nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-019-0188-9

Structure determination protocol for transmembrane domain oligomers

Qingshan Fu[1,3], Alessandro Piai[1,3], Wen Chen[1,3], Ke Xia[2] and James J. Chou[1*]

The transmembrane (TM) anchors of cell surface proteins have been one of the 'blind spots' in structural biology because they are generally very hydrophobic, sometimes dynamic, and thus difficult targets for structural characterization. A plethora of examples show these membrane anchors are not merely anchors but can multimerize specifically to activate signaling receptors on the cell surface or to stabilize envelope proteins in viruses. Through a series of studies of the TM domains (TMDs) of immune receptors and viral membrane proteins, we have established a robust protocol for determining atomic-resolution structures of TM oligomers by NMR in bicelles that closely mimic a lipid bilayer. Our protocol overcomes hurdles typically encountered by structural biology techniques such as X-ray crystallography and cryo-electron microscopy (cryo-EM) when studying small TMDs. Here, we provide the details of the protocol, covering five major technical aspects: (i) a general method for producing isotopically labeled TM or membrane-proximal (MP) protein fragments that involves expression of the protein (which is fused to TrpLE) into inclusion bodies and releasing the target protein by cyanogen bromide (CNBr) cleavage; (ii) determination of the oligomeric state of TMDs in bicelles; (iii) detection of intermolecular contacts using nuclear Overhauser effect (NOE) experiments; (iv) structure determination; and (v) paramagnetic probe titration (PPT) to characterize the membrane partition of the TM oligomers. This protocol is broadly applicable for filling structural gaps of many type I/II membrane proteins. The procedures may take 3–6 months to complete, depending on the complexity and stability of the protein sample.

Introduction

Mounting structural and functional data have indicated that what are commonly designated as the TM anchors of many signaling receptors actually play critical roles in receptor signaling, and the diversity of the mechanisms with which the TM regions can promote signaling is beyond the traditional scope of receptor biology. For example, mature assembly of the TCR/CD3 complex is primarily mediated by interactions among the TMDs[1,2]. Different modes of TM helix dimerization appear to contribute to the 'on' and 'off' states of the epidermal growth factor (EGF) receptor[3,4]. TM domain trimerization is required for the Fas death receptor, a member of the tumor necrosis factor (TNF) receptor family, to signal, and this appears to apply as well to other members of the TNF receptor superfamily[5]. These examples are only the tip of the iceberg, as there remain a vast number of type I and II TM receptors whose membrane regions are unknown. Revealing the structures of these membrane regions is required to gain a thorough understanding of receptor activation for many of the immune co-stimulators currently being targeted for cancer immunotherapy. The membrane regions of cell surface proteins, however, have been difficult targets for crystallography, because they are generally very hydrophobic and often dynamic; they are also too small for cryo-EM at the moment.

As a versatile spectroscopic tool capable of determining atomic-resolution structures, solution NMR has often been the go-to method for tackling small TMDs of type I and II membrane proteins. The application of solution NMR to TM helix oligomers was demonstrated more than two decades ago on the TMD of glycophorin A[6]. In that study, the NMR structure, solved in dodecylphosphocholine micelles, revealed the structural role of the GXXXG signature sequence in mediating TM helix dimer formation; recently, this NMR structure was independently validated by a crystal structure of the same TMD determined in the lipidic cubic phase (LCP)[7]. NMR has since been widely applied to investigate the structures of small TMDs in detergent micelles and small bicelles[1,3,5,8–15]. Despite the powerful utility, a general and robust protocol for determining TM structures by NMR is

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA. [2]Department of Chemistry and Chemical Biology, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY, USA. [3]These authors contributed equally: Qingshan Fu, Alessandro Piai, Wen Chen. *e-mail: james_chou@hms.harvard.edu

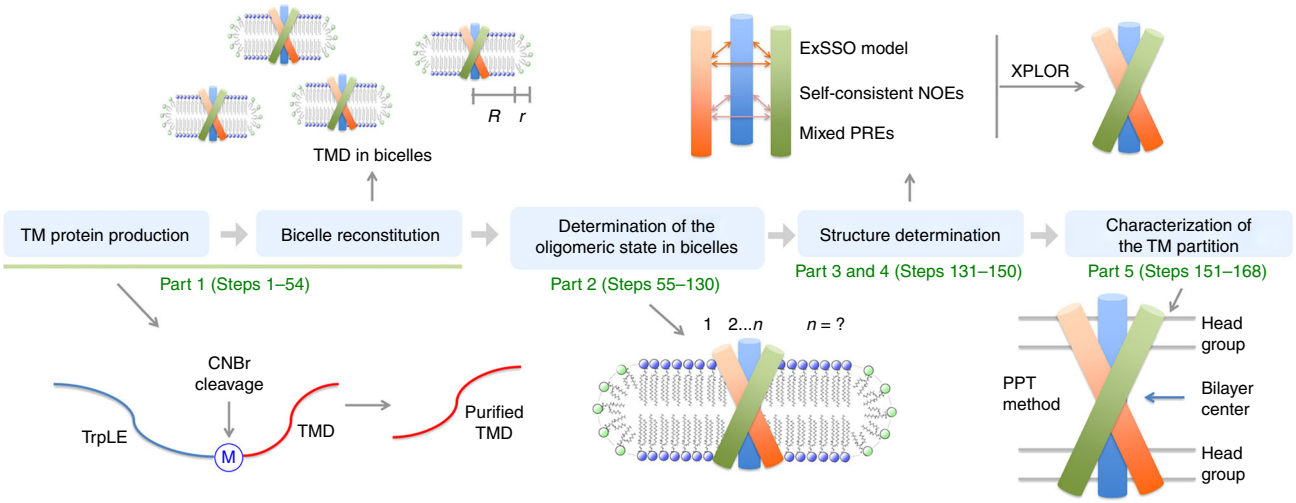


Fig. 1 | Protocol overview. The protocol consist of the following parts: TM protein production, bicelle reconstitution, oligomeric state determination, structure determination, and characterization of the TM partition. M, methionine.

still lacking. We have integrated the most effective and practical methods from a series of our recent applications to generate a protocol for general structural characterization of the small membrane-embedded and MP regions of TM proteins in bicelles that are sufficiently large to mimic a lipid bilayer.

Overview of the protocol

This protocol addresses the major technical challenges associated with structural analysis of small TM/MP complexes, including the production of highly hydrophobic protein fragments with suitable isotope labeling for NMR, protein reconstitution in appropriate membrane-mimetic media, determination of the protein oligomerization state, detection of interchain contacts for structure determination, and characterization of the protein partition in lipid bilayer. A conceptual overview of the protocol is shown in Fig. 1.

Specifically, we show that by using targeted expression into inclusion bodies in bacteria, along with highly optimized affinity and high-pressure reverse-phase liquid chromatography, hydrophobic peptides with a range of sizes (10–200 residues) and hydrophobicity (grand average of hydropathicity (GRAVY) score = 0.5–2)[16,17] can be produced. The peptides can be uniformly isotopically labeled for NMR measurements. In the protocol, we reconstitute the peptides into bicelles that are sufficiently large to mimic the membrane, because the small TM/MP domains often require a lipid bilayer environment to form stable oligomeric complexes. We found that when using bicelles with molar ratio of lipid/detergent (q) ≥0.5 (e.g., $0.5 \leq q \leq 0.6$; the bilayer region size is between 44 and 50 Å), at which the size and structure of the bicelles are very close to those of lipid discs[18–21], the TM proteins still yield high-quality NMR spectra feasible for structure determination. A challenge associated with the use of bicelles is the accurate determination of the oligomeric state of the small TMDs, as the dynamic self-assembly of bicelles is incompatible with standard methods for measuring molecular mass of protein complexes such as size-exclusion chromatography and equilibrium sedimentation. We note that the use of NMR relaxation parameters to infer molecular mass is potentially misleading, because membrane proteins usually exhibit very heterogeneous dynamics. Hence, our approach is to use intermolecular paramagnetic resonance enhancement (PRE) analysis to confirm multimeric assembly and the oligomer label (OG-label) method to determine the oligomeric state of the TM complex.

NMR-based structure determination of small membrane proteins has been rather controversial, as different sources of structural information such as NOE, PRE, or residual dipolar coupling (RDC) have been used to solve structures. This protocol focuses almost entirely on the use of the NOE for structure determination, because the NOE remains the most direct NMR probe of internuclear distances in a molecule; it makes full use of the isotope-labeling strategy and strand-selective experiments for detecting interchain NOEs that define the oligomeric structure. Finally, probably the

most unique feature of the protocol is the incorporation of the PPT method, which allows accurate characterization of the membrane partition of TM proteins. This method is based on the notion that when bicelles are sufficiently wide ($q \geq 0.5$), so that the protein resides essentially in the bilayer region of the bicelles, simple titration of either water-soluble or lipophilic paramagnetic probes, such as gadolinium (III) 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetate (Gd-DOTA) or nitroxide-labeled fatty acids (e.g., 2-(14-carboxytetradecyl)-2-ethyl-4,4-dimethyl-3-oxazolidinyloxy (16-DSA)), respectively, can be used to accurately determine residue-specific depth immersion along the bicelle normal[19,22–24]. When the structure of a TM oligomer is known, the PPT data can thus be used to determine the membrane partition of the protein. The above techniques constitute a comprehensive analysis of the structure, oligomerization state, and membrane partition of the TM and MP regions of cell surface proteins.

Advantages in comparison to alternative methods

In addition to solution NMR, other techniques have been used to determine structures of TM oligomers. For example, in several cases, LCP crystallization generated crystals of small TM fragments that were good enough for high-resolution structure determination[7,25–27]. A more recent crystallographic study also managed to capture the entire structure of the viral envelope protein (including the membrane region) from herpes simplex virus[28]. Solid-state NMR is another attractive alternative, as it allows structural study in a completely lipid bilayer environment. Achieving atomic-resolution structures by solid-state NMR relies on proteoliposome samples that yield NMR spectra with high resolution, and this is often case dependent[29,30]. Small TMDs are obviously too small for current cryo-EM applications; they must be analyzed by cryo-EM as a part of the much larger full-length proteins, which would be the best approach, if feasible, for studying TM structures. The success of this type of cryo-EM application, however, depends on whether the large extra- or intra- cellular domains are rigidly connected to the TMD, which is case dependent.

