Accepted Manuscript

Bacterial expression, purification and biophysical characterization of the smallest plant reticulon isoform, RTNLB13

Michael Chow, Meropi Sklepari, Lorenzo Frigerio, Ann M. Dixon
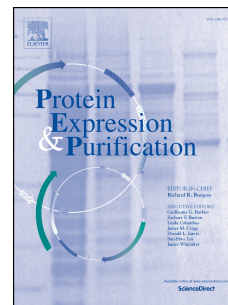
Please cite this article as: M. Chow, M. Sklepari, L. Frigerio, A.M. Dixon, Bacterial expression, purification and biophysical characterization of the smallest plant reticulon isoform, RTNLB13, *Protein Expression and Purification* (2018), doi: 10.1016/j.pep.2018.06.015.

Bacterial expression, purification and biophysical characterization of the smallest plant

reticulon isoform, RTNLB13

Michael Chow [+], Meropi Sklepari[§], Lorenzo Frigerio[‡], and Ann M. Dixon[†*]

Running Title: Expression and characterization of RTNLB13

[+] MOAC Doctoral Training Centre, § Warwick Centre for Analytical Science, ‡ School of Life Sciences and † Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK.

*To whom correspondence should be addressed: Dr Ann Dixon, Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK, Telephone: +44 2476 150037; FAX: +44 2476 524112; email: ann.dixon@warwick.ac.uk

ABSTRACT

Reticulons are a large family of integral membrane proteins that are ubiquitous in eukaryotes and play a key role in functional remodelling of the endoplasmic reticulum membrane. The reticulon family is especially large in plants, with the *Arabidopsis thaliana* genome containing twenty-one isoforms. Reticulons vary in length but all contain a conserved C-terminal reticulon homology domain (RHD) that associates with membranes. An understanding of the structure and membrane interactions of RHDs is key to unlocking their mechanism of function, however no three-dimensional structure has been solved. We believe that this is, in part, due to difficulties in obtaining reticulon proteins in yields sufficient for structural study. To address this, we report here the first bacterial overexpression, purification, and biophysical investigation of a reticulon protein from plants, the RTNLB13 protein from *A. thaliana*. RTNLB13 is the smallest plant reticulon and is made up of a single RHD. We used circular dichroism, SDS-PAGE and analytical ultracentrifugation to reveal that RTNLB13 is 45% α-helical in a number of detergent environments, monomeric at low concentrations, and capable of self-association at higher concentrations. We used solution-state NMR to screen the effect of detergent type on the fold of isotopically-enriched RTNLB13, and found that ~60% of the expected protein peaks were broadened due to slow dynamics. This broadening points toward a large network of protein-membrane interactions throughout the sequence. We have interpreted our results in light of current literature and suggest a preliminary description of RTNLB13 structure and topology.

KEYWORDS

INTRODUCTION

The endoplasmic reticulum (ER) is the entry of point of the secretory pathway in eukaryotic cells [1]. It is a very dynamic, polymorphic network of membrane sheets and tubules [2], which is anchored to the cytoskeleton and undergoes constant remodelling [3]. In plants, the ER is the key site of synthesis of storage proteins and lipids. In some species, the ER itself serves as a protein storage compartment [4]. The unique morphology of the ER depends on proteins that can shape its membrane into tubules or sheets [5].

Reticulons are a large family of integral membrane proteins that have been demonstrated to be both necessary and sufficient to induce the formation of ER tubules, both *in vivo* [6, 7] and *in vitro* [8]. Reticulons have been shown to be a requirement for embryonic viability in *Caenorhabditis elegans* [9], and are ubiquitous in eukaryotes, suggesting a general requirement for reticulons. In higher eukaryotes, including humans, reticulons appear to have a role in neurological disorders with the best studied example, RTN4-A or Nogo, re-igniting interest in this protein family twelve years ago [10, 11]. They have also been implicated in apoptosis, neurite growth and interacting with the BACE enzymes in mammalian systems [12]. Recently, GTPases such as RHD3 and atlastins were shown to interact with the reticulons and to further contribute to ER network formation [8]. The reticulon gene family is remarkably large in higher plants [10], referred to as reticulon-like proteins family B (RTNLB). Indeed there appear to be many more isoforms in plants than are found in mammalian or yeast systems, with the *Arabidopsis thaliana* genome containing at least 21 reticulon isoforms [5] as compared to four in mammalian systems (RTN1-4)[11] and two in yeast. When overexpressed in plant cells, reticulons are able to remodel the ER membrane by introducing constrictions in the tubular ER network, which limit diffusion of soluble proteins in its lumen [6, 13]

Reticulons vary considerably in length (from 200–1200 residues), but share a conserved C-terminal region called the reticulon homology domain (RHD) which facilitates association with the membrane via two large hydrophobic regions [10]. The shortest plant reticulon isoform, RTNLB13 from *A. thaliana* (206 amino acids), is composed almost exclusively of a single RHD and has been shown to remodel large ER cisternae into tubules *in vivo* [14]. While the topology of the RHD has been the subject of some debate (i.e. whether it forms a two transmembrane (TM) domain V-shape vs. a four TM domain W-shape in the membrane, Fig. 1A), *in vivo* data suggest that five reticulon isoforms (including RTNLB13) from *A. thaliana* each contain four TM domains within the RHD, forming a W shape in the ER membrane and placing the N- and C-termini and the loop between TM domains 2 and 3 in the cytosol [13]. This topology has been further supported in a recent structural study by Brady and co-workers of the RHD-containing Yop1p protein from *Saccharomyces cerevisiae* [15].

The tubule-forming activity of reticulons has been proposed to derive from a combination of wedging – i.e. the ability of reticulons to cause higher displacement of the outer leaflet of the ER membrane – and scaffolding, which is caused by several wedge-shaped reticulons forming low-mobility oligomeric assemblies within the membrane [9]. The wedging property of reticulons seems to be a function of both their topology and the unusual length of their TM regions. Both shortening the TM domains of a plant reticulon [14] and lengthening the TM domains of a mammalian reticulon [16] lead to a loss of both tubule-forming function and the capacity of these proteins to oligomerise. The general conclusion from these experiments is that the length and number of the TM domains is crucial to reticulon function.

Despite these studies, as yet there is no three-dimensional structure elucidating the exact TM toplogy of reticulons or their mode of function. This is in part due to the significant challenges of

protein production. Indeed, over the past 10 years very few laboratories have been able to demonstrate successful overexpression, purification and reconstitution of recombinant RHDs in yields suitable for structural or functional measurements, with only two reports in *E. coli* [15, 17] and two in yeast [7, 8]. The few high-resolution structural studies reported have focussed primarily on the soluble loops and N- and C-termini of mammalian isoforms, specifically human RTN1-C [18], human RTN3 [19] and human RTN4 (also known as the neurite outgrowth inhibitor or Nogo) [20-22]. These studies have revealed metal-binding properties of the C-terminus of RTN1-C [18], the intrinsic disorder of the N- and C-terminal tails of human RTN3 and RTN4[19, 21] (proposed to confer multi-functionality) and the stable folding of the extracellular loop of human RTN4 (called Nogo-66) on the surface of membrane mimetics [20, 22]. Only twice have high-resolution structural analyses been reported for the entire RHD (i.e. including TM domains): the first was for human RTN1-C [17] and more recently the Yop1p RHD from *S. cerevisiae* was investigated [15]. In both studies, the RHDs were shown to contain significant α-helical content and form oligomers. Brady and coworkers went further to suggest a structural model for Yop1p RHD based on NMR chemical shift and relaxation measurements [15] which included a putative C-terminal amphipathic helix thought to play a key role in tubule formation *in vitro*. Such an amphipathic helix has also been proposed at the C-terminus of the plant reticulon RTNLB13 [23].

As RTNLB13 is both the smallest and best-studied plant reticulon isoform, it is an ideal candidate for further structural investigation of the RHD. Toward this end, we report here the first bacterial overexpression, purification, and preliminary biophysical characterization of RTNLB13. We have optimized expression of this protein in yields suitable for structural characterization, developed a protocol for reconstitution in a range of membrane mimetic environments, and have used circular dichroism and analytical ultracentrifugation to characterize

5

the secondary structure and oligomeric state of the protein. We have also used solution NMR to screen the effect of detergent type on the quality of NMR spectra for isotopically-enriched RTNLB13 samples in order to identify conditions for future structural investigations. Our results suggest common features between RHDs from plants, yeast and mammals (despite very low sequence conservation) and propose a preliminary model to describe RTNLB13 structure and topology.

EXPERIMENTAL METHODS

Protein expression and membrane harvesting.

The *A. thaliana* RTNLB13 gene was codon optimized for *E. coli* and an N-terminal octahistidine tag was added at the N-terminus. The optimized His_8-RTNLB13 gene was synthesized by Genscript (Piscataway, NJ), amplified using PCR and digested with *NcoI* and *BamHI*. The digested PCR product was ligated into a similarly digested pET28a vector (Novagen) and transformed into *E. coli* expression strains BL21 (DE3), C41 (DE3) or C43 (DE3). Cells expressing unlabeled protein were grown at 37°C to $OD_{600} = 0.6–0.7$ before induction with 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), at which point the temperature was reduced to 30°C and overexpression was carried out for 4–6 hours. To prepare ^{15}N-labelled protein, cells were grown at 37 °C to $OD_{600} = 0.6–0.7$. Cells were then harvested gently by centrifugation ($1500 \times g$), washed twice with $1 \times$ M9 minimal salts medium (0.4% D-glucose, 0.1% NH_4Cl, 50 mM $Na_2HPO_4 \cdot 7H_2O$, 50 mM KH_2PO_4, 5 mM NaCl and 10 mM $MgSO_4$), and resuspended in M9 containing 2 g/L glucose, 1 g/L ^{15}N ammonium chloride, $1 \times$ Basal Medium Eagle (BME) vitamins, and 100 µM iron (III) chloride. The resuspension volume was either half or one quarter of the Luria broth (LB) volume to achieve a two-fold (or four-fold)

6

concentration in cell density according to a method published by Marley and coworkers [24]. Cells were then incubated for 1 hour at 20°C, induced with 1 mM IPTG, and overexpression was carried out over 24 hours at 180 rpm. Triply (^{13}C, ^{15}N, and ^{2}H) labeled protein was prepared as above with the addition of uniformly ^{13}C-labelled glucose (4 g/L) to the M9 media and substitution of 100% deuterium oxide for water. Expression levels were monitored by Western blotting against the His$_8$ tag, and a band near the theoretical molecular weight of the His$_8$-RTNLB13 product (24.9 kDa) indicated protein expression.

