvature since highly curved membranes are expected to be 2-4 Å thinner (Zemel et al., 2008) and this effect may partially explain why lengthening the transmembrane regions of RHDs results in defective localisation (Tolley et al., 2010, Zurek et al., 2011). Delineation of the Yop1p transmembrane domains also facilitates a comparison with the reticulon RHDs. Taking into account knowledge of the Nogo-66 domain (GrandPré et al., 2000, Vasudevan et al., 2010) and the Yop1p secondary structure allows alignment of Yop1p with the human reticulons and indicates similar transmembrane domain lengths (Figure 6.17).

5.5 Dynamics

To evaluate the internal dynamics of Yop1p, the backbone amide ^{15}N R_1 and R_2 relaxation rates were measured at 600 and 950 MHz (Figure 5.12). The ^{15}N heteronuclear NOE was also measured at 600 MHz (Farrow et al., 1994) (Figure 5.13). In Figure 5.14, the product of the ^{15}N relaxation rates (R_1*R_2) is plotted against their ratio (R_2/R_1) (Kneller et al, 2002) to observe potential clusters in the dynamics of the identified structural regions. Increases and decreases in R_1*R_2 and R_2/R_1 are associated with slow (microsecond to millisecond) and fast (picosecond to nanosecond) time scale motions on the NMR time scale, respectively, although the R_2/R_1 ratios can depend also on any rotational anisotropy. The data are plotted for 600 (Figure 5.14, top) and 950 MHz (Figure 5.14, bottom). The values at 600 MHz form a much tighter cluster than the 950 MHz data indicative of chemical exchange for residues at the termini and in the loop regions. This data also highlights a possible pitfall of high field NMR since certain peaks such as those deviating from the central clusters may be broadened by exchange processes that are in the intermediate NMR time scale for the given field. Interestingly, Kalverda et al. (2014) recently reported lower SNR for methyl-TROSY spectra collected at 900 MHz *vs* 750 MHz. This is at odds with what one might expect given the Boltzmann distribution, $\frac{N_{upper}}{N_{lower}} = e^{\frac{-hv}{kT}}$ (where N is the number of spins in the *upper* or *lower* energy levels, h is Planck's constant, v is the $lamor\ frequency$, k is Boltzmann's constant and T is the temperature in Kelvin (K)), since one would expect a larger population of NMR observable spins at higher fields.

Data for residues in the relatively rigid transmembrane regions define a central region of the plots (red, cyan, and blue circles). By contrast, the most flexible residues, those in the C-terminus (yellow circles), cluster in the lower left hand corner of the 600 MHz plot while shifting towards the upper left for the 950 MHz data. Compared with the C-terminus, the next most flexible region is that of the 10 residue loop connecting TM2 to TM3. The luminal loop connecting TM1 with TM2 is more structured (Gly56 and Gly57), and none of the residues identified as likely TM3/TM4 luminal loop residues deviate significantly from the values found for other residues in the TM3/TM4 region. Thus, the luminal loops are relatively rigid compared with the cytosolic loop. In addition, the region connecting the end of TM4 and the start of the APH was not highly flexible apart from Thr133 and Gly134, which are close to the APH.

The amphipathic helix displays relatively favourable spectral characteristics with increased signal to noise and improved lineshapes. The τ_c for this region (≈12 ns) is comparable to that of the TMs suggesting that it is associated with the micelle/bicelle but probably not packed with the TMs. Helical NOEs but no NOEs to water

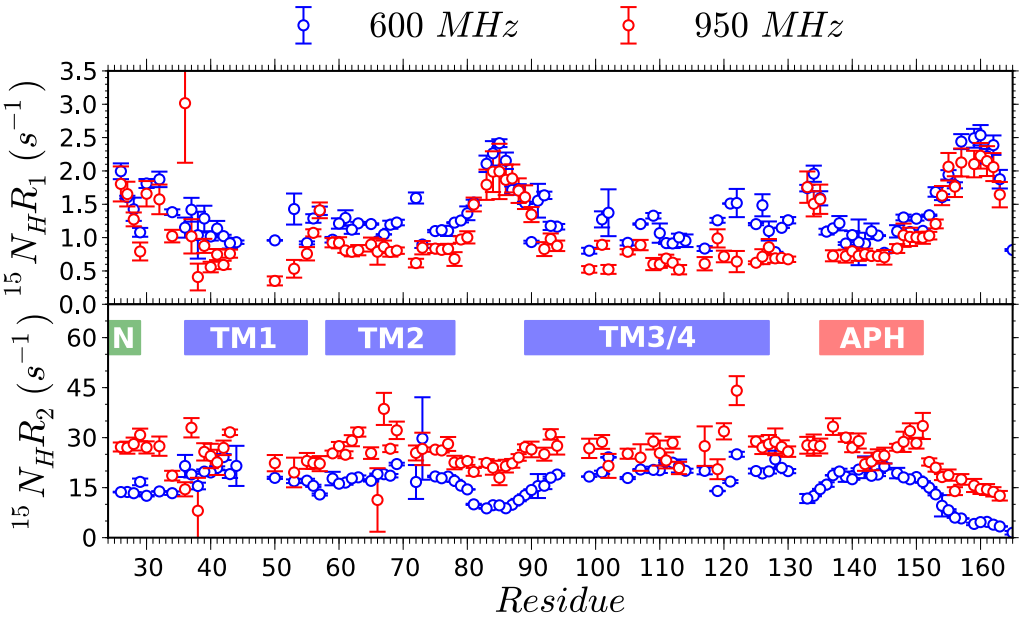


FIGURE 5.12: ^{15}N R_1 and R_2 values at 600 MHz (blue) and 950 MHz (red). The delay times used for each experiment are listed in Section 2.14.3 and the data were fit to decaying exponentials with Analysis 2.4. Errors were calculated using the covariance method and significantly overlapped peaks were omitted from the analysis.

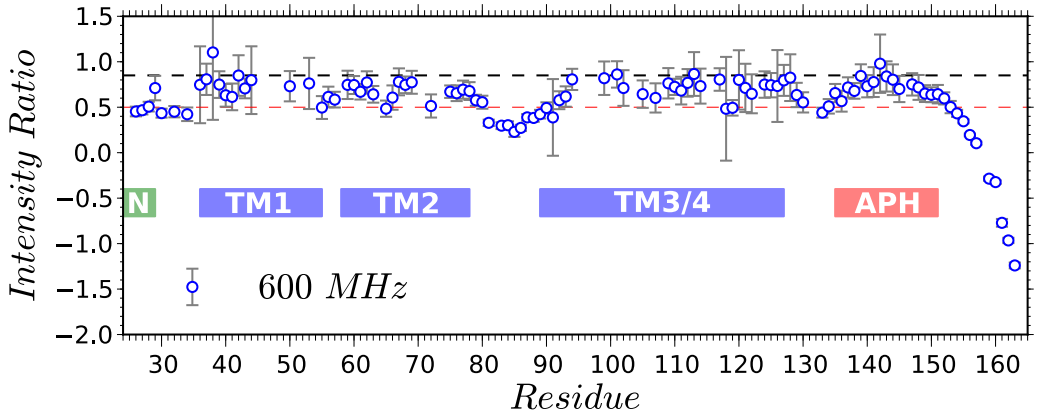


FIGURE 5.13: Heteronuclear NOE for Yop1_ΔC2 at 42 °C at 600 MHz. The errors were calculated according to Section 2.14.3 and Eq. (2.6). The theoretical maximum of 0.8 is marked by a black dotted line. None of the residues display values that are significantly above this maximum relative to their associated errors.

are observed and the CLEANEX data for this region (Figure 5.10) shows no significant exchange with water. Together, this characterises the APH as a stable helical region which is likely to be buried within the micelle/bicelle headgroup region.

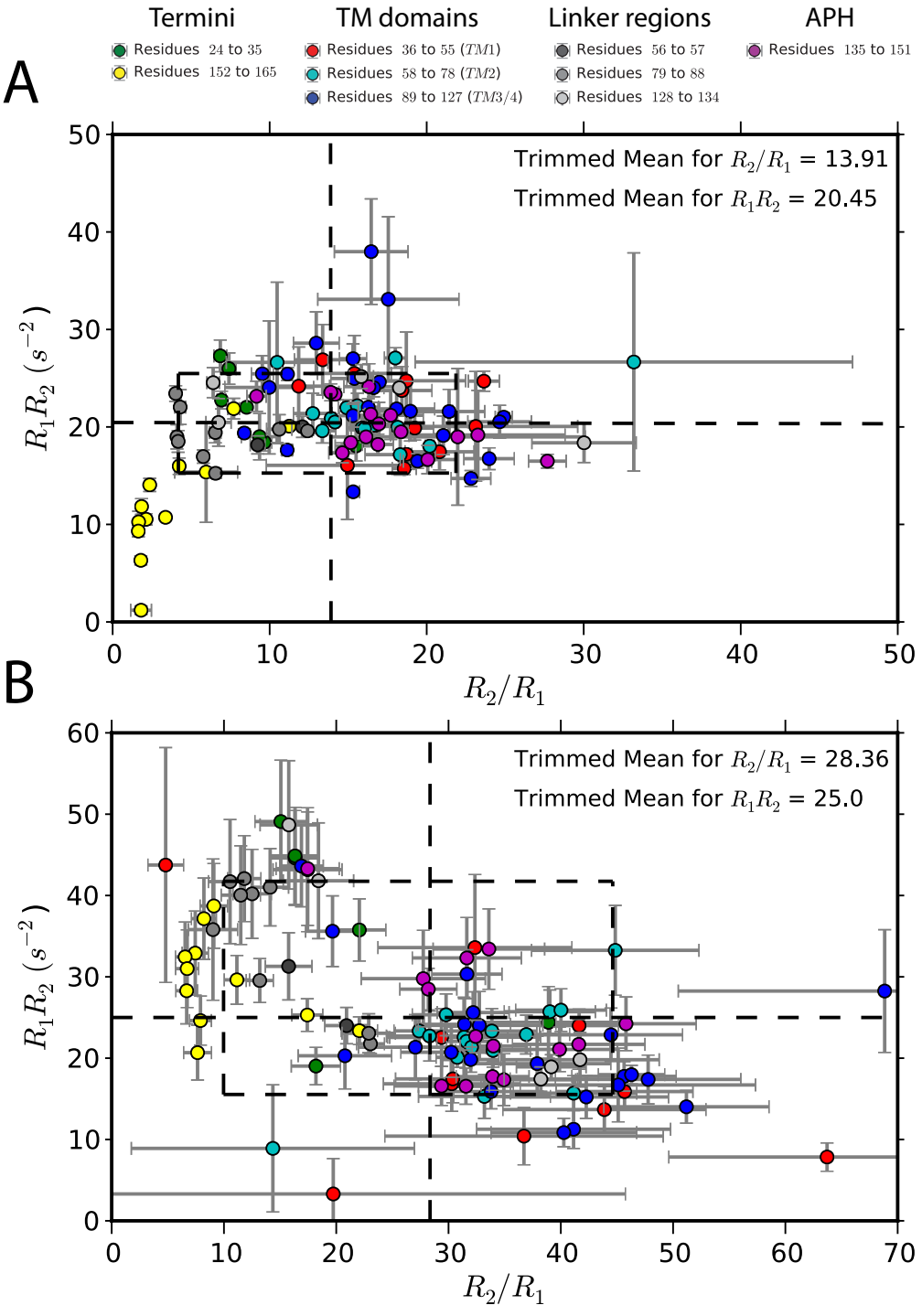
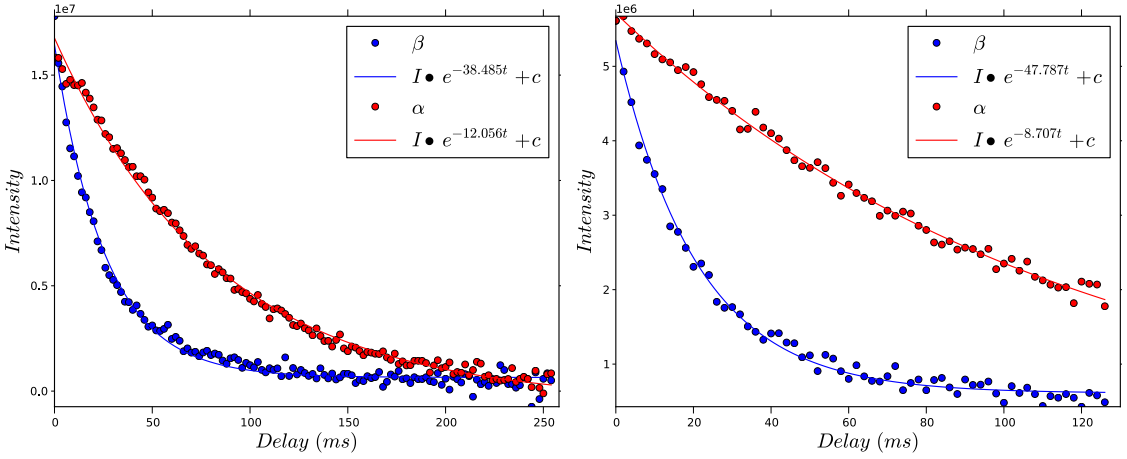


FIGURE 5.14: The product and ratio of ^{15}N R_1 and R_2 values at 600 (A) and 950 Mhz (B). According to Kneller et al. (Kneller et al., 2002), the product negates the effects of motional anisotropy with high values indicative of slow ms-μs timescale exchange and low values are indicate ns-ps timescale motions. The ratio (R_1/R_2) contains information about the tumbling of the molecule (τ_c). The 10% trimmed mean excludes data points from the 10^{th} and 90^{th} percentiles. The 950 MHz data displays far more variations in motional anisotropy and chemical exchange. Evidently membrane proteins display heterogeneous dynamic properties.

TABLE 5.1: τ_c values from R_1/R_2 relaxation data for Yop1_ΔC2 in 42 mM LMPG at 42 °C collected at 600 and 950 MHz. Values were calculated according to Section 2.14.3.

Field	TM1	TM2	TM3/TM4	APH
τ_c at 600 MHz	13.2 ± 1.4 ns	12.3 ± 1.8 ns	12.8 ± 2 ns	12.5 ± 1.6 ns
τ_c at 950 MHz	11.6 ± 3 ns	11.4 ± 1.3 ns	12.2 ± 2 ns	11.2 ± 1.2 ns



(A) Micelle (42 mM LMPG) (B) Bicelle (q = 0.5 LMPG:DPPC)

Sample	Temp °C	Field (MHz)	$R\alpha$ (s^{-1})	$R\beta$ (s^{-1})	η_{xy}	τ_c (ns)
Yop1_ΔC2 (micelle)	42	950	12.05	38.49	13.22	7.89
Yop1_ΔC2 (bicelle)	42	950	8.71	47.79	19.54	11.87

FIGURE 5.15: $^{15}N^{1}H$-TRACT experiments were collected at 950 MHz at 42 °C in 20 mM KH_2PO_4; pH 6.5. The α and β state selecting experiments were acquired in an interleaved fashion. For (A) 128 points for each state up to 256 ms were collected with 2 ms intervals. For (B) 64 points per state were collected at 2 ms intervals up to 128 ms.

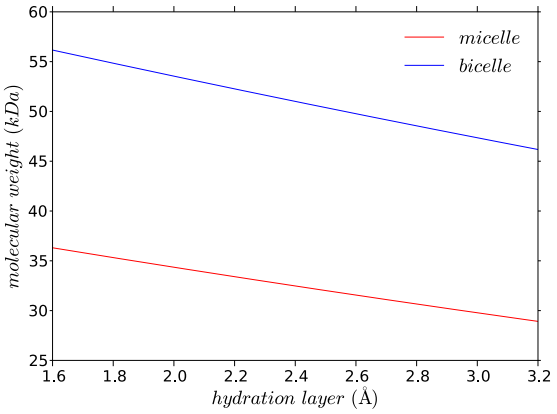


FIGURE 5.16: Apparent molecular weights back calculated from τ_c values from Figure 5.15. Molecular weight was calculated for a range of hydration layer values between 1.6 and 3.2 Å using Eqs. (5.1) and (5.3). The molecular weight could represent a slight overestimate due to the increased viscosity due to detergent and lipid molecules. The viscosity of water was calculated according to Eq. (5.2) where η_0 = 0.001792 $N\ s\ m^2$, a = -1.94, b = -4.80, c = 6.74, T_0 = 273.16 K and T is the temperature at which the experiment was carried out, in this case 315 K.

$$r_H = \sqrt[3]{\frac{3\tau_c kT}{4\pi\eta}} \tag{5.1}$$

$$\eta = \eta_0 \cdot e^{\left(a+b\left(\frac{T_0}{T}\right)+c\left(\frac{T_0}{T}\right)^2\right)} \tag{5.2}$$

$$M_w = \frac{(r_H - r_w)^3 \cdot 4\pi N_a}{3\rho} \tag{5.3}$$

TABLE 5.2: Parameters used for molecular weight calculations. The value for η is for water at 315 K and represents a source of error since the viscosity will be affected by the presence of detergent or lipids. r_H was calculated from Eq. (5.1).

Constant	Symbol	Value	Units
Boltzmann	k	$1.3806488e^{-23}$	$m^2\ kg\ s^{-2}\ K^{-1}$
Avagadro	N_a	$6.0221413e^{23}$	mol^{-1}
Viscosity	η	$637e^{-6}$	$N\ s\ m^2$
Specific volume	ρ	$0.73e^{-6}$	$m^3\ g^{-1}$
Hydration radius	r_w	1.6-3.2e^{-10}	m
Hydrodynamic radius	r_H	$23.5_{micelle}/26.9_{bicelle}e^{-10}$	m

According to the τ_c values calculated from the ^{15}N^{1}H-TRACT experiments in Figures 5.15 and 5.16 the apparent molecular weight of Yop1_ΔC2 is ≈30-35 kDa in micelles and ≈50-55 kDa in bicelles (q=0.5). The aggregation number of micellar LMPG has been empirically determined by SAXS to be 63-69 (Oliver et al., 2013) making LMPG micelles between 30 and 33 kDa in size. Klammt et al. (2012) also measured LMPG micelle masses from 36 to 50 kDa using SEC-UV/LS/RI which combines UV absorbance at 280 nm with refractive index and static scattering at 3 angles (45°, 90° and 135°). At first glance my value for the molecular weight of Yop1_ΔC2 in LMPG micelles appears inconsistent with the published data since the micelle would be 14-20 kDa in size. However, this lower than expected tumbling time may reflect tumbling of the protein within the detergent micelle. Additionally, it is not possible to extrapolate from empty micelles to protein-detergent micelles since the number of protein associated detergent molecules will vary from protein to protein. The calculation of τ_c also assumes a spherical system which is most likely not the case since LMPG micelles have been shown to have an oblate structure (Oliver et al., 2013). The molecular weight values presented for Yop1_ΔC2 reflect the behaviour of the protein in solution rather than an absolute mass for the system. One would conclude that Yop1_ΔC2 is behaving like a protein 2-3 times its size due to the constraints of a micellar environment.

The discrepancy between τ_c values calculated from R_1 and R_2 can be explained by the fact that above ≈30 kDa R_2 measurements become inaccurate due to intramolecular motions (Luginbühl and Wüthrich, 2002). The $^{15}N^1H$-TRACT data also provides a lower estimate for the tumbling times due to the rigid body assumptions made in the τ_c calculation (Lee et al., 2006). The $^{15}N^1H$-TRACT experiment also has the benefit of abrogating dipole-dipole (DD) relaxation induced by remote protons and chemical exchange processes that contribute strongly to relaxation rates of larger molecules (Lee et al., 2006). It would be preferable in future to use the 2D version of the $^{15}N^1H$-TRACT experiment (Lee et al., 2006), the practical application of which was demonstrated by Lakomek et al. (2013) in their characterisation of gp41 from HIV-1. The TROSY-HNCO based relaxation experiments presented by Chill et al. (2006) could also be highly applicable since they would circumvent problems associated with overlapping signals. Various other exclusively TROSY-based methods for measurement of R_1, R_2 and $R_{2\rho}$ have also been reported (Lakomek et al., 2012). The use of time equivalent NUS to increase signal to noise and resolution (Hyberts et al., 2013) could enable more accurate and reliable measurement of R_1, R_2 values, especially in conjunction with recent software released by Aoto et al. (2014). This enables selection of Poisson gap sampling schedules which will most likely result in faithful reconstruction of peak intensities thus circumventing one of the weaknesses of NUS: accurate peak intensity measurement.

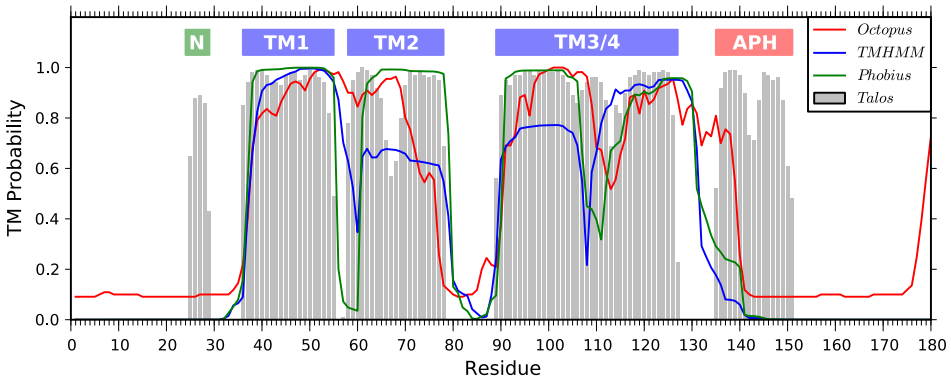


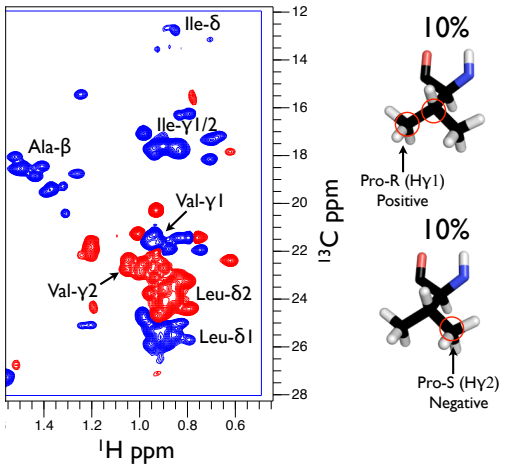
FIGURE 5.17: Comparison of three different TM predictions (Phobius (Käll et al., 2007), TMHMM (Krogh et al., 2001) and Octopus (Viklund and Elofsson, 2008)) with the experimentally determined secondary structure of Yop1_ΔC2.

Figure 5.17 shows a comparison between my TALOS-N derived secondary structure and the outputs from three transmembrane prediction algorithms, Octopus (Viklund and Elofsson, 2008), Phobius (Käll et al., 2007) and TMHMM (Krogh et al., 2001). For TM1 and 2 the predictions are in strong agreement with the experimental data. The region of TM3/4 is more ambiguous, with the predicted break points spanning residues 108 to 113. Nonetheless, a break within this region would be expected (see Figure 5.4).

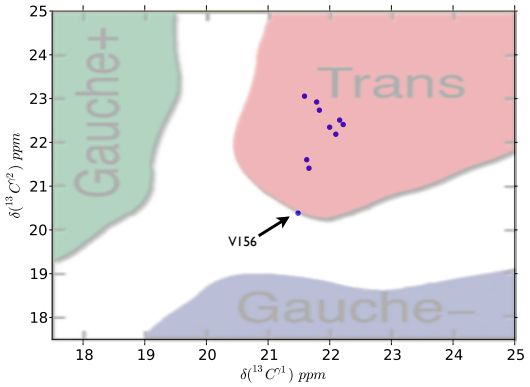
The linker between TM4 and the APH is predicted by Octopus to be transmembrane —this is in conflict with my experimental data showing high levels of solvent accessibility in this region (see Figure 5.9).

5.6 Stereospecific assignments

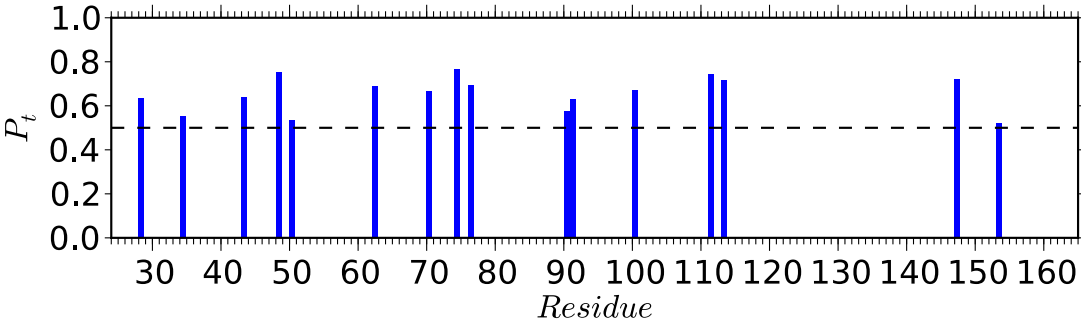
Side-chain assignments were obtained using ^{13}C and ^{15}N-edited NOESY experiments collected at 950 MHz. Stereospecific assignment of Leucine Hδ and Valine Hγ protons was achieved using the method of Neri et al. (Neri et al., 1989) and the most likely rotameric states of Ile, Val, and Leu side-chains were determined according to their chemical shifts (Hansen and Kay, 2011, Hansen et al., 2010, Mulder, 2009). All Ile, Val and Leu sidechains were determined to be in *trans* conformation which is consistent with Yop1p being a helical membrane protein. This data is summarised in Figure 5.18 and will be used in future to constrain sidechain packing interactions during structure calculations.



(A) Neri et al. (1989) labelling of Yop1_ΔC2



(B) Valine rotamers are primarily in *trans*



(C) Probability of *trans* rotamer (P_t) for Leu methyls ($\Delta\delta(^{13}C^{\delta}) = {}^{13}C^{\delta 1} - {}^{13}C^{\delta 2} = -5 + 10 \cdot P_t$)

FIGURE 5.18: Stereospecific assignment of ILV methyl rotamers using the method of Neri et al. (1989). (A) Leuδ_2 and Valγ_2 methyls are negative while Leuδ_1 and Valγ_1 methyls are positive. To the right is the expected labelling pattern for a Valine side chain. The red circles represent ^{13}C labelled nuclei with the same pattern expected for Leu-δ carbons. (B) Valine rotamers were predicted from chemical shifts according to Hansen and Kay (2011). The only outlier is C-terminal V156 which is expected to be dynamic. (C) Leucine rotamers were calculated according to Mulder (2009) and found to be primarily in *trans*. All Ileδ_1 rotamers were predicted to be in *trans* according to Hansen et al. (2010).

6

Yop1p Contains an Amphipathic Helix Required for Tubule Formation

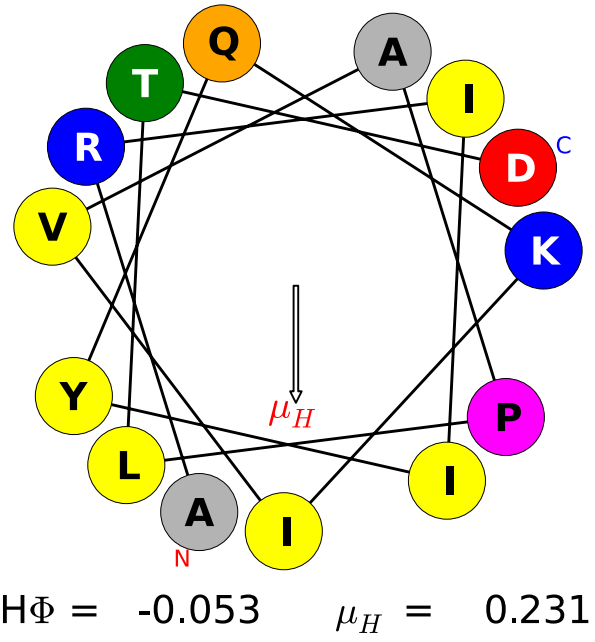


FIGURE 6.1: Helical wheel for Yop1p residues 136-149 plotted using MatPlotLib (Hunter, 2007) (see Section 2.13.2). The arrow points in the direction of the overall hydrophobic moment. HΦ is hydrophobicity and μ_H is hydrophobic moment

6.1 An unexpected finding

In light of the literature, the initial focus of my attention was on the TM regions of Yop1p which, through their "wedge" like insertion in the bilayer and oligomerisation, were proposed to drive membrane tubule formation. However, upon chemical shift assignment and secondary structure determination, the uncovering of a C-terminal helical region with intriguing sequence characteristics piqued my interest. This region possesses amphipathic characteristics and is the subject of this chapter. To begin, I will introduce the field of amphipathic helices and then go on to discuss the implications of their presence in Yop1p. What follows will be the proposal of a novel mechanism for membrane curvature generation and stabilisation by RHD proteins.