The main strength of the reported protocol is that it provides a practical, robust, and general solution for obtaining TM structures in a near-native environment. First, this protocol provides a comprehensive solution for determining TM oligomer structures in bicelles that are sufficiently large to mimic a lipid bilayer, including the robust OG-label method for determining protein oligomeric states in bicelles. Second, the use of ideal bicelles (i.e., bicelles in which the lipid and detergent are well segregated) enables the implementation of the PPT method to determine the membrane partition of the TMDs. This is a distinct advantage over previous NMR methods, as this protocol provides not only the TM structure, but also information about how the TM protein resides in membrane. Third, the protocol is very rigorous from sample preparation to structure determination; it provides practical strategies for directly detecting intermolecular contacts that are critical for structure determination of TM oligomers. Although intended for solution NMR studies, the NMR sample preparation protocol is potentially applicable to other magnetic resonance techniques, such as solid-state NMR and electron paramagnetic resonance (EPR). For example, the bicelle samples used in the current protocol can be easily converted to proteoliposome samples suitable for solid-state NMR by completely removing the detergent.

Limitations

The key limitation of the protocol—that the TMDs cannot be studied in the context of the full-length TM proteins—is due to the fundamental molecular mass limitation of the solution NMR technique. Empirically, we found that TM/MP oligomers with monomeric chains of ~10 kDa, reconstituted in bicelles with $q = 0.5$, approach the size limitation within which NMR-based structure determination is feasible, but larger oligomers can be used, depending on the protein's oligomeric state and dynamics. Furthermore, bicelle solutions still contain free detergent, e.g., the concentration of free 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC) in 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC)/DHPC bicelles is ~5 mM[31], and this amount of free detergent, although not affecting the TMD, could potentially generate structural artifacts in the MP regions.

Future applications

Our protocol should be useful for studying TMD oligomerization of many immunoreceptors and receptor tyrosine kinases, because, for many of these type I/II membrane proteins, the TMD plays an essential role in receptor assembly and possibly in receptor clustering as well. Detailed structural information of TMD oligomerization in a membrane-like environment provides valuable clues for

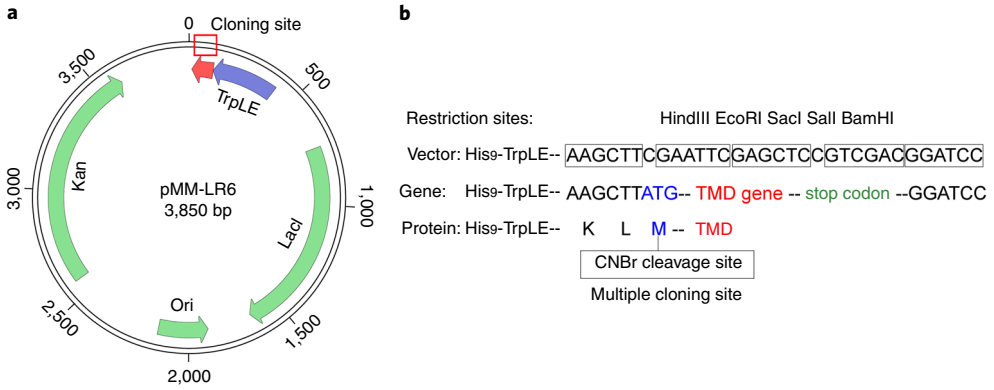


Fig. 2 | Schematic of the pMM-LR6 vector. a, Schematic diagram of the pMM-LR6 vector. **b**, Multiple cloning sites are designed for optimal TM gene insertion. The methionine codon (ATG) is added between the TrpLE and target gene for CNBr cleavage during protein purification; stop codons (TAA, TGA, TAG) are added at the end of the target gene. TrpLE peptide amino acid sequence: KAIFVLKGSLDRDLDSRIELELRTDHKELSEHLLLVDLARNDLARIATPGSRY-VADLTKVDRYSYVLHLVSRVVGELRHDLDALHAYRAALNLGTLSGAPKVRAKL. Kan, kanamycin resistance gene; LacI, lactose operon repressor; Ori, the origin of replication; TrpLE, fragment gene from anthranilate synthase, which cannot properly fold and guides the target protein to inclusion bodies.

elucidating how the connected intracellular signaling domains are clustered to activate the downstream signaling. The protocol is equally applicable to the unknown membrane regions of the many viral membrane fusion proteins. Although our earlier NMR applications focused primarily on the TMDs, the current protocol has also proven effective in revealing the structures of MP domains, as demonstrated recently for the MP external region (MPER) of the HIV-1 gp41 fusion protein[24]. In addition to type I/II membrane proteins, the protocol is in principle applicable to TMDs that span the membrane multiple times, as long as the size of the overall protein complex in bicelles is within the limitation of solution NMR. For example, bacteria have developed two-component systems (receptor histidine kinases) for sensing all kinds of environmental factors, such as cellular cytokines, osmotic pressure, pH, and membrane curvature[32–35]. These receptors have a small TMD, usually consisting of two TM helices, that is believed to physically transmit signals from extracellular environmental interactions to activate intracellular kinase. We believe our protocol can be used effectively on these interesting systems as well.

Experimental design

TM protein production

Transmembrane domains of cell surface proteins are very hydrophobic, which makes protein expression and purification difficult. We express the small hydrophobic TMDs in *Escherichia coli* as C-terminal fusions to the TrpLE sequence, which is a fragmented protein that drives inclusion body formation upon synthesis. As such, the hydrophobic TMDs, usually toxic to the cells per se, are sequestered in the inclusion bodies and thus can be expressed to high quantities. This is done using a pMM-LR6 plasmid, originally from Blacklow[36,37]. Based on pMM-LR6, a His_9-tag is added to the N terminus of TrpLE to facilitate purification. A methionine is inserted between the TrpLE and the TMD to enable cleavage at this position with CNBr so that the two fragments can be separated (Fig. 2). Hence, the TM sequence cannot contain any methionine. The purification of the TMD can be achieved with three major steps, all under denaturing conditions (Fig. 3): (i) purification of the TrpLE–TMD fusion from inclusion bodies by Ni-NTA affinity, (ii) CNBr cleavage to separate TrpLE and TMD, and (iii) purification of the TMD by reverse-phase HPLC.

For certain NMR experiments, a very high level of deuteration (e.g., >98%) is needed. We note that when using the above expression vector in the BL21(DE3) *E. coli* cell line, this level of deuteration is achievable by using 99% deuterium oxide (D_2O) and deuterated (98% ^2H) glucose (Box 1).

For each target TMD, we suggest performing a sequence analysis to check for methionines and cysteines, and to evaluate the overall hydrophobicity of the TMD. As methionine is the CNBr cleavage site, additional methionines, if any, must be mutated to an amino acid of similar hydropathicity. Cysteines might form non-native disulfides during purification. If they are not conserved,

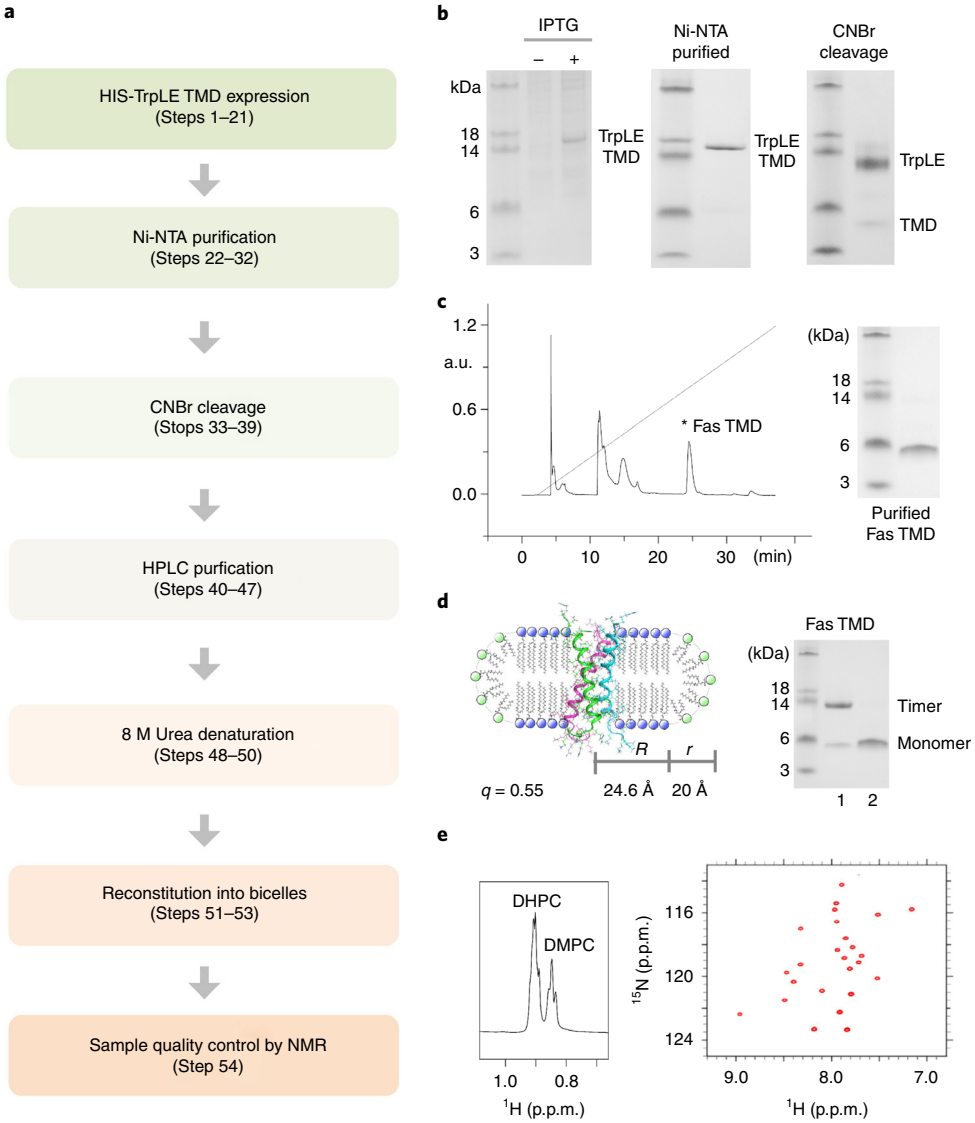


Fig. 3 | Expression, purification, and bicelle reconstitution of TMDs. The protocol was demonstrated for the Fas TMD. **a**, Flowchart of the steps required for TM/MP domain expression, purification and reconstitution into bicelles with $0.5 \leq q \leq 0.6$. **b**, SDS-PAGE analysis of protein expression and purification. Left: cell lysate before (−) and after (+) induction with IPTG; middle: the TrpLE TMD fusion protein after Ni-NTA affinity purification; right: purified TrpLE TMD after CNBr cleavage. **c**, HPLC chromatogram of cleaved TrpLE TMD (left) and SDS-PAGE analysis of the Fas TMD elution peak (right). **d**, Schematic illustration (left) of the Fas TMD reconstituted in bicelles with $q = 0.55$ (q is the molar ratio of DMPC to DHPC). The radii of the planar region of the bicelle (R) and the DHPC rim (r) are 24.6 and 20 Å, respectively. Right, SDS-PAGE analysis of the Fas TMD after bicelle reconstitution (1) and the Fas TMD before reconstitution (2; dried powder in loading buffer). **e**, The ^{1}H NMR spectrum of a typical bicelle-reconstituted Fas TMD sample, showing the DHPC and DMPC methyl peaks used for calculating the bicelle q (left), and the corresponding 2D ^{1}H-^{15}N TROSY-HSQC spectrum (right). The spectra were acquired at 600 MHz at 303 K. **c**–**e** adapted with permission from ref. [5], Elsevier.