Cell pellets from 1 L of culture were resuspended in lysis buffer (20 mg/mL phosphate buffer or 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mg/mL lysozyme, 5 ul DNAse (2 U/ul), 1 × protease inhibitor cocktail tablet (Roche, UK)) and incubated on ice for 30 minutes. The cells were lysed by sonication, and unbroken cells and cell debris were isolated from buoyant membrane fragments by a low speed centrifugation step (15,000 × g, 12 min., 4°C). Ultracentrifugation (150,000 × g, 45 mins.) was used to pellet the cell membranes from the supernatant. The membrane pellet was resuspended in phosphate buffered saline (25 mM Na$_2$HPO$_4$, 137 mM NaCl, pH 7.3) containing 10% glycerol and 1% n-dodecyl-β-d-maltopyranoside (DDM) detergent, and rotated overnight at 4°C. Insoluble membranes were pelleted by ultracentrifugation at 150,000 × g. The supernatant and the pellet were then analyzed by SDS-PAGE and immunoblotting against the His$_8$ tag. The resulting bands were analyzed using ImageJ [25] to quantify the fraction of protein solubilized.

Purification of RTNLB13 from membranes.

Purification was performed using immobilized metal affinity chromatography (IMAC) by incubating 1 mg/mL Talon cobalt resin (BD Biosciences, UK) with clarified, DDM-solubilized

membranes for 30 minutes at 4°C. Resin was separated by centrifugation (700 × g, 5 min.), resuspended in 5 mL 50 mM Na_2HPO_4, 300 mM NaCl, 0.1% detergent, pH 7.0 (wash buffer), and loaded into a 5 mL disposable column (Thermo Scientific Pierce, UK). The wash and elution steps were performed at 4°C in buffer containing 0.1 % w/v detergent and either 10 mM or 150 mM imidazole, respectively, according to the manufacturer's instructions. Fractions containing protein were pooled and concentrated using an Amicon Ultra centrifugal filter unit (Millipore) with a molecular weight cutoff (MWCO) of 5–10 kDa to a final volume 500 µl. A second purification step and detergent exchange into one of a panel of detergents was carried out using size exclusion chromatography (SEC), specifically a Superdex 200 10/300 GL gel filtration column connected to an Akta purification system (GE Healthcare/Amersham Biosciences, UK), equilibrated with 25 mM Na_2HPO_4, 150 mM NaCl, 0.05% detergent, pH 7.0. The panel of detergents consisted of 1,2-diheptanoyl-sn-glycero-3-phosphocholine (C_7-DHPC), 1,2-dihexanoyl-sn-glycero-3-phosphocholine (C_6-DHPC), sodium-n-dodecyl-sulphate (SDS), dodecylphosphocholine (DPC), 14:0 lyso-phosphatidylglycerol (LMPG), and 16:0 lyso-phosphatidylglycerol (LPPG) (all obtained from Sigma Aldrich and Avanti Polar lipids). Following SEC, His-RTNLB13 was concentrated using a 30 kDa MWCO Amicon concentrator unit. The identity and purity of the protein was confirmed by peptide mass fingerprinting analyses carried out at the School of Life Sciences, University of Warwick using electrospray ionization mass spectrometry (ESI-MS). Protein concentration estimation was performed using either the bicinchoninic acid (BCA) assay or UV-Vis absorbance at 280 nm and the theoretical extinction coefficient of RTNLB13 ($\varepsilon_{280} = 50,085$ M^{-1} cm^{-1}).

Sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE)

The purity of all protein fractions at each stage of preparation was assessed using SDS-PAGE (visualized using both Coomasie Blue R250 and silver nitrate) and immunoblotting against the His$_8$ tag. SDS-PAGE gels (12%) were prepared according to the Laemmli protocol [26]. Pre-cast 12% NuPAGE Bis-Tris gels (Invitrogen, UK) were also used according to the manufacturer's instructions. Samples for SDS-PAGE were prepared by mixing RTNLB13 samples with SDS-PAGE loading buffer (125 mM pH 6.8, 20% glycerol, 4% SDS, 0.02 % bromophenol blue, 5 mM dithiothreitol (DTT)) or NuPAGE LDS sample buffer (Invitrogen) and heating at 37°C for 10 minutes or 70°C for 10 minutes, and were run alongside See blue Plus 2 (Invitrogen, UK) or Color-plus 7-175 kDa (New England Biolabs, UK) prestained markers.

Circular Dichroism

CD spectra were collected on Jasco J-815 or J-720 spectropolarimeters (Jasco UK, Great Dunmow, UK) equipped with Peltier temperature control and xenon light sources. Samples typically contained 0.14–0.24 mg/mL (6–10 µM) His$_8$-RTNLB13, 20 mM sodium phosphate (pH 7.0), 50 mM NaCl, and the detergent of interest at concentrations ranging from 15–20 mM. Spectra were recorded between 190 nm and 280 nm, with a bandwidth of 2 nm and a data pitch of 0.2 nm. Buffer spectra were also recorded for baseline subtraction. Protein concentrations were approximated via the Beer Lambert law using absorbance of the protein at 280 nm (A$_{280}$). The machine units of mdeg were converted to mean residue ellipticity (MRE) to normalize to protein concentration. CD data were evaluated using a self-organizing map algorithm, Secondary Structure Neural Network (SSNN), which provided a prediction of the protein's secondary structure and an independent estimate of the samples' concentration [27]. Data below 196 nm were

9

excluded as the quality of the spectra in this region was less reliable and would negatively impact the accuracy of structure predictions. SSNN was retrained for this work using parameters according to the developer's instructions [28], and CD data were fit using the newly trained neural network after applying scaling factors from 0.05–0.8 in steps of 0.05 or 0.01 to account for inaccurate protein concentration estimates. The best fit of the experimental data was selected based on the root mean squared deviation and visual evaluation.

Analytical Ultracentrifugation

Analytical ultracentrifugation was performed on a Beckman XL-I ultracentrifuge fitted with interference and absorbance optics, at the School of Biological Sciences, University of Birmingham. Three protein concentrations (8–20 µM), corresponding to A_{280} values of 1.0, 0.7 and 0.4, were prepared in 50 mM Tris HCl (pH 7.2) containing 100 mM NaCl, 15 mM DPC, and 52.5% D_2O. These samples were loaded into a 6 cell center-piece outfitted with sapphire windows alongside buffer reference blanks, and centrifuged at 20,000, 28,000, and 35,000 rpm for 36 hours at 4°C. Absorbance data were analysed and processed with WinReed and WinMatch, before fitting in WinNonlin [29] as described previously [30].

Nuclear magnetic resonance experiments

Samples for NMR analyses were prepared by concentrating purified protein to 180 µL with an Amicon concentrator, either with a 3 or 10 kDa cutoff. If required, detergent was added to the sample following removal from the concentrator unit and sonicated for 5 minutes. The sample pH was checked and, if needed, adjusted with 0.5 M acetic acid, prior to pipetting the sample into a 3 mm NMR tube. Typical protein concentrations were 0.1–0.2 mM; at pH values

between 6–7, the buffer used was 10–20 mM Na_2HPO_4, 40 mM NaCl, 90% H_2O, 10% D_2O, with

the required concentration of detergent. For experiments performed between pH 5–6, 20 mM

sodium acetate buffer, containing 40 mM NaCl, 90% H_2O, 10% D_2O was used. NMR spectra

were recorded with 2048×128 data points or 2048×256 data points on either an Avance 700

MHz spectrometer or an Avance 600 MHz spectrometer (Bruker Biospin), the former equipped

with a triple resonance inverse cryoprobe with Z-gradients. The 600 MHz instrument was fitted

with a TBI inverse, room temperature, triple resonance probe with Z-gradients. Data was

processed in Topspin 2.1 (Bruker Biospin).

RESULTS AND DISCUSSION

E.coli expression, purification, and characterization of RTNLB13 from A. thaliana

Three *E.coli* strains were tested for expression levels of His_8-RTNLB13, namely

BL21(DE3), and the C41 and C43 strains optimized for membrane protein expression and

referred to as the Walker strains[31]. As shown in Fig. 1B, upon testing two expression

temperatures and two IPTG concentrations, the BL21 (DE3) strain was found to produce the

largest amount of protein when grown in LB medium. Western blotting revealed two intense

bands migrating at the expected monomeric and dimeric molecular weights (predicted

monomeric MW = 23.8 kDa), and also showed evidence of aggregated protein in the wells of the

gel. Differential centrifugation indicated that the bulk of His_8-RTNLB13 was located in the

membranes (Fig. 1C). When the medium was changed to M9 minimal medium, containing ^{15}N

ammonium chloride (1 g/L) as the primary nitrogen source and glucose (2 g/L) as the sole carbon

source, protein expression levels fell below detection. Because low levels of trace metals can

substantially affect cell health, and it has been shown that addition of 100 µM iron can increase

expression to similar levels achieved when a 'full' trace metal mixture was utilized [32], 100 µM iron (III) chloride was added to the media. As shown in Fig. 1D, expression was restored upon addition of iron and 1 mM IPTG. The glucose concentration was also varied, however had a negligible effect on expression levels of singly ^{15}N-labelled protein and was therefore maintained at 2 g/L. The glucose concentration was increased to 4 g/L (or even 8 g/L) to express doubly (^{13}C/^{15}N) and triply (^{13}C/^{15}N/^{2}H) labeled His$_8$-RTNLB13, although His$_8$-RTNLB13 expression levels were severely diminished upon addition of ^{13}C glucose and deuterium oxide (data not shown) making further investigation of the doubly and triply labeled protein intractable. Subsequent analyses were carried out on the uniformly ^{15}N-labeled protein.