6.2 What is an amphipathic helix?

An alpha helix with both hydrophobic and hydrophilic faces relative to the helical axis is defined to be amphipathic. A formal mathematical definition was provided in seminal papers by David Eisenberg (Eisenberg et al., 1982a,b, 1984) where amphipathicity was defined as the vector sum of the hydrophobicities of the individual amino acids comprising a helix.

Formal mathematical definition of hydrophobic moment

The angle in radians per amino acid in a helix is given by, $\theta = \frac{2\pi}{m}$, where m is the period of the helix, H is the hydrophobicity of the n^{th} amino acid.

$$\mu_H = \sqrt{\left(\sum_{n=1}^{n} H_n(sin\theta n)\right)^2 + \left(\sum_{n=1}^{n} H_n(cos\theta n)\right)^2}$$

The value μ_H is then divided by the number of amino acids over which it was calculated to nomalise for sequence length.

6.3 Why are amphipathic helices interesting?

Amphipathic amino acid sequences are ubiquitous in nature and assume diverse and important functions, from ion-channel pore formation (Raghuraman and Chattopadhyay, 2007) to membrane curvature sensing (Antonny, 2011, Pranke et al., 2011), stabilisation,

generation (Drin and Antonny, 2010) and membrane fission/fusion (Liu et al., 2012). The specific sequence properties governing their assumed folds and functions is an area of intense scientific interest with great pertinence to human health and disease. For example, α-synuclein, a causal agent of Parkinson's disease, assumes an amphipathic helical fold which is thought to the comprise the basic building block for the pathogenic species (Winner et al., 2011). Antimicrobial peptides are commonly amphipathic in nature, while the APHs from BAR domains are essential for their membrane curving properties (Peter et al., 2004, Suarez et al., 2014). The APH from the Influenza M2 ion channel is essential for scission of viral buds (Rossman et al., 2010).

6.4 How amphipathic is a helix?

The greater the hydrophobic moment of a helix the more polar one face is relative to the other; however, the values depend on the hydrophobicity scale used for the calculation. There are many scales for defining hydrophobicity which generally derive free energies of transfer from a polar to non-polar environment using semi-experimental or theoretical information. A consensus scale was defined in 1982 by Eisenberg et al. (1982b) which combined 5 known scales (Chothia, 1976, Janin, 1979, Nozaki and Tanford, 1971, Segrest and Feldmann, 1974, Von Heijne and Blomberg, 1979, Wolfenden et al., 1981) by defining normalised hydrophobicites for each scale as the number of standard deviations that each amino acid's hydrophobicity lay above or below the mean of its corresponding scale. These values were then averaged and converted to $kcal\ mol^{-1}$ by multiplying by the standard deviation of the Janin (1979) scale and adding the values to the mean of that scale. Another widely use hydrophobicity scale is that of Kyte and Doolittle (Kyte and Doolittle, 1982). This scale combines water-vapour transfer energies with the hydrophobic burial scale from Chothia (1976). In their paper (Kyte and Doolittle, 1982) they introduced a computer program for producing hydropathy plots based on summing amino acid hydrophobicities with a sliding window over the target sequence. Both Janin (1979) and Chothia (1976) scales were derived from the interior-exterior distributions of amino acids based on soluble (globular) proteins of known structure.

In 1996 Wimley and White (Wimley and White, 1996) defined a purely experimental hydrophobicity scale based on the free energy of partitioning peptides into either POPC or octanol. A key feature of their scale is that it takes into account the energetics of inserting a peptide bond into a hydrophobic environment and explains why hydrogen bonded secondary structure is highly favoured for transmembrane stretches of amino acids.

6.5 Yop1p has an amphipathic helix

The structural data presented in the previous chapter defined a C-terminal helix in Yop1p comprising residues 135 to 151. This region was also predicted to be helical by several algorithms (JPRED (Cole et al., 2008), PSIPRED (Jones, 1999) and CF-SSP (Chou and Fasman, 1974)). The helix exhibits a large hydrophobic moment and has a theoretical pI of 9.7. In the Yop1p(24-165) construct the APH displayed more favourable spectral characteristics than the transmembrane regions including increased SNR and more symmetric lineshapes. However, the τ_c estimated from ^{15}N relaxation measurements (Figure 5.14, Table 5.1) for this region was comparable to that of the transmembrane regions suggesting that it does not have significant internal motion relative to the rest of protein. In addition, the amide resonances for residues 135-145 are strongly protected from solvent (Figure 5.9). Thus, a significant portion of the N-terminus of the APH likely interacts with the anionic micelle surface. Intramolecular helical NOEs but no NOEs to water protons were observed for the first two-thirds of the helix (up to P145) while NOEs to water protons were observed for the C-terminal one-third (residues 146-151), consistent with the pattern of exposure to Mn^{2+}EDDA (Figure 5.9). Together, this indicates that the APH is a stable helical region mostly bound within the lipid and detergent headgroup region.

6.6 APH from Yop1p displays classical characteristics

To further characterise the APH, Peter Smith and I expressed and purified a peptide corresponding to residues 129-154 (Yop1p(129-154)). Upon reconstitution of the peptide into detergent-free aqueous solution (20 mM KH_2PO_4; pH 6.5) it displayed random coil characteristics as shown by low proton chemical shift dispersion (Figure 6.3A) and low absorbance in far UV circular dichroism (CD) (Figure 6.3B). However, the isolated APH interacted strongly with LMPG micelles, resulting in greatly increased helicity as determined by CD (Figure 6.3B). In the presence of LMPG, helical NOEs were observed for residues 136 to 142 and 145 to 150 (Figure 6.2), with the break in helicity corresponding to those residues immediately N-terminal to P145. The same pattern was observed for this region in the full-length protein (Figure 5.2).

By contrast, the presence of the zwitterionic lysolipid LPPC resulted in only weak interactions as indicated by CD (Figure 6.3B). In order to probe specifically for membrane interactions, the CD spectrum of the isolated APH was studied in the presence of liposomes (Figure 6.3C). The APH adopted α-helical structure in the presence of liposomes

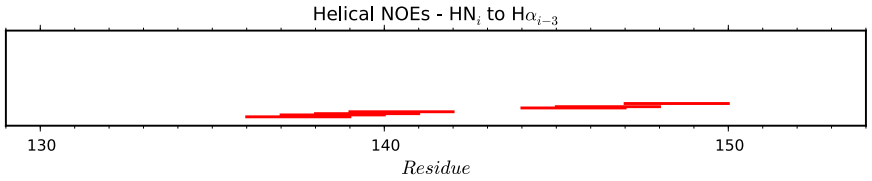


FIGURE 6.2: Helical HN$_i$-Hα_{i-3} NOEs observed for the residues 129-154 in 42mM LMPG at pH 6.5 at 37°C. Loss of helical NOEs between residues K142 and A145 can be explained the lack of amide group at Pro 146.

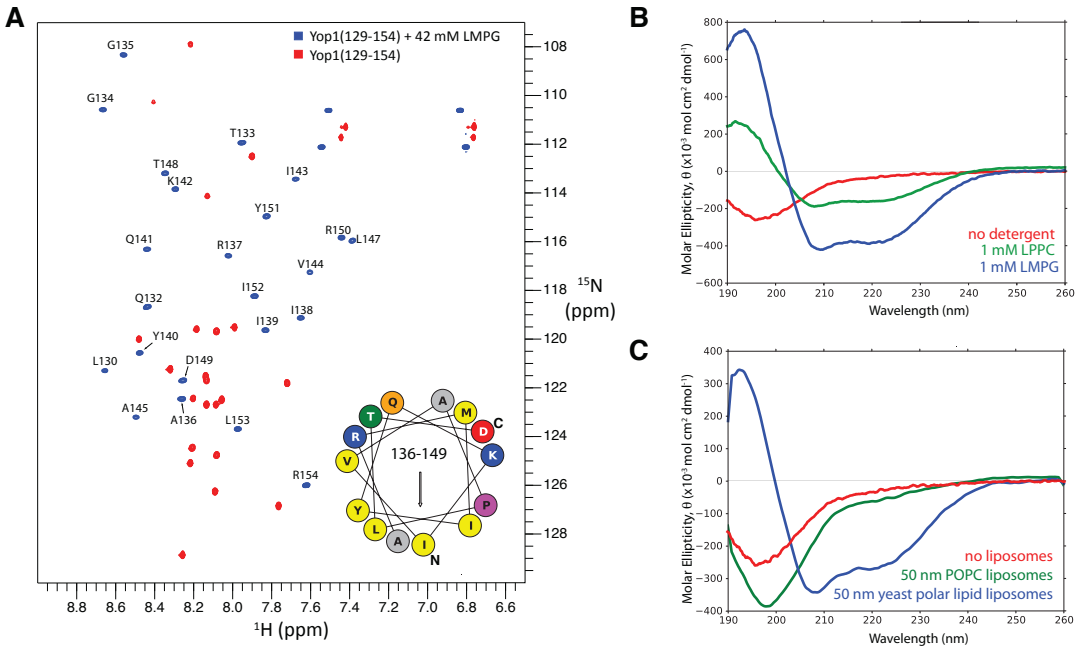


FIGURE 6.3: (A) Assigned spectrum of APH. The resonances show greater chemical shift dispersion upon addition of LMPG indicating secondary formation (most likely α-helical structure since β-sheet would result in larger dispersion). Both HSQC spectra were collected at 950 MHz at 37 °C using NUS with a poisson gap sampling schedule and processed using 400 iterations of IST. 25% of 192 indirect points were collected. Assignments for the 42 mM LMPG sample (blue) were obtained from an NUS ^{15}N-edited NOESY collected under the same conditions collecting 25% of 48 and 196 points for ^{15}N and ^{1}H indirect dimensions, respectively. (B) CD spectra of the isolated Yop1p APH with and without the addition of the anionic lysolipids LMPG or the zwitterionic 1-palmitoyl-2-hydroxy-*sn*-glycero-3-phosphocholine (LPPC). (C) CD spectra of the isolated Yop1p APH with and without the addition of liposomes formed from POPC lipid or yeast polar lipid extract.

formed from yeast polar lipid extracts, whereas little or no helical structure was observed in the presence of liposomes formed from POPC. Together these results suggest that the APH membrane interactions were favoured by anionic headgroups, consistent with the high pI of the APH. However, the isolated APH was not sufficient for tubule formation (Figure 6.5C), indicating that both the APH and the transmembrane domains were required.

6.6.1 Importance of interaction with small negatively charged headgroups

The CD spectra collected for the WT APH suggest that specific interactions with small negatively charged headgroups could be important. Preliminary results show that DPPC liposomes containing at least 30% LαPA or DPPG induced helical structure, whereas liposomes containing PI, PE and PC headgroups alone were unable to induce helicity. Yeast Polar Lipid extract along with *E. coli* Polar Lipids, both containing high proportions of negatively charged headgroups, induced helicity.

6.7 APH is necessary but not sufficient for tubule formation

A C-terminal truncation of Yop1p after residue 137 was previously shown to be unable to generate tubules *in vitro* or rescue tubule formation *in vivo* pointing to the functional importance of this protein region (Hu et al., 2008). To increase confidence that loss of tubule formation was attributable to the APH, a variant of Yop1p in which the APH region (residues 136-154) was deleted (Yop1p(ΔAPH)) was expressed and purified following the same protocols as for the WT protein. Crosslinking results for Yop1p(ΔAPH) with glutaraldehyde displayed similar oligomerisation characteristics to the full length protein (Figure 6.4A), suggesting that Yop1p oligomerisation is primarily mediated by the transmembrane domains. Furthermore, the presence of an SDS resistant dimer was observed, similar to that observed previously for the full length protein (Hu et al., 2008, Shibata et al., 2008) (Figure 4.11), although the significance of this dimer is unknown. Figure 6.4B shows a spectral overlay between the WT and truncated construct. The lack of major chemical shift perturbations more distant to the site of truncation (Figure 6.4C) suggests that the two constructs have the same overall structure and further supports the idea that the APH interacts primarily with detergent or lipid, rather than other regions of the protein. In addition, reconstitution of Yop1p(ΔAPH) into *E. coli* polar lipids did not result in tubule formation consistent with the importance of this region in Yop1p function (Figure 6.5).

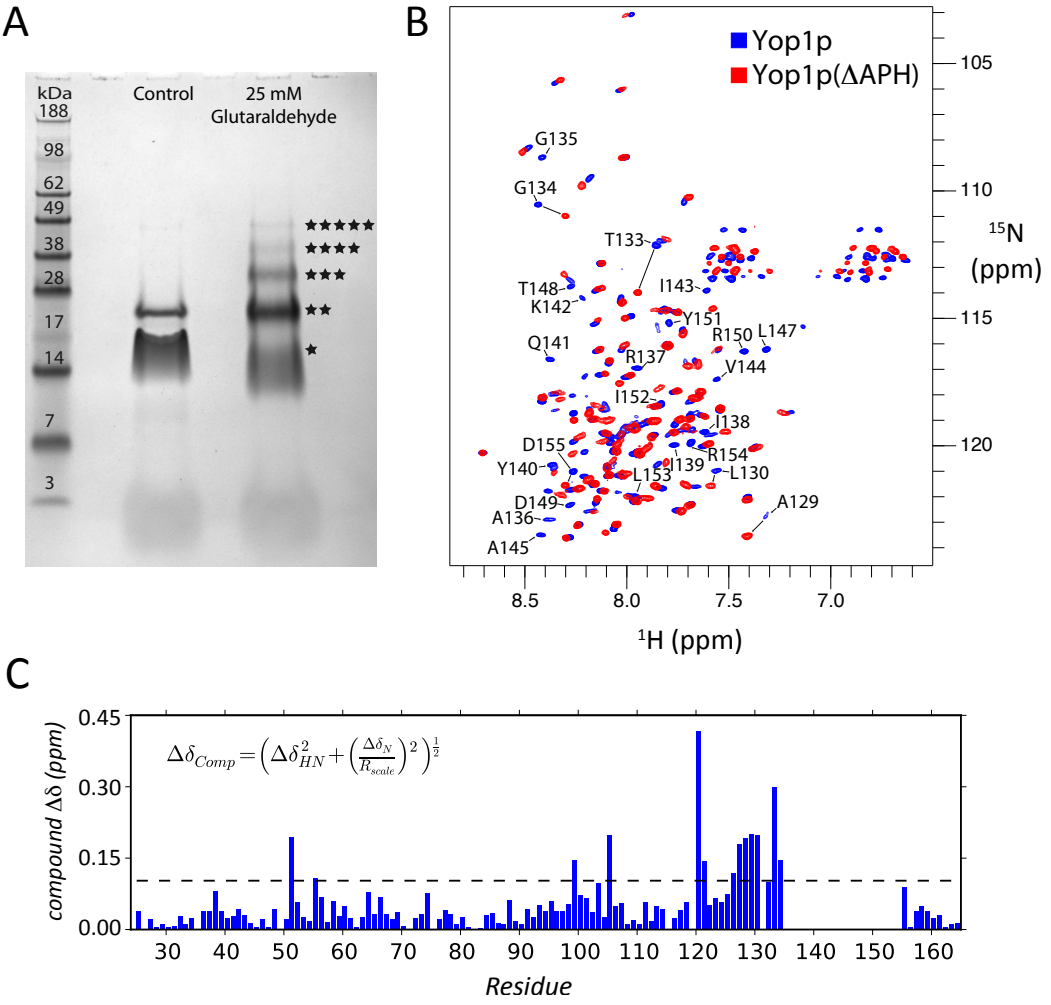


FIGURE 6.4: (A) Chemical crosslinking of Yop1p(ΔAPH) reconstituted into EPL according to Sections 2.8 and 2.11 with the indicated glutaraldehyde concentrations. The control has no glutaraldehyde and the gel was stained with Coomassie Brilliant Blue. The stars to the right of the star indicate the apparent oligomerisation state of each easily observable band. (B) Overlay of ^{1}H-^{15}N TROSY-HSQC spectra for Yop1p (residues 24-165) and Yop1p(ΔAPH). Each spectrum was acquired with NUS collecting 50% of 192 real points in the indirect dimension and reconstructed with 400 iterations of IST. Yop1p was reconstituted into q=0.5 LMPG:DPPC mixed micelles and Yop1p(ΔAPH) in LMPG. Both spectra were recorded at 950 MHz (^{1}H) at pH 6.5 in 20 mM KH_2PO_4. (C) Combined ($^{1}H_N$ and $^{15}N_H$) chemical shift difference plot (Mulder et al., 1999) between Yop1p and Yop1p(ΔAPH)

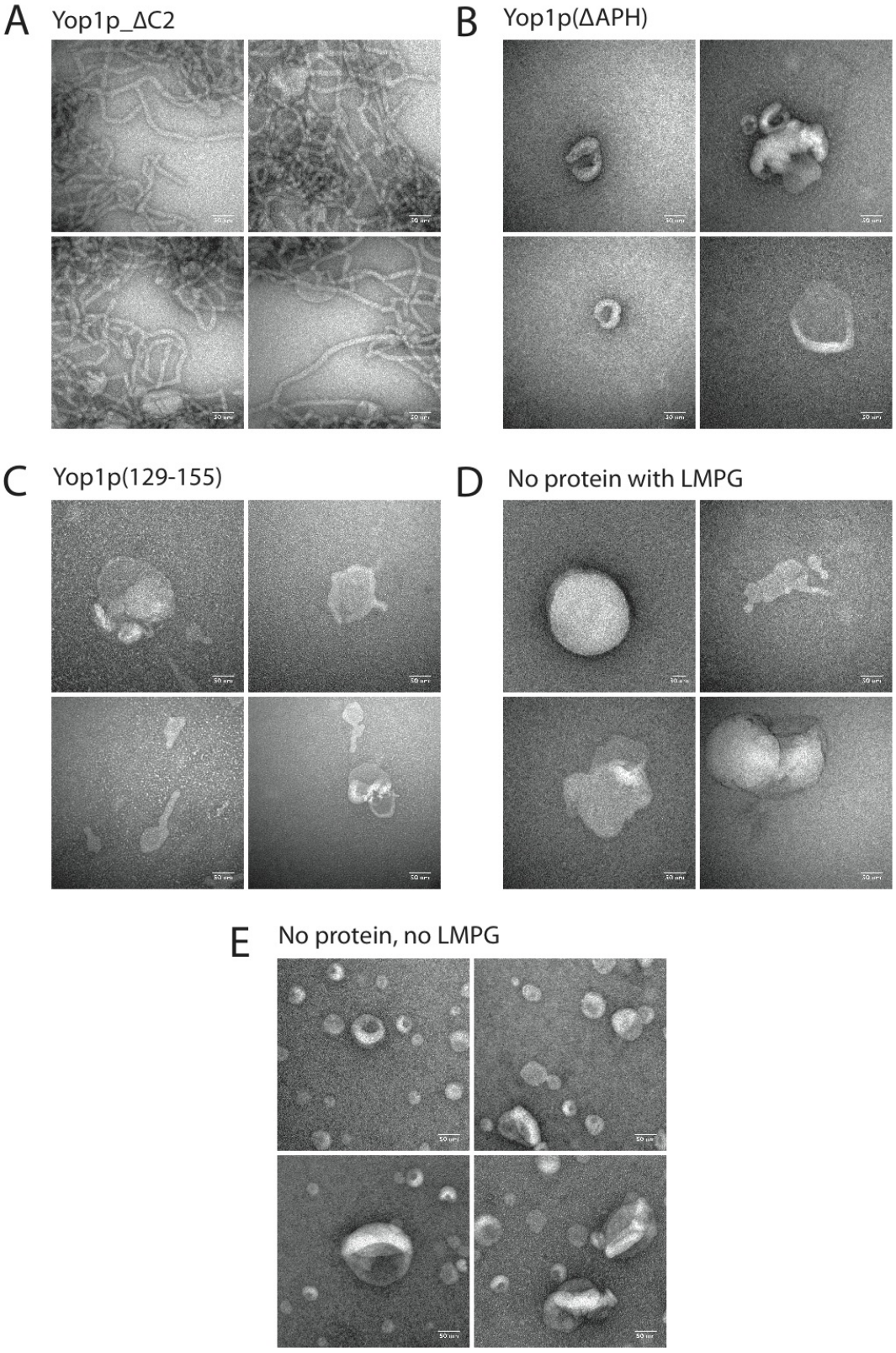
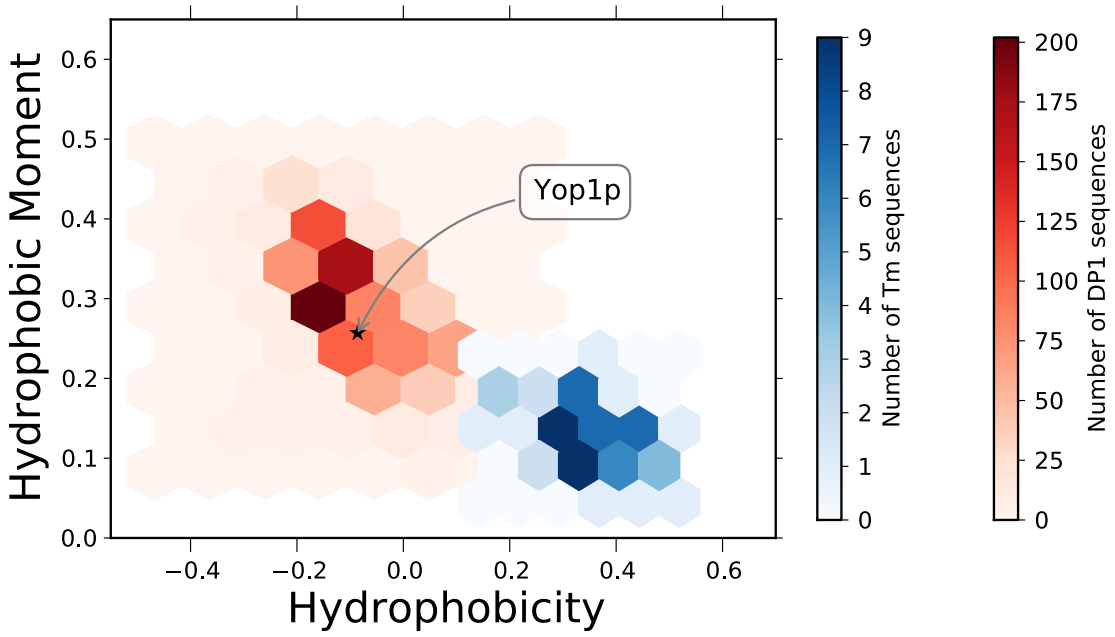


FIGURE 6.5: (A) Electron micrographs of Yop1_ΔC2 reconstituted into E. coli polar lipids from 0.01% LMPG showing extensive tubule formation. Micrographs of samples containing Yop1p(ΔAPH) (B), Yop1p(129-154) (C), no protein (D), or no protein or LMPG (E) indicated liposome formation but no tubules. Yop1_ΔC2 and Yop1p(ΔAPH) samples were prepared with 2 mg/mL protein at ratio to lipid of 1:1 (w:w); a 1 in 50 dilution was placed on to each grid. Yop1p(129-154) samples were prepared with 1 mg mL^{-1} protein at a ratio to lipid of 1:1 (w:w); 1 in 25 dilutions were placed onto each grid. Scale bars represent 50 nm

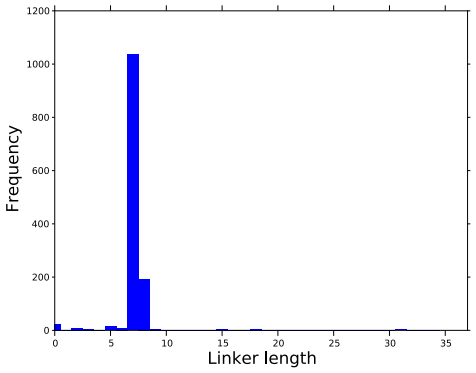
6.8 RHD proteins contain conserved amphipathic helices

Since amphipathic helices are often implicated in membrane curvature sensing and generation (Antonny, 2011) the conservation of the APH amongst Yop1p homologs was evaluated. Conservation of sequence identity in this region was low (average <30%), and thus more general properties of the sequences were explored. All sequences of the DP1 family (PF03134; 1350 sequences) were analysed for both hydrophobicity and hydrophobic moment to evaluate whether the membrane surface seeking properties of the helix (Eisenberg et al., 1982a) may be conserved. For Yop1p, the 14 helical residues from A136 to D149 have the largest hydrophobic moment (0.257). Thus, in calculating the regions of highest hydrophobic moment for other DP1 family sequences, the sequences were assumed to be helical and a 14-residue averaging window was used.

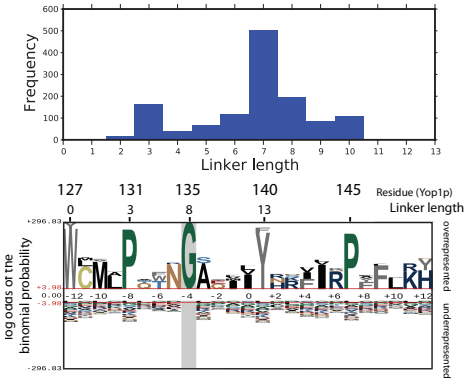
Analyses indicated that all DP1 family members contain a potential amphipathic helix with highest hydrophobic moment C-terminal to the transmembrane domains and that their hydrophobic moments and hydrophobicities are consistent with them being membrane seeking helices (Figure 6.6A). Since both the transmembrane domains and the APH are necessary for membrane tubule formation by Yop1p (Figure 6.5), I tested whether the distance in amino acids between them is conserved. Indeed, the vast majority of transmembrane domain-APH linkers are 7 or 8 residues in length (Figure 6.6B-C), suggesting that the transmembrane domains and the APH may function together to generate membrane curvature. Figure 6.7A further demonstrates that the vast majority of sequences contain a region of high hydrophobic moment between 7 to 8 residues from the end of TM4.



(A) Data from 1284 sequences from the DP1 PFAM entry

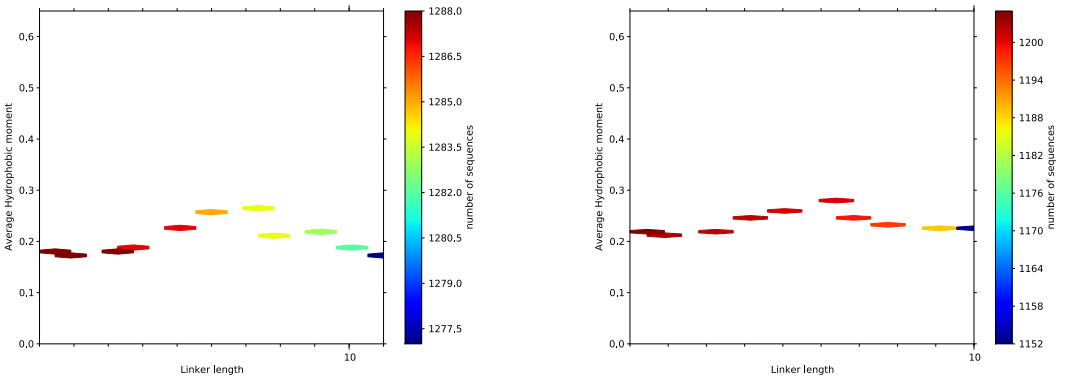


(B) Linker length between Y127 and G135 relative to the aligned sequences



(C) Linker length between Y127 and point of highest hydrophobic moment

FIGURE 6.6: (A) 2D histogram of hydrophobic moment versus hydrophobicity (Terwilliger, 1982) for the APH region of the DP1 family (PF03134;1350 members, 1288 of which were analysed (see Section 2.13.2)) aligned using Clustal Omega (Sievers et al., 2011). Data for the DP1 family members are plotted in red, and selected transmembrane sequences from yeast membrane proteins (Sharpe et al., 2010) are shown in blue for comparison. (B) Histogram for all members of the DP1 family showing the linker length between the end of the last transmembrane domain (equivalent to position Y127 in Yop1p) and the start of the conserved APH region (G135). (C, top) Histogram for linker lengths starting from Y127 to position of highest hydrophobic moment. Hydrophobic moments were determined using the python program described in the methods section (see Section 2.13.2). (C, bottom) pLogo (O'Shea et al., 2013) visualisation of sequence conservation for a subset of 341 sequences containing no alignment gaps relative to Yop1p (residues Y127-Y151 using Yop1p numbering). Plots were generated with MatPlotLib (Hunter, 2007).