we suggest mutating them to serine or alanine, which would simplify the purification and reconstitution procedures. The overall hydrophobicity of the TMD sequence can be evaluated using the program ProtParam (https://web.expasy.org/protparam/), which calculates the GRAVY score. The GRAVY score for a typical TMD is 0.5–1.5; <0.5 means it is hydrophilic, and >1.5 means it is extremely hydrophobic. If a TMD has a very high GRAVY score (>1.5), and an attempt to purify it failed, one can consider the option of mutating non-conserved hydrophobic residues to decrease the GRAVY score. For TMDs having internal methionines that cannot be mutated, alternative expression and purification methods must be used. For example, the Pro-Asp sequence can be cleaved under acidic conditions (pH ~1) and high temperature (~80 °C)[38,39]. Alternatively, hydrophobic fragments

Box 1 | Quantification of the sample deuteration level

The expression/purification strategy described in part 1 of our protocol allows researchers to achieve high levels of protein deuteration, as shown for the Fas TMD in Fig. 6b by measuring its deuteration rate and purity (~98.3% deuteration). The high level of deuteration allows the use of mixed samples for exclusive detection of interchain NOEs in NOESY experiments (see part 3, Fig. 6c), ensuring that no intrachain NOEs from residual protonation of aliphatic groups are present in the spectra.

The procedure to quantify the deuteration level of a TMD (e.g., the Fas TMD in Fig. 6b) is reported below:

Procedure

1 Dissolve 100 μg of Fas TMD powder in 100 μL of 0.2% FA.
2 Connect a BioBasic 18 LC column directly to the standard electrospray ionization source of an LTQ-Orbitrap XL Fourier transform mass spectrometer.
3 Inject 5 μL of sample through an Agilent 1200 autosampler. Set the gradient from 5% MS elution buffer (98% (vol/vol) ACN, 0.2% FA) to 80% MS elution buffer over 20 min at a flow rate of 250 μL/min.
4 Set the mass spectrometer parameters as follows: spray voltage = 4.5 kV, capillary voltage = 26 V, tube lens voltage = 120 V, capillary temperature = 275 °C, sheath flow rate = 40, auxiliary gas flow = 25.
5 Acquire all Fourier transform mass spectra at resolution of 60,000 with 300- to 2,000-Da mass ranges. Mass accuracy is >3 p.p.m. after external calibration.
6 De-convolute the mass spectra using Xcaliber.

can be expressed as fusions to other proteins, such as the maltose binding protein and thioredoxin, but the purification method will be completely different because cleavage is typically done with enzymes under non-denaturing conditions[40].

In Steps 1–54, we describe the general protocol for generating and reconstituting the following samples, which can be used for subsequent analyses as described:

- A ~1:1 (wt/wt) mixture of lyophilized ^{15}N-labeled protein and a Cys mutant of the same protein that can be used for the intermolecular PRE analysis in Steps 55–76.
- A His_6-tag-containing protein can be used for the OG-label analysis in Steps 121–130.
- A ~1:1 (wt/wt) mixture of ^{2}H,^{15}N-labeled protein (perdeuterated or partially deuterated) and ^{13}C-labeled protein that can be used for the intermolecular NOE analysis in Steps 131–140.
- A ^{15}N-labeled protein that can be used for the PPT analysis in Steps 151–168.

Reconstitution in bicelles

The target TMDs are reconstituted in bicelles that are sufficiently large to mimic the bilayer environment of the membrane. When the lipid/detergent ratio (q) is >0.5, DMPC/DHPC bicelles are known to become disc-like, in which case the lipids and detergents become largely segregated[18,19,41]. When a protein is reconstituted in such bicelles, typically with $0.5 \leq q \leq 0.6$, it is essentially in a lipid bilayer environment while still amenable to high-resolution solution NMR spectroscopy. This approach has been successful for several important systems, including the TMDs of HIV-1 Env[12,24], Fas[5], and the intact p7 channel of HCV[23]. Although DMPC has been the most common lipid for bicelles, other lipids, such as POPC (1-palmitoyl-2-oleoyl-glycero-3-phosphocholine), POPG (1-hexadecanoyl-2-(9Z-octadecenoyl)-sn-glycero-3-phospho-(1'-rac-glycerol)), and POPE (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine), are equally compatible with DHPC in forming bicelles. A general protocol for incorporating TMDs into bicelles is to first denature and completely solubilize the protein in the presence of lipid and detergent and then slowly remove the denaturant to allow self-assembly of bicelles around the protein. Specifically, the purified and lyophilized TMD is dissolved in 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) with a suitable amount of lipid, followed by drying of the solution under a nitrogen stream to achieve thin films. The thin films are then dissolved in 8 M urea solution containing a calculated amount of detergent. Reconstitution begins as the denaturant is removed by dialysis. Some detergents are lost during dialysis and therefore must be added back to the sample to maintain the desired bicelle q value. The q value of the sample can be accurately measured using 1D ^{1}H NMR.

Determination of oligomeric state in bicelles

Characterizing the oligomeric state of small TMDs remains difficult, especially in dynamically assembled systems such as lipid/detergent bicelles. The two-component nature of bicelles makes it extremely difficult to perform experiments such as size-exclusion chromatography–multi-angle light scattering (SEC-MALS) and sedimentation equilibrium. When membrane proteins are reconstituted in the lipid bilayer, the TM regions are mostly buried and/or their primary amine

groups are often not in ideal position for direct crosslinking, resulting in nonspecific ladder patterns. We have developed an OG-label method to solve this problem[23]. Each protomer in an oligomer is paired with a small soluble protein through affinity interaction; the soluble protein can then be crosslinked to determine the oligomeric state of the membrane protein of interest. GB1 protein (molecular weight (MW) = 8.4 kDa) is used as the soluble crosslinkable protein (SCP) in the method. Its N terminus is covalently linked to a tri-nitrilotriacetic acid (TriNTA) molecule and charged with Ni^{2+}. A His_6-tag sequence should be included in the target TM protein sequence. The Ni^{2+}-charged TriNTA complex binds the His_6-tag with high affinity (20 ± 10 nM)[42], specifically recruiting the GB1 to the individual protomers of the oligomeric membrane protein in lipid bilayer. The GB1 molecules tethered stoichiometrically to the membrane protein oligomer are much more crosslinkable than the free GB1 molecules in solution. The crosslinked GB1 can be released from the oligomer by addition of EDTA or imidazole and analyzed by SDS–PAGE to determine the oligomerization number.

Interprotomer structural restraints and structural calculation

Structure determination of homo-oligomers is challenging because NOEs between structurally equivalent subunits having the same chemical shifts are needed as interprotomer distance restraints. To solve this problem, we use mixed samples in which half of the monomers are (^{15}N,^{2}H)-labeled and the other half ^{13}C-labeled, which allows us to detect exclusively NOEs between the ^{15}N-attached protons of one subunit and the ^{13}C-attached protons of the neighboring subunits using a ^{1}H-^{13}C scalar coupling (J_{CH})-modulated NOE experiment[5,24]. These restraints are then used to build a model of the oligomeric TMD, and, for this, we need a fast and efficient way to assemble the symmetric TM oligomer and calculate all the conformations that fit the confirmed NOEs. We developed a new program, named exhaustive search for symmetric oligomer (ExSSO), to achieve this goal[43]. This program performs an exhaustive search to find all oligomeric assemblies that satisfy the interprotomer NOE restraints. If a unique oligomeric packing solution has converged, a representative structure is used as the starting model for further structural refinement in a standard program such as XPLOR-NIH[44]. This refinement process involves (i) identifying self-consistent backbone–side-chain and side-chain–side-chain NOE restraints in conventional ^{15}N- and ^{13}C-edited NOESY (nuclear Overhauser effect spectroscopy) experiments, and (ii) updating the structure with the new NOE restraints. The above two steps are performed iteratively until the desired root mean square deviation (RMSD) of the structural ensemble is reached.

TM partition

An important aspect of a TM or MP structure is its interaction with the membrane, which can provide clues to its structural and functional roles. It is well known that TMDs have the ability to modulate the thickness of a lipid bilayer[23,45,46], whereas MP regions can cause membrane curvature or deformation to facilitate their function[47,48]. Therefore, accurate determination of the protein membrane partition is very important. We have developed a method, PPT[19,22], that provides PRE data that globally reflect the protein partition in the bicelles. To be applicable, the method assumes that the protein structure is known and requires the protein to be reconstituted in wide or 'ideal' bicelles ($q \geq 0.5$). This requirement, in addition to providing a near-membrane environment for the protein, greatly simplifies the data analysis by ensuring that the measurable PRE is proportional only to the residue position along the bicelle normal, i.e., the axis perpendicular to the bicelle plane. As such, the sample is titrated with either a water-soluble or a lipophilic paramagnetic agent (e.g., Gd-DOTA or 16-DSA, respectively) while recording a 2D ^{1}H-^{15}N TROSY-HSQC (transverse relaxation optimized spectroscopy–heteronuclear single-quantum coherence) spectrum at each titration point. The analysis of the peak intensity decay versus the paramagnetic probe concentration derives residue-specific PRE amplitudes (PRE_{amp} values), which are reporters of residue-specific membrane immersion depths. By exploiting the knowledge of the protein structure, which provides the relative position of each residue along the protein symmetry axis, which is parallel to the bilayer normal, it is then possible to determine the position of the protein relative to the bilayer center that yields the best fit to the experimental PRE_{amp} values, and such placement represents the membrane partition of the protein. This partition analysis can address which protein regions are outside the membrane, measure variations in the membrane thickness around the protein, and, in cases of complex TM proteins, can provide information on the relative orientation of different protein segments.

Materials

Biological materials

- BL21(DE3) competent cells (New England BioLabs, cat. no. C2527). Store them at -80 °C for up to 1 year. **! CAUTION** The cell lines used in your research should be regularly checked to ensure they are authentic and are not infected with mycoplasma.
- Expression plasmid (pMM-LR6 vector, constructed as shown in Fig. 2) coding for the target protein fused to the C terminus of the TrpLE tag. The plasmid carries antibiotic resistance to kanamycin. Store it at -20 °C for up to 1 year.
- Expression plasmid (pET15b vector) coding for GB1 (constructed in our lab and available upon request). A cysteine, additional lysines, and a Strep-tag sequence (CKDKDKWSHPQFEK) are added to the N terminus of the GB1 sequence. The plasmid carries antibiotic resistance to ampicillin. Store it at -20 °C for up to 1 year.
- Expression plasmid (pET15b vector) coding for Foldon (constructed in our lab and available upon request). A His_6-tag is added to the C terminus of the protein sequence. The plasmid carries antibiotic resistance to ampicillin. Store it at −20 °C for up to 1 year.