Dodecyl-β-D-maltoside (DDM) detergent was used to extract His$_8$-RTNLB13 from cell membranes. DDM demonstrated good efficiency in solubilizing His$_8$-RTNLB13 and is considered to be a "mild" detergent that is commonly used in membrane protein studies and thought to maintain protein structure during purification. Furthermore, DDM does not interfere with charged binding interfaces as employed by IMAC-based technologies, preventing a reduction in recovery. His$_8$-RTNLB13 was initially purified by application of DDM-solubilized membranes to cobalt-charged NTA resin followed by elution with imidazole into a 0.1% w/v detergent solution. SDS-PAGE and immunoblotting were used to monitor binding and elution (Fig. 1E). Protein bands were only visible upon Western blotting against the poly-His tag: Coomassie blue staining (see upper panel of Fig. 1E) was consistently ineffective at detecting His$_8$-RTNLB13. To resolve these detection issues, silver staining was used, however silver nitrate was also unable to bind to His$_8$-RTNLB13. This behavior has been reported in the past, for example Newstead and coworkers [33] were only able to detect overexpression of some membrane protein-GFP fusions through fluorescence detection because Coomassie staining was

12

not effective. For His_8-RTNLB13, the cause of this poor staining is unknown, but is not due to a lack of basic residues (Coomassie primarily binds to arginine, lysine and histidine) since there are 30 such residues in the protein. Nevertheless, we were able to use a combination of Coomassie staining and Western blotting to confirm the presence and purity of overexpressed His_8-RTNLB13.

Identification of optimal detergent for structural analyses of RTNLB13

As demonstrated previously for the yeast reticulon-like protein Yop1p[15], NMR has the potential to shed new light on the topology and three-dimensional structure of RHD-containing proteins in membrane mimetics such as detergent micelles. However, while DDM is a suitable detergent for purification, the micelles formed by this non-ionic detergent are very large (50–70 kDa) and intractable for solution-state NMR studies. The current lack of a 'universal' detergent that optimally solubilizes all membrane proteins and has properties suitable for solution NMR structural studies means that empirical studies must be carried out to ascertain which detergents yield well-folded and concentrated membrane protein samples. To this end, we screened a panel of detergents for their ability to solubilize NMR-viable concentrations of His_8-RTNLB13 in a folded state. To select detergents for testing, solution NMR-based structural studies curated on the Membrane Protein NMR Database [http://www.drorlist.com/nmr/MPNMR] were analyzed, revealing that membrane protein structures have been successfully solved in DPC, C_6- and C_7-DHPC, LMPG, LPPG and SDS. Therefore, these detergents were investigated here for their ability to solubilize and stabilize His_8-RTNLB13. The properties of these detergents are summarized in Table 1 and demonstrate that all of the detergents tested (apart from LPPG) have a small micelle size (micellar molecular weights (MW_{mic}) ≤ 26 kDa) and either zwitterionic or

charged head groups. They have a variety of critical micelle concentrations (CMC = 18 μM – 14 mM) and acyl chain lengths (6 – 16 carbons in length), but all detergents tested have fully saturated acyl chains.

Because DPC and LMPG are the detergents in which the largest number of membrane protein structures have been solved in solution, His$_8$-RTNLB13 was initially exchanged from DDM into each of these detergents on a size exclusion chromatography (SEC) column, equilibrated with the appropriate buffer and detergent. This also acted as a further purification step producing highly pure protein.

Effect of detergent concentration, temperature, and pH on His$_8$-RTNLB13 NMR spectra

His$_8$-RTNLB13 was reconstituted into both DPC and LMPG at a variety of detergent concentrations to investigate the role of this parameter in the quality of solution state NMR spectra (i.e. peak number, intensity, width, and chemical shift dispersion). All spectra were acquired with 128×2048 data points and 40 scans for samples containing $0.1 - 0.13$ mM His$_8$-RTNLB13. We selected detergent concentrations that yielded micelle : protein (M/P) molar ratios of 1, 8, 16 and 32, calculated using the values for aggregation number and CMC given in Table 1[34-40]. As shown in the ^{1}H-^{15}N HSQC spectra of DPC-solubilized His$_8$-RTNLB13 in Figs. 2A-C, increasing the detergent concentration from 10 mM (M/P = 1) to 200 mM (M/P = 32), at pH 6.8 and 25°C, did not lead to a large increase in the number of peaks observed (provided in parentheses) but did lead to sharpening of the peaks and more consistency in peak intensities across the spectrum. A similar improvement of the spectrum was observed in LMPG from concentrations between 60 mM (M/P = 8) and 230 mM (M/P = 32) (see Fig. S1), except that in this detergent the number of protein peaks did increase with detergent concentration.

14

The temperature at which NMR spectra are collected can also affect the number and width of observed peaks. Often a high temperature (> 37°C) results in improvement of the spectrum due to an increase in the tumbling rate of the membrane protein – micelle complex and subsequent sharpening of the peaks. In this study spectra were acquired for His_8-RTNLB13 solubilized in either 200 mM DPC or 230 mM LMPG at temperatures ranging from 25-45°C. Comparison of Fig. 2C and 2D demonstrates that several new peaks (we estimate seven) were observed in the DPC-solubilized His-RTNLB13 spectrum upon increasing the temperature from 25°C to 37°C. This effect was observed in both detergents (although data is only shown for DPC), however further increases in temperature (40-45°C) lead to protein precipitation and / or aggregation. Therefore 37°C was selected as the optimal temperature for further analyses.

Finally, the effect of solution pH on NMR spectra was investigated as this parameter can have a large impact on the quality of the spectrum. Reducing the pH below physiological pH is beneficial as it reduces the amide proton exchange rate and leads to sharpening of amide peaks. Therefore, pH values as low as pH 4.0 are regularly used for structure determination by NMR. ^1H-^{15}N HSQC spectra of His_8-RTNLB13 solubilized in either 200 mM DPC or 230 mM LMPG were acquired at pH 6.8 and pH 5.0 (37°C), and the results are shown in Figs. 2E (DPC) and S1 (LMPG). In both detergents, the widths of the observed peaks were not significantly affected by changing the pH, nor was there a noticeable increase in chemical shift dispersion. However, the intensity and number of peaks observed in the HSQC spectrum increased by 20-30% as the pH was decreased to 5.0.

Detergent screen using optimized detergent concentration, pH and temperature

To extend the number of detergents tested to include the six major detergents used thus far to solve membrane protein solution structures, His_8-RTNLB13 was also reconstituted in C_6- and C_7-DHPC, SDS, and LPPG in addition to DPC and LMPG. All samples were prepared using the optimal conditions determined from the experiments described above, i.e. detergent concentrations at or above 100 mM, a low pH (pH 5.0) and a temperature of 37°C, and ^1H-^{15}N HSQC spectra were acquired with 128×2048 data points and 40 scans. The one exception was the sample prepared in SDS, in which precipitation was observed at low pH therefore the sample was prepared at pH 6.8 and the spectrum was acquired at 15°C.

The resulting HSQC spectra are shown in Fig. 3 and the number of observed NMR peaks in each is summarized in Table 1. Spectra acquired in C_6- and C_7-DHPC yielded very few peaks (38 and 26, respectively), and had the lowest signal-to-noise ratio of all the data collected. This is likely due to the low stability of His_8-RTNLB13 in these detergents. Indeed, protein precipitation was observed in these samples over days (not weeks) suggesting slow protein aggregation. This may be linked to the short acyl chain length in these detergents (Table 1) failing to fully solubilize the protein. SDS yielded a slightly improved spectrum in terms of signal-to-noise ratio and number of observable peaks (48 peaks), but poor chemical shift dispersion (likely caused by the denaturing effects of this anionic detergent) eliminated this sample as a candidate for further study. The spectrum acquired in LPPG yielded 65 peaks, equivalent to that observed in DPC and LMPG, but the variable intensity of peaks in this spectrum appeared more pronounced. Therefore, our results suggest that DPC and LMPG yield the most stable samples and the highest quality NMR spectra. With this in mind, we acquired HSQC spectra with a larger number of points (2048×256) and scans (ns = 160) to increase the

16

signal to noise ratio, and the number of observable NMR peaks increased to 107 in 200 mM DPC and 114 in 230 mM LMPG (see values in parentheses in Table 1).