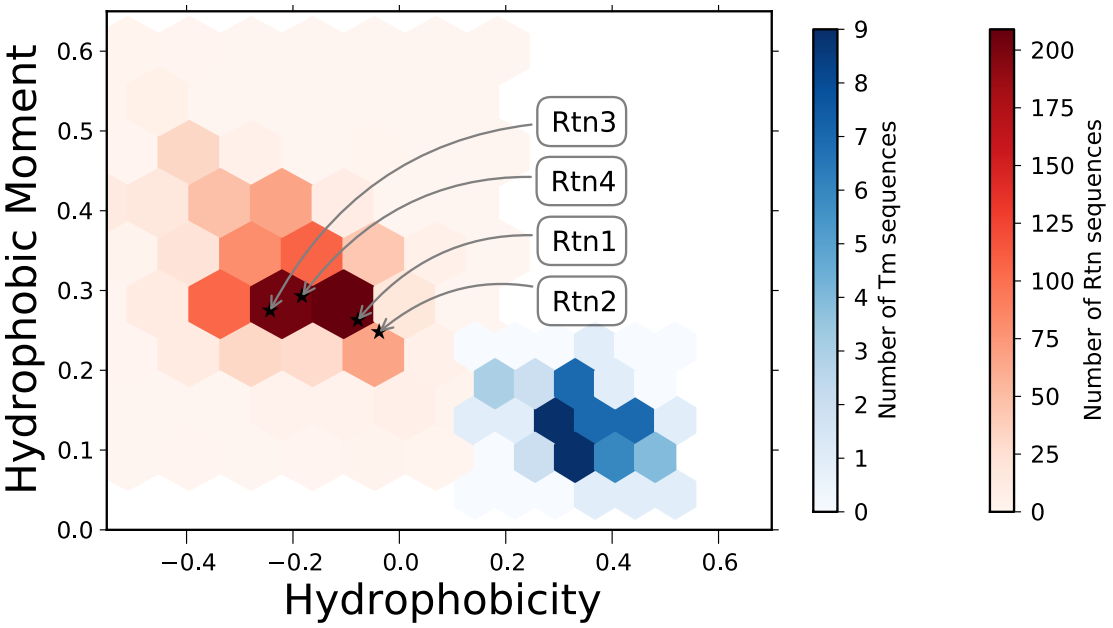
(A) DP1 family sequences (B) Rtn family sequences

FIGURE 6.7: The linker length is plotted against mean hydrophobic moment with the colour scheme representing the number of sequences that contributed to each average value.

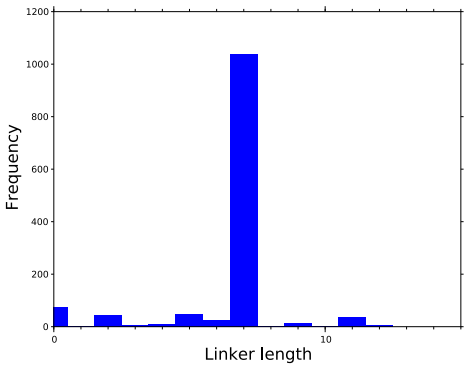
Next, I asked whether other RHD containing proteins such as the reticulon family may contain amphipathic helices. Rtn1-4 sequences were all predicted to contain amphipathic helices C-terminal to their RHDs and also to cluster in the surface seeking region of a plot of hydrophobicity versus hydrophobic moment (Figure 6.8A). As shown in Figure 6.7B the vast majority of sequences have a linker length of 6 to 7 residues from the predicted end of the fourth TM and the point of highest hydrophobic moment (see Figure 6.17 for sequence context.). Table 6.1 summarises the data for Yop1p and the human reticulon sequences which are also labelled in Figure 6.6A and Figure 6.8, respectively.

TABLE 6.1: Hydrophobic moment data for Yop1p and human reticulons

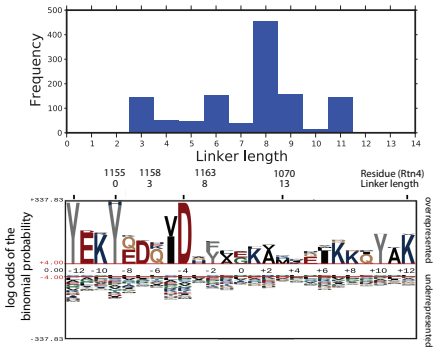
Protein	Uniprot ID	Species	Residues	Hydrophobicity	Hydrophobic moment
Yop1p	Q12402	*S. cerevisiae*	136-149	-0.086	0.257
Rtn1c	Q16799-3	*H. sapiens*	179-192	-0.078	0.263
Rtn2c	O75298-3	*H. sapiens*	162-175	-0.038	0.248
Rtn3c	O95197-3	*H. sapiens*	206-219	-0.243	0.275
Rtn4c	Q9NQC3-3	*H. sapiens*	169-182	-0.183	0.293



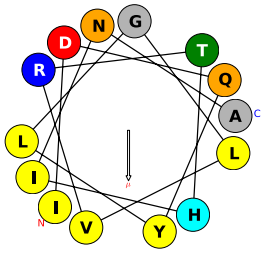
(A) Data from 1203 sequences from the Rtn PFAM entry



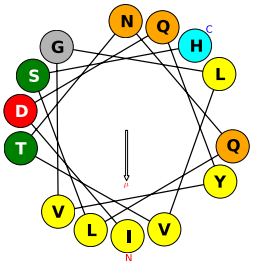
(B) Linker lengths between Y1155 and I1162 relative to Rtn4



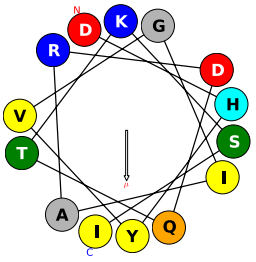
(C) Linker length from Y1155 to point of highest hydrophobic moment



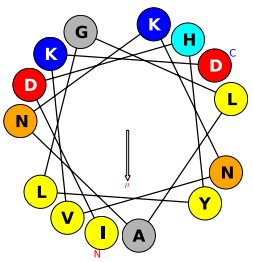
(D) Rtn1 (746-759)



(E) Rtn2 (502-515)



(F) Rtn3 (1002-1015)



(G) Rtn4 (1162-1175)

FIGURE 6.8: Potential conservation of amphipathic character in reticulon family proteins. (A) 2D histogram of hydrophobic moment and hydrophobicity plotted for a Clustal Omega alignment of 1203 sequences from the reticulon family (PF02453; 1346 members). The positions of the putative Rtn APHs are marked. (B) Histogram of linker lengths from the predicted end of the 4th TM (Y1155, Figure 6.17 -Rtn4 numbering) of the reticulon RHD to the beginning of the APH region (I1162). (C, top) Histogram of linker lengths based on the point of maximum hydrophobic moment for the same set of sequences. (C, bottom) pLogo (O'Shea et al., 2013) of a subset of 620 sequences containing no gaps relative to the human reticulons (residues I1155-A1179 using human rtn4 numbering). (D-G) Helical wheel plots of human reticulons 1-4 at the point of highest hydrophobic moment towards the C-terminus of the RHD. The arrows point in the direction of the resultant hydrophobic moment vector and the corresponding values are listed in Table 6.1. RHDs were defined using the Octopus transmembrane helix prediction server (Viklund and Elofsson, 2008) and the region C-terminal to each RHD with maximum hydrophobic moment was determined using the python program described in the methods section (see Section 2.13.2)

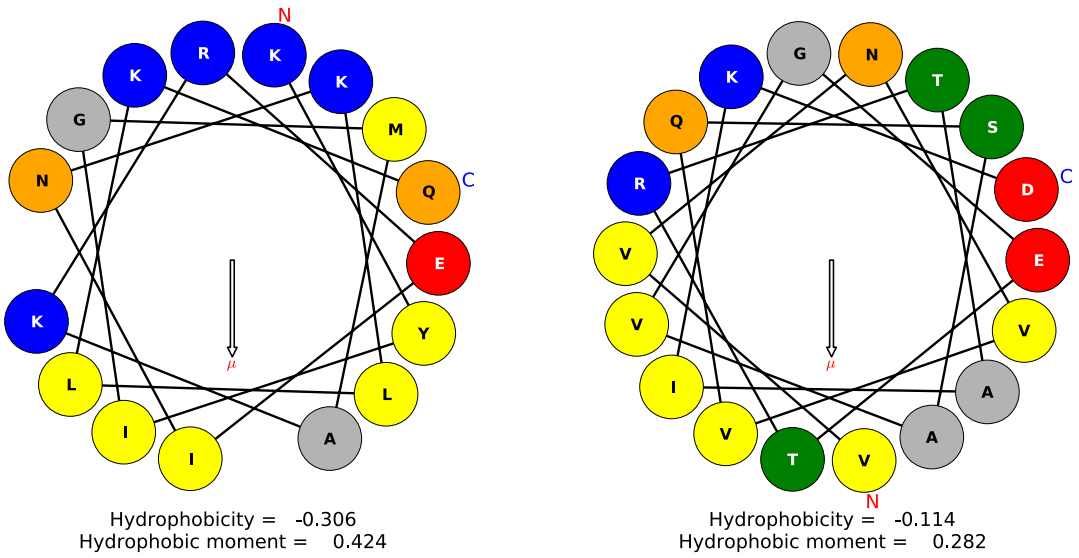
6.9 Possible roles for the APH in RHD mediated membrane curvature

The conservation of amphipathic character and positioning relative to the RHD transmembrane domains (Figure 6.6B), coupled with the finding that without it Yop1p is unable to form tubules *in vivo* or *in vitro* (Hu et al., 2008) (Figure 6.5), strongly implicates the APH as a functionally essential component of the DP1 family RHD. Furthermore, bioinformatic analyses of the reticulons suggest the presence of amphipathic motifs in these proteins as well (Figure 6.8). Although a direct test of the presence of an amphipathic helix in reticulons is lacking, strong membrane interactions in the C-terminal region of reticulons could provide an explanation for the intermediate level of labelling with maleimide PEG seen for the C-terminal region of Rtn4c that was found here to contain the amphipathic motif (Voeltz et al., 2006).

Other families involved in ER membrane remodelling also contain C-terminal amphipathic helices, although their functional relationship to that of Yop1p is not known. In atlastin proteins, a C-terminal amphipathic helix is required for ER membrane fusion and it is likely that Sey1p, atlastin's functional homolog in *S. cerevisiae*, also has a C-terminal amphipathic helix (Liu et al., 2012). In addition, Arl6lP1, an ER resident protein found in multicellular organisms that localizes to the tubular ER and is capable of generating tubules *in vitro* (Yamamoto et al., 2014), also appears to contain a sequence capable of forming an amphipathic helix at its C-terminus (residues 182-197). Helical wheel plots of the predicted amphipathic regions of both Arl6lP1 and Sey1p are shown in Figure 6.9.

My data suggest that the Yop1p APH has a preference for negatively charged lipid headgroup types, consistent with the high pI of the helix (9.7). This is interesting since previously reported tubule formation in pure PC lipids was less efficient (Hu et al., 2008) while tubules were efficiently formed in any lipid mixture containing significant amounts of anionic lipids such as *E. coli* polar lipids (67% PE, 23% PG, and 9.8% CA) and egg PC/PA. Thus, it could be that Yop1p requires specific interactions with negatively charged lipid head groups, which is a behaviour commonly seen in amphipathic helices involved in membrane curving. For example, N-, F- and I-BAR H0 helices (Peter et al., 2004) all bind preferentially to negatively charged lipids while α-synuclein (Pranke et al., 2011), for instance, specifically binds curved membranes containing negatively charged lipids.

ALPS motifs have been shown to specifically bind highly curved membranes with little sensitivity to lipid composition (Drin et al., 2007). Since a hallmark of ALPS motifs is



(A) Arl6lP1 residues 182-197

(B) Sey1p residues 730-747

FIGURE 6.9: Helical wheel schematics for Sey1p and Arl6lP1 showing potential amphipathic sequences towards their C-termini

their abundance of polar residues and lack of charge on their polar faces the Yop1p APH would probably not fall into the pure curvature sensing category. It is interesting to note that, like the Yop1p APH, the ALPS motifs and α-synuclein are both unstructured in solution and assume their helical conformation upon membrane binding.

The highly conserved central proline residue of the Yop1p APH (P146) may be functionally important in forming the correct lipid interactions for sensing or stabilising membrane curvature, since prolines are frequently found in functionally related proteins (Antonny, 2011, Drin and Antonny, 2010). It has been observed, for example, that mutation of the amphipathic helix proline in SpoVM from *B. subtilis*, which localises to highly curved membranes during spore formation, results in the peptide becoming insensitive to membrane curvature (Ramamurthi et al., 2009). Prolines have been found to be important also in the amphipathic helices of antimicrobial peptides (Dempsey et al., 1991, Lee et al., 2013, Yang et al., 2006).

6.9.1 Disrupting the APH

The laboratory based work for this section was carried out by a Part II student, Peter Smith, under my supervision. The experiments and constructs used in this study were

designed by me. NMR data collection and subsequent data analysis was performed by me.

In order to dissect the functional properties of the Yop1p APH, mutants were designed to disrupt the hydrophobic face of the helix. Only mutant I143N (Figure 6.10) was produced due to time constraints but a list of potential mutations, some of which already cloned, are shown in Figure 6.11. I143N was designed to reduce the hydrophobic moment to determine whether its membrane interactions could be knocked out.

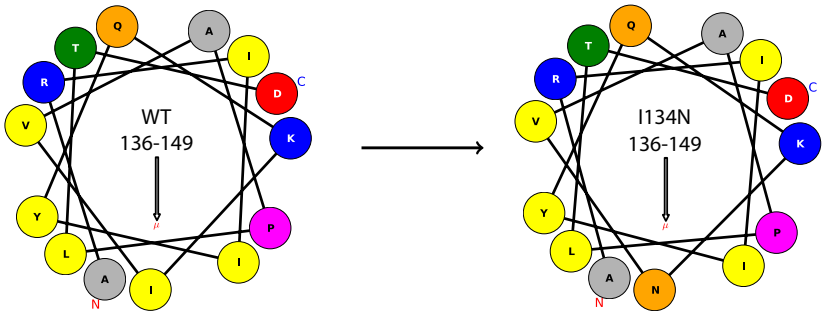
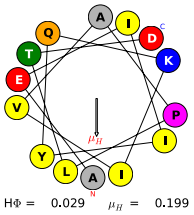


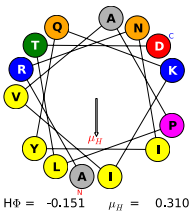
FIGURE 6.10: Disruption of amphipathic character was achieved by mutation of I143 to N reducing the hydrophobicity from -0.053 to -0.151 and the hydrophobic moment from 0.231 to 0.134. The arrows within the plots point in the direction of the overall hydrophobic moment, μ.

The mutant displayed similar behaviour to the WT in that it was fully soluble in aqueous detergent free solution and displayed random coil characteristics (Figure 6.12B). However, upon titration of LMPG or lipids the protein displayed lower helical propensity relative to the WT as evidenced by weaker α-helical CD signals (Figure 6.12) and lack of observable helical NOEs. The overall helical content predicted by TALOS-N (Shen and Bax, 2013a) was also lower compared to the WT helix (Figure 6.14); however TALOS's predictive power is limited by the absence of Cα, Cβ or C' shifts. The most striking chemical shift differences between the WT and I143N are at positions that would be expected to be on the same face of an α-helix (Figure 6.13).

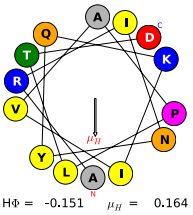
In conclusion, the mutant peptide interacts with lipids and detergents in a similar fashion to the WT protein and likely assumes a largely helical structure. However, the lower helical propensity and absence of helical NOEs suggests that the lipid interaction is disrupted by the presence of a polar amino acid (N) on the hydrophobic face. In order to completely abolish the lipid or detergent interactions and understand their molecular basis it will be necessary to carry out a more extensive mutagenesis study (Figure 6.11). The mutations should then be tested for their effects on membrane tubule formation in Yop1_ΔC2.



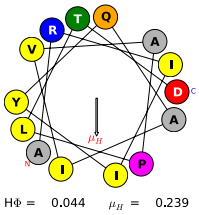
(A) R137E



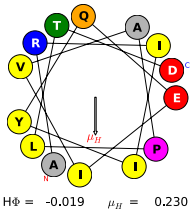
(B) I138N



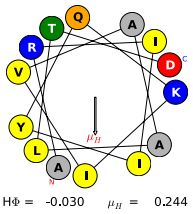
(C) I139N



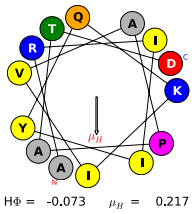
(D) K142A



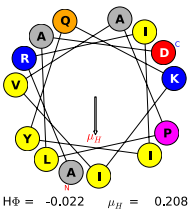
(E) K142E



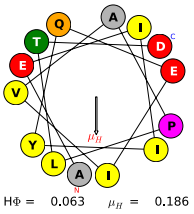
(F) P146A



(G) L147A



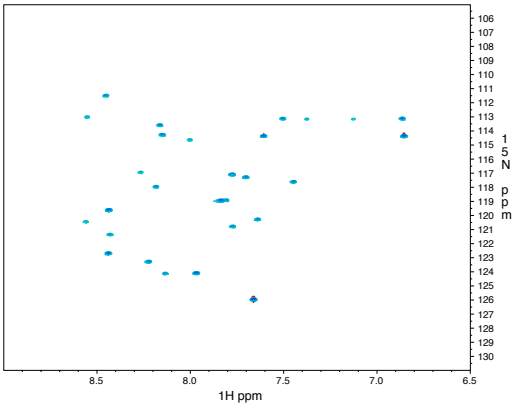
(H) T148A



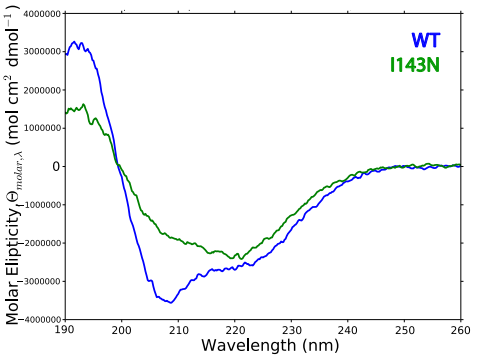
(I) R137E + K142E

	Mutation	Hydrophobicity	Hydrophobic moment	Cloned
WT	M138	-0.086	0.257	✓
A	R137E	0.029	0.199	-
B	I138N	-0.151	0.310	✓
C	I139N	-0.151	0.164	✓
D	K142A	0.044	0.239	-
E	K142E	-0.019	0.230	-
F	P146A	-0.030	0.244	-
G	L147A	-0.073	0.217	✓
H	T148A	-0.022	0.208	-
I	R137E,K142E	0.063	0.186	-

FIGURE 6.11: These proposed mutations and combinations thereof could be used to interrogate properties of the Yop1p APH such as specific lipid binding. These mutations could also be transferred to the full length protein to interrogate the effects on tubule formation. The positive to negative mutations may be able to disrupt interactions with negatively charged lipids.



(A) NMR spectrum of Yop1p(129-154) I143N with 42 mM LMPG at 37°C at 950 MHz



(B) Far CD spectra of WT and I143N Yop1p(129-154) in the presence of extruded YPL liposomes

FIGURE 6.12: (A) The HSQC spectrum I143N Yop1p(129-154) has similar dispersion to the WT helix indicating the presence of residual α-helical structure. (B) Comparison of far CD spectra for WT and mutant (I143N) APH in the presence of 0.15 mg mL^{-1} Yeast polar lipid extract (YPL) extruded through 200 nm polycarbonate membranes. The mutant construct displays reduced helical character relative to the WT.

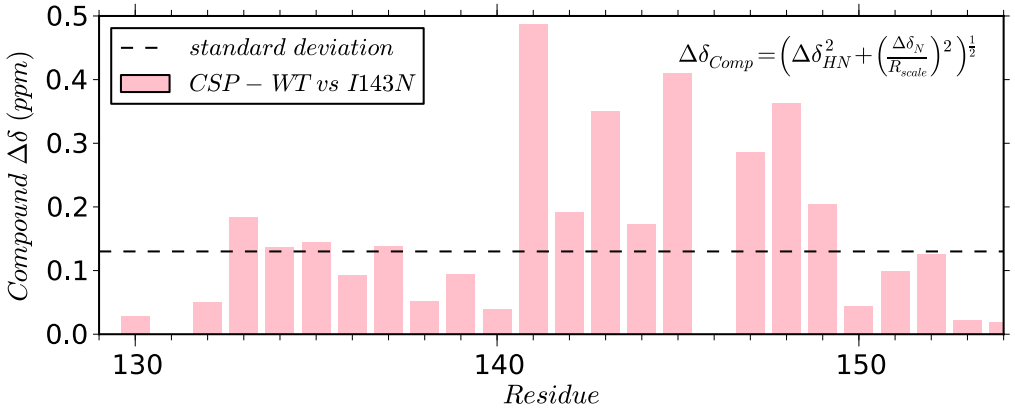


FIGURE 6.13: Amide chemical shift perturbation (CSP) between WT and I143N mutant of Yop1p(129-154) in the presence of 42 mM LMPG at 37 °C, pH 6.5. As expected the major shifts occur near the mutation site. CSP was calculated according to Mulder et al. (1999) where R_{scale} = 6.5 (see Section 2.14.4). The corrected standard deviation was not used here since the data set was too small.

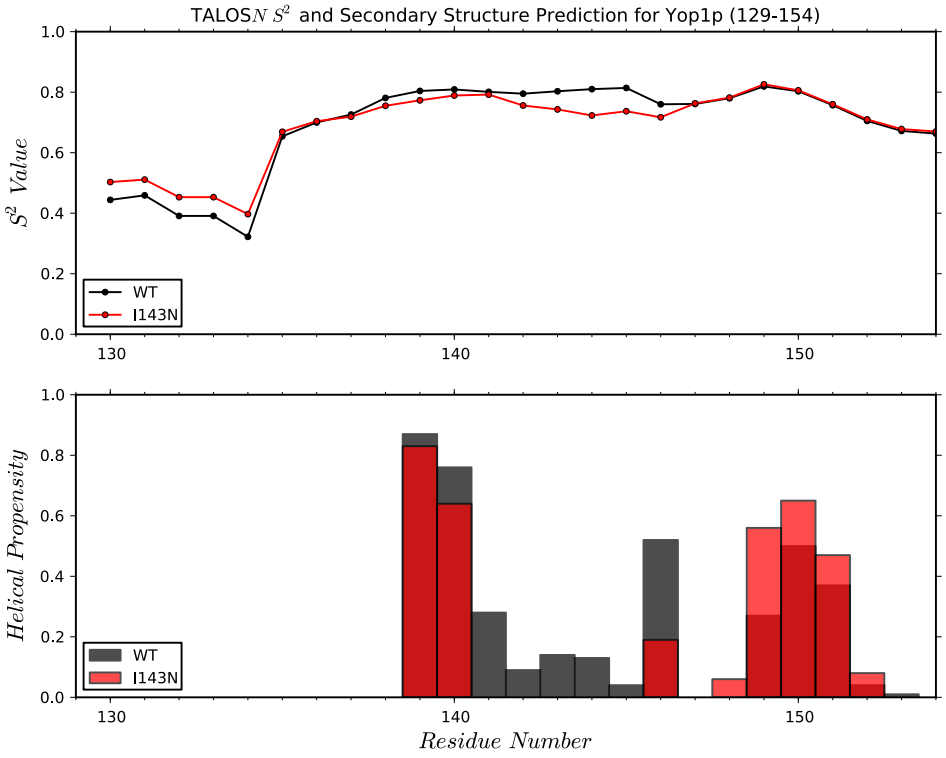
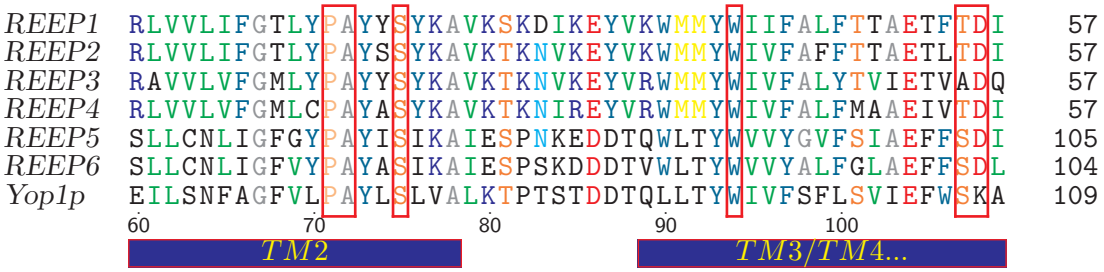


FIGURE 6.14: TALOS-N results for WT and I143N mutant of Yop1p residues 129-154 in the presence of 42 mM LMPG. The prediction is based on $^{1}H_N$, ^{15}N and $^{1}H\alpha$ chemical shifts which show a slight decrease in overall helical propensity for the mutant peptide

6.10 Disease related mutations

Yop1p is a potentially useful model protein for studying the effects of disease related mutations in human REEPs. Most strikingly, many point mutations in REEP1 (SPG31) that are known to cause HSP occur at highly conserved positions with respect to Yop1p (see Figure 6.15)(Falk et al., 2014). Therefore, exploring the effects of these mutations on Yop1p structure and function may enhance our understanding of the molecular basis of HSP. It is interesting to note that, as yet, no disease related mutations have been localised to the APH region. This could reflect the robustness of the amphipathic region to mutations such as I143N. Although, it would be interesting to demonstrate the effects of APH mutations *in vivo*.



Mutation	Yop1 Numbering	Phenotype (Falk et al., 2014)	Source
P19R	71	Mislocalisation	Beetz et al. (2008)
P19L	71	Mislocalisation	Goizet et al. (2011)
A20E	72	Mislocalisation	Beetz et al. (2008)
S23F	75	Mislocalisation	Goizet et al. (2011)
W42R	94	Mislocalisation	Goizet et al. (2011)
T55K	107	Ambiguous	Schlang et al. (2008)
D56N	108	Ambiguous	Goizet et al. (2011)

FIGURE 6.15: Top, highlighted in red are positions of HSP associated mutations in REEP1. The sequence numbering is with respect to Yop1p. A highly conserved region towards the C-terminus of TM2 and a conserved tryptophan residue at the N-terminus of TM3 contain mutations that cause pure HSP. Strikingly mutations in the putative break point between TM3 and 4 have also been shown to cause HSP further implicating the region as structurally important. Bottom, table of mutations with references. Numbering is with respect to human REEP1. The *in vivo* phenotypes of the mutations were recently characterised by Falk et al. (2014) however, the two mutations around the TM3/4 break did not show any significant phenotypic changes. Mislocalisation, in this case, refers to lack of colocalisation with an ER marker. An additional mutation, L107P has been found in the microtubule binding domain of REEP1 (Schlang et al., 2008).

6.11 A new model for RHD structure and function

The work discussed in the preceding chapters and sections has enabled the proposal of a new model for RHD structure and function. Yop1p has a C-terminal amphipathic helix that is necessary but not sufficient for tubule formation *in vitro* and sequence analysis suggests that C-terminal amphipathic sequences may be a general feature of RHDs (Figures 6.6 and 6.8). Based on structural data and sequence conservation, a single break near W106 is proposed in which TM3 and TM4 are able to span the bilayer at least to the opposing leaflet headgroup region (Figure 6.16). Furthermore, this structural data has informed the first convincing alignment of a DP1 family protein (Yop1p) with the reticulon proteins (Figure 6.17). Despite the low sequence identity between Yop1p and the human reticulons (∼16%) this alignment highlights their likely similar topological features and C-terminal amphipathic character. Surprisingly many of the residues at the starts and ends of the proposed TM regions are conserved.

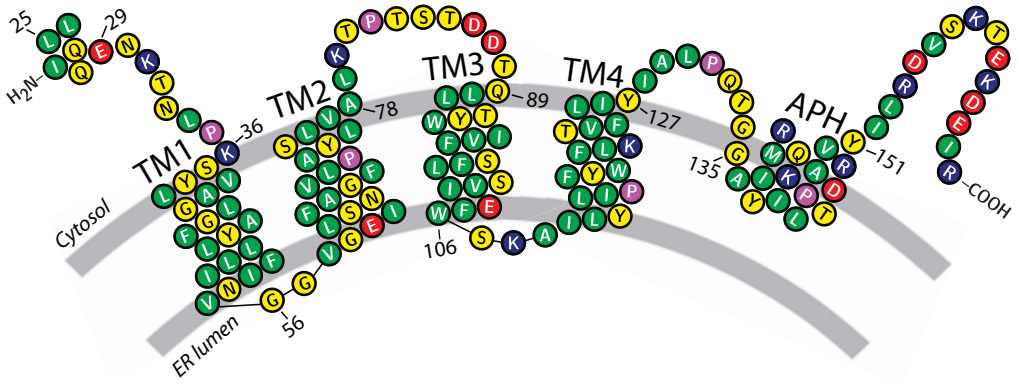


FIGURE 6.16: Proposed topology of Yop1p. Figure was produced by modification of TeXtopo output

Regardless of the RHD model, the amphipathic helix of Yop1p interacts strongly with membranes, is required for membrane tubule formation, and both its amphipathic character and positioning relative to the RHD are conserved. Moreover, transmembrane-mediated oligomerisation could provide a scaffold for the correct orientation and insertion depth of the APH in order for membrane curvature to be stabilised. Thus, it is possible to propose a model in which the transmembrane domains of DP1/reticulons provide a membrane embedded scaffold for the congregation of amphipathic helices with specific orientations and insertion depths. This model is reminiscent of that for the N-BAR domains, in which an oligomerisation and scaffolding domain orients APHs in the membrane to stabilise the membrane curving. However, in this case the scaffolding domain is in the membrane rather than on the membrane surface. Like in the BAR

domains, oligomerisation via the RHD transmembrane domains may provide a curved scaffold for curvature sensing and stabilisation.