Reagents

- 1,2-dimyristoyl-d54-*sn*-glycero-3-phosphocholine (deuterated DMPC, d54, 99% (wt/wt); Avanti Polar Lipids, cat. no. 860345)
- 16-DSA (free radical; Sigma-Aldrich, cat. no. 253596)
- (1-Oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)methanethiosulfonate (MTSL; Santa Cruz Biotechnology, cat. no. sc-208677) **! CAUTION** MTSL is light sensitive.
- 2-(*N*-Morpholino)ethanesulfonic acid hydrate (MES; Oakwood Chemical, cat. no. M05729)
- 2,2,2-Trifluoroethanol (TFE, 99% (vol/vol); STREM Chemicals, cat. no. 09-7310)
- 4-(1,1,3,3-Tetramethylbutyl)phenyl-polyethylene glycol (Triton X-100; Sigma-Aldrich, cat. no. X100)
- Acetonitrile (ACN; EMD Millipore, cat. no. EM-AX0151-1)
- Ammonium chloride (NH_4Cl; Sigma-Aldrich, cat. no. A9434)
- Ammonium chloride (^{15}N, 99% (wt/wt); Cambridge Isotope Laboratories, cat. no. NLM-467)
- Ampicillin (Sigma-Aldrich, cat. no. A9393)
- BS3 crosslinker (Thermo Fisher, cat. no. 21580)
- Calcium chloride ($CaCl_2$; Sigma-Aldrich, cat. no. C1016)
- CNBr (Sigma-Aldrich, cat. no. 57654055) **! CAUTION** CNBr is fatal if swallowed, upon contact with skin or if inhaled. It causes severe skin burns and eye damage. Contact with acids liberates very toxic gas. A cyanide antidote kit must be rapidly available and its ingredients replaced every 1–2 years to ensure freshness.
- D_2O (99.96% (vol/vol); Cambridge Isotope Laboratories, cat. no. DLM-6-PK)
- *d*-Desthiobiotin (Sigma-Aldrich, cat. no. D1411)
- D-Glucose (1,2,3,4,5,6,6-D7, 98% (wt/wt); Cambridge Isotope Laboratories, cat. no. DLM-2062)
- D-Glucose (U-$^{13}C6$, 99% (wt/wt); Cambridge Isotope Laboratories, cat. no. CLM-1396)
- D-Glucose (Sigma-Aldrich, cat. no. G8270)
- DHPC (Avanti Polar Lipids, cat. no. 850305)
- Dichloromethane (99.9% (vol/vol); EMD Millipore, cat. no. 75-09-2)
- Dimethyl sulfoxide (DMSO; Sigma-Aldrich, cat. no. 276855)
- Disodium hydrogen phosphate (Na_2HPO_4; Sigma-Aldrich, cat. no. S3264)
- DMPC (Avanti Polar Lipids, cat. no. 850345)
- DTT (VWR, cat. no. 97061) **! CAUTION** DTT may cause skin and eye irritation.
- EDTA disodium salt dehydrate (EDTA; Sigma-Aldrich, cat. no. E6635)
- Formic acid (FA, 90% (vol/vol); VWR International, cat. no. JT0129-1) **! CAUTION** FA is a combustible liquid. It causes severe skin burns and eye damage and is toxic if inhaled.
- Gd-DOTA (Macrocyclics, cat. no. M-147) **! CAUTION** Gd-DOTA may cause respiratory and skin irritation.
- Glutaraldehyde (Sigma-Aldrich, cat. no. G7776)
- Guanidine hydrochloride (VWR International, cat. no. 71003)
- HEPES (Sigma-Aldrich, cat. no. H3375)
- HFIP (99.5% (vol/vol), Oakwood Products, cat. no. 003409)
- HisPur Ni-NTA resin (Thermo Fisher, cat. no. 88223)
- Hydrochloric acid (HCl, 37% (vol/vol); Sigma-Aldrich, cat. no. 320331)

- Imidazole (Sigma-Aldrich, cat. no. I5513)
- IPTG (Sigma-Aldrich, cat. no. 329815691)
- Isopropyl alcohol (IPA; EMD Millipore, cat. no. EM-PX1838P-1)
- Kanamycin sulfate (Sigma-Aldrich, cat. no. 60615)
- L-Ascorbic acid (Sigma-Aldrich, cat. no. A92902)
- Liquid nitrogen (Medical-Technical Gases, cat. no. UN1977)
- Luria–Bertani (LB) agar (granulated; RPI Research Products International, cat. no. L24033-500)
- LB broth (VWR Life Science, cat. no. J106)
- Magnesium sulfate ($MgSO_4$; Sigma-Aldrich, cat. no. M7506)
- NHS-activated agarose resin (Thermo Fisher, cat. no. 26196)
- Nitrogen gas (Airgas, cat. no. UN1066)
- Pierce protease inhibitor tablets (EDTA-free; Thermo Fisher, cat. no. A32965)
- Potassium dihydrogen phosphate (KH_2PO_4; Sigma-Aldrich, cat. no. P5655)
- SM(PEG)2 (PEGylated succinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate (SMCC) crosslinker; Thermo Fisher, cat. no. 22103)
- SOC medium (New England BioLabs, cat. no. B9020S)
- Sodium bicarbonate ($NaHCO_3$; Sigma-Aldrich, cat. no. S5761)
- Sodium chloride (NaCl; Sigma-Aldrich, cat. no. S7653)
- Sodium hydroxide (NaOH; Sigma-Aldrich, cat. no. S8045)
- Sodium phosphate, monobasic (NaH_2PO_4; Sigma-Aldrich, cat. no. S3139)
- Sulfo-NHS-acetate (Thermo Fisher, cat. no. 26777)
- Triethylammonium acetate (TEAA; Calbiochem, cat. no. 625718)
- Trifluoroacetic acid (TFA; EMD Millipore, cat. no. EM-TX1275-3) **! CAUTION** TFA causes severe skin burns and eye damage. Harmful if inhaled.
- TriNTA (synthesized by Medicilon (Shanghai, China) upon request). The detailed description of the synthesis was previously published[42].
- Tris base (Tris; Calbiochem, cat. no. 648310)
- Urea (Thermo Fisher Scientific, cat. no. 15505050)

Equipment

- Amicon Ultra-15 centrifugal filter unit (EMD Millipore, cat. no. UFC900324)
- Autosampler (Agilent, model no. 1200)
- Dialysis cassettes (0.5–3.0-mL and 3.0–12.0 mL, 3,500 MWCO; Thermo Fisher, cat. nos. 87723 and 66110)
- Eppendorf tube (Thermo Fisher, cat. no. AM12450)
- Falcon tubes (15 and 50 ml; Thermo Fisher, cat. nos. 339650 and 339652)
- Fourier transform mass spectrometer (Thermo Fisher Scientific, LTQ-Orbitrap XL model)
- FPLC instrument (e.g., AKTA pure protein purification system; GE Healthcare, model no. 29018226) consisting of multiple pumps, sampler, detectors, and fraction collector
- HPLC instrument (e.g., Bio-Rad Duo Flow system; Bio-Rad, model no. 7600037) consisting of a degasser, sampler, pumps, and detectors (to measure conductivity and UV absorbance at 214 and 280 nm)
- Hydrophilic nylon membrane, 0.2-μm pore size, 47-mm diameter (Millipore Sigma, cat. no. GNWP04700)
- LC column (BioBasic 18 LC column, 2.1 × 100 mm, 5-μm particle size, 300 Å; Thermo Fisher Scientific, cat. no. 72105-102130)
- NMR spectrometer fully equipped for triple-resonance experiments with biological macromolecules, equipped with a cryogenically cooled probe head with *z*-field gradients. For measurement in NOESY experiments, the use of strong magnetic field is recommended (800–900 MHz)
- NanoDrop 2000 Spectrophotometer (Thermo Fisher, model no. ND-2000)
- NMR tube (Shigemi, cat. no. BMS-005B)
- PD-10 column (GE Healthcare, cat. no. 17085101)
- StrepTrap HP column (GE Healthcare, cat. no. 28907547)
- Superdex S75 16/60 column; GE Healthcare, cat. no. 28989334)
- Puradisc syringe filter FP30 (GE Healthcare, cat. no. 10462610)
- Tissue grinder (50 mL, glass; VWR, model no. 47732-450)
- Vortex mixer (VWR, model no. 10153-836)

- Zorbax SB-C3 column (9.4 × 250 mm; Agilent Technologies, cat. no. 880995-209)
- Zorbax SB-C18 column (9.4 × 250 mm; Agilent Technologies, cat. no. 880995-202)

Software

- Data analysis software (e.g., Origin (https://www.originlab.com/), MATLAB (https://www.mathworks.com/products/matlab.html))
- ExSSO program (http://www.csbio.sjtu.edu.cn/bioinf/ExSSO/), see also Supplementary Software 1 and 2
- Molecular graphics software (e.g., PyMOL (https://pymol.org/2/), UCSF Chimera (https://www.cgl.ucsf.edu/chimera/))
- NMR data-processing software (e.g., TopSpin (https://www.bruker.com/service/support-upgrades/software-downloads/nmr.html), NMRPipe (https://www.ibbr.umd.edu/nmrpipe/))
- NMR structure-calculation software (e.g., XPLOR-NIH (https://nmr.cit.nih.gov/xplor-nih/), CYANA (http://www.cyana.org/), ARIA (http://aria.pasteur.fr/))
- ProtParam (https://web.expasy.org/protparam/)
- TALOS+ (https://spin.niddk.nih.gov/bax/software/TALOS/)
- Visualization and analysis software for NMR spectra (e.g., TopSpin (https://www.bruker.com/service/support-upgrades/software-downloads/nmr.html), CARA (http://www.cara.nmr.ch/), Sparky (https://www.cgl.ucsf.edu/home/sparky/), VnmrJ (http://openvnmrj.org/), CCPN (https://www.ccpn.ac.uk/))

Reagent setup

16-DSA, 20 mM

Dissolve 1.0 mg of 16-DSA in a bicelle solution of $q = 0.5$. Store the solution at room temperature (RT; 25 °C) and use for the duration of the NMR titration (3 d–1 week, Steps 155–157). ▲ **CRITICAL** The solubilization of 16-DSA in the bicelles can be aided by flash-freezing and then thawing the solution a couple of times, as well as by gentle sonication.