Oligomerization of His-RTNLB13 in detergent micelles

The maximum number of peaks (114 peaks) observed in the HSQC spectrum of His_8-RTNLB13 accounted for only 40% of the number of peaks predicted from the amino acid sequence of His_8-RTNLB13 (214 backbone peaks and 70 side chain = 284 peaks). The greatly reduced number of observable NMR peaks prevented assignment of the protein, and could be due to oligomerization of His_8-RTNLB13 and subsequent relaxation broadening of protein signals. The immunoblot shown in Fig. 1E (bottom panel) indicated the formation of SDS-resistant dimers and higher-order His-RTNLB13 oligomers. The elution fraction containing the highest concentration of His_8-RTNLB13 (elution fraction 3) yielded bands at the monomeric molecular weight (band labeled *M* at ~25 kDa), dimeric molecular weight (band labeled *D* at ~50 kDa) and higher molecular weights (band labeled *n* at > 64 kDa). This oligomerization is unsurprising as RTNLB13 has been reported to self-associate *in vivo* [9, 13]. To further investigate the strength of RTNLB13 oligomerization, sedimentation equilibrium analytical ultracentrifugation experiments were performed for the His_8-RTNLB13 protein solubilized in the detergent DPC. DPC is highly amenable to AUC measurements and has been used several times in the past for membrane protein analyses [30, 41, 42]. Samples contained $8 - 20$ µM His_8-RTNLB13 and 15 mM DPC to achieve micelle: protein molar ratios similar to those used in NMR analyses (M/P = $9 - 24$). The concentration versus radial distance profiles collected for His_8-RTNLB13 at three protein concentrations and three speeds are shown in Fig. 4. Data were initially fit to a monomer model as well as models of higher oligomer formation (e.g monomer-dimer), and the

resulting variance and residuals of the fit were used as a measure of the goodness-of-fit of a given model. Surprisingly, the model which best represented the DPC-solubilized His_8-RTNLB13 data was the simple monomer model, and the fit curve is shown with the experimental data in Fig. 4. Addition of higher order oligomeric states to the model resulted in a poor fit of the data and obvious trends in the residuals. Thus the higher order oligomers observed in Western bolts were not observed using analytical ultracentrifugation. These results do not rule out formation of RTNLB13 oligomers *in vivo* (indeed, we see oligomers of the protein consistently on gels), but do suggest that the protein is predominantly monomeric under NMR conditions and the absence of NMR peaks is not likely due to protein oligomerization.

Secondary structure of His-RTNLB13 in various detergents using CD

The secondary structure content of RTNLB13 has not yet been reported, and may shed light on the structure and topology of the protein and complement the level of information available from NMR. With this in mind, we used both theoretical and experimental tools to study the secondary structure of RTNLB13. The Jpred3 [43] and Psipred [44] algorithms were used initially to predict the secondary structure content from the primary sequence, and yielded very similar values (~ 68% α-helical/5% β-sheet, with predicted helical regions shown in Fig. 5A). CD was then used to estimate the secondary structure content of purified His_8-RTNLB13 in the panel of detergents studied here. Because of the poor NMR spectra obtained in C_6- and C_7-DHPC, these detergents were not included in this analysis. Fig. 5B shows the resulting CD spectra upon solubilization of 6–10 μM His_8-RTNLB13 in a variety of detergent types (typically at a concentration of 15–20 mM). All spectra contain negative maxima at 210 nm and 222 nm, and a positive peak at 197 nm, suggesting the protein has significant α-helical content in each detergent

18

tested. Data below 195 nm were not considered due to light scattering and high absorbance reducing the quality of the data. The CD data were fit using the SSNN method [27, 28] (see Fig. S2 for fits) in order to estimate the secondary structure content in each detergent environment, and the results are summarized in Fig. 5C. DPC, LMPG, and SDS-solubilized His_8-RTNLB13 yielded very similar secondary structure content, with between 44–46% α-helix, 9–12% β-sheet, and 13–20% turns, suggesting the protein is well-solubilized and adopts a similar fold in these detergents. DDM-solubilized His_8-RTNLB13 had a slightly higher helical content (54%) while LPPG yielded the lowest helical content (31%). Broadly speaking, the CD data suggested a helical content for reconstituted His_8-RTNLB13 that was much lower than the estimates obtained from the primary sequence (i.e. Jpred3 and Psipred). However, when protein topology was used as an estimate of helical content, predicted and experimental values were in much closer agreement. Here we used a panel of seven topology-prediction algorithms[44-51] to analyze the primary sequence of RTNLB13. The majority suggested that the protein contains four TM regions of approximately 18–24 residues each (Fig. 6A), with an average of 85 residues embedded in the membrane. If one assumes that all of the predicted membrane-spanning residues are present as α-helices, this would suggest a helical content of approximately 40% (see bar labeled "predicted", Fig. 5C) for the 214-residue His_8-RTNLB13. This range is in good agreement with the experimental results from CD, which indicated 44–46% α-helical content for RTNLB13 solubilized in LMPG, DPC or SDS.

CONCLUSIONS

We report here the first heterologous expression of a plant reticulon, RTNLB13 from *A. thaliana,* in *E. coli* for biophysical and structural analyses. Protein expression was achieved

without a fusion partner, and the resulting protein was readily soluble in detergent micelles. The CD, SDS-PAGE and AUC results demonstrated that our recombinant RTNLB13 protein was highly α-helical and capable of forming oligomers, but predominantly monomeric under NMR conditions making it a suitable candidate for further study using high-resolution NMR methods. We carried out an initial optimization of sample conditions for NMR studies, screening a number of detergent concentrations and types, as well as pH and temperature. There was little correlation between physical properties of the detergent (e.g. chain length, head group polarity, CMC) and the ability to solubilize RTNLB13, further highlighting the requirement to test a range of detergents when working with membrane proteins. Although steady improvement in the quality of the NMR spectrum was achieved, the dynamics of His_8-RTNLB13 in detergent micelles lead to severe broadening of over 60% of the expected peaks thus impeding attempts at assignment. Further work is required to establish solution conditions amenable to full backbone assignment.

Apart from demonstration of methods for efficient bacterial overexpression and reconstitution of RTNLB13 for further structural analysis, we sought to shed light on the structure, membrane interactions and topology of this plant reticulon. Topology prediction and secondary structure estimates from CD data lead us to suggest there are likely four TM regions in His_8-RTNLB13 (and scope for additional helical structure, which will be discussed later). The residues spanning the membrane bilayer (or detergent micelle, in this case) will be highly structured and will reorient more slowly in solution than residues present in the soluble regions of the protein. The reduced dynamics of the TM residues would lead to broad, undetectable signals in solution NMR experiments while the signals from the more mobile, soluble protein regions are easily detectable. This phenomenon has been exploited in the past to map the location of membrane-spanning regions for a variety of membrane proteins [52], and offers a potential

explanation for the reduced number of peaks observed in the HSQC spectra of His$_8$-RTNLB13 described here. The topology prediction analyses of His$_8$-RTNLB13 suggested that an average of 129 residues comprise the soluble (i.e. non-TM) regions of the protein such as the loops and C- and N-terminal tails. The 114 observed peaks in the optimized HSQC spectrum are therefore likely due to residues in these soluble regions, but still cannot account fully for the number of residues predicted in these regions.

The reduced number of NMR signals obtained suggests that other residues in His$_8$-RTNLB13, apart from those in the TM domains, display reduced dynamics in the presence of detergent micelles leading to a disappearance of peaks in the NMR spectrum. This reduction in dynamics may be due to association of other regions of RTNLB13 with the membrane mimetic. We have recently proposed the presence of a 16-residue membrane-bound amphipathic helix (APH) in RTNLB13 spanning residues E$_{160}$-K$_{175}$[23], similar to that found in the RHD-containing protein Yop1p[15]. If we assume that the 16 residues in this APH are highly helical, this would increase the total predicted helical content to 47%, a value that agrees closely with the 44-46% range estimated from CD data. Likewise, if we assume that the 16 APH residues display slow dynamics due to association with the micelle, as was shown for the Yop1p APH[15], the total number of observable peaks estimated in the HSQC spectrum would decrease to 113, a value that is very close to the maximum number of observed peaks in this study (114). Together, our preliminary CD and NMR data point toward a surprisingly self-consistent model for the protein that is shown in Fig. 6B. This model is highly speculative, but accounts for the helical content observed here as well as the reduced dynamics of 60% of the residues as would be expected for a predominantly membrane embedded/associated protein. This is in keeping with the very poor staining of RTNLB13 by Coomassie and silver nitrate reported here, which also indicates that

21

the bulk of the protein is inaccessible, and supports the W-shape topology for reticulons proposed in the literature.

The data reported here support results obtained *in vivo* regarding the transmembrane topology of RTNLB13, and extend our understanding to regions outside the TM domains. Specifically, our results are in agreement with recent reports indicating that nearly all members of the reticulon family carry an amphipathic helix distally from the RHD [15, 23]. Given that the length of the transmembrane domains seems to be a crucial feature for the tubule-forming function of reticulons, it will be interesting to test whether the APH also plays a role in stabilizing ER tubule curvature *in vivo*.