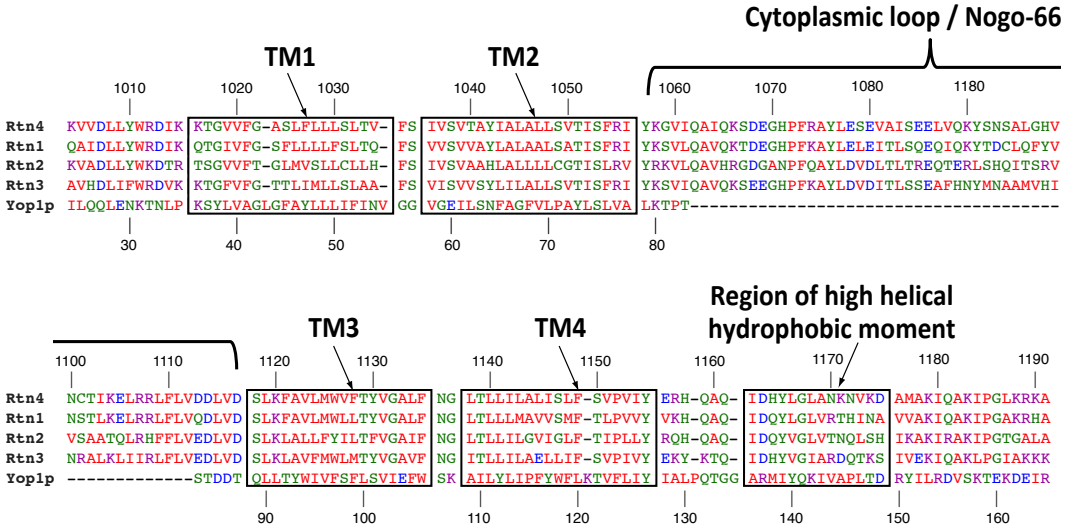


FIGURE 6.17: Sequence alignment of the RHD regions of the human reticulons (Rtn 1-4) with Yop1p. The alignment was generated manually (by Jason Schnell) based on the Yop1p secondary structure determined here, and the previously determined location of the Nogo-66 domain of the reticulons (GrandPre et al, 2000; Vasudevan et al, 2010). After inclusion of a gap in Yop1p for the Nogo-66 domain, few additional gaps were required to optimise the alignment, although the homologous TM2 region of the reticulons is 18 amino acids in length rather than the 20 amino acids that was determined for Yop1p. Residue numbering above the sequences corresponds to that of Nogo-A and residue numbering below the sequences corresponds to that of Yop1p.

There remain many outstanding questions about the mechanism of RHD action, including which regions of the protein are involved in oligomerisation, and which are involved in interactions with other proteins such as spastin and the atlastins. A high resolution tertiary structure of Yop1p or a related protein remains an important goal that will help significantly in focusing these questions. The demonstration here of the feasibility of obtaining high-resolution structural information on Yop1p may lead to a better understanding of the molecular mechanism of membrane curving in RHD-containing proteins.

In the Chapter 7 I will discuss the current state of progress towards solving the 3D structure of Yop1p using NMR. But first I will highlight some experiments of potential use for further elucidating functional aspects of RHD proteins.

6.12 Future work

It would be potentially fruitful to experimentally determine the existence of amphipathic helices in other RHD proteins and fully understand their roles in ER structure and function. To begin with, a CD study of synthetically produced putative APH regions from the reticulon family could provide experimental verification of their amphipathic character and lipid interactions. This work could then be carried further into the study of full length RHDs. Below in Table 6.2 are the peptides to be synthesised.

TABLE 6.2: Proposed peptides for study of amphipathic character in the reticulon family

Reticulon	Species	Residues	Sequence
Rtn1_a	*H. sapiens*	741-765	KHQAQIDQYLGLVRTHINAVVAKIQ
Rtn2_a	*H. sapiens*	496-520	RQHQAQIDQYVGLVTNQLSHIKAKI
Rtn3_a	*H. sapiens*	996-1020	KYKTQIDHYVGIARDQTKSIVEKIQ
Rtn4_a	*H. sapiens*	1057-1081	RHQAQIDHYLGLANKSVKDAMAKIQ

APH disrupting mutations such as those shown in Figure 6.11 should be introduced into Yop1_ΔC2 in order to see if tubule formation can be abrogated. This information could also be used to design Yop1p mutants to be tested *in vivo*. A precedent for such experiments was set by Gallop et al. (2006) who showed that mutation of positively charged residues to negatively charged residues in the endophilin H0 helix abrogated the proteins membrane curving properties by disrupting its lipid interactions.

Further information on the tertiary and oligomeric structure of Yop1p could be used in simulations to propose an exact mechanism for tubule formation in the ER.

To test the importance of linker length between the fourth TM and the APH region, constructs should be produced with truncations or extensions of the region and tested for their ability to form tubules *in vitro*.

With the age of "big data" and huge computational resources in full swing it is possible to screen sequence databases and potentially discover new classes of proteins that contain amphipathic helices or show conservation of features across known protein classes. However, despite these resources the main bottleneck remains our knowledge of exactly which sequence features contribute to specific characteristics. For instance, the propensity of a sequence to fold upon interaction with a hydrophobic surface such as a membrane. Thus, correctly identifying APHs and predicting their functions remains a problem. Gautier et al. (2008) have made huge progress in this area with the release of HELIQUEST which scans databases to find potential amphipathic sequences in yet uncharacterised proteins across genomes. However, the success of the search relies strongly

on the input parameters, and hence prior knowledge, and the program failed to find an APH in the RDH family proteins.

Programs such as HELIQUEST could potentially be improved with machine learning techniques such as support vector machines (SVM). By training against data sets of amphipathic helices with known functions and characteristics it may be possible to build a program capable of finding amphipathic sequences and also predicting their functions. However, I think this kind of approach may require a lot more data than is currently available.

An approach for finding amphipathic sequences based on sequence data alone remains difficult and a challenge beyond the scope of this thesis. However, for specific instances where amphipathic sequences are defined (such as Yop1p), experimental tests to probe lipid specificity, curvature sensing and structural propensities are relatively straightforward to carry out, requiring only basic biophysical methods.

7

Towards a Tertiary Structure Determination of Yop1p

"That is the exploration that awaits you; not mapping stars and studying nebulae, but charting the unknown possibilities of existence."

Q

7.1 The hunt for long-range NOEs

By NMR, the most convincing and accurate way to define the fold and packing interactions in proteins is to observe as many long, intermediate and short range intramolecular NOEs as possible. The short and intermediate range NOEs are used to define the secondary structure in conjunction with chemical shift based dihedral angles or J-couplings. Long-range NOEs are then used to define the overall protein fold and are challenging to observe in membrane proteins due to their slow tumbling times, degenerate amino acid compositions and frequent presence of exchange broadening due to detergent micelles.

Thus far I have been unable to unambiguously assign any long-range NOEs for Yop1p. I have primarily been using data from NUS ^{15}N-^{13}C-aliphatic and ^{13}C-aromatic edited 3D NOESYs collected at 950 MHz.

Owing to the r^{-6} (where r is the distance between two nuclei) dependency of the NOE, amide to amide NOEs are rarely observed in membrane proteins . This is due to the fact that helix-helix packing generally positions inter helical amides ≈10 Å apart (NOE's for distances >5 Å are difficult to observe). Therefore the hunt for long range NOEs should probably begin from the side chains.

7.1.1 State of progress with Yop1p

Currently, methyl assignments for Yop1_ΔC2 are incomplete and in some cases ambiguous since they were achieved using 3D ^{13}C and ^{15}N edited NOESY experiments (Figure 7.1).

I have produced a deuterated ^{15}N sample of Yop1_ΔC2 that was selectively protonated and ^{13}C labelled at Ileδ_1, Leu-δ and Val-γ (ILV) methyls. This labelling was achieved following the SPP protocol as described in Section 2.2 but at the stage of induction the 10 L growth was condensed into 1 L of deuterated M9 media containing 2 g of ^{12}C-^{2}H-Glucose, 0.5 g of $^{15}NH_4Cl$, 35 mg 2-ketobutyric acid-4-^{13}C,3,3-d_2 sodium salt hydrate and 60 mg 2-keto-3-methyl-^{13}C-butyric-4-^{13}C d_3 acid sodium salt (compounds d and e, respectively, from Figure 1 in Tugarinov et al. (2006)). Figure 7.2 shows an overall improvement in signal homogeneity with reduced linewidths in proton and improved signal to noise for the ILV labelled sample relative to a protonated sample. The slight overall increase in ^{13}C linewidths was unexpected but could be explained by differences in experimental set up such as shimming or the fact that the SNR of longer evolution time points would be lower giving the appearance of lower resolution.

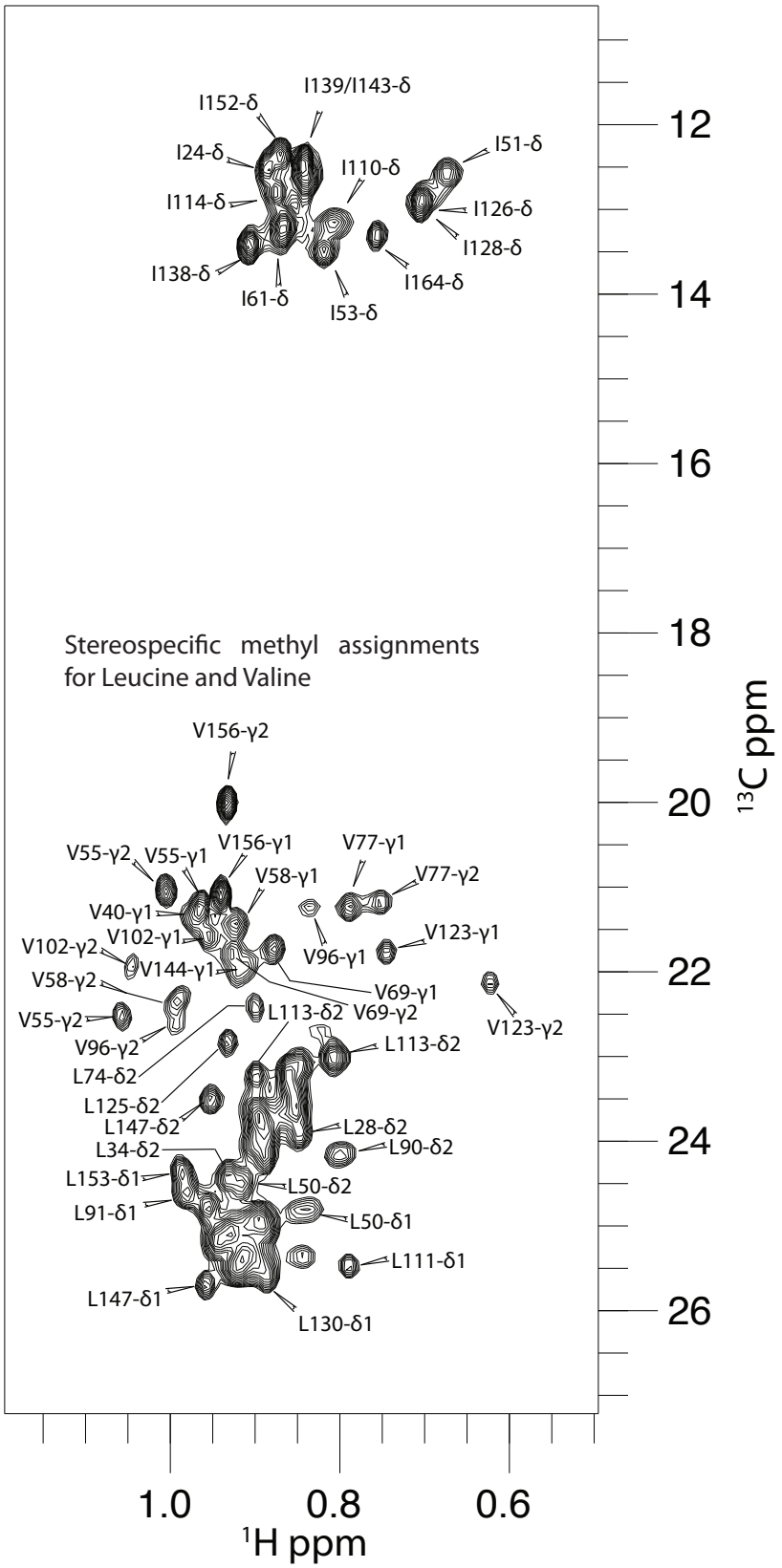
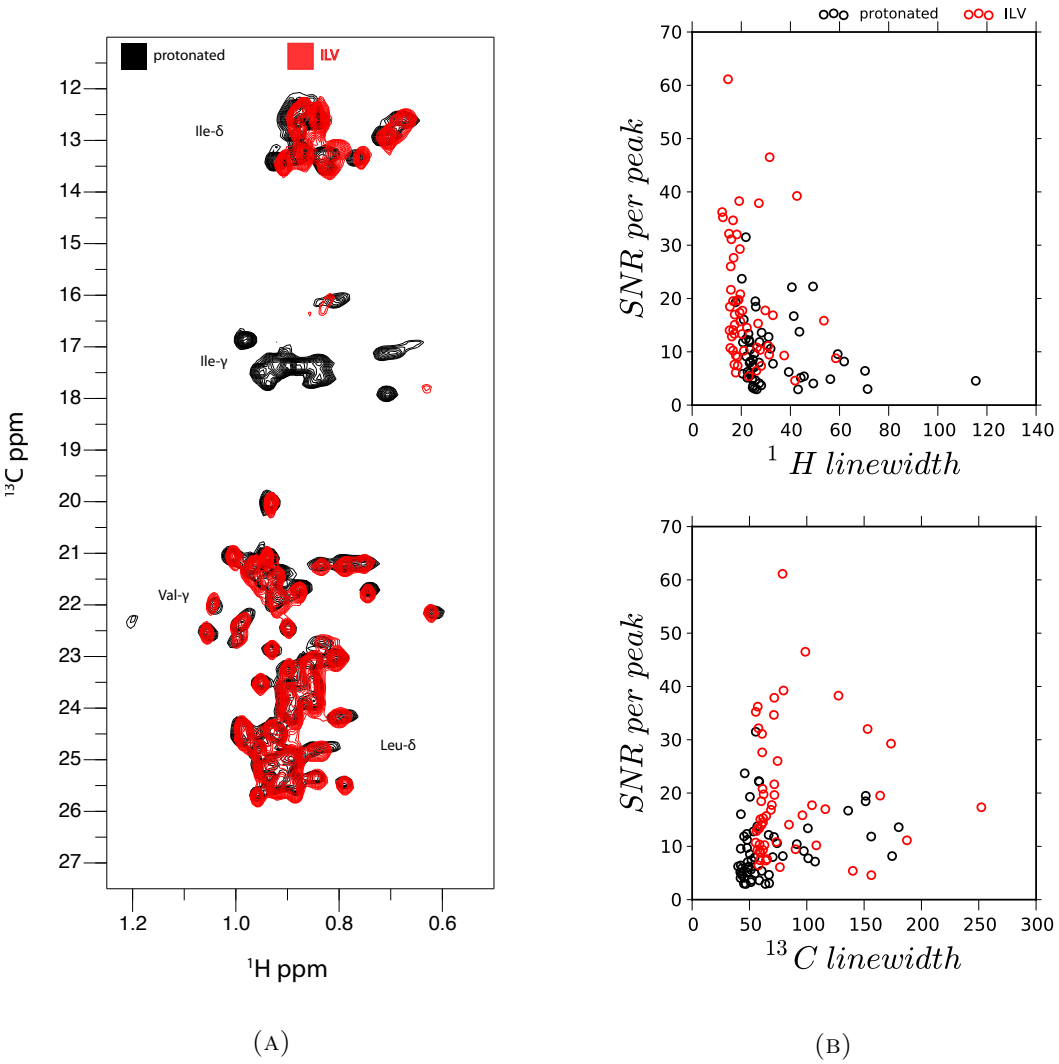


FIGURE 7.1: Stereospecific methyl assignments for ^{13}C-CT-HSQC collected at 950 MHz using 16 scans with 128 complex points in carbon and 80 μs dwell time. 1024 complex points in the direct dimension with a spectral width of 12500 Hz (~13 ppm)



(A) (B)

FIGURE 7.2: (A) Overlay of CT-^{13}C-HSQCs collected at 950 MHz at 42°C on Yop1_ΔC2 in 42 mM d_{27}-LMPG. (black) 400 μM protonated ^{13}C-^{15}N labelled sample collected using 80 scans with 128 complex points in the indirect (carbon) dimension and dwell time of 80 μs. (red) 250 μM ILV labelled sample using 16 scans with 128 complex points in carbon and 80 μs dwell time. Both spectra were collected with 1024 complex points in the direct dimension and a spectral width of 12500 Hz (~13 ppm). Both spectra were processed identically. (B) Signal to noise ratio (SNR) per peak against linewidth for ILV and protonated sample. Since the protonated sample was collected with 5 times more scans and 1.6 times higher protein concentration the values for the protonated signals were divided by 1.6 and resulting SNR was divided by $\sqrt{5}$. The peak parameters and noise values were calculated by Analysis 2.2.

The resolution and sensitivity gains of ILV labelling could be improved further by taking advantage of the methyl TROSY effect with a ^{13}C-HMQC pulse sequence such as described in Tugarinov et al. (2003). Collection of ∼50% of the points using Poisson gap NUS and sampling substantially longer evolution times would also significantly improve the resolution and signal to noise per unit time relative to the protonated sample. Unfortunately, this sample was produced before I managed to implement NUS on the 950 MHz spectrometer.

I attempted collection of long range NOEs using a uniformly sampled ^{13}C-3D-NOESY-HSQC pulse sequence without success. A caveat in the collection of such data is that the high intensity of methyl signals generally leads to truncation artefacts near the diagonal which shroud the cross peaks close to the diagonal. A simple solution to this problem would be to use NUS to sample longer evolution times and reduce truncation artefacts. Also the relatively low sample concentration (250 μM) was not sufficient for signal to noise purposes.

In the following section, I would like to discuss some of the advances in NMR labelling strategies and pulse sequences. These advances should enable observation of long-range NOEs for 3D structure determination on difficult systems such as Yop1_ΔC2 and other membrane proteins. Many of these approaches were developed primarily for large soluble protein systems >25 kDa and are likely to be highly applicable to membrane protein systems.

7.2 NMR of "big" things

The main obstacles to NMR studies of membrane proteins and large biomolecules (>25 kDa) are their poor relaxation properties and signal overlap. The primary reasons for the former are the combination of dipole-dipole (DD) interactions and chemical shift anisotropy (CSA), which as proteins become larger (or tumble slower) and at higher fields become more pronounced, respectively. One way to circumvent these problems is to either fully or selectively deuterate proteins at non-exchangeable sites. Deuteration reduces the number of observable proton signals and largely removes the effects of ^{1}H-^{1}H and ^{1}H-^{13}C/^{15}N heteronuclear relaxation pathways that cause line broadening and loss of resolving power in protonated samples. Although the gyromagnetic ratio of ^{2}H is 6.5 times smaller than that of the proton, deuterium still gives rise to broadening of signals from the heteronucleus (Browne et al., 1973, Gardner and Kay, 1998). The use of ^{2}H decoupling schemes reduces line broadening caused by the dipole-dipole coupling between the deuteron and the heteronucleus (Grzesiek et al., 1993) and further improves the resolution attainable from a deuterated protein sample. The improved relaxation

properties of the remaining protons in deuterated proteins can be exploited by NMR pulse sequences that have enabled inflation of the upper size limit for proteins studied by NMR to 100 kDa and beyond. A major break though in this regard was the discovery of transverse relaxation optimised spectroscopy (TROSY) (Pervushin et al., 1997).

TROSY relies on the fact that interference of the cross-correlated CSA and DD relaxation pathways impart disparate relaxation rates on each of the singlet peaks of a spectrum collected without heteronuclear decoupling. In a TROSY experiment the singlet peak for which DD and CSA most efficiently cancel (i.e. the slow relaxing component) is selected resulting in higher sensitivity and resolution spectra. An interesting point is that the ratio of CSA to DD is largely independent of molecular weight meaning that TROSY experiments can be applied to small and large systems. However, the choice of field strength is important; since CSA is proportional to the square of field strength, the optimal TROSY effect generally occurs with field strengths between 20-25 T (≈850 MHz to 1 GHz).

Application of TROSY to triple resonance experiments has substantially increased their sensitivity (Salzmann et al., 1998) enabling assignment of larger proteins as shown by Yang and Kay (1999a) for the 370 residue maltose binding protein (MBP) at 5°C. Here ^{13}C',^{13}Cα,^{15}N and ^{1}H chemical shifts were obtained by use of 4D HNCOCA and HN-CACO pulse sequences where the slow relaxing components of the ^{1}H-^{15}N magnetisation are selected through a modified TROSY scheme introduced by Yang and Kay (1999b). Slight modifications on these pulse sequences were also used in the assignment of the 723 residue malate synthase G (MSG) from *E. coli*, further extending the upper size limit for NMR studies of protein structure and dynamics (Tugarinov et al., 2002). Around this time Wüthrich and coworkers demonstrated the feasibility of obtaining assignments for a 121 residue chain in a 110 kDa octomeric complex (Salzmann et al., 2000). They also determined the backbone structure of the β-barrel membrane protein OmpX (Fernández et al., 2001) with a full NOE based structure coming 3 years later (Fernández et al., 2004). Ironically the NOEs for this structure were actually observed using conventional ^{15}N-edited 3D NOESY experiments, likely due to the wide dispersion of shifts and multitude of amides in close proximity to other amide and Hα protons (<5 Å) found in β-barrel membrane proteins.

To further increase the number of NMR observable probes for dissection of protein structure and function in larger systems selective labelling has been applied. One of the first examples of selective protonation strategies applied to proteins expressed in *E. coli* BL21(DE3) cells is that of Rosen et al. (1996) in which they added protonated ^{13}C-pyruvate incrementally throughout induction as the sole carbon source on a deuterated background. This enabled the production of samples enriched in ^{15}N-^{13}C-^{2}H but with

selectively protonated Ile-δ_1, Val-γ, Leu-δ (ILV) and Ala-β methyl groups. Draw backs of this method were low labelling efficiencies and the presence of methyl isotopomers ($CH_2{}^2H$ and CH^2H_2) arising from the exchange of pyruvate protons with the surrounding deuterons. These isotomopers give rise to multiple peaks per methyl group, reduce signal to noise and increase the spectral complexity. Nonetheless, methyl isotopomers such as CH^2H_2 have been used as probes for protein dynamics, an example being a study of the gating mechanism in the prokaryotic proteosome (Baldwin et al., 2010). However, the method of production of these samples involves use of α-ketobutyrate and α-isovalerate (described below). In general, unless one is studying side chain dynamics, fully protonated methyls are the preferred labelling strategy giving >2 times better signal to noise than the other isotopomers (Ollerenshaw et al., 2005).

The solution to the isotopomer problem was provided incrementally by the Kay group. Initially ^{13}C 2-ketobutyrate was used to produce selectively protonated Ileδ_1 methyls and ^{15}N-^{13}C labelled Valine to produce protonated Val and Leu methyls (Gardner and Kay, 1997). However, this method only resulted in 85% Valγ and 75% Leuδ methyl protonation. The problem of low yields was solved by Goto et al. (1999) where they showed that supplementation of $^{13/12}C^2H$-glucose containing media with the α-ketobutyrate and α-isovalerate resulted in uniform selective protonation of ILV methyls. This method is also highly flexible due the diverse set of labelling strategies available from combinations of ^{13}C and ^{2}H labelling patterns in the α-ketobutyrate and α-isovalerate precursors (Tugarinov et al., 2006). This labelling strategy is also relatively economical.

There are numerous other labelling strategies and this is an area where scientific creativity flows. Recently, selective methyl protonation of artificially introduced methionine residues was used to probe the interaction surface of a 62 kDa protein complex (Stoffregen et al., 2012). This approach was also applied to GPCRs, where a combination of natural and mutant Met methyls were used to probe conformational changes during activation of the human β_2 adrenergic receptor (Nygaard et al., 2013). This labelling was achieved by growing insect (Sf9) cells in methionine-deficient media supplemented with ^{13}C methyl-labeled methionine. A more unconventional scheme was used by Bokoch et al. (2010) where ^{13}C-methylation of lysine side chains in the human β_2 adrenergic receptor was achieved by addition of sodium cyanoborohydride followed by ^{13}C-formaldehyde . The samples were then used to probe dynamics of ligand binding on the extracellular surface (ECS) loops, an often poorly resolved region in the crystal structures. Kalverda et al. (2014) recently demonstrated the feasibility of studying ligand binding to a 12 TM α-helical sugar transporter from *E. coli*, GalP, by TROSY NMR methods. They first used an *E. coli* strain auxotrophic for Tryptophan to introduce ^{2}H-^{15}N-Trp into a ^{2}H-^{14}N background and also used ILV labelling to observe chemical shift

changes upon addition of an inhibitor. This was done using the 'preservation of equivalent pathways' (PEP) HMQC experiment (Guo et al., 2008) to reduce interference from detergent methylene groups.

These studies highlight one of the strengths of NMR in bringing the "static" crystal structures to life and providing insights into the mysterious conformational changes that these proteins undergo upon ligand binding.

A frequently cited argument for the utility of ILV labelling schemes in soluble proteins is a paper by Janin et al. (1988) who demonstrated a strong bias for ILV amino acid burial within the hydrophobic core of proteins. This makes ILV methyls prime candidates for observation of structural restraints in soluble proteins. Whether this remains true for membrane proteins is not clear. Recent work by Kim et al. (2014) suggested that polar-polar interactions are more important than hydrophobics for defining membrane protein folds. Work on the GPCR, sensory rhodopsin II (pSRII) from *Natronomonas pharaonis*, would suggest that NOE restraints derived from aromatic side chains (Phe, Tyr and Trp) are of crucial importance for the definition of membrane protein folds (Gautier et al., 2010). However, in my hands standard 3D NOESY experiments for aromatics are not able to sufficiently resolve signals due to strong coupling between the aromatic ring carbons. The indirect proton dimension is also badly affected by the resultant peak splitting.

Implementation of a recently developed 4D aliphatic-aromatic NOESY experiment could provide unambiguous distance restraints between aromatic and aliphatic protons (Stanek et al., 2013). There are now methods for production of phenyl-α-ketoacid precursors for phenylalanine and tyrosine which can be used to introduce protons and ^{13}C at specific locations on otherwise deuterated ^{12}C aromatic rings (Lichtenecker, 2014, Lichtenecker et al., 2013). These precursors (an example of which is shown in Figure 7.3) could be used in conjunction with ILV labelling strategies and potentially facilitate structure determination of membrane proteins. One caveat is that the precursors have to be synthesised since they are not currently commercially available. However, Lichtenecker (2014) provides detailed protocols for precursor synthesis.

Alanine methyls provide another potentially fruitful NMR probe. The first efficient method for production and incorporation of methyl protonated alanine on a deuterated background was provided by Isaacson et al. (2007). This protocol was later improved upon by enabling production of methyl protonated alanine using traditional M9-media supplemented with labelled alanine (produced following Isaacson's method) and deuterated metabolic precursors (succinate-d_4, α-keto-isovalerate-d_7 and L-isoleucine-d_{10}) to suppress the scrambling reactions that lead to off target incorporation of protons (Ayala et al., 2009). This progress was finally tied together by Godoy-Ruiz et al. (2010)

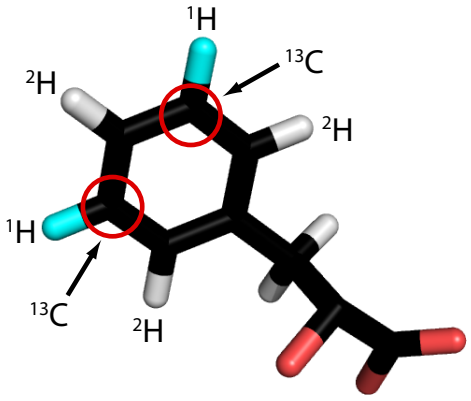


FIGURE 7.3: Example phenyl-α-ketoacid precursor for production of selectively protonated and ^{13}C labelled phenylalanine (Lichtenecker et al., 2013). Positions of ^{13}C nuclei are circled in red. Oxygens are red, deuterons grey and protons cyan.

with their demonstration of Ala-β-ILV methyl protonated samples by replacing the α-ketoisovalerate-d_7 and L-isoleucine-d_{10} with the α-keto-isovalerate and α-ketoisobutyrate precursors required for the desired ILV labelling pattern. Ala-β methyl protonation has also proven useful for membrane protein structure determination, for example in a recent and controversial study of p7 from Hepatitis C virus (OuYang et al., 2013).