Ascorbic acid, 0.5 M

Dissolve 0.9 g of ascorbic acid in 10 mL of previously degassed NMR buffer (pH 6.7). Adjust the pH with NaOH (1 M) to match that of the NMR buffer. Prepare fresh and discard after use. ▲ **CRITICAL** Adjust the pH of the solution to be as close as possible to that of the final NMR sample.

Bicelle solution

Prepare a bicelle solution of $q = 0.5$ by dissolving 9 mg of DMPC and 12.04 mg of DHPC in 130 µL of NMR buffer (pH 6.7). Store the solution at 25 °C for up to 1 week.

$CaCl_2$, 1 M

Dissolve 14.7 g of $CaCl_2$ in 100 mL of dH_2O. Store the solution at RT for up to 1 year. ▲ **CRITICAL** For perdeuterated samples, dissolve $CaCl_2$ in 99.96% D_2O, instead of dH_2O and store at RT for up to 6 months.

DHPC (100 mg/mL)

Dissolve 100 mg of DHPC (protonated or deuterated) in 1 mL of NMR buffer (pH 6.7). Split the solution into 50-µL aliquots and store them at -20 °C for up to 1 year.

DTT, 1 M

Dissolve 1.6 g of DTT in dH_2O to reach a final volume of 10 mL. Split the solution into 1-mL aliquots and store them at -20 °C. Use the solution within 1 year.

GB1 storage buffer

GB1 storage buffer is 25 mM HEPES, pH 7.2, 10 mM DTT. Dissolve 5.95 g of HEPES and 1.54 g of DTT in 1 L of dH_2O. Adjust the pH to 7.2 with 37% (vol/vol) HCl. Make the buffer fresh.

Gd-DOTA, 600 mM

Dissolve 97.9 mg of Gd-DOTA in 250 µL of NMR buffer (pH 6.7). Split the solution into 25-µL aliquots and store them at -20 °C. Use the solution within 1 year.

Gd-DOTA, 200 mM
Dilute the 600 mM Gd-DOTA solution 1:3 with NMR buffer (pH 6.7) (e.g., 10 µL of 600 mM Gd-DOTA in 20 µL of NMR buffer). Store at RT and use for the duration of the NMR titration (Steps 155–157); then discard.

Guanidine buffer
Guanidine buffer is 6 M guanidine–HCl, 50 mM Tris, pH 8.0, 200 mM NaCl, 1% (vol/vol) Triton X-100. Dissolve 573 g of guanidine–HCl, 6 g of Tris base, and 11.8 g of NaCl in dH_2O to reach a final volume of 1 L. Adjust the pH to 8.0 with 37% (vol/vol) HCl. Add 10 mL of Triton X-100. Store the solution at RT for up to 6 months.

HEPES buffer
HEPES buffer is 25 mM HEPES, pH 7.5. Dissolve 5.95 g of HEPES in 1 L of dH_2O. Adjust the pH to 7.5 with 37% (vol/vol) HCl. Store the solution at RT for up to 1 year.

HPLC buffer A
HPLC buffer A is 5% (vol/vol) IPA, 95% (vol/vol) H_2O, 0.1% (vol/vol) TFA. For 1 L of solution, mix 50 mL of IPA, 950 mL of dH_2O, and 1 mL of TFA. Before use, degas and filter the solution with a 0.2-µm pore size, hydrophilic nylon membrane. Store the solution at RT for up to 6 months.

HPLC buffer B
HPLC buffer B is 75% (vol/vol) IPA, 25% (vol/vol) ACN, 0.1% (vol/vol) TFA. For 1 L of solution, mix 750 mL of IPA, 250 mL of ACN, and 1 mL of TFA. Before use, degas and filter the solution with a 0.2-µm pore size, hydrophilic nylon membrane. Store the solution at RT for up to 6 months.

HPLC buffer C
HPLC buffer C is 95% (vol/vol) IPA, 5% (vol/vol) H_2O, 0.1% (vol/vol) TFA. For 1 L of solution, mix 950 mL of IPA, 50 mL of dH_2O, and 1 mL of TFA. Before use, degas and filter the solution with a 0.2-µm pore size, hydrophilic nylon membrane. Store the solution at RT for up to 6 months.

HPLC buffer D
HPLC buffer D is 5% (vol/vol) ACN, 0.1 M TEAA. For 1 L of solution, mix 50 mL of ACN, 100 mL of 1M TEAA, and 850 mL of dH_2O. Before use, degas and filter the solution with a 0.2-µm pore size, hydrophilic nylon membrane. Store the solution at RT for up to 1 year.

HPLC buffer E
HPLC buffer E is 60% (vol/vol) ACN, 0.1 M TEAA. For 1 L of solution, mix 600 mL of ACN, 100 mL of 1M TEAA, and 300 mL of dH_2O. Before use, degas and filter the solution with a 0.2-µm pore size, hydrophilic nylon membrane. Store the solution at RT for up to 1 year.

IPTG, 1 M
Dissolve 2.4 g of IPTG in 10 mL of dH_2O. Split the solution into 1-mL aliquots and store them at -20 °C for up to 1 year. ▲ **CRITICAL** For perdeuterated samples, dissolve IPTG in 99.96% D_2O instead of dH_2O.

LB agar plates
The LB agar for the plates contains 50 µg/mL kanamycin sulfate. Approximately ten plates are obtained per 200 mL of prepared solution. Dissolve the LB agar in the desired volume of dH_2O so that its concentration is 37 g/L. Autoclave the solution and let it cool under gentle shaking to prevent solidification of the agar. When the solution temperature reaches ~50 °C, add kanamycin sulfate to a final concentration of 50 µg/mL and mix well. Operate close to the flame to avoid buffer contamination. Divide the medium and pour it into sterile plates. After gel coagulation, store the plates at 4 °C for up to 1 month.

LB medium (K 50 µg/mL)
LB medium contains 50 µg/mL kanamycin sulfate. Dissolve 25 g of LB broth in 1 L of dH_2O and autoclave the solution. Once the solution has cooled to RT, add 50 mg of kanamycin sulfate

(50 μg/mL). Operate close to the flame to avoid contamination of the medium. Store the solution at RT for up to 1 month.

Lysis buffer
Lysis buffer is 50 mM Tris, pH 8.0, 200 mM NaCl. Dissolve 6.1 g of Tris base and 11.8 g of NaCl in 1 L of dH_2O. Adjust the pH to 8.0 with 37% (vol/vol) HCl. Store the solution at RT for up to 6 months.

M9 medium (1 L)
Dissolve 6 g of Na_2HPO_4, 3 g of KH2PO_4, 0.5 g of NaCl, and 1 g of NH_4Cl (^{15}N, 99%, for ^{15}N-labeled samples) in 1 L of dH_2O (or 99.96% D_2O, for perdeuterated samples, as specified in Steps 10, 12, 14, and 15). Autoclave the solution, let it cool to RT, and then add 4g of D-glucose (or 99% U-$^{13}C6$, for ^{13}C-labeled samples; or 98% 1,2,3,4,5,6,6-D7, for perdeuterated samples), 2 mL of 1 M $MgSO_4$ (in 99.96% D_2O, for perdeuterated samples), 100 μL of 1 M $CaCl_2$ (in 99.96% D_2O, for perdeuterated samples), and 50 mg of kanamycin sulfate (50 μg/mL). Mix well until complete dissolution of the chemicals. Store the medium at RT and use it within a couple of days. ▲ **CRITICAL** When using 99.96% D_2O (perdeuterated samples), do not autoclave the medium to prevent moisture contamination; autoclave only the empty flask before preparing the solution.

$MgSO_4$, 1 M
Dissolve 24 g of $MgSO_4$ in 200 mL of dH_2O. Store the solution at RT for up to 1 year. ▲ **CRITICAL** For perdeuterated samples, dissolve $MgSO_4$ in 99.96% D_2O, instead of dH_2O. Store at RT for up to 6 months.

MTSL, 100 mM
Dissolve 5.3 mg of MTSL in 200 μL of DMSO. Split the solution into 20-μL aliquots and store them at -20 °C. Use the solution within 1 year. ! **CAUTION** Store the aliquots in dark vials or cover them with aluminum foil.

Ni-NTA binding buffer
Ni-NTA binding buffer is 25 mM Na_2HPO_4, pH 7.4, 100 mM NaCl. Dissolve 3.55 g of Na_2HPO_4 and 5.85 g of NaCl in 1 L of dH_2O. Adjust the pH to 7.4 with 37% (vol/vol) HCl. Store the solution at RT for up to 1 year.

Ni-NTA elution buffer
Ni-NTA elution buffer is 25 mM Na_2HPO_4, pH 7.4, 100 mM NaCl, 500 mM imidazole. Dissolve 3.55 g of Na_2HPO_4, 5.85 g of NaCl, and 34 g of imidazole in 1 L of dH_2O. Adjust the pH to 7.4 with 37% (vol/vol) HCl. Store the solution at RT for up to 1 year.

NMR buffer (pH 6.7)
The NMR buffer can be optimized for the specific sample under study. Typically, it consists of ~20–50 mM phosphate or MES buffer at close to physiological pH, i.e., pH 6.7 (e.g., dissolve 5.97 g of Na_2HPO_4 and 3.83 g of NaH_2PO_4 in 1 L of H_2O to make a 50 mM phosphate buffer at pH 6.7; dissolve 9.76 g of MES free acid in 1 L of H_2O and adjust the pH to 6.7 with 1 M NaOH to make a 50 mM MES buffer). Store at RT for up to 1 year.

NMR buffer (pH 6.2)
Prepare a buffer to the identical to the NMR buffer (pH 6.7), but at pH 6.2. Store at RT for up to 1 year. Degas right before use.

OG-label buffer
OG-label buffer is 25 mM Na_2HPO_4, pH 7.5, 200 mM NaCl. Dissolve 3.55 g of Na_2HPO_4 and 11.7 g of NaCl in 1 L of dH_2O. Adjust the pH to 7.5 with 37% (vol/vol) HCl. Store the solution at RT for up to 1 year.

Pi buffer, 400 mM (pH 7.5)
Dissolve 4.6 g of Na_2HPO_4 and 1 g of NaH_2PO_4 in 100 mL of dH_2O. Adjust the pH to 7.5 with 37% (vol/vol) HCl. Store the solution at RT for 1 year.

Phosphate NMR buffer
Phosphate NMR buffer is 25 mM Na_2HPO_4, pH 7.2. Dissolve 3.55 g of Na_2HPO_4 in 1 L of dH_2O. Adjust the pH to 7.2 with 37% (vol/vol) HCl. Store the solution at RT for up to 1 year.