ACKNOWLEDGEMENTS

REFERENCES

1.   Stefano, G., Hawes, C., and Brandizzi, F. ER - the key to the highway. Curr. Opin. Plant Biol. 2014; 22: 30-38.
2.   Lin, C., Zhang, Y., Sparkes, I., and Ashwin, P. Structure and dynamics of ER: minimal networks and biophysical constraints. Biophys. J. 2014; 107: 763-772.
3.   Griffing, L. R., Gao, H. T., and Sparkes, I. ER network dynamics are differentially controlled by myosins XI-K, XI-C, XI-E, XI-I, XI-1, and XI-2. Front Plant Sci 2014; 5: 218.
4.   Ibl, V., and Stoger, E. The formation, function and fate of protein storage compartments in seeds. Protoplasma 2012; 249: 379-392.
5.   Sparkes, I., Hawes, C., and Frigerio, L. FrontiERs: movers and shapers of the higher plant cortical endoplasmic reticulum. Curr. Opin. Plant Biol. 2011; 14: 658-665.
6.   Tolley, N., Sparkes, I. A., Hunter, P. R., Craddock, C. P., Nuttall, J., Roberts, L. M., Hawes, C., Pedrazzini, E., and Frigerio, L. Overexpression of a Plant Reticulon Remodels the Lumen of the Cortical Endoplasmic Reticulum but Does not Perturb Protein Transport. Traffic 2008; 9: 94-102.
7.   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 2006; 124: 573-586.
8.   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 2008; 319:
9.   Shibata, Y., Voss, C., Rist, J. M., Hu, J., Rapoport, T. A., Prinz, W. A., and Voeltz, G. K. The reticulon and DP1/Yop1p proteins form immobile oligomers in the tubular endoplasmic reticulum. J. Biol. Chem. 2008; 283: 18892-18904.
10. Oertle, T., Klinger, M., Stuermer, C. A. O., and Schwab, M. E. A reticular rhapsody: phylogenic evolution and nomenclature of the RTN/Nogo gene family. FASEB J 2003; 17: 1238-1247.
11. Oertle, T., and Schwab, M. E. Nogo and its paRTNers. Trends Cell Biol. 2003; 13: 187-194.
12. He, W. X., Lu, Y. F., Qahwash, I., Hu, X. Y., Chang, A. S., and Yan, R. Q. Reticulon family members modulate BACE1 activity and amyloid-beta peptide generation. Nature Medicine 2004; 10: 959-965.
13. Sparkes, I., Tolley, N., Aller, I., Svozil, J., Osterrieder, A., Botchway, S., Frigerio, L., and Hawes, C. P. C.-. Five plant reticulon isoforms share ER location, topology and membrane shaping properties. Plant Cell 2010; 22: 1333-1343.
14. Tolley, N., Sparkes, I., Craddock, C., Eastmond, P., 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. Plant J. 2010; 64: 411-418.
15. Brady, J. P., Claridge, J. K., Smith, P. G., and Schnell, J. R. A conserved amphipathic helix is required for membrane tubule formation by Yop1p. Proc. Natl. Acad. Sci. U.S.A 2015; 112: E639-E648.
16. Zurek, N., Sparks, L., and Voeltz, G. Reticulon short hairpin transmembrane domains are used to shape ER tubules. Traffic 2011; 12: 28-41.

17. Fazi, B., Melino, S., DiSano, F., Cicero, D. O., Piacentini, M., and Paci, M. Cloning, expression, and preliminary structural characterization of RTN-1C. Biochem. Biophys. Res. Comm. 2006; 342: 881-886.

18. Nepravishta, R., Polizio, F., Pacia, M., and Melino, S. A metal-binding site in the RTN1-C protein: new perspectives on the physiological role of a neuronal protein. Metallomics 2012; 4: 480-487.

19. Liu, J., Zhu, W., Qin, H., and Song, J. NMR studies reveal a novel mode for hFADD to bind with the unstructured hRTN3 which initiates the ER-stress activated apoptosis. Biochem. Biophys. Res. Comm. 2009; 383: 433-439.

20. Alhoshani, A., Vithayathil, R., Bandong, J., Chrunyk, K. M., Moreno, G. O., Weiss, G. A., and Cocco, M. J. Glutamate provides a key structural contact between reticulon-4 (Nogo-66) and phosphocholine. Biochim. Biophys. Acta 2014; 1838: 2350-2356.

21. Li, M., and Song, J. The N- and C-Termini of the Human Nogo Molecules are Intrinsically Unstructured: Bioinformatics, CD, NMR Characterization, and Functional Implications. Proteins 2007; 68: 100-108.

22. 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. Proc. Natl. Acad. Sci. U.S.A 2010; 107: 6847-6851.

23. Breeze, E., Dzimitrowicz, N., Kriechbaumer, V., Brooks, R., Botchway, S. W., Brady, J. P., Hawes, C., Dixon, A. M., Schnell, J. R., Fricker, M. D., and Frigerio, L. A C-terminal amphipathic helix is necessary for the in vivo tubule-shaping function of a plant reticulon. Proc. Natl. Acad. Sci. U.S.A 2016; 113: 10902-10907.

24. Marley, J., Lu, M., and Bracken, C. A method for efficient isotopic labeling of recombinant proteins. J. Biol. NMR 2001; 20: 71-75.

25. Abramoff, M. D., Magelhaes, P. J., and Ram, S. J. Image Processing with ImageJ. Biophotonics International 2004; 11: 36-42.

26. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970; 227: 680-685.

27. Hall, V., Sklepari, M., and Rodger, A. Protein Secondary Structure Prediction from Circular Dichroism Spectra Using a Self-Organizing Map with Concentration Correction. Chirality 2014; 26: 471-482.

28. Hall, V., Nash, A., and Rodger, A. SSNN, a method for neural network protein secondary structure fitting using circular dichroism data. Anal. Methods 2014; 6: 6721.

29. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. Biophys. J. 1981; 36: 575-588.

30. Dixon, A. M., Stanley, B. J., Matthews, E. E., Dawson, J. P., and Engelman, D. M. Invariant Chain Transmembrane Domain Trimerization: A Step in MHC Class II Assembly. Biochemistry 2006; 45: 5228-5234.

31. 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. Proc. Natl. Acad. Sci. U.S.A 2008; 105: 14371-14376.

32. Studier, F. W. Protein production by auto-induction in high-density shaking cultures. Protein Expr. Purif. 2005; 41: 207-234.

33. Newstead, S., Kim, H., von Heijne, G., Iwata, S., and Drew, D. High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.A 2007; 104: 13936-13941.

34. Arora, A., and Tamm, L. K. Biophysical approaches to membrane protein structure determination. Curr. Opin. Struct. Biol. 2001; 11: 540-547.

35. Chou, J. J., Baber, J. L., and Bax, A. Characterization of phospholipid mixed micelles by translational diffusion. J. Biomol. NMR 2004; 29: 299-308.

36. Doughty, D. A. Isopiestic compositions of aqueous ionic surfactant systems as a measure of preferential interactions. Application to the determination of micelle aggregation numbers by equilibrium ultracentrifugation. J. Phys. Chem. 1979; 83: 2621-2628.

37. Stafford, R. E., Fanni , T., and Dennis, E. A. Interfacial properties and critical micelle concentration of lysophospholipids. Biochemistry 1989; 28: 5113-5120.

38. Strop, P., and Brunger, A. T. Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. Prot. Sci. 2005; 14: 2207-2211.

39. Tausk, R. J., van Esch, J., Karmiggelt, J., Voordouw, G., and Overbeek, J. T. Physical chemical studies of short-chain lecithin homologues. II. Micellar weights of dihexanoyl- and diheptanoyllecithin. Biophys. Chem. 1974; 1: 184-203.

40. Warschawski, D. E., Arnold, A. A., Beaugrand, M., Gravel, A., Chartrand, E., and Marcotte, I. Choosing membrane mimetics for NMR structural studies of transmembrane proteins. Biochim. Biophys. Acta 2011; 1808: 1957-1974.

41. Jenei, Z. A., Borthwick, K., Zammit, V. A., and Dixon, A. M. Self-association of Transmembrane Domain 2 (TM2), but Not TM1, in Carnitine Palmitoyltransferase 1A: Role of GXXXG(A) motifs. Journal of Biological Chemistry 2009; 284: 6988-6997.

42. Oates, J., King, G., and Dixon, A. M. Strong oligomerization behavior of PDGF receptor transmembrane domain and its regulation by the juxtamembrane regions. Biochim. Biophys. Acta 2010; 1798: 605-615.

43. Cole, C., Barber, J. D., and Barton, G. J. The Jpred 3 secondary structure prediction server. Nucl. Acids Res. 2008; 36: W197-W201.

44. McGuffin, L. J., Bryson, K., and Jones, D. T. The PSIPRED protein structure prediction server. Bioinformatics 2000; 16: 404-405.

45. Bernsel, A., Viklund, H., Falk, J., Lindahl, E., von Heijne, G., and Elofsson, A. Prediction of membrane-protein topology from first principles. Proc. Natl. Acad. Sci. USA 2008; 105: 7177-7181.

46. Bernsel, A., Viklund, H., Hennerdal, A., and Elofsson, A. TOPCONS: consensus prediction of membrane protein topology. Nucl. Acids Res. 2009; 37: W465-468.

47. Käll, L., Krogh, A., and Sonnhammer, E. L. An HMM posterior decoder for sequence feature prediction that includes homology information. Bioinformatics 2005; 21: i251-i257.

48. Reynolds, S. M., Käll, L., Riffle, M. E., Bilmes, J. A., and Noble, W. S. Transmembrane topology and signal peptide prediction using dynamic bayesian networks. PLoS Comput Biol. 2008; 11:

49. Tusnády, G. E., and Simon, I. Principles Governing Amino Acid Composition of Integral Membrane Proteins: Applications to Topology Prediction. J. Mol. Biol 1998; 283: 489-506.

50. Viklund, H., and Elofsson, A. A method that improves topology prediction for transmembrane proteins by using two-track ANN-based preference scores and an improved topological grammar. Bioinformatics 2008; 24: 1662-1668.

51. von Heijne, G. Membrane Protein Structure Prediction, Hydrophobicity Analysis and the Positive-inside Rule. J. Mol. Biol. 1992; 225: 487-494.

52. Son, W. S., Park, S. H., Nothnagel, H. J., Lu, G. J., Wang, Y., Zhang, H., Cook, G. A., Howell, S. C., and Opella, S. J. q-Titration' of long-chain and short-chain lipids differentiates between structured and mobile residues of membrane proteins studied in bicelles by solution NMR spectroscopy. J. Magn. Reson. 2012; 214: 111-118.

TABLES

Table 1: Properties of the detergents tested in this study and resulting number of ^{1}H-^{15}N HSQC NMR peaks observed for ^{15}N-labelled His$_8$-RTNLB13. Published values for the aggregation number (N_A) and critical micelle concentration (CMC) have been taken from the literature for DDM[37], DPC[33], C6- and C7-DHPC[38, 39], SDS[35], LMPG[36] and LPPG[34, 36]. The final column lists the number of His$_8$-RTNLB13 peaks observed in each detergent in spectra acquired with 2048×128 data points and 40 scans per plane (the number of observable peaks upon increasing the number of points to 2048×256 and the number of scans to 160-180 is shown in parentheses).