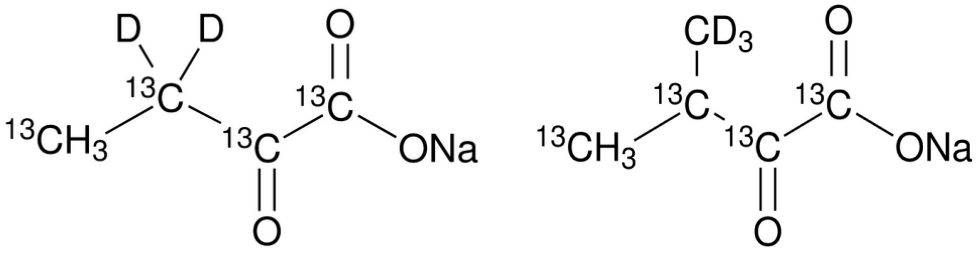
The methods discussed above have advantages over labelling schemes such as stereo array isotope labelling (SAIL) (Kainosho et al., 2006) which require use of cell free expression systems and are extremely expensive.

7.2.1 Deuteration of lipids and detergent

NMR experiments that channel magnetisation through protons will be affected by high concentrations of protons due to the presence of detergent. In Hagn et al. (2013) they state that the deuteration of lipids is highly important for NMR experiments involving transfer of magnetisation through side chains, for example HNCACB type triple resonance experiments. In my hands this is certainly true and collection of ^{13}C-NOESY experiments is not possible in the presence of protonated lipids or detergent since most of the magnetisation is transferred through the acyl chain protons. Impressively however, the pSRII structure from Gautier et al. (2010) was solved in the presence of protonated 1,2-diheptanoyl-sn-glycero-3-phosphocholine (D(C_7)HPC) since the deuterated form was commercially unavailable at the time. On the other hand, collection of useful spectra on the M2 proton channel in protonated 1,2-dihexanoyl-sn-glycero-3-phosphocholine (D(C_6)HPC) was not possible and the deuterated form had to be used (personal communication with Prof. Jason Schnell).

7.3 Ideal assignment and structure determination strategy

Firstly, production of a deuterated $U^{13}C$-$U^{15}N$ sample with selective protonation at Ileδ_1, Leu-δ and Val-γ using the precursors in Figure 7.4 followed by collection of the COSY-based experiments described by Tugarinov and Kay (2003) should enable unambiguous assignment of the methyl groups.



(A) Ileδ_1 (B) Valγ and Leuδ

FIGURE 7.4: ILV labelling is achieved by supplementation of deuterated ^{13}C-^{2}H-Glucose and $^{15}NH_4Cl$ containing induction media with (A) 2-Ketobutyric acid-$^{13}C_4$,3,3-d_2 for selective protonation of Ileδ_1 methyls. (B) 2-Keto-3-(methyl-d_3)-butyric acid-1,2,3,4-$^{13}C_4$ can be used to selectively protonate 1 of each of the methyl groups of Val and Leu side chains. See Tugarinov et al. (2006).

Following unambiguous assignment this sample could be used to attempt collection of methyl-methyl NOE restraints. However, production of another ILV labelled sample where only the methyls are ^{13}C labelled would be preferable to avoid peak splittings due to coupled ^{13}C nuclei.

Alternatively, 4D methyl-methyl TROSY-NOESY experiments in conjunction with NUS (Tugarinov et al., 2005b) solves the issues associated with intense diagonal peaks and also reduces the ambiguity of NOE assignment with the addition of another dimension. This experiment has been used to solve the backbone structure of the 723-residue (82-kDa) enzyme malate synthase G (Tugarinov et al., 2005a) and also by the Wagner group in conjunction with Poisson gap sampling and hmsIST (Hyberts et al., 2012). The 4D

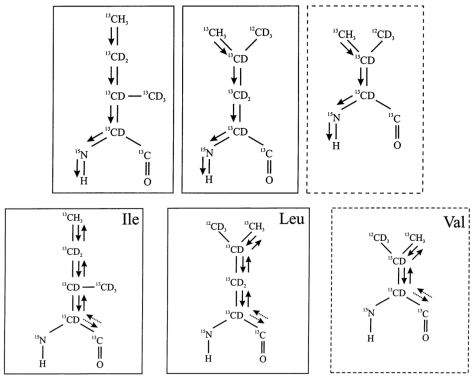


FIGURE 7.5: The pulse sequences developed by Tugarinov and Kay (2003) should be used to assign the ILV side chains. The valine schematic is surrounded by a dotted line since some pulses are omitted for collection of valine methyl chemical shifts. The images were taken from Tugarinov and Kay (2003).

aromatic-aliphatic NOESY (Stanek et al., 2013) discussed in Section 7.2 could also be used in conjunction with the novel selective aromatic labelling (Lichtenecker et al., 2013) to collect aromatic-methyl restraints.

In my view, the advent of user friendly NUS methodologies and fast reconstruction algorithms (such as hmsIST) will make collection of high resolution 4D spectra a routine process in the future. Thus, NMR will be successfully applied to structural studies of large biomolecules on a regular basis.

Observation of as many long range NOEs as possible from the experiments listed above should be the primary objective for a future structure determination of Yop1p. These initial NOE based models could then be refined with PRE and RDC data (as discussed below).

It is worth noting that the experiments and methodologies discussed above were primarily developed on systems with tumbling times in the order ≈30-400 ns which is much larger than Yop1_ΔC2 (τ_c ≈12 ns in bicelles, see Figure 5.15).

7.4 Paramagnetic relaxation enhancements

PREs provide long range distance restraints up to 25 Å. One of their attractions is that the data can be collected relatively rapidly in the form of 2D experiments and their and downstream analysis is greatly simplified.

Currently, my best attempt at PRE collection is displayed in Figure 5.11. Significant trends were observed but there was also a lot of scatter which hindered attempts to use the data as structural restraints. The fact that relaxation enhancements occur at both the N- and C-termini of the APH region suggests that oligomeric behaviour may be contributing to the observed PRE which, in turn, makes the data highly complicated to analyse. The apparently raised base line also suggests that there may be residual MTSL associated with the LMPG micelles.

Table 7.1 is a list of mutants I cloned and produced in order to perform PRE studies on Yop1_ΔC2. Labelling positions were chosen to be near the ends of TMs according to my chemical shift based secondary structure data. Chen et al. (2011a) proposed an optimal strategy for spin labelling of membrane proteins suggesting that data from as few as 1-2 spin labelled sites should be sufficient to define the topology and helix-helix packing of 4-7 TM helix bundles. However, this relies on having a good starting model to predict the optimal sites which is often lacking for membrane proteins. Gottstein et al. (2012) recommended that at least one spin label should be placed on each TM helix and

TABLE 7.1: Mutants for PRE labelling. Ticks and crosses signify whether or not the mutants having cloned, expressed or purified.

Mutant	expressed	purified	labelled
S37C	✔	✔	✔
V55C	✔	✔	✖
V58C	✔	✔	✖
S84C	✔	✔	✖
S107C	✔	✔	✖
T148C	✖	✖	✖

that they should be distributed evenly throughout the sequence. In my view this is a more sensible approach for defining an accurate structure: the more restraints that it is possible to obtain the better. Berardi et al. (2011) demonstrated the power of combining PREs with RDCs and the molecular fragment replacement (MFR) method (Delaglio et al., 2000, Kontaxis et al., 2005) in their determination of the backbone structure of a mitochondrial uncoupling protein, UCP2, a six TM protein (Berardi et al., 2011). The authors used PRE restraints from four separate labels (collected using TROSY-HNCO experiments) to guide the assembly of fragments selected using the RDCs with MFR.

An example of the dangers involved in such heavy reliance on PRE restraints is the structure of diacylglycerol kinase (DAGK) (Van Horn et al., 2009). This structure was initially solved using only PRE and RDC restraints to define the global fold. The later published crystal structure from lipidic cubic phase (Li et al., 2013) demonstrated fairly convincingly that the domain swapping observed between monomers in the NMR structure most likely represents a mis-interpretation of the NMR data. It would therefore be interesting to see a rigorous reanalysis of the level of agreement between the NMR data and the crystal structure. Notably the helices in the crystal structure are generally straighter with fewer distortions than the solution structure. This could reflect a lack of restraints in the NMR structure or the effects of solubilisation by detergent systems. The intensities of short to medium range NOE peaks used to define the secondary structure could also represent a significant source of error. In my view NMR based membrane protein structures would benefit from refinement against the Rosetta membrane force-field (Yarov-Yarovoy et al., 2005) to impose more optimal packing and geometries on the system.

Use of PRE experiments for structure determination of Yop1p is currently fraught with difficulties including removal of free spin label and the complications arising from oligomerisation (even if transient). It would be preferable to pursue a primarily NOE based structure determination introducing the longer range and more ambiguous PRE restraints ($\leq$25 Å) for later refinement of the global fold.

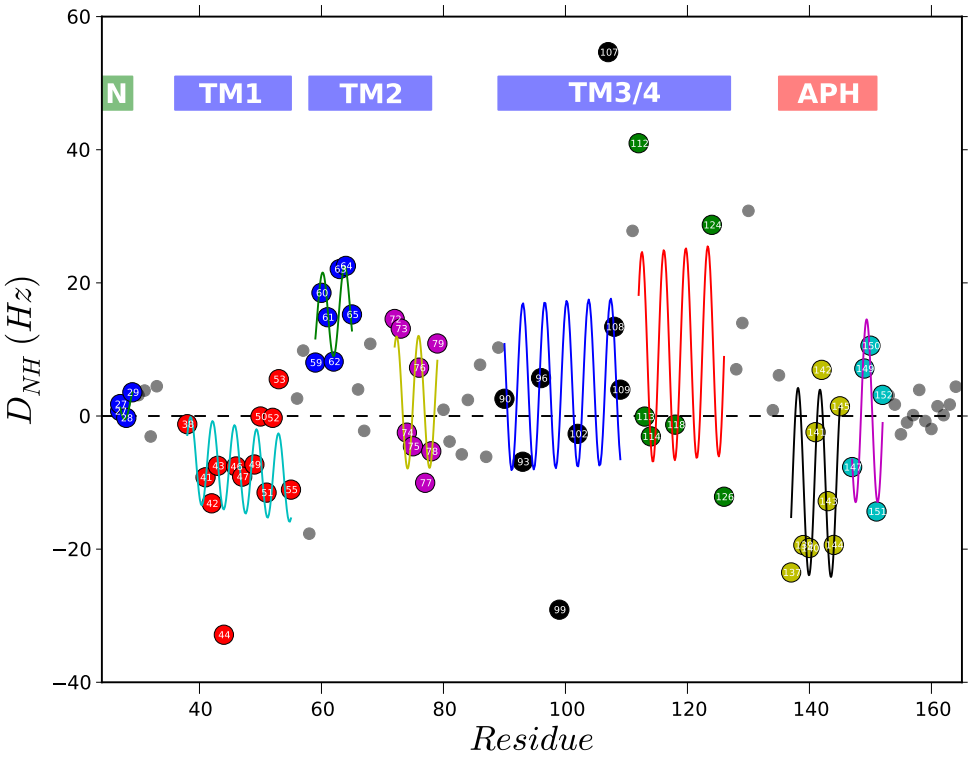


FIGURE 7.6: RDCs for Yop1_ΔC2 soaked into a 4.5% stretched polyacrylamide gel (Section 2.14.6). Fitting to sine waves with a period of 3.6 was attempted using a python script that allows selection of specific regions. The residues used for each fit are colour coded and numbered. Residues omitted from fitting are shown in grey. The choice of fitted points was based on the TALOS-N based secondary structure and visual inspection. If a proline was present within a well defined helix, fits were split either side of the proline position. e.g. at P71 and P146.

7.5 Residual dipolar couplings

RDCs can provide global bond orientation restraints that are conventionally useful for refinement of protein structures. They are observed by calculating the difference in chemical shifts between an anisotropically tumbling sample (one that is weakly aligned to the magnetic field due to the presence of an external medium) and one that is tumbling isotropically.

I have successfully soaked Yop1_ΔC2 into polyacrylamide gels using the method described by Chou et al. (2001) and in Section 2.14.6. The differences in ^{15}N frequencies between a decoupled HSQC and TROSY pair were calculated as described by Kontaxis et al. (2000). However, I did not use the suggested temperature compensation to account for heating due to the decoupling scheme in the HSQC.

The data that I collected did not appear particularly reliable, and fitting of sine waves to the α-helical regions was also largely unsuccessful (Mesleh and Opella, 2003, Mesleh et al., 2003) (Figure 7.6).

For future reference, once soaked into a gel the Yop1_ΔC2 samples tumble too slowly for observation of many peaks by HSQC methods. Therefore the use of TROSY-based methods such as the ARTSY pulse sequence (Fitzkee and Bax, 2010) and also NUS could improve the accuracy of chemical shift measurements. TROSY-based HNCO pulse sequences for measurement of ^{15}N-^{13}CO, ^{1}HN-^{13}CO, ^{13}CO-^{13}Cα and ^{1}HN-^{13}Cα dipolar couplings could also be extremely useful (Yang et al., 1999) since they reduce signal overlap and orthogonal data sets can be used to define bond angles more unambiguously.

Other methods for collection of RDCs such as charged polyacrylamide gels (Cierpicki and Bushweller, 2004, Meier et al., 2002) and G-tetrad DNA (Lorieau et al., 2008) could enable collection of orthogonal sets of RDC. Also, the labelling of proteins with lanthanide tags has also been demonstrated as a promising method for collection of RDCs (Kamen et al., 2007). Furthermore, this would have the added potential for collection of pseudo contact shifts which can also be used to infer structural information. Addition of filamentous phage (Hansen et al., 1998) was also attempted but they were not compatible with LMPG. No splitting of the deuterium signals was observed and the filaments had clearly been denatured by the presence of detergent.

A reliable RDC data set could be used in future to refine or help solve the structure of Yop1p using MFR or CS-Rosetta in combination with Xplor-NIH (Schwieters et al., 2003).

7.6 CS-Rosetta

Rosetta is a knowledge based program for *ab initio* protein structure prediction (Simons et al., 1997, 1999). It uses secondary structure information from sequence, chemical shifts (Bowers et al., 2000, Shen et al., 2008) or proteins of known structure to select a biased set of 3-mer and 9-mer fragments which are then used to build structural models using a Monte Carlo simulated annealing protocol. One of the strengths of Rosetta is its forcefields which combine sequence specific terms such as hydrophobic burial, electrostatics and disulphide bonds with general terms for packing and assembly of helices and beta sheets (Simons et al., 1999). A gyration radius term also ensures that the resulting structures are realistic and compact. The fragment based approach also improves the computational efficiency of the structure calculations.

Membrane protein NMR spectra are notoriously bad in terms of signal to noise and resonance overlap. This makes unambiguous assignment and observation of long-range NOEs very challenging. The introduction of a membrane forcefield into Rosetta (Yarov-Yarovoy et al., 2005) in combination with chemical shift based fragment selection and structure selection has been presented as a promising approach to structure determination of membrane proteins (Barth et al., 2007, 2009). However, without specific constraints such as NOEs, PREs or RDCs the program can only reliably be applied to proteins <15 kDa in size. This places Yop1p near the upper limit.

3-mer and 9-mer fragments were generated using CS-rosetta (Shen et al., 2008) which selects fragments based on chemical shift derived secondary structure from TALOS (Figure 5.2). The fragment library was then used to generate 10,000 centroid based structures using Rosetta-membrane *ab initio* protocol (Yarov-Yarovoy et al., 2005). The 1000 lowest energy structures were then relaxed in the Rosetta-membrane forcefield and the resulting structures were clustered by backbone RMSD to the lowest energy structure. Figure 7.8 shows the lowest RMSD models overlayed using PyMol (Schrödinger, LLC, 2010) in combination with an MDAnalysis script (Michaud-Agrawal et al., 2011) for aligning the structures based on specific secondary structure elements.

From the cluster analysis of Figure 7.7 it can be seen that Rosetta-membrane did not converge on a specific fold for Yop1p with most of the structures having RMSD > 10 Å from the lowest energy structure. This could be a reflection of Yop1p being too large to efficiently sample the folding space for convergence on a set of low RMSD low energy structures.

Recently, an implicit membrane forcefield based on knowledge of the types of interactions important in membrane protein folding (Kim et al., 2014) has been published. Such forcefields will useful in future for refining membrane protein structures *in silico* and represent an advance from the general status quo for protein structure refinement in either water or vacuum.

The Rosetta calculations presented here were performed using the SBGrid Consortium computing network and the script used to run the Rosetta program is listed below.

```
# Script to run Rosetta with Membrane forcefield
${rosetta3}\ # Path to rosetta executable.
-in::file::fasta ./t000_.fasta\
-in::file::spanfile ./m138i.span\ # TM prediction from Octopus
-in::file::lipofile ./m138.lips4\ # Lipophilicity parameters
-in::file::frag3 ./aat000_03.200_R3\ # 3-mer fragments
-in::file::frag9 ./aat000_09.200_R3\ # 9-mer fragments
-in::path::database ${rosetta3DB}\ # Rosetta database
-abinitio::membrane\ # Choosing membrane forcefield
-score::find_neighbors_3dgrid\ # Scoring structures
-membrane::no_interpolate_Mpair\
-membrane::Menv_penalties\
-membrane::normal_cycles 40 \
-membrane::normal_mag 15\
-membrane::center_mag 2\
-nstruct 10000\ # Number of structures to generate
-out::file::silent ./m138i_silent.out\
-mute all\
```

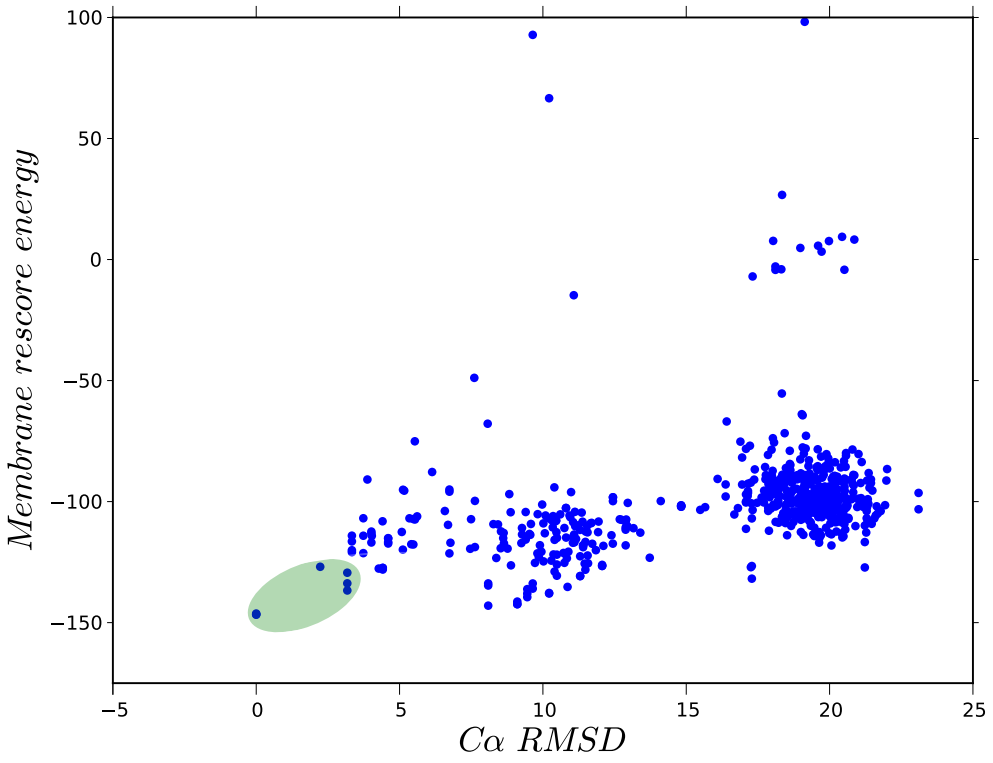


FIGURE 7.7: 1000 lowest scoring decoys clustered by backbone RMSD (Å) to the lowest energy model. The models used to produce Figure 7.8 are shaded in green.

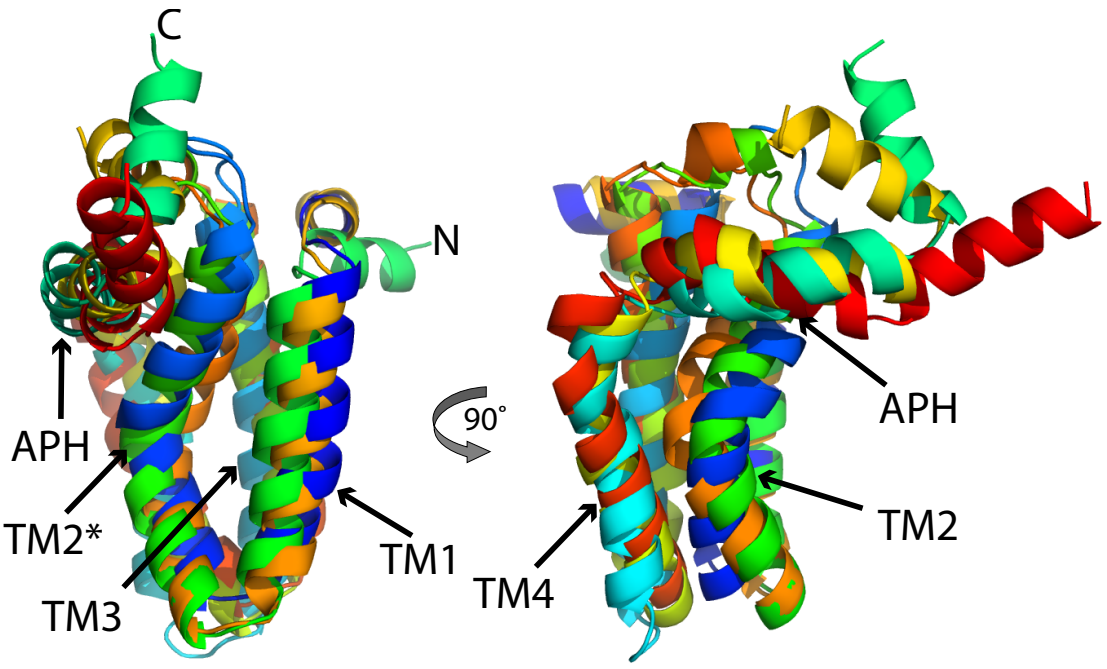
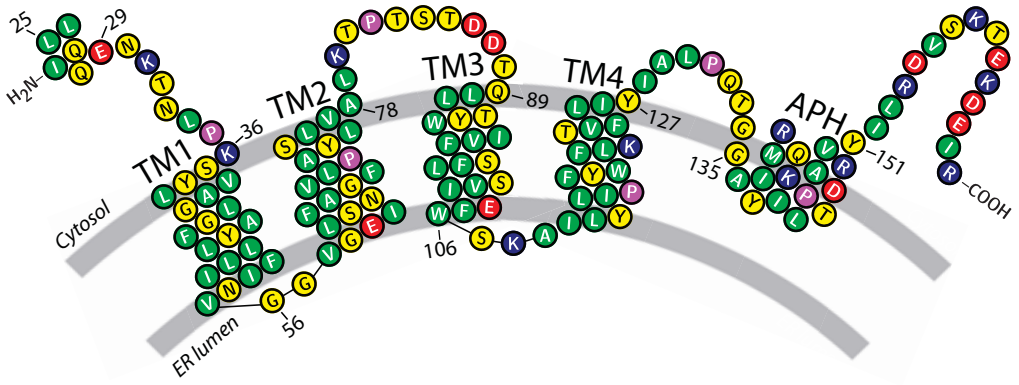


FIGURE 7.8: Rendered lowest score/RMSD models from shaded region of Figure 7.7. Secondary structure elements are labelled. The kink in TM2 (starred) can be seen with the APH packed against the C-terminal end of TM2. Rosetta predicts TM3 to end at Y112 and TM4 to begin at F116.

8

Conclusion



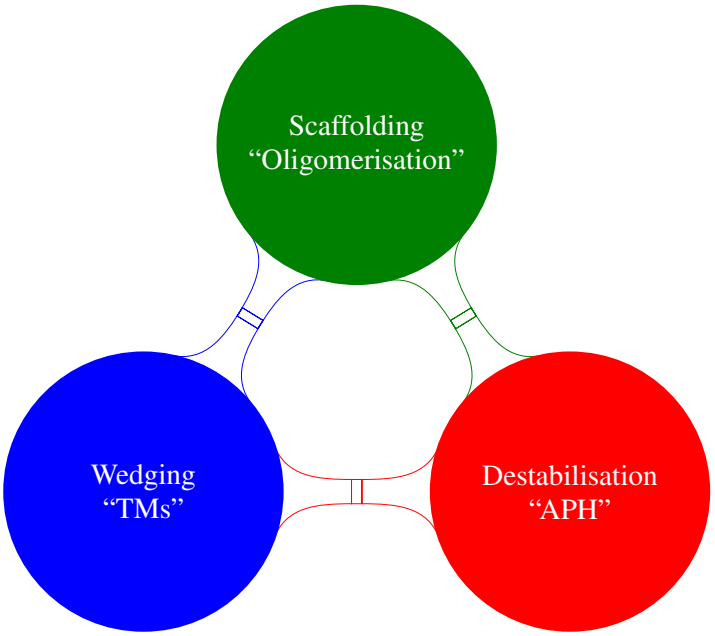


FIGURE 8.1: Schematic of key processes likely to be involved in membrane curvature by RHD proteins.

The prevailing idea for the mechanism of membrane curvature stabilisation by RHD proteins in the ER has been that their short TM regions insert into the lipid bilayer in concert with oligomerisation to both wedge and scaffold the membrane into a curved architecture (Hu et al., 2008, Voeltz et al., 2006).

The most important finding of this thesis is that membrane curvature in the ER likely requires a third component: amphipathic helices (Figure 8.1). I have addressed the role of the APH in this mechanism by showing that it is required for tubule formation *in vitro*. The likely conservation of the APH region amongst other RHD proteins also suggests that it may have a generally important role to play. However, in order to fully understand the underlying mechanisms of membrane curvature by RHD proteins, efforts must be made to solve a representative structure. To begin with, a monomeric structure would be useful in order to model the effects of TM insertion on bilayer curvature. An altogether more challenging aspect to understand is the role of oligomerisation and how it is mediated. It would be very interesting to determine whether the scaffolds assume curved architectures and also probe their dynamics and the origins of their polydispersity.

Recent advances with regard to labelling strategies and pulse sequence development make NMR a well placed technique for tackling these questions. I have also developed protocols for the study of Yop1p (a representative RHD protein) which, in turn, should facilitate further studies.