StrepTrap binding buffer
StrepTrap binding buffer is 100 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, 10 mM DTT. Dissolve 12.1 g of Tris base, 8.8 g of NaCl, 0.372 g of EDTA, and 1.54 g of DTT in 1 L of dH_2O. Adjust the pH to 8.0 with 37% (vol/vol) HCl. Make the buffer fresh.

StrepTrap elution buffer
StrepTrap elution buffer is 100 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, 10 mM DTT, 2.5 mM desthiobiotin. Dissolve 12.1 g of Tris base, 8.8 g of NaCl, 0.372 g of EDTA, 1.54 g of DTT, and 0.535 g of desthiobiotin in 1 L of dH_2O. Adjust the pH to 8.0 with 37% (vol/vol) HCl. Store the solution at 4 °C for up to 1 month.

TE buffer
TE buffer is 10 mM Tris, pH 8.0, 1 mM EDTA. Dissolve 1.21 g of Tris base and 0.372 g of EDTA in 1 L of dH_2O. Adjust the pH to 8.0 with 37% (vol/vol) HCl. Store the solution at RT for up to 1 year.

TriNTA reaction buffer (100 mM sodium bicarbonate, pH 8.0)
TriNTA reaction buffer is 100 mM $NaHCO_3$, pH 8.0. Dissolve 8.4 g of $NaHCO_3$ in 1 L of dH_2O. Adjust the pH to 8.0 with 37% (vol/vol) HCl. Make the buffer fresh and discard after use.

Tris buffer (1 M Tris, pH 7.5)
Dissolve 121 g of Tris base in 1 L of dH_2O. Adjust the pH to 7.5 with 37% (vol/vol) HCl. Store the solution at RT for up to 1 year.

Urea, 8 M
Dissolve 480.5 g of urea in dH_2O to reach a final volume of 1 L. The solution can be heated to aid the urea dissolution. Store the solution at RT for up to 1 year.

Equipment setup

'Cleaning' HPLC method
- Isocratic flow: HPLC buffer A (100%, (vol/vol)). Flow rate: 2 mL/min. Total volume: 30 mL.
- Linear gradient: from 100% (vol/vol) HPLC buffer A/0% (vol/vol) HPLC buffer B to 0% (vol/vol) HPLC buffer A/100% (vol/vol) HPLC buffer B. Flow rate: 2 mL/min. Total volume: 50 mL.
- Isocratic flow: 100% (vol/vol) HPLC buffer B. Flow rate: 2 mL/min. Total volume: 30 mL.

'Sample purification' HPLC method
- Isocratic flow: 100% (vol/vol) HPLC buffer A. Flow rate: 2 mL/min. Total volume: 15 mL.
- Linear gradient: from 100% (vol/vol) HPLC buffer A/0% (vol/vol) HPLC buffer B to 70% (vol/vol) HPLC buffer A/30% (vol/vol) HPLC buffer B. Flow rate: 2 mL/min. Total volume: 5 mL.
- Linear gradient: from 70% (vol/vol) HPLC buffer A/30% (vol/vol) HPLC buffer B to 0% (vol/vol) HPLC buffer A /100% (vol/vol) HPLC buffer B. Flow rate: 2 mL/min. Total volume: 150 mL.
- Isocratic flow: 100% (vol/vol) HPLC buffer B. Flow rate: 2 mL/min. Total volume: 20 mL.

'Protein refinement' program
- Connect the Superdex S75 16/60 column to the FPLC system.
- Connect the pump to the GB1 storage buffer or Ni-NTA binding buffer (for purifying GB1 or Foldon, respectively).
- Connect the 5-mL sample loop.
- Empty the loop. Flow rate: 1 mL/min. Total volume: 10 mL.
- Wash the Superdex S75 16/60 column with GB1 storage buffer or Ni-NTA binding buffer (for purifying GB1 or Foldon, respectively). Flow rate: 1 mL/min. Total volume: 150 mL.
- Collect fractions, starting from the 10th mL; collect 1 mL per fraction.

'SMCC-TriNTA purification' program

- Connect the C18 column to the HPLC system.
- Connect the HPLC buffer D to pump A and the HPLC buffer E to the pump B.
- Isocratic flow to empty the sample loop: 100% (vol/vol) HPLC buffer D. Flow rate: 3 mL/min. Total volume: 15 mL.
- Linear gradient: from 100% (vol/vol) HPLC buffer D/0% (vol/vol) HPLC buffer E to 50% HPLC buffer D/50% (vol/vol) HPLC buffer E. Flow rate: 3 mL/min. Total volume: 120 mL.
- Linear gradient: from 50% (vol/vol) HPLC buffer D/50% (vol/vol) HPLC buffer E to 0% (vol/vol) HPLC buffer D/100% (vol/vol) HPLC buffer E. Flow rate: 3 mL/min. Total volume: 30 mL.
- Isocratic flow: 100% (vol/vol) HPLC buffer B. Flow rate: 3 mL/min. Total volume: 30 mL.

Procedure

Part 1: Expression, purification, and reconstitution of TMDs

Protein expression

▲ **CRITICAL** The following protocol (Steps 1–21) describes the procedure for expression of perdeuterated proteins. For protonated samples, Steps 10–13 can be omitted and dH_2O should be used instead of D_2O in Steps 14 and 15.

1 Take one vial (50 µL) of BL21(DE3) competent *E. coli* cells and allow the vial to thaw on ice.

2 Add 1 µL of plasmid coding for the target protein to the competent cells and mix well. Incubate the mixture on ice for 30 min.

3 Move the mixture into a warm bath or incubator at 42 °C. Incubate for 50 s.

4 Transfer the mixture onto ice and incubate for 2 min.

5 Add 200 µL of SOC medium to the cells and move them into an incubator shaker at 37 °C for 30 min.

6 Spread the cells on an LB agar plate (50 µg/mL kanamycin sulfate), using a sterile glass pipette or plating beads.

7 Incubate the plate at 37 °C for 16 h.
■ **PAUSE POINT** This step is typically carried out overnight. The plate can then be stored at 4 °C, but the *E. coli* colonies should be grown within 1–2 d after transformation.

8 Pick a single colony from the plate and inoculate it into 5 mL of LB medium (50 µg/mL kanamycin sulfate). Move the tube into an incubator shaker at 37 °C and shake at 220 r.p.m. for 8 h.

9 Centrifuge the cells at 4 °C at 5,000*g* for 5 min.

10 Discard the supernatant (by pouring it off and then pipetting out the residual), inoculate the cells into 100 mL of M9 medium (H_2O), and incubate the culture at 37 °C with shaking at 220 r.p.m. until the OD_{600} value is close to 0.4 (~2 h).

11 Centrifuge the cells at 4 °C at 5,000*g* for 5 min.

12 Discard the supernatant (by pouring it off and then pipetting out the residual), inoculate the cells into 10 mL of M9 medium (50% (vol/vol) H_2O, 50% (vol/vol) D_2O), and incubate the culture at 37 °C with shaking at 220 r.p.m. until the OD_{600} value is close to 1.0 (~2 h).

13 Centrifuge 2 mL of culture at 4 °C at 5,000*g* for 5 min to collect cells.

14 Discard the supernatant (by pouring it off and then pipetting out the residual), inoculate the cells into 100 mL of M9 medium (D_2O), and incubate the culture at 37 °C with shaking at 220 r.p.m. until the OD_{600} value is close to 1.0 (~8 h).

15 Inoculate the entire culture into 1 L of M9 medium (D_2O). Incubate the culture at 37 °C with shaking at 220 r.p.m. until the OD_{600} value reaches to 0.6–0.8 (~6 h). Collect 100 µL of culture, measure its OD_{600} value and store it at 4 °C (for up to 1 week), to be used later (in Step 21) as a control for evaluating the protein expression level.

16 Change the incubator shaker temperature to the desired temperature for protein expression (e.g., 22 °C). Wait ~30 min until the culture reaches the desired temperature.

17 Induce protein expression by adding 0.15 mM IPTG from the 1 M stock solution.

18 Grow the culture with shaking at 220 r.p.m. for ~16 h at 22 °C. Collect 100 µL of culture, measure its OD_{600} value and store it at 4 °C (for up to 1 week), to be used later (in Step 21) to evaluate the protein expression level.

19 Collect the cells by centrifuging at 8,000*g* for 30 min at 4 °C.

20 Discard the supernatant (by pouring it off) and suspend the cell pellet in 50 mL of lysis buffer at RT.
■ **PAUSE POINT** The cell suspension can be stored at −80 °C for up to 1 week.

21 Evaluate the protein expression level by SDS–PAGE analysis[49] of the samples collected at Steps 15 (before induction) and 18 (after expression). For a better comparison, use the measured OD_{600} values to determine the amount of each sample to be used for the analysis so that all samples contain a comparable number of cells in the gel.
? TROUBLESHOOTING

Ni-NTA purification

22 Disrupt the cell suspension from Step 20 by sonicating it for 10 min at intervals of 1 s; apply pulses of 1 s at 40% of the maximum power. Keep the sample on ice during the procedure.
23 Centrifuge the suspension at 35,000*g* for 20 min at 4 °C.
24 Take 10 µL of supernatant and precipitate from Step 23 for SDS–PAGE analysis[49] to verify that the target protein is located in the precipitate.
25 Discard the remaining supernatant from Step 23 (by pouring it off) and dissolve the precipitate (from Step 23) in 40 mL of guanidine buffer, using a 50-mL glass tissue grinder.
26 Centrifuge the solution at 35,000*g* for 30 min at 4 °C.
27 Transfer the supernatant to a 50-mL Falcon tube and add 4 mL of HisPur Ni-NTA resin. Mix them well by continuously, but gently, stirring the mixture on a rotator for ~2–8 h at RT.
▲CRITICAL STEP Before use, wash the HisPur Ni-NTA resin several times with dH_2O to completely remove the ethanol used for its storage.
■PAUSE POINT The mixture can be left stirring overnight.
28 Transfer the mixture to a glass chromatography column for gravity-flow purification. The target protein, bound to the Ni-NTA resin, remains trapped in the column. Discard the flow-through.
29 Wash the column twice with 50 mL of 8 M Urea. Discard the flow-through.
30 Wash the column twice with 20 mL of dH_2O. Discard the flow-through. Do not let the resin dry.
31 To release the target protein from the Ni-NTA resin, add 4 mL of 90% FA. Let the acid react at RT for ~2–5 min. Collect the eluate in a 50-mL Falcon tube.
32 Repeat Step 31 two more times, eluting the protein in a total volume of 12 mL of 90% FA.