Detergent	Chain length	MW (Da)	Charge	N_A	MW$_{mic}$ (kDa)	CMC (mM)	NMR peaks
DDM	12	511	nonionic	140	71.5	0.17	n/d
DPC	12	351	zwitterionic	70-80	24.6	1.5	65 (107)
C$_7$-DHPC	7	482	zwitterionic	25	12.1	1.5	26
C$_6$-DHPC	6	453	zwitterionic	35	15.9	14	38
SDS	12	288	anionic	50	14.4	8	48
LMPG	14	478	anionic	55	26.3	0.16	67 (114)
LPPG	16	507	anionic	125	63.4	0.018	65

FIGURE LEGENDS

Figure 1. Bacterial overexpression of His$_8$-RTNLB13. A. Proposed topologies for the reticulon homology domain. **B.** Anti-His Western blot of His$_8$-RTNLB13 expression in three strains of *E. coli*, namely BL21, C41 and C43. Two temperatures (15°C and 30°C) and two IPTG concentrations (0.05 and 0.5 mM) were tested, and expression was carried out in LB medium. Induction was carried out for 24 and 48 hours at 15°C, however bands for His$_8$-RTNLB13 (at approximately 20 and 40 kDa) were observed only in BL21 after 48 hours (no protein expression was observed in C41 or C43). At 30°C, induction was carried out for 4 hours and expression was observed in all three strains. The highest expression levels were seen in BL21. Uninduced controls are shown for comparison. **C.** Anti-his immunoblots of soluble and insoluble fractions obtained after harvesting membranes by ultracentrifugation, demonstrating that the bulk of overexpressed His$_8$-RTNLB13 is localized to the membranes. **D.** Expression in M9 minimal medium (*E. coli* strain BL21) required the addition of 100 µM Iron (III) chloride and 1 mM IPTG to restore expression. **E.** Coomassie-stained SDS-PAGE gel (upper panel) and an anti-His Western blot (lower panel) are shown for fractions from the purification of His$_8$-RTNLB13 using Talon cobalt resin. Membranes containing His$_8$-RTNLB13 were solubilized in 0.1% DDM (lane *S*) and incubated with resin for 30 minutes before loading into a column. The flow-through (*FT*) was analyzed, as was the eluent after washing the column with DDM-containing buffer + 10 mM imidazole (*W*). The protein was eluted after addition of 150 mM imidazole (elution fractions are shown in lanes *1-9*). Protein is only visible in Western blots as Coomassie repeatedly failed to stain His$_8$-RTNLB13. Molecular weight marker is shown in lane *M*. Monomer (M), dimer (D), and higher order oligomer (*n*) bands are shown.

28

Figure 2. Effect of detergent concentration, temperature and solution pH on NMR data.
^{1}H-^{15}N HSQC spectra of 180 μM ^{15}N-labelled His$_8$-RTNLB13 solubilized in (**A.**) 10 mM, (**B.**)
100 mM and (**C.**) 200 mM DPC detergent. All three samples were prepared in 20 mM Na$_2$HPO$_4$,
40 mM NaCl, 90% H$_2$O, and 10% D$_2$O at pH 6.8, and spectra were collected at 25°C. A steady
improvement in peak shape and a reduction in variable peak intensities was observed as DPC
concentration was increased. The number of observable peaks is given in parentheses. **D.** ^{1}H-^{15}N
HSQC spectrum of sample shown in C. measured at 37°C, in which a number of new peaks
become visible. **E.** The sample analysed in D. was prepared at pH 5.0, substituting sodium
phosphate buffer for 20 mM sodium acetate buffer and holding all other conditions constant, and
the spectrum was measured at 37°C. All spectra are shown at a contour level at which baseline
noise is just becoming visible. Samples prepared in LMPG show a similar trend (Fig. S1).

Figure 3. Effect of detergent type on His$_8$-RTNLB13 NMR spectrum. ^{1}H-^{15}N HSQC spectra
of (from top left to bottom right): 110 μM His$_8$-RTNLB13 solubilized in 100 mM C$_6$-DHPC; 110
μM His$_8$-RTNLB13 solubilized in 100 mM C$_7$-DHPC; 100 μM His$_8$-RTNLB13 solubilized in
100 mM SDS; 95 μM His$_8$-RTNLB13 solubilized in 100 mM LPPG; 180 μM His$_8$-RTNLB13
solubilized in 200 mM DPC; 130 μM His$_8$-RTNLB13 solubilized in 230 mM LMPG. All
protein samples were ^{15}N-labeled and prepared at pH 5.0, and all spectra were recorded at 37°C
with the exception of the sample containing SDS, which was prepared at pH 6.8 and analyzed at
15°C. All spectra shown were acquired with 2048 × 128 data points for ease of comparison.

Figure 4. Analytical ultracentrifugation analyses of His$_8$-RTNLB13 oligomerization.
Sedimentation equilibrium data obtained for His$_8$-RTNLB13 protein reconstituted into 15 mM

29

DPC and density matched with 52% D_2O. Data were collected at three speeds (■ 20,000, ○ 28,000 and ● 35,000 r.p.m.) and three protein concentrations ranging from 8-20 μM. In the upper panel, symbols represent experimental data and solid curves display the best fit upon global analysis of all nine data sets, in the case of His8-RTNLB13 the data were well represented by a monomeric fit. Below each plot are the residuals of the fitting process.

Figure 5. Secondary structure of detergent-solubilized His8-RTNLB13. A. Schematic depicting the results of secondary structure prediction from the RTNLB13 sequence using Psipred and Jpred3 algorithms. Regions predicted to be primarily α-helical are shaded. **B.** Circular dichroism spectra of His8-RTNLB13 (6-10 μM) dissolved in a range of detergent micelles (15-20 mM). Data are given in units of mean residue ellipticity (MRE, mdeg cm^2 $dmol^{-1}$). **C.** His8-RTNLB13 secondary structure content as estimated from CD data in each of the detergents tested, as well as from topology prediction algorithms (bar labeled "predicted").