A

Scripts

A.1 Calculation of τ_c from ^{15}N-TRACT data

```
''' Equations taken from Lee et al. (2006)  '''
import numpy as np
''' Constants  '''
uo    = 1.2566370614e-6 # Permeability of free space (m kg s^-2 A^-2)
gH    = 267.513e6       # Proton gyromagnetic ratio (rad s^-1 T^-1)
gN    = -27.126e6       # Nitrogen gyromagnetic ratio (rad s^-1 T^-1)
planck = 6.62606957e-34 # Planck's constant (m^2kg/s)
r_hn  = 1.02e-10        # HN internuclear distance (m)
ddN   = 160.0e-6        # ppm
''' Field values '''
mg_rat  = -9.868831583   # Ratio between H and N
h_field = float(raw_input('What is the field in Hz? '))
n_field = h_field/mg_rat # this is in Hz
omega_n   = n_field*(2.0*np.pi) # this is in radians per second
print 'omega_n = %f'%omega_n
B0      = (h_field*2.0*np.pi)/gH
print 'B0 = %f'%B0
''' Value of 2eta_xy i.e. R-beta - R-alpha '''
two_eta_xy = float(raw_input('What is the value of Rb-Ra in Hz? '))
theta     = 17.0
theta_rad = np.deg2rad(theta)
print 'theta_rad = %f'%theta_rad
''' Equation 4 '''
rho = (uo*gH*gN*planck)/(16.0*np.square(np.pi)*np.sqrt(2.0)*np.power(r_hn,3))
print 'rho = %f'%rho
''' Equation 5 '''
delta_n = (gN*B0*ddN)/(3.0*np.sqrt(2.0))
print 'delta = %f'%delta_n
cos_term = (3.0*(np.square(np.cos(theta_rad)))-1)*(2.0*rho*delta_n)
x = two_eta_xy/(3.0*(np.square(np.cos(theta_rad)))-1)*(2.0*rho*delta_n)
print 'x = %f'%x
''' Constants for cubic  '''
a   = (cos_term *1.6 *np.square(omega_n))
b   = (-two_eta_xy*np.square(omega_n))
c   = (cos_term*2.8)
d   = (-two_eta_xy)
'''Solving cubic equation '''
coefficients=[a,b,c,d]
tauc = np.roots(coefficients)
print 'The tumbling time is %.3f ns'%(np.real(tauc)[0] * 1e9)
```

A.2 Corrected standard deviation

```
import numpy as np
import sys as s
import numpy.ma as ma

def SDCorr(data,num_of_SDs=3):
  sd = np.std(data)
  avg = np.mean(data)

  v1 = num_of_SDs
  v2 = (sd*v1)+avg
  data = np.array(data)

  original_data = data

  print "s.d =", sd
  print "mean =", avg
  print "Cutoff =", v2
  while np.any(data[:] >= v2):
    mask = ma.less_equal(data, v2)
    print data
    print mask
    data = data[mask]
    sd = np.std(data)
    avg = np.mean(data)
    v2 = (sd*v1)+avg
    print 'Iterating'
    print data

  print 'Complete'
  print data
  corrSD = (v2/v1)*2+avg
  ''' Test to check whether any data was actually excluded  '''
  try:
    original_data - data
    print 'You do not have any values over 3 times the standard deviation, use SD instead'
    print 'SD = %8.3f'% sd
    return sd
  except:
    print 'The Corrected Standard Deviation is ...'
    print corrSD
    return corrSD
```

B

Chemical Shift Data

B.1 Micelle

Residue	**AA**	H_N	N_H	$H\alpha$	$C\alpha$	$C\beta$	C'
24	I	-	-	3.929	-	-	-
25	L	9.188	111.245	4.14	56.57	-	177.523
26	Q	8.757	119.752	4.068	58.343	28.318	177.466
27	Q	8.444	117.551	4.148	57.946	28.018	177.192
28	L	7.791	118.464	4.187	56.036	40.688	177.99
29	E	8.107	119.372	4.066	58.204	29.488	176.978
30	N	8.131	116.142	4.666	53.905	38.469	175.928
31	K	7.93	118.498	4.377	56.214	32.603	176.233
32	T	7.732	109.565	4.338	61.915	69.394	173.786
33	N	8.144	119.041	4.705	53.169	38.328	174.587
34	L	7.968	121.494	4.518	53.791	41.003	174.605
35	P	-	-	-	63.443	-	177.359
36	K	8.515	122.541	3.966	59.03	-	177.988
37	S	8.326	112.87	4.133	60.784	62.382	176.461
38	Y	7.717	121.14	4.397	59.957	37.867	177.4
39	L	7.606	119.01	4.079	57.916	39.016	178.233
40	V	8.043	116.812	3.624	66.002	31.305	178.177
41	A	7.809	122.068	4.128	54.686	18.307	179.749
42	G	8.401	105.271	3.682	47.318	-	174.768
43	L	8.317	121.208	4.093	57.901	31.353	178.601
44	G	8.073	105.315	4.447	46.962	-	175.55
45	F	8.15	120.334	4.193	60.322	-	176.672
46	A	8.254	120.682	3.885	55.4	17.788	178.525
47	Y	8.184	115.737	3.702	60.772	-	177.319
48	L	8.382	118.104	3.905	-	-	-
49	L	7.722	117.676	-	57.501	-	-
50	L	7.787	117.303	3.996	57.416	-	179.087
51	I	7.894	117.91	3.747	62.721	-	177.914
52	F	7.969	118.747	3.97	60.251	-	177.074
53	I	8.04	115.694	3.891	62.687	-	176.228
54	N	7.904	118.612	4.649	57.416	-	176.164
55	V	8.039	118.761	4.054	63.299	-	176.849
56	G	8.225	108.869	-	45.745	19.426	174.936
57	G	8.051	108.118	3.918	45.835	-	175.54

Continued on next page...

Residue	**AA**	H_N	N_H	$H\alpha$	$C\alpha$	$C\beta$	C'
58	V	8.064	119.013	3.814	65.412	31.305	177.173
59	G	8.551	107.812	-	47.218	-	175.994
60	E	7.869	121.185	4.137	58.663	29.015	178.712
61	I	7.673	119.488	3.835	64.408	37.718	178.139
62	L	8.218	118.464	4.117	58.663	-	178.874
63	S	8.201	113.407	4.167	61.326	-	176.81
64	N	8.018	119.849	4.593	55.511	-	177.105
65	F	8.305	119.291	-	60.16	-	174.655
66	A	8.433	121.251	3.968	54.941	18.353	178.693
67	G	8.019	102.562	-	45.946	-	174.613
68	F	7.686	117.491	4.644	58.895	40.189	176.134
69	V	7.69	116.06	3.822	64.426	32.45	176.134
70	L	8.149	120.411	4.329	58.622	-	175.961
71	P	-	-	-	65.886	-	178.894
72	A	7.421	119.449	4.1	54.275	17.941	179.022
73	Y	8.152	117.978	4.133	60.745	-	176.699
74	L	8.391	117.453	3.926	57.181	-	179.34
75	S	7.804	114.209	4.237	61.153	62.911	175.41
76	L	7.725	121.71	4.121	57.146	41.84	178.443
77	V	7.764	114.838	3.718	64.515	-	177.04
78	A	7.447	121.483	4.212	53.043	18.476	178.262
79	L	7.585	118.096	4.214	55.67	41.84	177.141
80	K	7.734	118.272	4.35	55.428	32.755	175.886
81	T	7.849	115.509	4.189	59.913	68.758	173.066
82	P	-	-	4.49	63.314	-	176.818
83	T	8.176	112.295	4.373	61.496	69.51	174.532
84	S	8.188	116.715	3.915	57.994	63.979	174.895
85	T	8.212	114.741	4.337	62.217	68.83	174.444
86	D	8.178	121.221	4.337	54.521	-	176.266
87	D	8.107	119.632	4.619	54.971	-	176.742
88	T	8.054	113.501	4.134	63.606	68.896	175.554
89	Q	8.277	122.568	4.263	57.189	28.557	176.644
90	L	8.008	121.705	4.193	56.389	-	177.923
91	L	8.088	119.319	4.26	57.55	-	177.843
92	T	7.858	111.523	4.19	64.871	68.702	175.594
93	Y	7.909	120.288	4.152	60.481	-	177.053
94	W	8.01	118.982	4.446	59.974	30.538	179.021

Continued on next page...

Residue	**AA**	H_N	N_H	$H\alpha$	$C\alpha$	$C\beta$	C'
95	I	8.192	121.154	-	64.127	-	176.977
96	V	7.841	119.135	4.168	66.283	-	177.013
97	F	8.34	118.291	-	60.534	-	178.52
98	S	8.206	116.711	4.366	60.85	-	176.4
99	F	7.662	118.461	4.182	60.643	-	176.798
100	L	8.307	117.548	4.162	57.195	-	179.193
101	S	7.837	114.218	4.253	61.249	-	175.836
102	V	7.8	120.78	3.994	64.664	31.305	178.246
103	I	7.661	118.968	3.937	63.616	-	177.943
104	E	8.115	120.221	4.585	57.424	28.839	178.143
105	F	7.937	118.501	4.374	59.712	-	177.084
106	W	8.109	120.393	3.946	56.916	-	176.908
107	S	8.046	112.463	4.348	61.563	-	175.648
108	K	7.434	119.613	4.116	58.01	32.374	177.829
109	A	8.071	120.604	4.035	54.74	18.122	179.269
110	I	7.847	114.186	3.729	63.398	-	176.642
111	L	7.224	118.194	3.993	57.171	40.92	178.783
112	Y	7.597	115.795	4.349	59.115	38.328	176.194
113	L	7.765	116.057	4.377	55.021	-	177.807
114	I	7.729	118.936	3.958	65.725	-	177.206
115	P	-	-	-	63.832	-	177.386
116	F	7.942	118.985	4.353	60.006	-	176.83
117	Y	8.386	120.346	4.113	61.012	-	176.693
118	W	8.168	117.787	4.134	60.007	28.352	-
120	L	8.126	112.981	3.996	60.528	-	-
121	K	8.239	117.987	-	59.324	31.801	-
122	T	7.653	114.391	4.069	66.435	68.255	175.969
123	V	7.891	119.003	3.603	65.793	30.443	177.156
124	F	8.262	119.768	4.162	61.715	-	176.284
125	L	8.027	116.736	3.976	57.311	-	179.339
126	I	7.671	118.336	3.718	63.846	-	-
127	Y	7.804	118.873	3.997	56.169	-	-
128	I	7.892	114.38	3.829	61.857	-	175.644
129	A	7.333	122.01	4.307	51.773	19.163	177.822
130	L	7.586	120.455	4.389	54.348	-	175.68
131	P	-	-	-	63.439	-	177.058
132	Q	8.263	117.573	4.39	56.572	-	176.457

Continued on next page...

Residue	**AA**	H_N	N_H	$H\alpha$	$C\alpha$	$C\beta$	C'
133	T	7.903	111.596	4.361	61.837	69.889	175.381
134	G	8.478	109.954	-	46.091	-	175.767
135	G	8.46	108.16	3.833	46.593	-	174.732
136	A	8.456	122.363	4.051	54.9	18.094	178.775
137	R	8.004	116.45	4.074	59.219	29.107	178.295
138	I	7.665	119.032	3.944	63.527	-	177.769
139	I	7.813	119.441	3.681	64.629	-	178.357
140	Y	8.405	120.211	4.057	62.103	-	176.545
141	Q	8.427	116.153	3.849	58.706	28.706	177.706
142	K	8.259	113.636	-	57.362	32.984	176.704
143	I	7.65	113.378	4.317	61.815	38.703	175.193
144	V	7.602	116.865	4.035	66.124	-	177.415
145	A	8.453	122.968	4.133	55.752	15.615	176.507
146	P	-	-	-	65.598	-	179.262
147	L	7.359	115.671	4.263	57.166	41.993	178.892
148	T	8.297	113.208	4.316	65.995	68.438	176.801
149	D	8.333	121.885	4.426	56.84	-	177.112
150	R	7.446	115.748	4.03	57.621	30.694	177.134
151	Y	7.832	114.574	-	59.395	39.703	175.824
152	I	7.869	117.77	4.193	61.683	38.325	176.007
153	L	7.998	121.516	4.349	54.86	41.153	177.034
154	R	7.74	119.37	4.229	55.994	30.312	175.913
155	D	8.31	120.515	4.649	53.768	40.749	176.566
156	V	7.996	119.737	4.149	62.058	31.992	176.463
157	S	8.307	118.224	4.396	58.569	63.443	174.776
158	K	8.111	122.74	4.373	56.04	32.526	176.725
159	T	8.019	114.429	4.32	61.665	69.3	174.645
160	E	8.323	123.068	4.3	56.455	29.803	176.378
161	K	8.198	121.748	4.299	55.994	33.053	176.074
162	D	8.216	120.962	4.595	54.248	41.0	175.836
163	E	8.139	120.704	4.291	56.044	30.389	175.874
164	I	8.101	122.656	4.123	60.847	37.87	175.19
165	R	7.846	130.478	4.136	56.803	30.756	180.712

Table B.1

B.2 Bicelle

Residue	**AA**	H_N	N_H	$H\alpha$	$C\alpha$	$C\beta$	C'
24	I	-	-	3.929	-	-	-
25	L	9.188	111.245	4.14	56.57	-	177.523
26	Q	8.757	119.752	4.068	58.343	28.318	177.466
27	Q	8.444	117.551	4.148	57.946	28.018	177.192
28	L	7.791	118.464	4.187	56.036	40.688	177.99
29	E	8.107	119.372	4.066	58.204	29.488	176.978
30	N	8.131	116.142	4.666	53.905	38.469	175.928
31	K	7.93	118.498	4.377	56.214	32.603	176.233
32	T	7.732	109.565	4.338	61.915	69.394	173.786
33	N	8.144	119.041	4.705	53.169	38.328	174.587
34	L	7.968	121.494	4.518	53.791	41.003	174.605
35	P	-	-	-	63.443	-	177.359
36	K	8.515	122.541	3.966	59.03	-	177.988
37	S	8.326	112.87	4.133	60.784	62.382	176.461
38	Y	7.717	121.14	4.397	59.957	37.867	177.4
39	L	7.606	119.01	4.079	57.916	39.016	178.233
40	V	8.043	116.812	3.624	66.002	31.305	178.177
41	A	7.809	122.068	4.128	54.686	18.307	179.749
42	G	8.401	105.271	3.682	47.318	-	174.768
43	L	8.317	121.208	4.093	57.901	31.353	178.601
44	G	8.073	105.315	4.447	46.962	-	175.55
45	F	8.15	120.334	4.193	60.322	-	176.672
46	A	8.254	120.682	3.885	55.4	17.788	178.525
47	Y	8.184	115.737	3.702	60.772	-	177.319
48	L	8.382	118.104	3.905	-	-	-
49	L	7.722	117.676	-	57.501	-	-
50	L	7.787	117.303	3.996	57.416	-	179.087
51	I	7.894	117.91	3.747	62.721	-	177.914
52	F	7.969	118.747	3.97	60.251	-	177.074
53	I	8.04	115.694	3.891	62.687	-	176.228
54	N	7.904	118.612	4.649	57.416	-	176.164
55	V	8.039	118.761	4.054	63.299	-	176.849
56	G	8.225	108.869	-	45.745	19.426	174.936
57	G	8.051	108.118	3.918	45.835	-	175.54

Continued on next page...

Residue	**AA**	H_N	N_H	$H\alpha$	$C\alpha$	$C\beta$	C'
58	V	8.064	119.013	3.814	65.412	31.305	177.173
59	G	8.551	107.812	-	47.218	-	175.994
60	E	7.869	121.185	4.137	58.663	29.015	178.712
61	I	7.673	119.488	3.835	64.408	37.718	178.139
62	L	8.218	118.464	4.117	58.663	-	178.874
63	S	8.201	113.407	4.167	61.326	-	176.81
64	N	8.018	119.849	4.593	55.511	-	177.105
65	F	8.305	119.291	-	60.16	-	174.655
66	A	8.433	121.251	3.968	54.941	18.353	178.693
67	G	8.019	102.562	-	45.946	-	174.613
68	F	7.686	117.491	4.644	58.895	40.189	176.134
69	V	7.69	116.06	3.822	64.426	32.45	176.134
70	L	8.149	120.411	4.329	58.622	-	175.961
71	P	-	-	-	65.886	-	178.894
72	A	7.421	119.449	4.1	54.275	17.941	179.022
73	Y	8.152	117.978	4.133	60.745	-	176.699
74	L	8.391	117.453	3.926	57.181	-	179.34
75	S	7.804	114.209	4.237	61.153	62.911	175.41
76	L	7.725	121.71	4.121	57.146	41.84	178.443
77	V	7.764	114.838	3.718	64.515	-	177.04
78	A	7.447	121.483	4.212	53.043	18.476	178.262
79	L	7.585	118.096	4.214	55.67	41.84	177.141
80	K	7.734	118.272	4.35	55.428	32.755	175.886
81	T	7.849	115.509	4.189	59.913	68.758	173.066
82	P	-	-	4.49	63.314	-	176.818
83	T	8.176	112.295	4.373	61.496	69.51	174.532
84	S	8.188	116.715	3.915	57.994	63.979	174.895
85	T	8.212	114.741	4.337	62.217	68.83	174.444
86	D	8.178	121.221	4.337	54.521	-	176.266
87	D	8.107	119.632	4.619	54.971	-	176.742
88	T	8.054	113.501	4.134	63.606	68.896	175.554
89	Q	8.277	122.568	4.263	57.189	28.557	176.644
90	L	8.008	121.705	4.193	56.389	-	177.923
91	L	8.088	119.319	4.26	57.55	-	177.843
92	T	7.858	111.523	4.19	64.871	68.702	175.594
93	Y	7.909	120.288	4.152	60.481	-	177.053
94	W	8.01	118.982	4.446	59.974	30.538	179.021

Continued on next page...

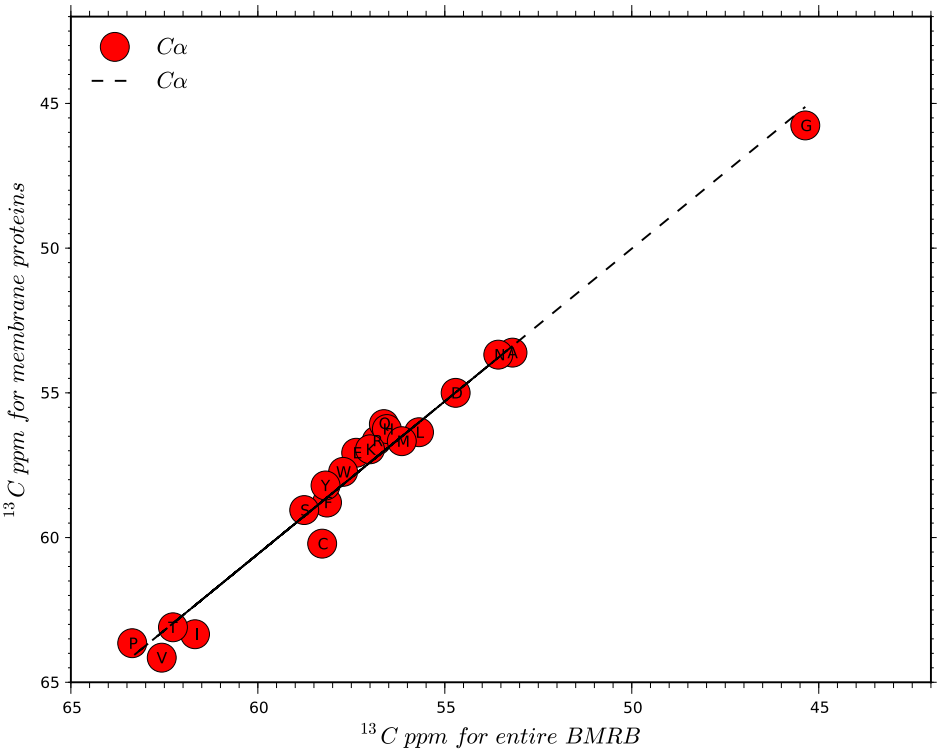
Residue	**AA**	H_N	N_H	$H\alpha$	$C\alpha$	$C\beta$	C'
95	I	8.192	121.154	-	64.127	-	176.977
96	V	7.841	119.135	4.168	66.283	-	177.013
97	F	8.34	118.291	-	60.534	-	178.52
98	S	8.206	116.711	4.366	60.85	-	176.4
99	F	7.662	118.461	4.182	60.643	-	176.798
100	L	8.307	117.548	4.162	57.195	-	179.193
101	S	7.837	114.218	4.253	61.249	-	175.836
102	V	7.8	120.78	3.994	64.664	31.305	178.246
103	I	7.661	118.968	3.937	63.616	-	177.943
104	E	8.115	120.221	4.585	57.424	28.839	178.143
105	F	7.937	118.501	4.374	59.712	-	177.084
106	W	8.109	120.393	3.946	56.916	-	176.908
107	S	8.046	112.463	4.348	61.563	-	175.648
108	K	7.434	119.613	4.116	58.01	32.374	177.829
109	A	8.071	120.604	4.035	54.74	18.122	179.269
110	I	7.847	114.186	3.729	63.398	-	176.642
111	L	7.224	118.194	3.993	57.171	40.92	178.783
112	Y	7.597	115.795	4.349	59.115	38.328	176.194
113	L	7.765	116.057	4.377	55.021	-	177.807
114	I	7.729	118.936	3.958	65.725	-	177.206
115	P	-	-	-	63.832	-	177.386
116	F	7.942	118.985	4.353	60.006	-	176.83
117	Y	8.386	120.346	4.113	61.012	-	176.693
118	W	8.168	117.787	4.134	60.007	28.352	-
120	L	8.126	112.981	3.996	60.528	-	-
121	K	8.239	117.987	-	59.324	31.801	-
122	T	7.653	114.391	4.069	66.435	68.255	175.969
123	V	7.891	119.003	3.603	65.793	30.443	177.156
124	F	8.262	119.768	4.162	61.715	-	176.284
125	L	8.027	116.736	3.976	57.311	-	179.339
126	I	7.671	118.336	3.718	63.846	-	-
127	Y	7.804	118.873	3.997	56.169	-	-
128	I	7.892	114.38	3.829	61.857	-	175.644
129	A	7.333	122.01	4.307	51.773	19.163	177.822
130	L	7.586	120.455	4.389	54.348	-	175.68
131	P	-	-	-	63.439	-	177.058
132	Q	8.263	117.573	4.39	56.572	-	176.457

Continued on next page...

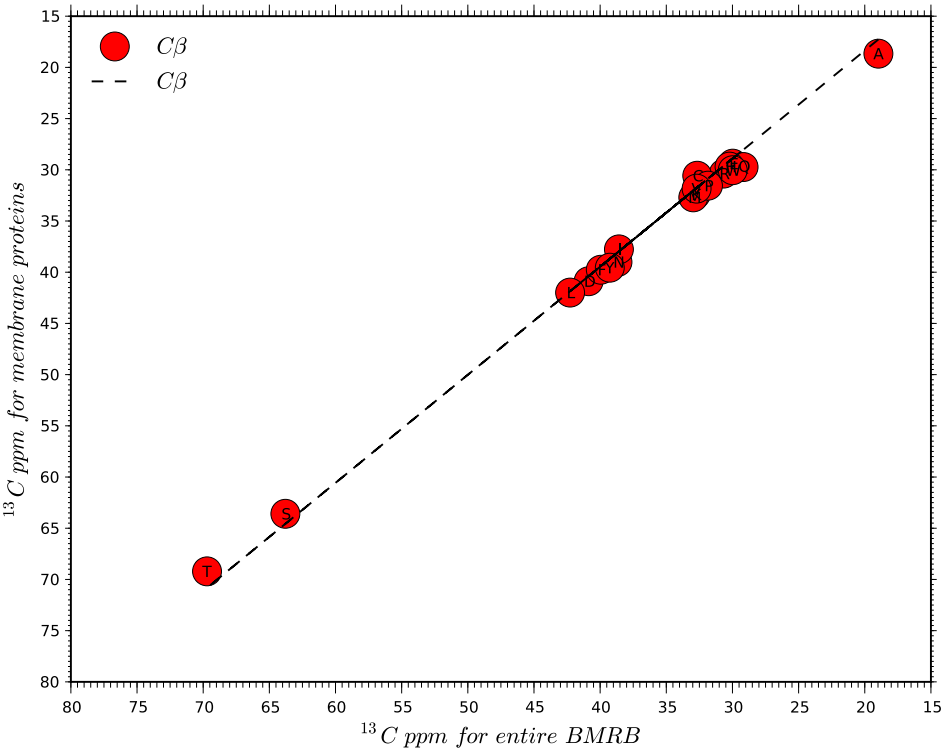
Residue	**AA**	H_N	N_H	$H\alpha$	$C\alpha$	$C\beta$	C'
133	T	7.903	111.596	4.361	61.837	69.889	175.381
134	G	8.478	109.954	-	46.091	-	175.767
135	G	8.46	108.16	3.833	46.593	-	174.732
136	A	8.456	122.363	4.051	54.9	18.094	178.775
137	R	8.004	116.45	4.074	59.219	29.107	178.295
138	I	7.665	119.032	3.944	63.527	-	177.769
139	I	7.813	119.441	3.681	64.629	-	178.357
140	Y	8.405	120.211	4.057	62.103	-	176.545
141	Q	8.427	116.153	3.849	58.706	28.706	177.706
142	K	8.259	113.636	-	57.362	32.984	176.704
143	I	7.65	113.378	4.317	61.815	38.703	175.193
144	V	7.602	116.865	4.035	66.124	-	177.415
145	A	8.453	122.968	4.133	55.752	15.615	176.507
146	P	-	-	-	65.598	-	179.262
147	L	7.359	115.671	4.263	57.166	41.993	178.892
148	T	8.297	113.208	4.316	65.995	68.438	176.801
149	D	8.333	121.885	4.426	56.84	-	177.112
150	R	7.446	115.748	4.03	57.621	30.694	177.134
151	Y	7.832	114.574	-	59.395	39.703	175.824
152	I	7.869	117.77	4.193	61.683	38.325	176.007
153	L	7.998	121.516	4.349	54.86	41.153	177.034
154	R	7.74	119.37	4.229	55.994	30.312	175.913
155	D	8.31	120.515	4.649	53.768	40.749	176.566
156	V	7.996	119.737	4.149	62.058	31.992	176.463
157	S	8.307	118.224	4.396	58.569	63.443	174.776
158	K	8.111	122.74	4.373	56.04	32.526	176.725
159	T	8.019	114.429	4.32	61.665	69.3	174.645
160	E	8.323	123.068	4.3	56.455	29.803	176.378
161	K	8.198	121.748	4.299	55.994	33.053	176.074
162	D	8.216	120.962	4.595	54.248	41.0	175.836
163	E	8.139	120.704	4.291	56.044	30.389	175.874
164	I	8.101	122.656	4.123	60.847	37.87	175.19
165	R	7.846	130.478	4.136	56.803	30.756	180.712

Table B.2

B.3 Comparison of membrane protein *vs* soluble protein chemical shifts



(A) $C\alpha_{sol}$ *vs* $C\alpha_{memb}$



(B) $C\beta_{sol}$ *vs* $C\beta_{memb}$

FIGURE B.1: Backbone chemical shifts from membrane proteins are correlated with soluble proteins.

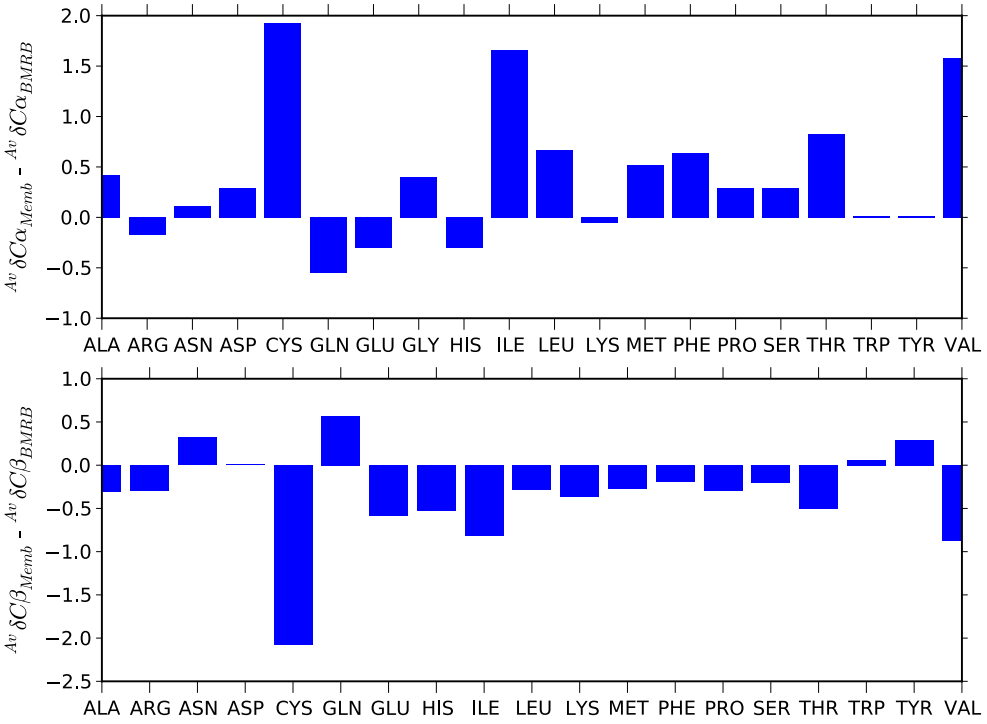


FIGURE B.2: Difference between membrane and soluble protein shifts. There is a slight downfield shift for the Cα's and a slight upfield shift for the Cβ's both of which are indicative of α-helical propensity. This could reflect the fact that membrane proteins are more commonly α-helical in nature.

C

Yop1p Constructs and Mutants

C.1 Yop1p clones and mutations

TABLE C.1: Clones were verified by Sanger sequencing (Source Biosciences) and are stored at -20 °C. Amp = Ampicillin resistance, Species = Species from which gene was derived and Vector = the vector into which the gene was cloned.