CNBr cleavage

33 Before use, let the CNBr warm to RT (~30 min).
! CAUTION CNBr is very toxic, so it should be always manipulated under a fume hood while wearing the appropriate personal protective equipment (PPE). All the equipment and waste contaminated by CNBr should be collected in a dedicated hazardous waste container for proper disposal.
34 Add ~1.5 g of CNBr to the 12-mL protein solution in 90% FA. Use a vortex mixer to completely dissolve the CNBr.
35 Cover the tube with aluminum foil to shield it from light and let the reaction occur under a gentle nitrogen gas stream for 1 h.
▲CRITICAL STEP Do not increase the reaction time, as this may increase the occurrence of side reactions, e.g., formylation.
? TROUBLESHOOTING
36 After the reaction has completed, transfer the solution to a 3.0- to 12.0-mL dialysis cassette (3,500 MWCO) and dialyze it versus 4 L of dH_2O for 40 min.
37 Dialyze the sample second time versus 4 L of dH_2O for 40 min.
38 Transfer the solution to a 50-mL Falcon tube and prepare the sample for lyophilization by making a few holes in the tube cap.
39 Freeze the solution in liquid nitrogen (-196 °C) for 10–15 min, making sure the solution is well-frozen. Transfer the tube to the lyophilizer (-80 °C, ~10–20 mBar) and lyophilize the protein (for ~1 d).
■PAUSE POINT After lyophilization, the protein powder can be stored at −20 °C for 1 month.

Reverse-phase HPLC

40 Switch on the HPLC instrument and annexed detectors. Connect a Zorbax SB-C3 column to the instrument.
41 Each time a new buffer is used, wash the system and related pump for 1–2 min (flow rate: 10 mL/min).
▲CRITICAL STEP Make sure that no bubbles are present in the system.

42 Wash and equilibrate the column with 20 mL of each of the following buffers sequentially, adjusting the flow rate depending on each buffer viscosity to maintain a column pressure of ~1,500–2,000 p.s.i.
- dH_2O
- HPLC buffer A
- HPLC buffer C
- Dichloromethane
- HPLC buffer C
- HPLC buffer B
- HPLC buffer A

43 Wash the column by injecting 5 mL of 90% FA and running the 'Cleaning' HPLC method (see 'Equipment setup' section).
▲ **CRITICAL STEP** Remove all bubbles before loading the 90% FA; remove the air from the loop before running the 'Cleaning' HPLC method.
? TROUBLESHOOTING

44 After the column is clean, equilibrate it with buffer A until a steady conductivity value (~5 mS/cm) is reached.

45 Dissolve the dried protein powder (from Step 39) in 4 mL of 90% FA. Load the sample (5 mL) into the column and run the 'Sample purification' HPLC method (see 'Equipment setup' section). Collect fractions of 2-mL volume.
? TROUBLESHOOTING

46 Save the fractions exhibiting UV_{280} and UV_{214} absorbances. For each elution peak, lyophilize a small aliquot (transfer 100–200 µL of sample to an Eppendorf tube; dilute 1:1 with dH_2O; flash-freeze the sample in liquid nitrogen; and then lyophilize it at −80 °C and ~10–20 mBar until completely dry) and use it for SDS–PAGE analysis to identify which peak represents the target protein.
? TROUBLESHOOTING

47 Lyophilize the fractions of the elution peak containing the purified target protein (identified in Step 46) as described in Steps 38 and 39. ACN or dH_2O can be added in a 1:2 ratio to the HPLC fractions to avoid melting of the solution in the lyophilizer.
■ **PAUSE POINT** After lyophilization, the protein powder can be stored at −20 °C for 1 month.

Protein reconstitution in DMPC/DHPC bicelles

48 Dissolve the lyophilized protein powder (1–2 mg) from Step 47 in 1 mL of HFIP. Mix it with ~9 mg of DMPC (protonated or deuterated) and 27 mg of DHPC (protonated or deuterated).

49 Dry the solution under a gentle nitrogen stream until a thin film is obtained, and then lyophilize it (−80 °C, ~10–20 mBar) overnight.
■ **PAUSE POINT** The dried mixture can be stored at −20 °C for 1 month.

50 Dissolve the dried solution in 3 mL of 8 M urea. If the solution is not clear, add additional DHPC (~5 mg). Mix well until the solution becomes clear. If needed, add more DHPC in small amounts (~2–5 mg).

51 Transfer the solution to a 0.5- to 3.0-mL dialysis cassette (3,500 MWCO) and dialyze it for 3 h versus 1 L of NMR buffer (pH 6.7). Stir gently. If the sample starts to become cloudy, add 50 µL of DHPC (100 mg/mL) directly to the dialysis cassette.
▲ **CRITICAL STEP** During the dialysis, the DHPC gradually diffuses outside the cassette, while the DMPC remains trapped inside. When approaching the liposome state, it is important to replenish the lost DHPC to prevent possible protein aggregation, which may occur especially with highly concentrated samples.
? TROUBLESHOOTING

52 Perform a second dialysis versus 1 L of NMR buffer (pH 6.7) for 3 h. If the sample starts to become cloudy, add 50 µL of DHPC (100 mg/mL) directly to the dialysis cassette. Adjust the DHPC amount so that the bicelle q is close to 0.5 (Step 54).

53 Transfer the sample to a 4-mL Amicon Ultra-15 centrifugal filter unit (3,000 MWCO) and concentrate it to a volume of ~300–350 µL by centrifuging at 6,000*g* for ~1 hour at RT. Add 10% (vol/vol) D_2O and transfer the sample to a Shigemi NMR tube. Use the sample immediately for further analysis/NMR experiments.
■ **PAUSE POINT** Alternatively, store at RT for few days or at -80 °C for up to 1 month.

54 Determine the exact bicelle q of the sample from Step 53 by recording a ^{1}H NMR spectrum (^{2}H NMR spectrum, if deuterated DMPC and DHPC are used). Quantify the relative amounts of DHPC and DMPC (~0.90 and ~0.85 p.p.m., respectively) as the ratio of the integral of their NMR signals[19]. If needed, adjust the q value to the desired value, e.g., $q = 0.5$ (see 'Troubleshooting' section).

▲ **CRITICAL STEP** The final DMPC concentration is 50 mM, and the protein concentration is in the range of 0.3–1.0 mM, depending on the requirements of each particular NMR experiment. Therefore, the lipid-to-protein ratio is between 50:1 and 167:1 in our sample system, which is comparable to that of natural cell membranes (estimated to be between 50:1 and 800:1 (refs. [50,51])).

? TROUBLESHOOTING

Part 2: Determination of the oligomeric state of TMDs in bicelles

Intermolecular PRE analysis: sample preparation

55 Dissolve the ~1:1 (wt/wt) mixture of lyophilized ^{15}N-labeled protein and Cys mutant of the same protein (expressed and purified as described in Steps 1–47 but with 5 mM DTT added to the lysis and guanidine buffers) in 1 mL of HFIP. Then add ~9 mg of DMPC and 27 mg of DHPC and mix.

56 Repeat Steps 49 and 50.

57 Add 30 µL of 1 M DTT to the solution. Mix well. Incubate for ~10 min, then gently centrifuge the solution at 2,000*g* for 2 min at RT.

58 Repeat Step 51, using 1 L of NMR buffer (pH 6.7) to which 1 mL of 1 M DTT has been added. After the dialysis, add 15 µL of 1 M DTT directly to the dialysis cassette.

59 Repeat Step 52, using 1 L of NMR buffer (pH 6.7) to which 1 mL of 1 M DTT has been added.

60 Perform a third dialysis versus 1 L of degassed NMR buffer (pH 6.2) for 3 h. When the sample starts to become cloudy, add 50 µL of DHPC (100 mg/mL) directly to the dialysis cassette.

61 During the third dialysis, use 3.5 mL of degassed NMR buffer (pH 6.2) (this volume corresponds approximately to the actual sample volume in the dialysis cassette) to determine the amount of 400 mM Pi buffer (pH 7.5) required to raise its pH to ~7.2–7.3. Note the volume added.

62 5–10 min before the end of the dialysis, thaw one aliquot of 100 mM MTSL solution.

▲ **CRITICAL STEP** Shield the vial from light.

63 After the dialysis is finished, transfer the sample to a 15-mL Falcon tube. Quickly raise the sample pH to ~7.2–7.3 by adding the same amount of 400 mM Pi buffer (pH 7.5) as determined in Step 61, then immediately add 10 µL of 100 mM MTSL (the MTSL concentration should be approximately ten times in excess with respect to that of the Cys mutant). Mix well, seal the tube with Parafilm, and cover it with aluminum foil. Let the reaction proceed overnight at RT.

▲ **CRITICAL STEP** Make sure that the tube is shielded from light during the reaction.

? TROUBLESHOOTING

■ **PAUSE POINT** The sample can be left overnight at RT.

64 Repeat Step 51, using 1 L of NMR buffer (pH 6.7).

▲ **CRITICAL STEP** Maintain the dialysis apparatus shielded from light.

65 Repeat Step 64 three additional times, always keeping the dialysis apparatus shielded from light.

▲ **CRITICAL STEP** The excess of free MTSL left from the reaction in Step 63 gradually diffuses outside the dialysis cassette together with the detergent. To ensure that the free MTSL has been completely removed from the sample, perform at least four dialysis steps (Steps 64 and 65) so that the amount of lost DHPC is ≥60 mg.

? TROUBLESHOOTING

66 Repeat Step 53.

▲ **CRITICAL STEP** Cover the NMR tube with aluminum foil to shield it from light.

Intermolecular PRE analysis: NMR experiments

67 Set the probe temperature to the desired value (e.g., 308 K). Insert the sample (from Step 66) into the magnet and wait ~15 min for the sample to achieve temperature stability. Lock, tune, and shim the magnet. Activate the auto-shim process.

68 Determine the pulse lengths of hard 90° pulses for ^{1}H and ^{15}N (in the direct and indirect acquisition mode setup, respectively).

69 Measure a 1D ^{1}H NMR spectrum to determine the actual bicelle q of the sample. If needed, adjust it to 0.5 as explained in Step 54.

70 Acquire a high-resolution 2D ^{1}H-^{15}N TROSY-HSQC spectrum of the paramagnetic state.
▲ **CRITICAL STEP** The recovery delay should be set to at least 3.5 s to prevent the occurrence of PRE of the NMR signals' longitudinal relaxation rates.

71 Process the spectrum with standard parameters[52]. Typically, raw data are multiplied by an apodization function, zero-filled and Fourier-transformed in both dimensions. Proper zero- and first-order phase correction and baseline correction in both dimensions are applied after the Fourier transformation.

72 Remove the sample from the magnet. Add 15 µL of 0.5 M ascorbic acid solution to the sample. Mix well.
▲ **CRITICAL STEP** Check that the sample pH remains identical after the addition of the ascorbic acid solution. If required, restore the pH to the previous value.

73 Repeat Steps 70 and 71 to acquire the 2D ^{1}H-^{15}N TROSY-HSQC spectrum of the diamagnetic state. Use acquisition and processing parameters identical to those used for the paramagnetic state.