Figure 6. Sequence analyses combined with experimental data yields putative model. A. The primary sequence of His8-RTNLB13 was analyzed using a variety of programs that predict membrane protein topology from sequence. The residues predicted to lie in each TM domain are listed, with all but two algorithms predicting four TM domains. The average values for the beginning and end of each TM domain are shown in the diagram (only values from 4 TM predictions were used to calculate the average). **B.** Preliminary structural model of RTNLB13 based on topology prediction, NMR, and CD data.
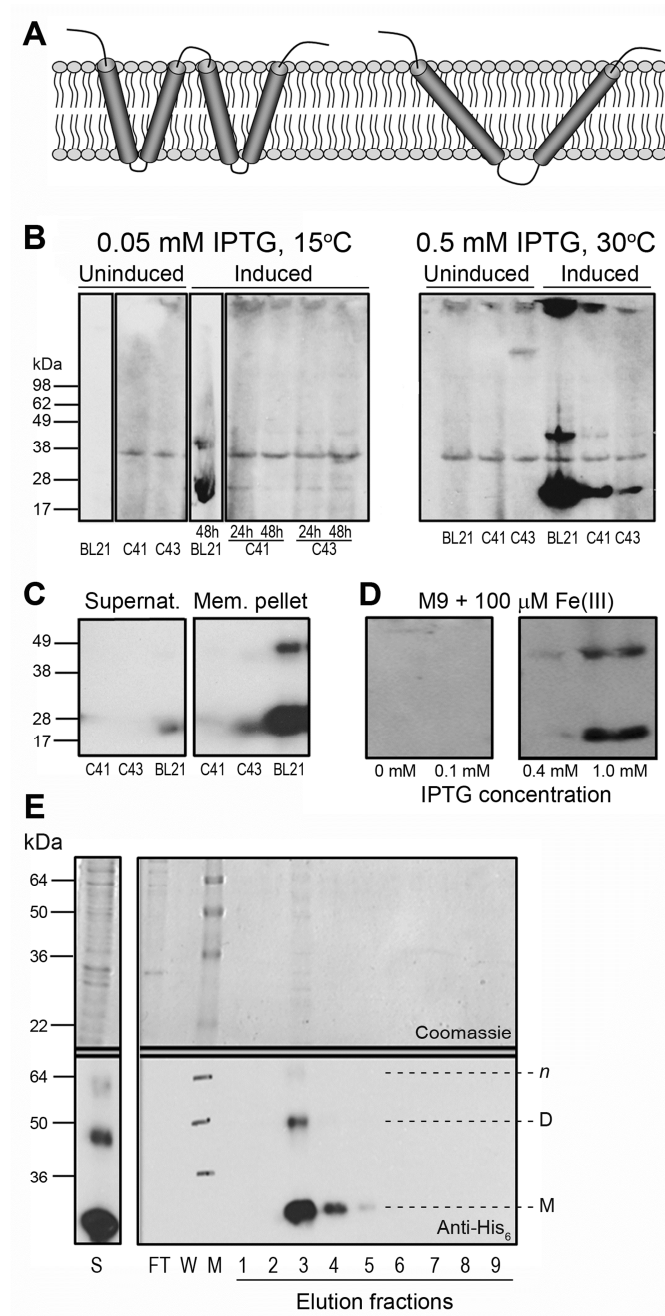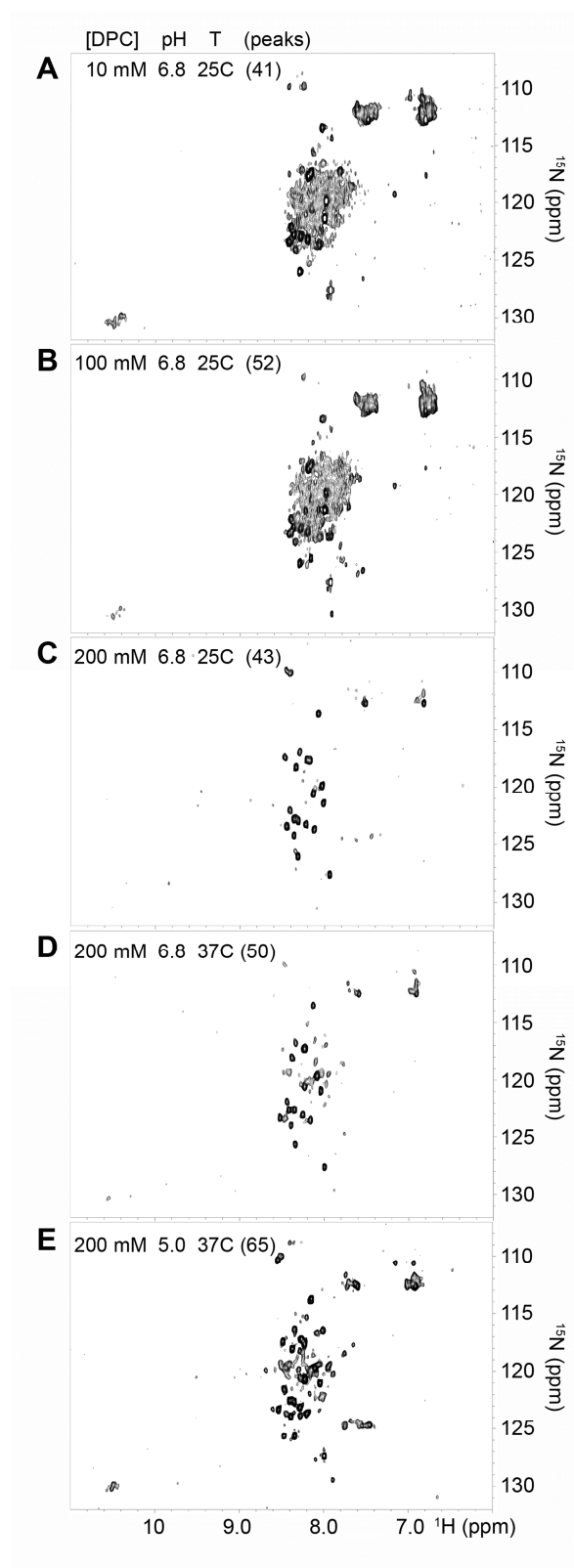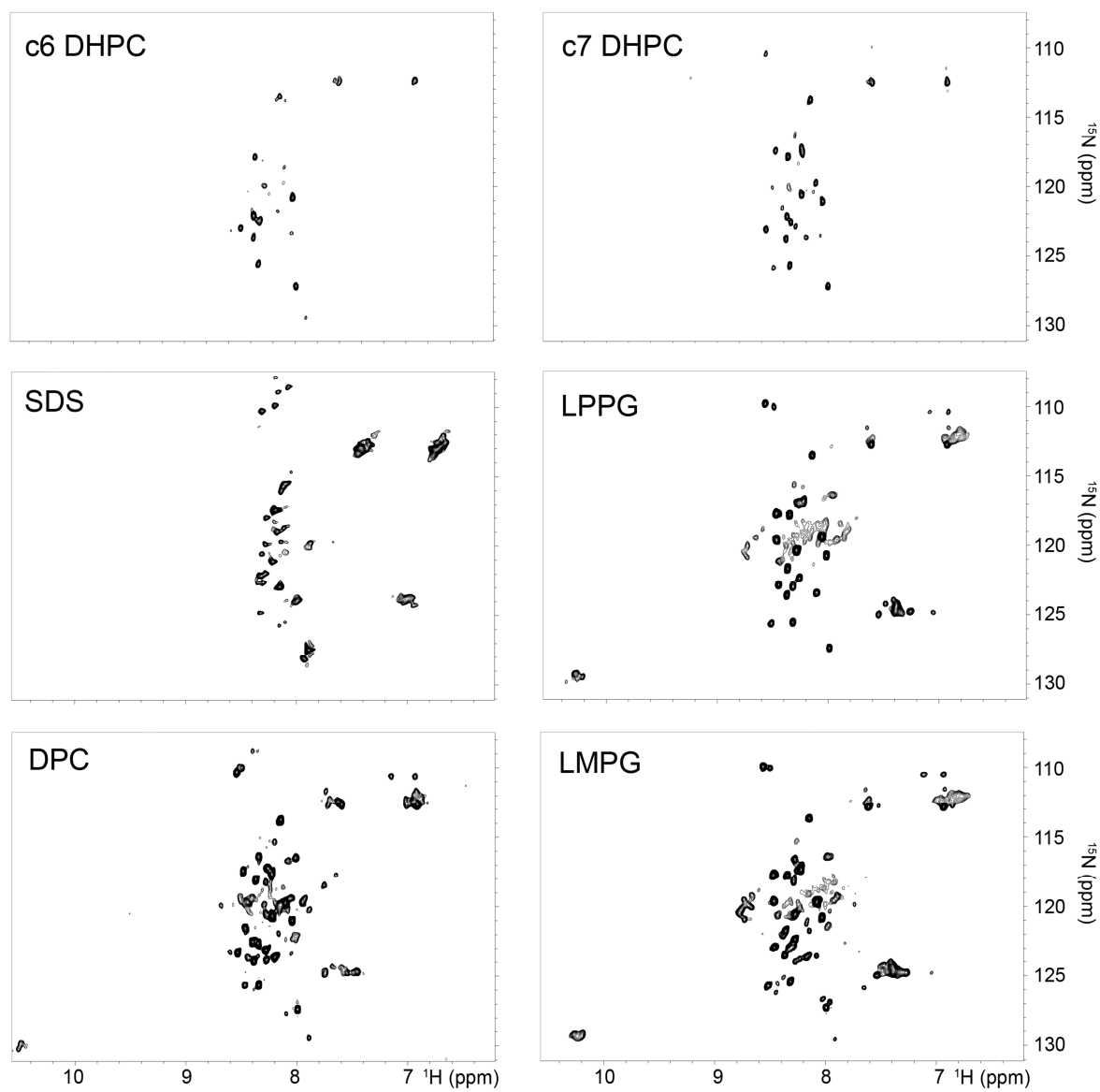
30

Figure 1.

A

B

0.05 mM IPTG, 15°C

0.5 mM IPTG, 30°C

Uninduced   Induced

Uninduced   Induced

kDa
98
62
49
38
28
17

48h
BL21   24h 48h   24h 48h
BL21 C41 C43   C41   C43

BL21 C41 C43   BL21 C41 C43

C

Supernat.   Mem. pellet

49
38
28
17

C41 C43 BL21   C41 C43 BL21

D

M9 + 100 µM Fe(III)

0 mM   0.1 mM   0.4 mM   1.0 mM

IPTG concentration

E

kDa

64
50
36
22

Coomassie

64
50

36

n
D

M

Anti-His$_6$

S   FT W M   1 2 3 4 5 6 7 8 9

Elution fractions

Figure 2.

Figure 3.

Figure 4.

Figure 5.

A

B

C
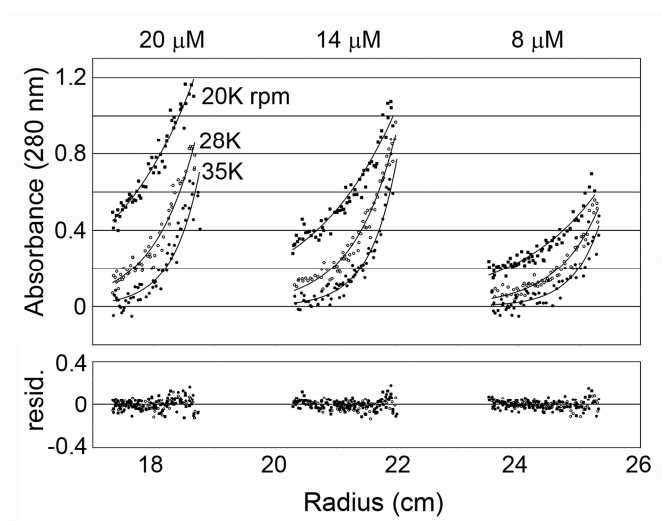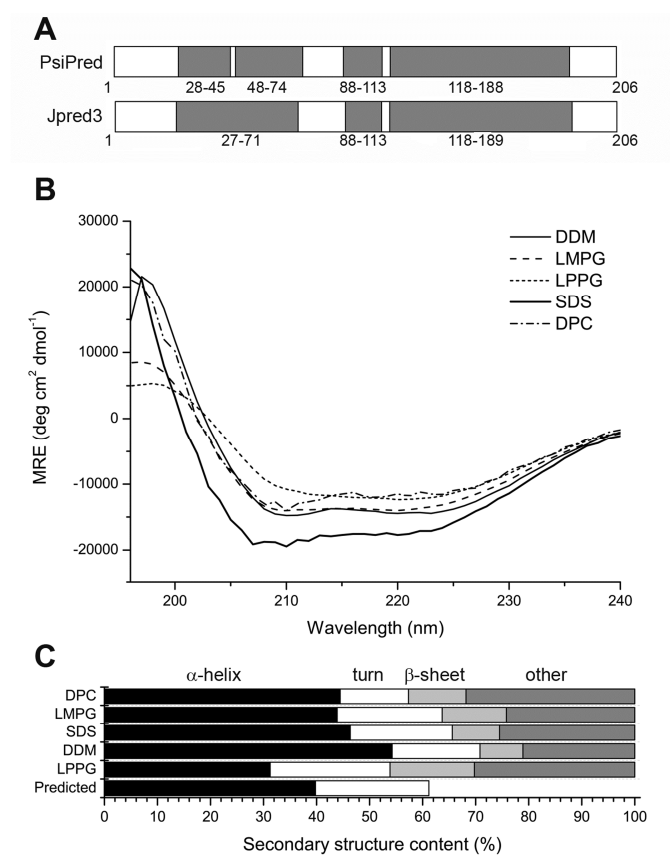
Figure 6.