Clone no.	Clone name	Species	Vector	Resistance
1	Yop1p (full length)	*S. cerevisiae*	pColdI	Amp
2	Yop1p (24-165)	*S. cerevisiae*	pColdI	Amp
3	Yop1p (24-165) [M138I]	*S. cerevisiae*	pColdI	Amp
4	Yop1p (24-165) [W118A]	*S. cerevisiae*	pColdI	Amp
5	Yop1p (24-86)	*S. cerevisiae*	pColdI	Amp
6	Yop1p (86-165)	*S. cerevisiae*	pColdI	Amp
7	Yop1p (1-86)	*S. cerevisiae*	pColdI	Amp
8	REEP1(full length)	*H. sapiens*	pColdI	Amp
9	Yop1p (24-165) [M138I/V55C]	*S. cerevisiae*	pColdI	Amp
10	Yop1p (24-165) [M138I/S84C]	*S. cerevisiae*	pColdI	Amp
11	Yop1p (24-165) [M138I/S107C]	*S. cerevisiae*	pColdI	Amp
12	Yop1p (24-165) [M138I/V58C]	*S. cerevisiae*	pColdI	Amp
13	Yop1p (24-165) [M138I/S37C]	*S. cerevisiae*	pColdI	Amp
14	Yop1p (24-165) [M138I/W118A]	*S. cerevisiae*	pColdI	Amp
15	Yop1p (24-165) [M138I/E60Q]	*S. cerevisiae*	pColdI	Amp
16	Yop1p (24-165) [M138I/P71A]	*S. cerevisiae*	pColdI	Amp
17	Yop1p (24-165) [M138I/P71L]	*S. cerevisiae*	pColdI	Amp
18	Yop1p (24-165) [M138I/P71R]	*S. cerevisiae*	pColdI	Amp
19	Yop1p (24-165) [M138I/A72E]	*S. cerevisiae*	pColdI	Amp
20	Yop1p (24-165) [M138I/W94A]	*S. cerevisiae*	pColdI	Amp
21	Yop1p (24-165) [M138I/K108E]	*S. cerevisiae*	pColdI	Amp
22	Yop1p (24-165) [I138N]	*S. cerevisiae*	pColdI	Amp
23	Yop1p (24-165) [M138I/I143N]	*S. cerevisiae*	pColdI	Amp
24	Yop1p (24-165) [M138I/G135A]	*S. cerevisiae*	pColdI	Amp
25	Yop1p (24-165) [M138I/E60K]	*S. cerevisiae*	pColdI	Amp
26	Yop1p (24-165) [M138I/W94R]	*S. cerevisiae*	pColdI	Amp
27	Yop1p (24-165) [M138*]	*S. cerevisiae*	pColdI	Amp
28	Yop1p Δ135-151	*S. cerevisiae*	pColdI	Amp
29	Yop1p Δ135-147	*S. cerevisiae*	pColdI	Amp
30	Yop1p Δ135-154	*S. cerevisiae*	pColdI	Amp

Bibliography

Altenbach, C., Greenhalgh, D. A., Khorana, H. G., and Hubbell, W. L. A collision gradient method to determine the immersion depth of nitroxides in lipid bilayers: application to spin-labeled mutants of bacteriorhodopsin. *Proceedings of the National Academy of Sciences*, 91(5):1667–1671, 1994.

Altmann, R. Die Elementarorganismen und ihre Beziehung zu den Zellen. *Leipzig: Veit & Comp*, 1890.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., et al. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410, 1990.

Antonny, B. Mechanisms of membrane curvature sensing. *Annual Review of Biochemistry*, 2011.

Aoto, P. C., Bryn Fenwick, R., Kroon, G. J., and Wright, P. E. Accurate scoring of non-uniform sampling schemes for quantitative NMR. *Journal of Magnetic Resonance*, 246:31–35, September 2014.

Ayala, I., Sounier, R., Usé, N., Gans, P., and Boisbouvier, J. An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *Journal of Biomolecular NMR*, 43(2):111–119, 2009.

Baldwin, A. J., Religa, T. L., Hansen, D. F., Bouvignies, G., and Kay, L. E. $^{13}CHD_2$ methyl group probes of millisecond time scale exchange in proteins by ^{1}H relaxation dispersion: an application to proteasome gating residue dynamics. *Journal of the American Chemical Society*, 132(32):10992–10995, August 2010.

Baldwin, A. J., Lioe, H., Robinson, C. V., Kay, L. E., and Benesch, J. L. αB-crystallin polydispersity is a consequence of unbiased quaternary dynamics. *Journal of Molecular Biology*, 413(2):297–309, 2011.

Barrett, P. J., Song, Y., Van Horn, W. D., Hustedt, E. J., Schafer, J. M., Hadziselimovic, A., Beel, A. J., and Sanders, C. R. The Amyloid Precursor Protein Has a Flexible Transmembrane Domain and Binds Cholesterol. *Science*, 336(6085):1168–1171, May 2012.

Barth, P., Schonbrun, J., and Baker, D. Toward high-resolution prediction and design of transmembrane helical protein structures. *Proceedings of the National Academy of Sciences*, 104(40):15682–15687, October 2007.

Barth, P., Wallner, B., and Baker, D. Prediction of membrane protein structures with complex topologies using limited constraints. *Proceedings of the National Academy of Sciences*, 106(5):1409–1414, February 2009.

Beetz, C., Schule, R., Deconinck, T., Tran-Viet, K. N., Zhu, H., Kremer, B. P. H., Frints, S. G. M., van Zelst-Stams, W. A. G., Byrne, P., Otto, S., Nygren, A. O. H., Baets, J., Smets, K., Ceulemans, B., Dan, B., Nagan, N., Kassubek, J., Klimpe, S., Klopstock, T., Stolze, H., Smeets, H. J. M., Schrander-Stumpel, C. T. R. M., Hutchinson, M., van de Warrenburg, B. P., Braastad, C., Deufel, T., Pericak-Vance, M., Schols, L., de Jonghe, P., and Zuchner, S. REEP1 mutation spectrum and genotype/phenotype correlation in hereditary spastic paraplegia type 31. *Brain*, 131(4):1078–1086, February 2008.

Beetz, C., Pieber, T. R., Hertel, N., Schabhüttl, M., Fischer, C., Trajanoski, S., Graf, E., Keiner, S., Kurth, I., Wieland, T., Varga, R.-E., Timmerman, V., Reilly, M. M., Strom, T. M., and Auer-Grumbach, M. Exome Sequencing Identifies a REEP1 Mutation Involved in Distal Hereditary Motor Neuropathy Type V. *American Journal of Human Genetics*, June 2012.

Beitz, E. TeXshade: shading and labeling of multiple sequence alignments using LaTeX2ϵ. *Bioinformatics*, 16(2):135–139, 2000.

Berardi, M. J., Shih, W. M., Harrison, S. C., and Chou, J. J. Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature*, 476 (7358):109–113, July 2011.

Björk, S., Hurt, C. M., Ho, V. K., and Angelotti, T. REEPs are membrane shaping adapter proteins that modulate specific G protein-coupled receptor trafficking by affecting ER cargo capacity. *PloS One*, 8(10):e76366, 2013.

Blackstone, C. Cellular Pathways of Hereditary Spastic Paraplegia. *Annual Review of Neuroscience*, 35:25–47, 2012.

Bokoch, M. P., Zou, Y., Rasmussen, S. G., Liu, C. W., Nygaard, R., Rosenbaum, D. M., Fung, J. J., Choi, H.-J., Thian, F. S., Kobilka, T. S., et al. Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature*, 463(7277):108–112, 2010.

Bowers, P. M., Strauss, C. E., and Baker, D. *De novo* protein structure determination using sparse NMR data. *Journal of Biomolecular NMR*, 18(4):311–318, December 2000.

Browne, D., Kenyon, G., Packer, E., Sternlicht, H., and Wilson, D. Studies of macromolecular structure by carbon-13 nuclear magnetic resonance. II. Specific labeling approach to the study of histidine residues in proteins. *Journal of the American Chemical Society*, 95(4):1316–1323, 1973.

Camilloni, C., De Simone, A., Vranken, W. F., and Vendruscolo, M. Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry*, 51(11):2224–2231, March 2012.

Campelo, F., McMahon, H. T., and Kozlov, M. M. The Hydrophobic Insertion Mechanism of Membrane Curvature Generation by Proteins. *Biophysical Journal*, 95(5): 2325–2339, September 2008.

Chen, C.-N., Chu, C.-C., Zentella, R., Pan, S.-M., and Ho, T.-H. D. AtHVA22 gene family in Arabidopsis: phylogenetic relationship, ABA and stress regulation, and tissue-specific expression. *Plant Molecular Biology*, 49(6):631–642, 2002.

Chen, H., Ji, F., Olman, V., Mobley, C. K., Liu, Y., Zhou, Y., Bushweller, J. H., Prestegard, J. H., and Xu, Y. Optimal Mutation Sites for PRE Data Collection and Membrane Protein Structure Prediction. *Structure/Folding and Design*, 19(4): 484–495, April 2011a.

Chen, J., Stefano, G., Brandizzi, F., and Zheng, H. Arabidopsis RHD3 mediates the generation of the tubular ER network and is required for Golgi distribution and motility in plant cells. *Journal of Cell Science*, 124(13):2241–2252, 2011b.

Chill, J. H., Louis, J. M., Baber, J. L., and Bax, A. Measurement of ^{15}N relaxation in the detergent-solubilized tetrameric KcsA potassium channel. *Journal of Biomolecular NMR*, 36(2):123–136, October 2006.

Chothia, C. The nature of the accessible and buried surfaces in proteins. *Journal of Molecular Biology*, 105(1):1–12, 1976.

Chou, J. J., Gaemers, S., Howder, B., Louis, J. M., and Bax, A. A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles. *Journal of Biomolecular NMR*, 21(4):377–382, 2001.

Chou, P. Y. and Fasman, G. D. Prediction of protein conformation. *Biochemistry*, 13 (2):222–245, 1974.

Cierpicki, T. and Bushweller, J. H. Charged gels as orienting media for measurement of residual dipolar couplings in soluble and integral membrane proteins. *Journal of the American Chemical Society*, 126(49):16259–16266, 2004.

Claridge, J. K., Aittoniemi, J., Cooper, D. M., and Schnell, J. R. Isotropic Bicelles Stabilize the Juxtamembrane Region of the Influenza M2 Protein for Solution NMR Studies. *Biochemistry*, 52(47):8420–8429, November 2013.

Cock, P. J., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11): 1422–1423, 2009.

Cole, C., Barber, J. D., and Barton, G. J. The Jpred 3 secondary structure prediction server. *Nucleic Acids Research*, 36(suppl 2):W197–W201, 2008.

Cordes, F. S., Bright, J. N., and Sansom, M. S. Proline-induced distortions of transmembrane helices. *Journal of Molecular Biology*, 323(5):951–960, 2002.

Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. Weblogo: a sequence logo generator. *Genome research*, 14(6):1188–1190, 2004.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR*, 6(3):277–293, November 1995.

Delaglio, F., Kontaxis, G., and Bax, A. Protein structure determination using molecular fragment replacement and NMR dipolar couplings. *Journal of the American Chemical Society*, 122(9):2142–2143, 2000.

Dempsey, C. E., Bazzo, R., Harvey, T. S., Syperek, I., Boheim, G., and Campbell, I. D. Contribution of proline-14 to the structure and actions of melittin. *FEBS Letters*, 281 (1):240–244, 1991.

Diaz, A., Wang, X., and Ahlquist, P. Membrane-shaping host reticulon proteins play crucial roles in viral RNA replication compartment formation and function. *Proceedings of the National Academy of Sciences*, 107(37):16291–16296, 2010.

Donaldson, L. W., Skrynnikov, N. R., Choy, W. Y., Muhandiram, D. R., Sarkar, B., Forman-Kay, J. D., and Kay, L. E. Structural characterization of proteins with an attached ATCUN motif by paramagnetic relaxation enhancement NMR spectroscopy. *Journal of the American Chemical Society*, 123(40):9843–9847, October 2001.

Drin, G. and Antonny, B. Amphipathic helices and membrane curvature. *FEBS Letters*, 584(9):1840–1847, 2010.

Drin, G., Casella, J.-F., Gautier, R., Boehmer, T., Schwartz, T. U., and Antonny, B. A general amphipathic α-helical motif for sensing membrane curvature. *Nature Structural & Molecular Biology*, 14(2):138–146, 2007.

Eisenberg, D., Weiss, R. M., and Terwilliger, T. C. The helical hydrophobic moment: a measure of amphiphilicity of a helix. *Nature*, 299(23):371–374, September 1982a.

Eisenberg, D., Weiss, R. M., Terwilliger, T. C., and Wilcox, W. Hydrophobic moments and protein structure. In *Faraday Symposia of the Chemical Society*, volume 17, pages 109–120. Royal Society of Chemistry, 1982b.

Eisenberg, D., Weiss, R. M., and Terwilliger, T. C. The hydrophobic moment detects periodicity in protein hydrophobicity. *Proceedings of the National Academy of Sciences*, 81(1):140–144, 1984.

Ejsing, C. S., Sampaio, J. L., Surendranath, V., Duchoslav, E., Ekroos, K., Klemm, R. W., Simons, K., and Shevchenko, A. Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry. *Proceedings of the National Academy of Sciences*, 106(7):2136–41, February 2009.

Falk, J., Rohde, M., Bekhite, M. M., Neugebauer, S., Hemmerich, P., Kiehntopf, M., Deufel, T., Hübner, C. A., and Beetz, C. Functional Mutation Analysis Provides Evidence for a Role of REEP1 in Lipid Droplet Biology. *Human Mutation*, 35(4): 497–504, 2014.

Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry*, 33(19):5984–6003, 1994.

Fernández, C., Adeishvili, K., and Wüthrich, K. Transverse relaxation-optimized NMR spectroscopy with the outer membrane protein OmpX in dihexanoyl phosphatidylcholine micelles. *Proceedings of the National Academy of Sciences*, 98(5):2358–2363, 2001.

Fernández, C., Hilty, C., Wider, G., Güntert, P., and Wüthrich, K. NMR structure of the integral membrane protein OmpX. *Journal of Molecular Biology*, 336(5):1211–1221, 2004.

Fitzkee, N. C. and Bax, A. Facile measurement of ^{1}H–^{15}N residual dipolar couplings in larger perdeuterated proteins. *Journal of Biomolecular NMR*, 48(2):65–70, 2010.

Ford, M. G., Mills, I. G., Peter, B. J., Vallis, Y., Praefcke, G. J., Evans, P. R., and McMahon, H. T. Curvature of clathrin-coated pits driven by epsin. *Nature*, 419 (6905):361–366, 2002.

Frost, A., Perera, R., Roux, A., Spasov, K., Destaing, O., Egelman, E. H., De Camilli, P., and Unger, V. M. Structural basis of membrane invagination by F-BAR domains. *Cell*, 132(5):807–817, 2008.

Frost, A., Unger, V. M., and De Camilli, P. The BAR domain superfamily: membrane-molding macromolecules. *Cell*, 137(2):191–196, 2009.

Gallop, J. L., Jao, C. C., Kent, H. M., Butler, P. J. G., Evans, P. R., Langen, R., and McMahon, H. T. Mechanism of endophilin N-BAR domain-mediated membrane curvature. *The EMBO Journal*, 25(12):2898–2910, 2006.

Gallop, J. L., Butler, P. J. G., and McMahon, H. T. Endophilin and CtBP/BARS are not acyl transferases in endocytosis or Golgi fission. *Nature*, 438(7068):675–678, 2005.

Gardner, K. H. and Kay, L. E. Production and incorporation of ^{15}N, ^{13}C, ^{2}H (^{1}H-δ1 methyl) isoleucine into proteins for multidimensional NMR studies. *Journal of the American Chemical Society*, 119(32):7599–7600, 1997.

Gardner, K. H. and Kay, L. E. THE USE OF 2H, 13C, 15N MULTIDIMENSIONAL NMR GTO STUDY THE STRUCTURE AND DYNAMICS OF PROTEINS. *Annual Review of Biophysics and Biomolecular Structure*, 27(1):357–406, 1998.

Gautier, A., Mott, H. R., Bostock, M. J., Kirkpatrick, J. P., and Nietlispach, D. Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nature Structural & Molecular Biology*, 17(6):768–774, May 2010.

Gautier, R., Douguet, D., Antonny, B., and Drin, G. HELIQUEST: a web server to screen sequences with specific alpha-helical properties. *Bioinformatics (Oxford, England)*, 24(18):2101–2102, September 2008.

Gerondopoulos, A., Bastos, R. N., Yoshimura, S.-i., Anderson, R., Carpanini, S., Aligianis, I., Handley, M. T., and Barr, F. A. Rab18 and a Rab18 GEF complex are

required for normal ER structure. *The Journal of Cell Biology*, 205(5):707–720, 2014. doi: 10.1083/jcb.201403026. URL `http://jcb.rupress.org/content/205/5/707.abstract`.

Godoy-Ruiz, R., Guo, C., and Tugarinov, V. Alanine methyl groups as NMR probes of molecular structure and dynamics in high-molecular-weight proteins. *Journal of the American Chemical Society*, 132(51):18340–18350, 2010.

Goizet, C., Depienne, C., Benard, G., Boukhris, A., Mundwiller, E., Solé, G., Coupry, I., Pilliod, J., Martin-Négrier, M.-L., Fedirko, E., Forlani, S., Cazeneuve, C., Hannequin, D., Charles, P., Feki, I., Pinel, J.-F., Ouvrard-Hernandez, A.-M., Lyonnet, S., Ollagnon-Roman, E., Yaouanq, J., Toutain, A., Dussert, C., Fontaine, B., Leguern, E., Lacombe, D., Durr, A., Rossignol, R., Brice, A., and Stevanin, G. REEP1 mutations in SPG31: Frequency, mutational spectrum, and potential association with mitochondrial morpho-functional dysfunction. *Human Mutation*, 32(10):1118–1127, September 2011.

Golgi, C. Intorno alla struttura delle cellule nervose. *Arch. Ital. Biol.*, 30:60–71, 1898.

Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C., and Kay, L. E. A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated ^{15}N-, ^{13}C-, ^{2}H-labeled proteins. *Journal of Biomolecular NMR*, 13(4):369–374, 1999.

Gottstein, D., Reckel, S., Dötsch, V., and Güntert, P. Requirements on Paramagnetic Relaxation Enhancement Data for Membrane Protein Structure Determination by NMR. *Structure/Folding and Design*, 20(6):1019–1027, June 2012.

Goyal, U. and Blackstone, C. Untangling the web: Mechanisms underlying ER network formation. *BBA - Molecular Cell Research*, pages 1–7, April 2013.

GrandPré, T., Nakamura, F., Vartanian, T., and Strittmatter, S. M. Identification of the Nogo inhibitor of axon regeneration as a Reticulon protein. *Nature*, 403(6768): 439–444, 2000.

Grzesiek, S., Anglister, J., Ren, H., and Bax, A. ^{13}C line narrowing by ^{2}H decoupling in ^{2}H/^{13}C/^{15}N-enriched proteins. Application to triple resonance 4D J connectivity of sequential amides. *Journal of the American Chemical Society*, 115(10):4369–4370, 1993.

Guo, C., Zhang, D., and Tugarinov, V. An NMR experiment for simultaneous TROSY-based detection of amide and methyl groups in large proteins. *Journal of the American Chemical Society*, 130(33):10872–10873, 2008.

Guo, Y., Sirkis, D. W., and Schekman, R. Protein Sorting at the trans Golgi Network. *Annual Review of Cell and Developmental Biology*, 30(1), 2014.

Hagn, F., Etzkorn, M., Raschle, T., and Wagner, G. Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins. *Journal of the American Chemical Society*, 2013.

Hansen, D. F. and Kay, L. E. Determining Valine Side-Chain Rotamer Conformations in Proteins from Methyl ^{13}C Chemical Shifts: Application to the 360 kDa Half-Proteasome. *Journal of the American Chemical Society*, 133(21):8272–8281, June 2011.

Hansen, D. F., Neudecker, P., and Kay, L. E. Determination of Isoleucine Side-Chain Conformations in Ground and Excited States of Proteins from Chemical Shifts. *Journal of the American Chemical Society*, 132(22):7589–7591, June 2010.

Hansen, M. R., Mueller, L., and Pardi, A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nature Structural & Molecular Biology*, 5(12):1065–1074, 1998.

Helmus, J. J. and Jaroniec, C. P. Nmrglue: an open source Python package for the analysis of multidimensional NMR data. *Journal of Biomolecular NMR*, 55(4):355–367, 2013.

Henne, W. M., Buchkovich, N. J., Zhao, Y., and Emr, S. D. The endosomal sorting complex ESCRT-II mediates the assembly and architecture of ESCRT-III helices. *Cell*, 151(2):356–371, October 2012.

Henne, W. M., Stenmark, H., and Emr, S. D. Molecular mechanisms of the membrane sculpting ESCRT pathway. *Cold Spring Harbor Perspectives in Biology*, 5(9):a016766, 2013.

Heymann, J. A. and Hinshaw, J. E. Dynamins at a glance. *Journal of Cell Science*, 122 (19):3427–3431, 2009.

Hiller, S., Ibraghimov, I., Wagner, G., and Orekhov, V. Y. Coupled decomposition of four-dimensional NOESY spectra. *Journal of the American Chemical Society*, 131 (36):12970–12978, 2009.

Hooke, R. Micrographia: or Some Physiological Descriptions of Minute Bodies. *London*, 1665.

Hu, J., Shibata, Y., Voss, C., Shemesh, T., Li, Z., Coughlin, M., Kozlov, M. M., Rapoport, T. A., and Prinz, W. A. Membrane proteins of the endoplasmic reticulum induce high-curvature tubules. *Science*, 319(5867):1247–1250, 2008.

Hu, J., Shibata, Y., Zhu, P.-P., Voss, C., Rismanchi, N., Prinz, W. A., Rapoport, T. A., and Blackstone, C. A Class of Dynamin-like GTPases Involved in the Generation of the Tubular ER Network. *Cell*, 138(3):549–561, August 2009.

Hu, J., Prinz, W. A., and Rapoport, T. A. Weaving the web of ER tubules. *Cell*, 147 (6):1226–1231, 2011.

Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science and Engineering*, 9(3):90–95, May-June 2007.

Hurt, C. M., Björk, S., Ho, V. K., Gilsbach, R., Hein, L., and Angelotti, T. REEP1 and REEP2 proteins are preferentially expressed in neuronal and neuronal-like exocytotic tissues. *Brain Research*, 1545:12–22, 2014.

Hwang, T. L., van Zijl, P. C., and Mori, S. Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *Journal of Biomolecular NMR*, 11(2):221–226, February 1998.

Hyberts, S. G., Takeuchi, K., and Wagner, G. Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. *Journal of the American Chemical Society*, 132(7):2145–2147, February 2010.

Hyberts, S. G., Milbradt, A. G., Wagner, A. B., Arthanari, H., and Wagner, G. Application of Iterative Soft Thresholding for Fast Reconstruction of NMR Data Non-uniformly Sampled with Multidimensional Poisson Gap Scheduling. *Journal of Biomolecular NMR*, 52(4):315, April 2012.

Hyberts, S. G., Robson, S. A., and Wagner, G. Exploring signal-to-noise ratio and sensitivity in non-uniformly sampled multi-dimensional NMR spectra. *Journal of Biomolecular NMR*, 55(2):167–178, February 2013.

Hyberts, S. G., Arthanari, H., Robson, S. A., and Wagner, G. Perspectives in magnetic resonance: NMR in the post-FFT era. *Journal of Magnetic Resonance*, 241:60–73, 2014.

Isaacson, R. L., Simpson, P. J., Liu, M., Cota, E., Zhang, X., Freemont, P., and Matthews, S. A new labeling method for methyl transverse relaxation-optimized spectroscopy NMR spectra of alanine residues. *Journal of the American Chemical Society*, 129(50):15428–15429, 2007.

Iwahara, J., Tang, C., and Marius Clore, G. Practical aspects of ^{1}H transverse paramagnetic relaxation enhancement measurements on macromolecules. *Journal of Magnetic Resonance*, 184(2):185–195, 2007.

Janin, J. Surface and inside volumes in globular proteins. *Nature*, 277:491–492, February 1979.

Janin, J., Miller, S., and Chothia, C. Surface, subunit interfaces and interior of oligomeric proteins. *Journal of Molecular Biology*, 204(1):155–164, 1988.

Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, 292(2):195–202, 1999.

Kainosho, M., Torizawa, T., Iwashita, Y., Terauchi, T., Ono, A. M., and Güntert, P. Optimal isotope labelling for NMR protein structure determinations. *Nature*, 440 (7080):52–57, 2006.

Käll, L., Krogh, A., and Sonnhammer, E. L. Advantages of combined transmembrane topology and signal peptide prediction–the Phobius web server. *Nucleic Acids Research*, 35(Web Server issue):W429–W432, 2007.

Kalverda, A. P., Gowdy, J., Thompson, G. S., Homans, S. W., Henderson, P. J., and Patching, S. G. TROSY NMR with a 52 kDa sugar transport protein and the binding of a small-molecule inhibitor. *Molecular Membrane Biology*, 31(4):131–140, 2014.

Kamen, D. E., Cahill, S. M., and Girvin, M. E. Multiple alignment of membrane proteins for measuring residual dipolar couplings using lanthanide ions bound to a small metal chelator. *Journal of the American Chemical Society*, 129(7):1846–1847, 2007.

Kay, L. E., Torchia, D. A., and Bax, A. Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, 28(23):8972–8979, 1989.

Kazimierczuk, K. and Orekhov, V. Y. Accelerated NMR spectroscopy by using compressed sensing. *Angewandte Chemie (International ed. in English)*, 50(24):5556–5559, June 2011.

Kazimierczuk, K., Stanek, J., Zawadzka-Kazimierczuk, A., and Koźmiński, W. High-Dimensional NMR Spectra for Structural Studies of Biomolecules. *ChemPhysChem*, 14(13):3015–3025, 2013.

Keller, R. L. J. *Optimising the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment.* PhD thesis, Swiss Federal Institute of Technology Zurich, 2004.

Kempf, A. and Schwab, M. E. Nogo-A represses anatomical and synaptic plasticity in the central nervous system. *Physiology*, 28(3):151–163, 2013.

Khademi, S., O'Connell, J., Remis, J., Robles-Colmenares, Y., Miercke, L. J., and Stroud, R. M. Mechanism of ammonia transport by Amt/MEP/Rh: structure of AmtB at 1.35 Å. *Science*, 305(5690):1587–1594, 2004.

Kim, B. L., Schafer, N. P., and Wolynes, P. G. Predictive energy landscapes for folding α-helical transmembrane proteins. *Proceedings of the National Academy of Sciences*, 111(30):11031–11036, 2014.

Klammt, C., Maslennikov, I., Bayrhuber, M., Eichmann, C., Vajpai, N., Chiu, E. J. C., Blain, K. Y., Esquivies, L., Kwon, J. H. J., Balana, B., et al. Facile backbone structure determination of human membrane proteins by NMR spectroscopy. *Nature Methods*, 9(8):834–839, 2012.

Kneller, J. M., Lu, M., and Bracken, C. An Effective Method for the Discrimination of Motional Anisotropy and Chemical Exchange. *Journal of the American Chemical Society*, 124(9):1852–1853, March 2002.

Kontaxis, G., Clore, G. M., and Bax, A. Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. *Journal of Magnetic Resonance*, 143(1):184–196, 2000.

Kontaxis, G., Delaglio, F., and Bax, A. Molecular fragment replacement approach to protein structure determination by chemical shift and dipolar homology database mining. *Methods in enzymology*, 394:42–78, 2005.

Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*, 305(3):567–580, 2001.

Kroncke, B. M. and Columbus, L. Identification and removal of nitroxide spin label contaminant: Impact on PRE studies of α-helical membrane proteins in detergent. *Protein Science*, 21(4):589–595, 2012.

Kyte, J. and Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1):105–132, 1982.

Laganowsky, A., Reading, E., Allison, T. M., Ulmschneider, M. B., Degiacomi, M. T., Baldwin, A. J., and Robinson, C. V. Membrane proteins bind lipids selectively to modulate their structure and function. *Nature*, 510(7503):172–175, 2014.

Lakomek, N.-A., Ying, J., and Bax, A. Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods. *Journal of Biomolecular NMR*, 53(3):209–221, 2012.

Lakomek, N.-A., Kaufman, J. D., Stahl, S. J., Louis, J. M., Grishaev, A., Wingfield, P. T., and Bax, A. Internal Dynamics of the Homotrimeric HIV-1 Viral Coat Protein gp41 on Multiple Time Scales. *Angewandte Chemie*, 125(14):4003–4007, 2013.

Lau, T.-L., Partridge, A. W., Ginsberg, M. H., and Ulmer, T. S. Structure of the Integrin β3 Transmembrane Segment in Phospholipid Bicelles and Detergent Micelles. *Biochemistry*, 47(13):4008–4016, 2008.

Lee, D., Hilty, C., Wider, G., and Wüthrich, K. Effective rotational correlation times of proteins from NMR relaxation interference. *Journal of Magnetic Resonance*, 178(1): 72–76, January 2006.

Lee, E., Marcucci, M., Daniell, L., Pypaert, M., Weisz, O. A., Ochoa, G.-C., Farsad, K., Wenk, M. R., and De Camilli, P. Amphiphysin 2 (Bin1) and T-tubule biogenesis in muscle. *Science*, 297(5584):1193–1196, 2002.

Lee, J. and Glover, K. J. The transmembrane domain of caveolin-1 exhibits a helix–break–helix structure. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1818(5): 1158–1164, 2012.

Lee, J. K., Gopal, R., Park, S.-C., Ko, H. S., Kim, Y., Hahm, K.-S., and Park, Y. A proline-hinge alters the characteristics of the amphipathic α-helical AMPs. *PloS One*, 8(7):e67597, 2013.

Leuwenhook, A. V. Arcana naturae detecta. 1680.

Li, D., Lyons, J. A., Pye, V. E., Vogeley, L., Aragão, D., Kenyon, C. P., Shah, S. T., Doherty, C., Aherne, M., and Caffrey, M. Crystal structure of the integral membrane diacylglycerol kinase. *Nature*, 497(7450):521–524, 2013.

Liang, B., Bushweller, J. H., and Tamm, L. K. Site-directed parallel spin-labeling and paramagnetic relaxation enhancement in structure determination of membrane proteins by solution NMR spectroscopy. *Journal of the American Chemical Society*, 128(13):4389–4397, April 2006.

Lichtenecker, R. Synthesis of aromatic $^{13}C/^{2}H$-α-ketoacid precursors to be used in selective phenylalanine and tyrosine protein labelling. *Organic & Biomolecular Chemistry*, 12(38):7551–7560, 2014.

Lichtenecker, R. J., Weinhäupl, K., Schmid, W., and Konrat, R. α-Ketoacids as precursors for phenylalanine and tyrosine labelling in cell-based protein overexpression. *Journal of Biomolecular NMR*, 57(4):327–331, 2013.

Liu, T. Y., Bian, X., Sun, S., Hu, X., Klemm, R. W., Prinz, W. A., Rapoport, T. A., and Hu, J. Lipid interaction of the C terminus and association of the transmembrane segments facilitate atlastin-mediated homotypic endoplasmic reticulum fusion. *Proceedings of the National Academy of Sciences*, 109(32):E2146–E2154, 2012.

Lorieau, J., Yao, L., and Bax, A. Liquid crystalline phase of G-tetrad DNA for NMR study of detergent-solubilized proteins. *Journal of the American Chemical Society*, 130(24):7536–7537, 2008.

Luginbühl, P. and Wüthrich, K. Semi-classical nuclear spin relaxation theory revisited for use with biological macromolecules. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 40(3):199–247, 2002.

Martens, S., Kozlov, M. M., and McMahon, H. T. How synaptotagmin promotes membrane fusion. *Science*, 316(5828):1205–1208, 2007.

Matsuo, H., Chevallier, J., Mayran, N., Le Blanc, I., Ferguson, C., Fauré, J., Blanc, N. S., Matile, S., Dubochet, J., Sadoul, R., et al. Role of LBPA and Alix in multivesicular liposome formation and endosome organization. *Science*, 303(5657):531–534, 2004.

Meier, S., Häussinger, D., and Grzesiek, S. Charged acrylamide copolymer gels as media for weak alignment. *Journal of Biomolecular NMR*, 24(4):351–356, 2002.

Mesleh, M. F. and Opella, S. J. Dipolar waves as NMR maps of helices in proteins. *Journal of Magnetic Resonance*, 163(2):288–299, 2003.

Mesleh, M. F., Lee, S., Veglia, G., Thiriot, D. S., Marassi, F. M., and Opella, S. J. Dipolar waves map the structure and topology of helices in membrane proteins. *Journal of the American Chemical Society*, 125(29):8928–8935, 2003.

Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., and Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry*, 32(10):2319–2327, 2011.

Mim, C., Cui, H., Gawronski-Salerno, J. A., Frost, A., Lyman, E., Voth, G. A., and Unger, V. M. Structural Basis of Membrane Bending by the N-BAR Protein Endophilin. *Cell*, 149(1):137–145, March 2012.

Mitra, K., Ubarretxena-Belandia, I., Taguchi, T., Warren, G., and Engelman, D. M. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proceedings of the National Academy of Sciences*, 101(12):4083–4088, March 2004.

Monné, M., Hermansson, M., and von Heijne, G. A turn propensity scale for transmembrane helices. *Journal of Molecular Biology*, 288(1):141–145, April 1999.

Mori, S., Abeygunawardana, C., Johnson, M., and Vanzijl, P. Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *Journal of Magnetic Resonance, Series B*, 108(1):94–98, 1995.

Mulder, F. A., Schipper, D., Bott, R., and Boelens, R. Altered flexibility in the substrate-binding site of related native and engineered high-alkaline Bacillus subtilisins. *Journal of Molecular Biology*, 292(1):111–123, September 1999.

Mulder, F. A. A. Leucine Side-Chain Conformation and Dynamics in Proteins from ^{13}C NMR Chemical Shifts. *ChemBioChem*, 10(9):1477–1479, June 2009.

Neri, D., Szyperski, T., Otting, G., Senn, H., and Wüthrich, K. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry*, 28(19):7510–7516, September 1989.

Nozaki, Y. and Tanford, C. The Solubility of Amino Acids and Two Glycine Peptides in Aqueous Ethanol and Dioxane Solutions: ESTABLISHMENT OF A HYDROPHOBICITY SCALE. *Journal of Biological Chemistry*, 246(7):2211–2217, 1971.

Nygaard, R., Zou, Y., Dror, R. O., Mildorf, T. J., Arlow, D. H., Manglik, A., Pan, A. C., Liu, C. W., Fung, J. J., Bokoch, M. P., et al. The Dynamic Process of β_2-Adrenergic Receptor Activation. *Cell*, 152(3):532–542, 2013.

Oertle, T. and Schwab, M. Nogo and its paRTNers. *Trends in Cell Biology*, 13(4): 187–194, 2003.

Oliphant, T. A Guide to NumPy, vol. 1. *Spanish Fork: Trelgol Publishing*, 2006.

Oliver, R. C., Lipfert, J., Fox, D. A., Lo, R. H., Doniach, S., and Columbus, L. Dependence of micelle size and shape on detergent alkyl chain length and head group. *PloS One*, 8(5):e62488, 2013.

Ollerenshaw, J. E., Tugarinov, V., Skrynnikov, N. R., and Kay, L. E. Comparison of $^{13}CH_3$, $^{13}CH_2D$, and $^{13}CHD_2$ methyl labeling strategies in proteins. *Journal of Biomolecular NMR*, 33(1):25–41, 2005.

Orekhov, V. Y. and Jaravine, V. A. Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 59(3):271–292, October 2011.

Ortega-Roldan, J. L., Ossa, F., and Schnell, J. R. Characterization of the human Sigma-1 Receptor chaperone domain structure and BiP interactions. *Journal of Biological Chemistry*, June 2013.

O'Shea, J. P., Chou, M. F., Quader, S. A., Ryan, J. K., Church, G. M., and Schwartz, D. pLogo: a probabilistic approach to visualizing sequence motifs. *Nature Methods*, 10(12):1211–1212, 2013.

OuYang, B., Xie, S., Berardi, M. J., Zhao, X., Dev, J., Yu, W., Sun, B., and Chou, J. J. Unusual architecture of the p7 channel from hepatitis C virus. *Nature*, 498(7455): 521–525, 2013.

Park, S. H. and Blackstone, C. Further assembly required: construction and dynamics of the endoplasmic reticulum network. *EMBO Reports*, 11(7):515–521, 2010.

Park, S. H., Zhu, P.-P., Parker, R. L., and Blackstone, C. Hereditary spastic paraplegia proteins REEP1, spastin, and atlastin-1 coordinate microtubule interactions with the tubular ER network. *Journal of Clinical Investigation*, 120(4):1097–1110, April 2010.

Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences*, 94(23):12366–12371, November 1997.

Peter, B. J., Kent, H. M., Mills, I. G., Vallis, Y., Butler, P. J. G., Evans, P. R., and McMahon, H. T. BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. *Science*, 303(5657):495–499, 2004.

Pielak, R. M., Oxenoid, K., and Chou, J. J. Structural investigation of rimantadine inhibition of the AM2-BM2 chimera channel of influenza viruses. *Structure*, 19(11): 1655–1663, 2011.

Poget, S. F. and Girvin, M. E. Solution NMR of membrane proteins in bilayer mimics: Small is beautiful, but sometimes bigger is better. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1768(12):3098–3106, December 2007.

Porter, K. R. Observations on a submicroscopic basophilic component of cytoplasm. *The Journal of Experimental Medicine*, 97(5):727–750, 1953.

Porter, K. R. and Palade, G. E. Studies on the endoplasmic reticulum III. Its form and distribution in striated muscle cells. *The Journal of Biophysical and Biochemical Cytology*, 3(2):269–300, 1957.

Pranke, I. M., Morello, V., Bigay, J., Gibson, K., Verbavatz, J.-M., Antonny, B., and Jackson, C. L. α-Synuclein and ALPS motifs are membrane curvature sensors whose contrasting chemistry mediates selective vesicle binding. *The Journal of Cell Biology*, 194(1):89–103, 2011.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. Numerical Recipes in C: The Art of Scientific Computing (New York). *Cambridge University Press*, page Section 15.6, 1992.

Raghuraman, H. and Chattopadhyay, A. Melittin: a membrane-active peptide with diverse functions. *Bioscience Reports*, 27:189–223, 2007.

Ramamurthi, K. S., Lecuyer, S., Stone, H. A., and Losick, R. Geometric Cue for Protein Localization in a Bacterium. *Science*, 323(5919):1354–1357, 2009.

Rieth, M. D., Lee, J., and Glover, K. J. Probing the caveolin-1 P132L mutant: critical insights into its oligomeric behavior and structure. *Biochemistry*, 51(18):3911–3918, 2012.

Roberts, K. L., Leser, G. P., Ma, C., and Lamb, R. A. The amphipathic helix of influenza A virus M2 protein is required for filamentous bud formation and scission of filamentous and spherical particles. *Journal of Virology*, 87(18):9973–9982, 2013.

Roseman, M. A. Hydrophobicity of the peptide C=O...H-N hydrogen-bonded group. *Journal of Molecular Biology*, 201(3):621 – 623, 1988. ISSN 0022-2836. URL `http://www.sciencedirect.com/science/article/pii/0022283688906420`.

Rosen, M. K., Gardner, K. H., Willis, R. C., Parris, W. E., Pawson, T., and Kay, L. E. Selective methyl group protonation of perdeuterated proteins. *Journal of Molecular Biology*, 263(5):627–636, 1996.

Rosenzweig, R., Moradi, S., Zarrine-Afsar, A., and Glover, J. R. Unraveling the Mechanism of Protein Disaggregation Through a ClpB-DnaK Interaction. *Science*, 2013.

Rossman, J. S., Jing, X., Leser, G. P., and Lamb, R. A. Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell*, 142(6):902–913, 2010.

Rossum, G. v. et al. Python programming language. *URL http://www. python. org*, 1989.

Roussel, B. D., Kruppa, A. J., Miranda, E., Crowther, D. C., Lomas, D. A., and Marciniak, S. J. Endoplasmic reticulum dysfunction in neurological disease. *The Lancet Neurology*, 12(1):105–118, 2013.

Rui, H., Root, K. T., Lee, J., Glover, K. J., and Im, W. Probing the U-Shaped Conformation of Caveolin-1 in a Bilayer. *Biophysical Journal*, 106(6):1371–1380, 2014.

Ruotolo, B. T., Benesch, J. L., Sandercock, A. M., Hyung, S.-J., and Robinson, C. V. Ion mobility–mass spectrometry analysis of large protein complexes. *Nature Protocols*, 3(7):1139–1152, 2008.

Saito, H., Kubota, M., Roberts, R. W., Chi, Q., and Matsunami, H. RTP family members induce functional expression of mammalian odorant receptors. *Cell*, 119(5): 679–691, November 2004.

Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wüthrich, K. TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proceedings of the National Academy of Sciences*, 95(23):13585–13590, November 1998.

Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wüthrich, K. NMR assignment and secondary structure determination of an octameric 110 kDa protein using TROSY in triple resonance experiments. *Journal of the American Chemical Society*, 122(31): 7543–7548, 2000.

Sambrook, J. and Russell, D. W. *Molecular cloning: a laboratory manual (3-volume set)*. Cold spring harbor laboratory press Cold Spring Harbor, New York:, 2001.

Schanda, P., Kupče, Ē., and Brutscher, B. SOFAST-HMQC Experiments for Recording Two-dimensional Heteronuclear Correlation Spectra of Proteins within a Few Seconds. *Journal of Biomolecular NMR*, 33(4):199–211, December 2005.

Schlang, K. J., Arning, L., Epplen, J. T., and Stemmler, S. Autosomal dominant hereditary spastic paraplegia: Novel mutations in the REEP1 gene (SPG31). *BMC Medical Genetics*, 9(1):71, 2008.

Schmidt, A., Wolde, M., Thiele, C., Fest, W., Kratzin, H., Podtelejnikov, A. V., Witke, W., Huttner, W. B., and SoÈling, H.-D. Endophilin I mediates synaptic vesicle formation by transfer of arachidonate to lysophosphatidic acid. *Nature*, 401(6749):133–141, 1999.

Schnell, J. R. and Chou, J. J. Structure and mechanism of the M2 proton channel of influenza A virus. *Nature*, 451(7178):591–595, January 2008.

Schrödinger, LLC. The PyMOL Molecular Graphics System. August 2010.

Schuldiner, M. and Schwappach, B. From rags to riches-the history of the endoplasmic reticulum. *Biochimica et Biophysica Acta*, 1833(11):2389–2391, 2013.

Schumann, F. H., Riepl, H., Maurer, T., Gronwald, W., Neidig, K.-P., and Kalbitzer, H. R. Combined chemical shift changes and amino acid specific chemical shift mapping of protein–protein interactions. *Journal of Biomolecular NMR*, 39(4):275–289, 2007.

Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Marius Clore, G. The Xplor-NIH NMR molecular structure determination package. *Journal of Magnetic Resonance*, 160(1):65–73, 2003.

Segrest, J. P. and Feldmann, R. J. Membrane proteins: amino acid sequence and membrane penetration. *Journal of Molecular Biology*, 87(4):853–858, 1974.

Sharpe, H. J., Stevens, T. J., and Munro, S. A Comprehensive Comparison of Transmembrane Domains Reveals Organelle-Specific Properties. *Cell*, 142(1):158–169, July 2010.

Shen, H., Pirruccello, M., and De Camilli, P. SnapShot: membrane curvature sensors and generators. *Cell*, 150(6), 2012.

Shen, Q., Uknes, S., and Ho, T. Hormone response complex in a novel abscisic acid and cycloheximide-inducible barley gene. *Journal of Biological Chemistry*, 268(31): 23652–23660, 1993.

Shen, Y. and Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *Journal of Biomolecular NMR*, 56(3):227–241, July 2013a.

Shen, Y. and Bax, A. Protein backbone and sidechain torsion angles predicted from nmr chemical shifts using artificial neural networks. *Journal of biomolecular NMR*, 56(3):227–241, 2013b.

Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J. M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K. K., Lemak, A., Ignatchenko, A., Arrowsmith, C. H., Szyperski, T., Montelione, G. T., Baker, D., and Bax, A. Consistent blind protein structure generation from NMR chemical shift data. *Proceedings of the National Academy of Sciences*, 105(12):4685–4690, 2008.

Shenkarev, Z. O., Lyukmanova, E. N., Paramonov, A. S., Shingarova, L. N., Chupin, V. V., Kirpichnikov, M. P., Blommers, M. J., and Arseniev, A. S. Lipid- Protein Nanodiscs as Reference Medium in Detergent Screening for High-Resolution NMR Studies of Integral Membrane Proteins. *Journal of the American Chemical Society*, 132(16):5628–5629, 2010.

Shibata, Y., Voss, C., Rist, J. M., Hu, J., Rapoport, T., Prinz, W. A., and Voeltz, G. K. The Reticulon and Dp1/Yop1p Proteins Form Immobile Oligomers in the Tubular Endoplasmic Reticulum. *Journal of Biological Chemistry*, 283(27):18892–18904, May 2008.

Shibata, Y., Hu, J., Kozlov, M. M., and Rapoport, T. A. Mechanisms shaping the membranes of cellular organelles. *Annual Review of Cell and Developmental*, 25:329–354, 2009.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1), 2011.

Simons, K. T., Kooperberg, C., Huang, E., and Baker, D. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *Journal of Molecular Biology*, 268(1):209–225, 1997.

Simons, K. T., Bonneau, R., Ruczinski, I., and Baker, D. *Ab initio* protein structure prediction of CASP III targets using ROSETTA. *Proteins: Structure, Function, and Bioinformatics*, 37(S3):171–176, 1999.

Stanek, J., Nowakowski, M., Saxena, S., Ruszczyńska-Bartnik, K., Ejchart, A., and Koźmiński, W. Selective diagonal-free ^{13}C,^{13}C-edited aliphatic–aromatic NOESY experiment with non-uniform sampling. *Journal of Biomolecular NMR*, 56(3):217–226, May 2013.

Stoffregen, M. C., Schwer, M. M., Renschler, F. A., and Wiesner, S. Methionine Scanning as an NMR Tool for Detecting and Analyzing Biomolecular Interaction Surfaces. *Structure*, 20(4):573–581, April 2012.

Suarez, A., Ueno, T., Huebner, R., McCaffery, J. M., and Inoue, T. Bin/Amphiphysin/Rvs (BAR) family members bend membranes in cells. *Scientific Reports*, 4, 2014.

Suzuki, M., Mao, L., and Inouye, M. Single protein production (SPP) system in Escherichia coli. *Nature Protocols*, 2(7):1802–1810, July 2007.

Tang, W.-F., Yang, S.-Y., Wu, B.-W., Jheng, J.-R., Chen, Y.-L., Shih, C.-H., Lin, K.-H., Lai, H.-C., Tang, P., and Horng, J.-T. Reticulon 3 binds the 2C protein of enterovirus 71 and is required for viral replication. *Journal of Biological Chemistry*, 282(8):5888–5898, 2007.

Terasaki, M., Shemesh, T., Kasthuri, N., Klemm, R. W., Schalek, R., Hayworth, K. J., Hand, A. R., Yankova, M., Huber, G., Lichtman, J. W., et al. Stacked endoplasmic reticulum sheets are connected by helicoidal membrane motifs. *Cell*, 154(2):285–296, 2013.

Terwilliger, T. C. The helical hydrophobic moment: a measure of the amphiphilicity of a helix. *Nature*, 299:371–374, 1982.

Tian, C., Vanoye, C. G., Kang, C., Welch, R. C., Kim, H. J., George, A. L., and Sanders, C. R. Preparation, functional characterization, and NMR studies of human KCNE1, a

voltage-gated potassium channel accessory subunit associated with deafness and long QT syndrome. *Biochemistry*, 46(41):11459–11472, 2007.

Tolley, N., Sparkes, I., Craddock, C. P., Eastmond, P. J., Runions, J., Hawes, C., and Frigerio, L. Transmembrane domain length is responsible for the ability of a plant reticulon to shape endoplasmic reticulum tubules *in vivo*. *The Plant Journal*, 64(3): 411–418, 2010.

Tugarinov, V. and Kay, L. E. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *Journal of the American Chemical Society*, 125(45):13868–13878, 2003.

Tugarinov, V., Muhandiram, R., Ayed, A., and Kay, L. E. Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase G. *Journal of the American Chemical Society*, 124(34):10025–10035, 2002.

Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E., and Kay, L. E. Cross-correlated relaxation enhanced ^{1}H-^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *Journal of the American Chemical Society*, 125(34):10420–10428, 2003.

Tugarinov, V., Choy, W.-Y., Orekhov, V. Y., and Kay, L. E. Solution NMR-derived global fold of a monomeric 82-kDa enzyme. *Proceedings of the National Academy of Sciences*, 102(3):622–627, 2005a.

Tugarinov, V., Kay, L. E., Ibraghimov, I., and Orekhov, V. Y. High-resolution four-dimensional ^{1}H-^{13}C NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. *Journal of the American Chemical Society*, 127 (8):2767–2775, 2005b.

Tugarinov, V., Kanelis, V., and Kay, L. E. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nature Protocols*, 1 (2):749–754, July 2006.

UniProt Consortium. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Research*, 42(D1):D191–D198, 2014.

Van Horn, W. D., Kim, H.-J., Ellis, C. D., Hadziselimovic, A., Sulistijo, E. S., Karra, M. D., Tian, C., Sönnichsen, F. D., and Sanders, C. R. Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. *Science*, 324(5935): 1726–1729, 2009.

van Meer, G. and de Kroon, A. I. Lipid map of the mammalian cell. *Journal of Cell Science*, 124(1):5–8, 2011.

Van Meer, G., Voelker, D. R., and Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nature Reviews Molecular Cell Biology*, 9(2):112–124, 2008.

Vasudevan, S. V., Schulz, J., Zhou, C., and Cocco, M. J. Protein folding at the membrane interface, the structure of Nogo-66 requires interactions with a phosphocholine surface. *Proceedings of the National Academy of Sciences*, 107(15):6847–6851, April 2010.

Veratti, E. Investigations on the fine structure of striated muscle fiber. *The Journal of Biophysical and Biochemical Cytology*, 10(4):1–59, 1961.

Viklund, H. and Elofsson, A. OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics*, 24(15):1662–1668, 2008.

Voeltz, G. K., Prinz, W. A., Shibata, Y., Rist, J. M., and Rapoport, T. A. A class of membrane proteins shaping the tubular endoplasmic reticulum. *Cell*, 124(3):573–586, February 2006.

Von Heijne, G. and Blomberg, C. Trans-membrane translocation of proteins. The direct transfer model. *European Journal of Biochemistry/FEBS*, 97(1):175–181, 1979.

Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: Structure, Function, and Bioinformatics*, 59(4):687–696, June 2005.

Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Hogbom, M., van Wijk, K. J., Slotboom, D. J., Persson, J. O., and de Gier, J.-W. Tuning *Escherichia coli* for membrane protein overexpression. *Proceedings of the National Academy of Sciences*, 105(38):14371–14376, 2008.

Wang, S., Romano, F. B., Field, C. M., Mitchison, T. J., and Rapoport, T. A. Multiple mechanisms determine ER network morphology during the cell cycle in Xenopus egg extracts. *The Journal of Cell Biology*, 203(5):801–814, 2013. doi: 10.1083/jcb.201308001. URL `http://jcb.rupress.org/content/203/5/801.abstract`.

Wang, T., Xiong, J.-Q., Ren, X.-B., and Sun, W. The role of Nogo-A in neuroregeneration: a review. *Brain Research Bulletin*, 87(6):499–503, 2012.

Weigert, R., Silletta, M. G., Spanò, S., Turacchio, G., Cericola, C., Colanzi, A., Senatore, S., Mancini, R., Polishchuk, E. V., Salmona, M., et al. CtBP/BARS induces fission

of Golgi membranes by acylating lysophosphatidic acid. *Nature*, 402(6760):429–433, 1999.

Willi, R. and Schwab, M. E. Nogo and Nogo receptor: relevance to schizophrenia? *Neurobiology of Disease*, 54:150–157, 2013.

Wimley, W. C. and White, S. H. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature Structural Biology*, 3(10):842–848, 1996.

Winner, B., Jappelli, R., Maji, S. K., Desplats, P. A., Boyer, L., Aigner, S., Hetzer, C., Loher, T., Vilar, M., Campioni, S., Tzitzilonis, C., Soragni, A., Jessberger, S., Mira, H., Consiglio, A., Pham, E., Masliah, E., Gage, F. H., and Riek, R. *In vivo* demonstration that α-synuclein oligomers are toxic. *Proceedings of the National Academy of Sciences*, 108(10):4194–4199, 2011. doi: 10.1073/pnas.1100976108. URL `http://www.pnas.org/content/108/10/4194.abstract`.

Wolfenden, R., Andersson, L., Cullis, P., and Southgate, C. Affinities of amino acid side chains for solvent water. *Biochemistry*, 20(4):849–855, 1981.

Wu, C.-H., Lee, S.-C., and Wang, C.-W. Viral protein targeting to the cortical endoplasmic reticulum is required for cell–cell spreading in plants. *The Journal of Cell Biology*, 193(3):521–535, 2011.

Wu, M.-J., Ke, P.-Y., Hsu, J. T.-A., Yeh, C.-T., and Horng, J.-T. Reticulon 3 interacts with NS4B of the hepatitis C virus and negatively regulates viral replication by disrupting NS4B self-interaction. *Cellular Microbiology*, 2014.

Yamamoto, Y., Yoshida, A., Miyazaki, N., Iwasaki, K., and Sakisaka, T. Arl6IP1 has the ability to shape the mammalian ER membrane in a reticulon-like fashion. *Biochemical Journal*, 458(1):69–79, 2014. doi: 10.1042/BJ20131186. URL `http://www.biochemj.org/bj/458/bj4580069.htm`.

Yang, D. and Kay, L. E. TROSY triple-resonance four-dimensional NMR spectroscopy of a 46 ns tumbling protein. *Journal of the American Chemical Society*, 121(11):2571–2575, 1999a.

Yang, D. and Kay, L. E. Improved ^{1}HN-detected triple resonance TROSY-based experiments. *Journal of Biomolecular NMR*, 13(1):3–10, 1999b.

Yang, D., Venters, R. A., Mueller, G. A., Choy, W., and Kay, L. E. TROSY-based HNCO pulse sequences for the measurement of ^{1}HN-^{15}N, ^{15}N-^{13}CO, ^{1}HN-^{13}CO, ^{13}CO-^{13}Cα and ^{1}HN-^{13}Cα dipolar couplings in ^{15}N, ^{13}C, ^{2}H-labeled proteins. *Journal of Biomolecular NMR*, 14(4):333–343, 1999.

Yang, S.-T., Lee, J. Y., Kim, H.-J., Eu, Y.-J., Shin, S. Y., Hahm, K.-S., and Kim, J. I. Contribution of a central proline in model amphipathic alpha-helical peptides to self-association, interaction with phospholipids, and antimicrobial mode of action. *The FEBS journal*, 273(17):4040–4054, September 2006.

Yang, Y. S. and Strittmatter, S. M. The reticulons: a family of proteins with diverse functions. *Genome Biol*, 8(12):234, 2007.

Yarov-Yarovoy, V., Schonbrun, J., and Baker, D. Multipass membrane protein structure prediction using Rosetta. *Proteins: Structure, Function, and Bioinformatics*, 62(4): 1010–1025, December 2005.

Zemel, A., Ben-Shaul, A., and May, S. Modulation of the spontaneous curvature and bending rigidity of lipid membranes by interfacially adsorbed amphipathic peptides. *The Journal of Physical Chemistry B*, 112(23):6988–6996, 2008.

Zeth, K., Diederichs, K., Welte, W., and Engelhardt, H. Crystal structure of Omp32, the anion-selective porin from *Comamonas acidovorans*, in complex with a periplasmic peptide at 2.1 Å resolution. *Structure*, 8(9):981–992, 2000.

Zhang, Y.-Z. *Protein and peptide structure and interactions studied by hydrogen exchange and NMR*. PhD thesis, University of Pennsylvania, PA,USA, 1995.

Zhao, X., Alvarado, D., Rainier, S., Lemons, R., Hedera, P., Weber, C. H., Tukel, T., Apak, M., Heiman-Patterson, T., Ming, L., et al. Mutations in a newly identified GTPase gene cause autosomal dominant hereditary spastic paraplegia. *Nature Genetics*, 29(3):326–331, 2001.

Zimmerberg, J. and Kozlov, M. M. How proteins produce cellular membrane curvature. *Nature Reviews. Molecular Cell Biology*, 7(1):9–19, January 2006.

Züchner, S., Wang, G., Tran-Viet, K.-N., Nance, M. A., Gaskell, P. C., Vance, J. M., Ashley-Koch, A. E., and Pericak-Vance, M. A. Mutations in the novel mitochondrial protein REEP1 cause hereditary spastic paraplegia type 31. *American Journal of Human Genetics*, 79(2):365–369, 2006.

Zurek, N., Sparks, L., and Voeltz, G. Reticulon short hairpin transmembrane domains are used to shape ER tubules. *Traffic*, 12(1):28–41, January 2011.