A

	TM1	TM2		TM3	TM4
Toppred	27-47	50-70		——	138-158
HMMtop	27-44	49-67		113-132	137-154
Octopus	27-48	49-70		123-154	——
Topcons	27-48	49-70		117-138	139-160
Philius	27-48	52-72		109-132	137-159
Polyphoibus	26-45	48-71		118-139	140-159
Scampi	27-48	51-72		117-138	139-160
Avg.	27-47	50-70		116-139	139-158

B
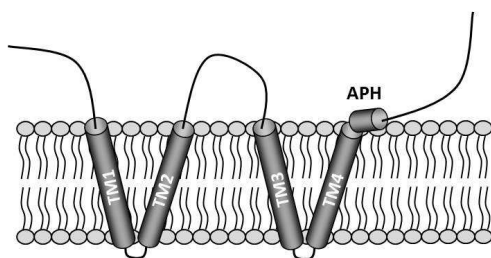
SUPPLEMENTARY INFORMATION FOR:

Bacterial expression, purification and biophysical characterization of the smallest plant reticulon isoform, RTNLB13

Michael Chow [+], Meropi Sklepari[§], Lorenzo Frigerio[‡], and Ann M. Dixon[†*]

Running Title: Expression and characterization of RTNLB13

Keywords: Reticulon; plant; transmembrane protein; bacterial expression; solution nuclear magnetic resonance

[+] MOAC Doctoral Training Centre, § Warwick Centre for Analytical Science, ‡ School of Life Sciences and † Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK.

*To whom correspondence should be addressed: Dr Ann Dixon, Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK, Telephone: +44 2476 150037; FAX: +44 2476 524112; email: ann.dixon@warwick.ac.uk

Figure S1. Effect of LMPG detergent concentration, temperature and solution pH on NMR data. ^{1}H-^{15}N HSQC spectra of 130 μM ^{15}N-labeled His-RTNLB13 solubilized in phosphate buffer (20 mM Na$_2$HPO$_4$, 40 mM NaCl, 90% H$_2$O, 10% D$_2$O, pH 6.8) containing either 63 mM LMPG (blue spectrum) or 230 mM LMPG (red spectrum) and collected at 25°C. Increasing the detergent concentration increased the number of observable peaks from 46 (blue spectrum) to 56 (red spectrum). A further increase in the number of observable peaks (to 67 peaks) was achieved upon reducing the pH to 5.0 and increasing the temperature to 37°C for the sample containing 230 mM LMPG (black spectrum).
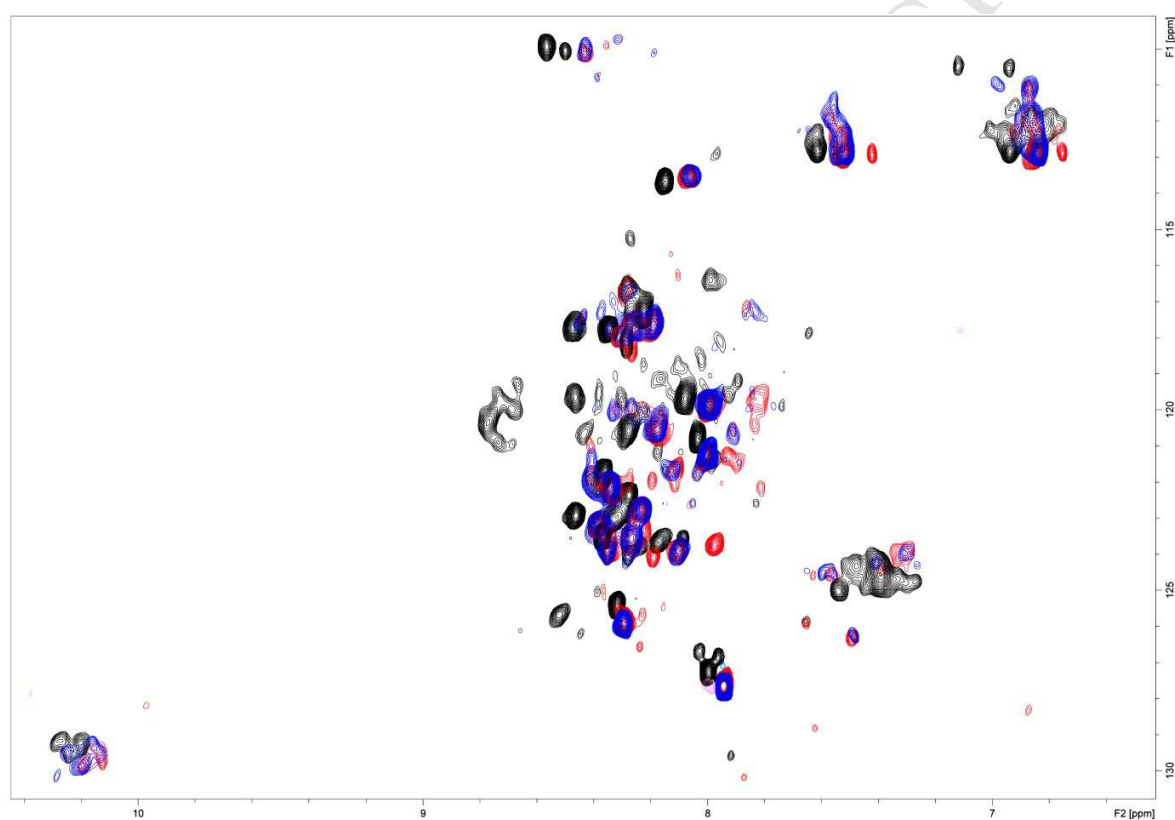
Figure S2. CD data fitting to obtain estimates of secondary structure content. Shown below are the original CD spectra (solid lines) obtained in (A) DPC, (B) LMPG, (C) SDS, (D) DDM, and (E) LPPG. Also shown are the best fits of the data using the SSNN software (dashed lines) after application of scaling factors (sf) to obtain optimal fit and provide an independent measure of protein concentration. The fit residuals are shown (dotted lines) as well as the resulting NRMSD of each fit.
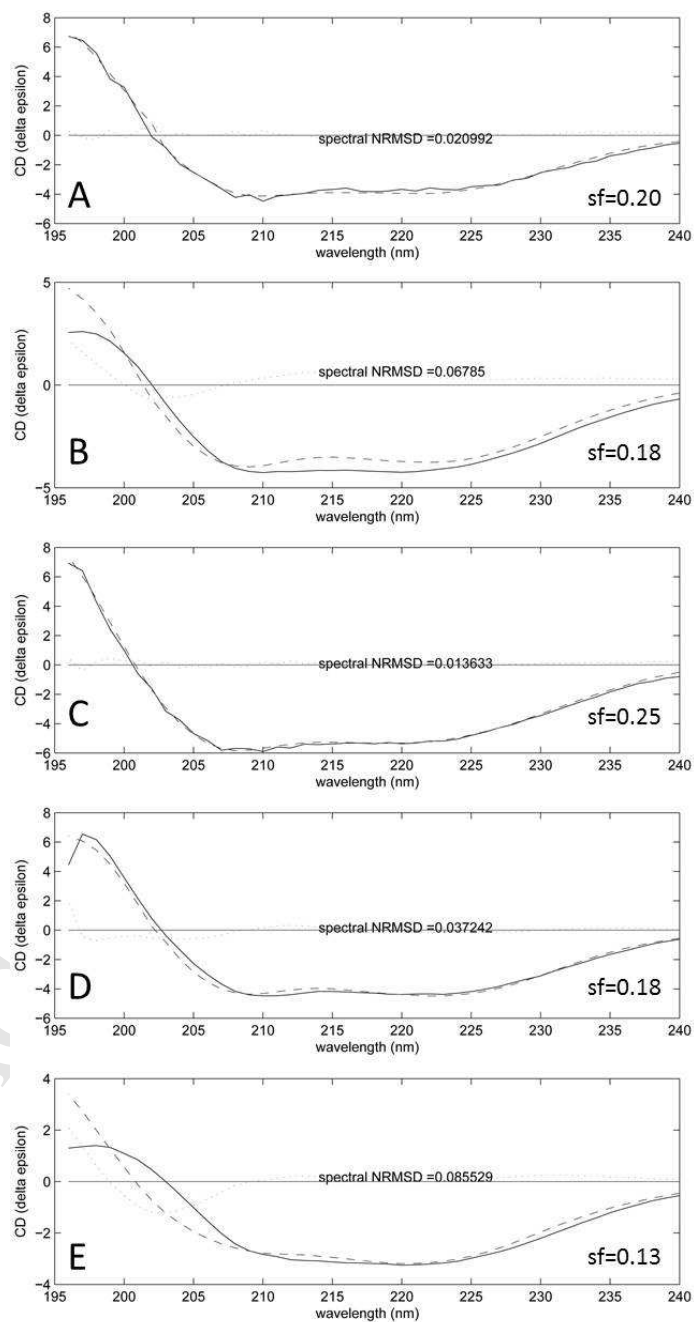
HIGHLIGHTS

- First heterologous expression of a plant reticulon, RTNLB13 from *A. thaliana,* in *E. coli.*

- Structure of detergent-solubilized protein consistent with current models in literature.

- Demonstrate a detergent concentration-dependent ability to oligomerize.