Intermolecular PRE analysis: data analysis

74 Load the assignment peak list on top of both 2D ^{1}H-^{15}N TROSY-HSQC spectra. Finely adjust the peak positions to exactly match the top of the NMR cross-peaks. Export the NMR peak intensities of the paramagnetic and diamagnetic states. Discard overlapping peaks from the analysis.

75 Divide the peak intensities of the paramagnetic state (I) by those of the diamagnetic state (I_0).

76 Plot I/I_0 versus residue number. For structure validation purposes, map the I/I_0 values on the protein structure.

OG-label for characterizing the oligomerization state: GB1 expression and purification

77 Dissolve 4 µg of GB1-pET15b plasmid in 20 µL of TE buffer.

78 Thaw one tube (50 µL) of BL21(DE3) competent cells on ice.

79 Add 1 µL of GB1-pET15b plasmid to 50 µL of competent cells, mix, and incubate the tube on ice for 30 min.

80 Repeat Steps 3 and 4.

81 Add 350 µL of SOC media to the tube and put it in an incubator at 37 °C with a 220-r.p.m. shaking speed for 1 h.

82 Add 60 µL of the culture from Step 81 to an LB agar plate (ampicillin, 100 µg/mL) and spread the medium.

83 Repeat Steps 7 and 8, using LB medium (ampicillin, 100 µm/mL) and incubating for ~16 h.
■ **PAUSE POINT** This step is typically carried out overnight.

84 Measure the OD_{600} value of the overnight culture, add the appropriate amount of overnight culture to 1 L of LB medium (ampicillin, 100 µm/mL) so that the starting OD_{600} value of the 1-L culture is ~0.1 (e.g., add 25 mL of culture having an OD_{600} value = 4 to 1 L of LB medium).

85 Incubate the 1-L culture in a 37 °C incubator with 220-r.p.m. shaking speed. Induce expression by adding 1 mL of 1 M IPTG to the culture when the OD_{600} value reaches 0.6.

86 After 4–6 h of growth at 37 °C, harvest the cells by centrifuging the culture at 4 °C at 5,000*g* for 30 min.
■ **PAUSE POINT** The cell pellet can be stored at −80 °C for up to 6 months.

87 Add one Pierce protease inhibitor tablet to 50 mL of StrepTrap binding buffer. Discard the supernatant from Step 86, and resuspend the cell pellet in the 50 mL of StrepTrap binding buffer; disrupt the cell suspension by sonicating the solution on ice for 5 min at intervals of 1.5 s at 60% of the maximum power. Repeat the sonication for another 5 min.

88 Centrifuge the cell suspension from Step 87 at 40,000*g* for 30 min at 4 °C, collect the supernatant and filter it via a 0.45-µm Puradisc syringe filter FP30.

89 Connect the StrepTrap HP column to the AKTA FPLC system and equilibrate the StrepTrap HP column by washing it with 5 column volumes of StrepTrap binding buffer at 5 mL/min.
▲ **CRITICAL STEP** Multiple StrepTrap columns can be used simultaneously to increase the yield of GB1.

90 Slowly inject (manually) the GB1 supernatant from Step 88 into the StrepTrap HP column.
▲ **CRITICAL STEP** Make sure that no bubbles are injected into the column.

91 Wash the StrepTrap HP column with 5 column volumes of StrepTrap binding buffer.
? TROUBLESHOOTING

92 Elute the GB1 protein, using 5 column volumes of StrepTrap elution buffer, and collect the elution.
■ **PAUSE POINT** The GB1 elution can be stored at 4 °C for no more than a week; we strongly recommend processing it as described in Steps 93 and 94 within 1–2 d.

93 Connect the Superdex 75 (16/60) column to the AKTA FPLC system and equilibrate the column with 150 mL of GB1 storage buffer.

94 Concentrate the GB1 elution from Step 92 to ~2.5 mL, using a 15-mL Amicon Ultra-15 centrifugal filter unit (MWCO = 3 kDa). Inject the concentrated GB1 elution into the AKTA injection loop and run the 'Protein refinement' program (see 'Equipment setup' section). Collect the GB1 and store it at -80 °C for up to 6 months.

OG-label for characterizing the oligomerization state: TriNTA-GB1 conjugation

95 Dissolve the TriNTA in the TriNTA reaction buffer to achieve a 1 mM TriNTA solution (e.g., dissolve 1.05 mg of TriNTA in 1 mL of buffer).
▲ **CRITICAL STEP** Use fresh TriNTA reaction buffer.

96 Add a four times excess of SM(PEG)2 with respect to the TriNTA solution (e.g., add 16 µL of 250 mM SM(PEG)2 to 1 mL of 1 mM TriNTA). Leave the reaction mixture at RT for 1 h or at 4 °C overnight.
? TROUBLESHOOTING

97 Connect the Zorbax C18 column to the HPLC system. Equilibrate the column with HPLC buffer D (wash for 5 column volumes).

98 Inject the reaction mixture from Step 96 into the HPLC sample loop, run the 'SMCC-TriNTA purification' program (see 'Equipment setup' section) and collect all fractions.

99 Save the fractions exhibiting strong UV_{280} and UV_{214} absorbances and lyophilize all the different fractions.

100 Dissolve 20% of the dried powder of each fraction in phosphate NMR buffer and acquire a 1D ^{1}H NMR spectrum for each fraction to identify fractions that contain the SMCC-TriNTA product. The expected SMCC-TriNTA is eluted when the gradient reaches 18% of HPLC buffer E.

101 Leave half of the SMCC-TriNTA product at RT for immediate use.
■ **PAUSE POINT** The rest of SMCC-TriNTA dry powder can be stored at −80 °C for up to 6 months.

102 Load the GB1 protein (Step 94) into a PD-10 desalting column. Remove the DTT using HEPES buffer following the instructions provided by the manufacturer. Elute the GB1 protein into a 50-mL Falcon tube using the amount of HEPES buffer indicated in the manufacturer's instructions.

103 Immediately after elution of the GB1 into a 50-mL Falcon tube (from Step 102), add a twofold molar excess of SMCC-TriNTA powder from Step 101 (SMCC-TriNTA/GB1 ~ 2:1). Degas the mixture for 30 min and incubate the reaction for an additional 30 min at RT.
▲ **CRITICAL STEP** The reaction mixture must be degassed to prevent GB1 from dimerizing via formation of disulfide bonds.
? TROUBLESHOOTING

104 Transfer the reaction solution to a dialysis cassette (7,000 MWCO) and dialyze it for 3 h versus 1 L of Ni-NTA binding buffer, stirring the solution. Perform two additional dialyses to achieve complete removal of the non-reacted SMCC-TriNTA in excess.

105 Measure the concentration of the TriNTA-GB1–GB1 mixture via a NanoDrop 2000 spectrophotometer, and then add a threefold molar excess of Ni^{2+} to the TriNTA-GB1 conjugate (Ni^{2+}/TriNTA-GB1 ~ 3:1).
■ **PAUSE POINT** The mixture can be stored at −80 °C for up to 3 weeks.

OG-label for characterizing the oligomerization state: conjugated TriNTA-GB1 purification

▲ **CRITICAL** Free GB1 (not conjugated with TriNTA) must be removed from the mixture before performing the OG-label experiment. A His-tag column is used to separate the conjugated from the free GB1. The column is obtained by covalent linking of a His-tag-containing Foldon protein to a NHS-activated agarose resin as described in the following section.

106 Express the Foldon protein, using the same procedures and experimental conditions as for the GB1 protein (Steps 77–86).

107 Resuspend the Foldon pellet in the Ni-NTA binding buffer and use the same sonication protocol as in Step 87 to lyse the cells.

108 Centrifuge the suspension at 40,000*g* for 30 min at 4 °C. Collect and filter the supernatant via a 0.45-µm Puradisc syringe filter FP30.

109 Add HisPur Ni-NTA resin to a chromatography column (1 mL of HisPur Ni-NTA per liter of expressed Foldon culture). Wash the Ni-NTA resin with 5 resin volumes of Ni-NTA binding buffer.

110 Add the supernatant from Step 108 to the HisPur Ni-NTA resin and incubate the mixture on a rotator for 30 min.

111 Wash the Ni-NTA resin with Ni-NTA binding buffer (ten times the resin volume), and then elute the Foldon protein with Ni-NTA elution buffer (five times the resin volume).

112 Use a 3-kDa MWCO Amicon centrifugal filter to concentrate the Foldon solution to 2.5 mL and purify the concentrated Foldon with a Superdex S75 26/60 column, using the 'Protein refinement' program (see 'Equipment setup' section) with Ni-NTA binding buffer.

113 Collect the Foldon fractions (identified by SDS–PAGE analysis) and measure the protein centration via a NanoDrop 2000 spectrophotometer.

■ PAUSE POINT The fractions can be stored at −4 °C for up to 1 month.

114 Take an aliquot of the Foldon solution containing 20 mg of protein. Add 600 mg of NHS-activated agarose to the solution and adjust the volume to ~8 mL with Ni-NTA binding buffer. Mix the reaction end-over-end for 1 h at 100 r.p.m.

115 Add 2 mL of 1 M Tris (pH 7.5) to quench the reaction. Incubate for 20 min at RT.

116 Wash this Foldon–resin mixture with 10 resin volumes of Ni-NTA binding buffer. Store the resin in Ni-NTA binding buffer at 4 °C for up to 6 months.

117 Add the TriNTA-GB1–GB1 mixture from Step 105 to the Foldon–resin mixture. Incubate the mixture on a rotator at RT for 30 min. Wash the resin with 10 resin volumes of Ni-NTA binding buffer, and then elute the TriNTA-GB1 using 5 resin volumes of Ni-NTA elution buffer. Collect all the wash and elution fractions.

? TROUBLESHOOTING

118 Confirm the purity of the TriNTA-GB1 by SDS–PAGE analysis.

119 Transfer the fraction containing the TriNTA-GB1 to a dialysis cassette (7,000 MWCO) and dialyze it for 3 h versus 1 L of OG-label buffer, always stirring the solution. Repeat the dialysis step three more times.

120 Use a 3-kDa MWCO Amicon centrifugal filter to concentrate the TriNTA-GB1 to ~125 μM.

■ PAUSE POINT The purified TriNTA-GB1 can be stored at −80 °C for up to 3 weeks.

OG-label for characterizing the oligomerization state: OG-label

121 Reconstitute the His_6-tag-containing TM protein (expressed and purified as described in Steps 1–47) in OG-label buffer as indicated in Steps 48–52.

122 Add a 100-fold molar excess of sulfo-NHS-acetate to the TM protein sample. Incubate the mixture for 1 h at RT.

▲ CRITICAL STEP Primary amines of the reconstituted TM protein must be blocked with sulfo-NHS-acetate before performing the OG-label experiment, to prevent unspecific crosslinking between the TM protein and GB1.

123 Add 5% (vol/vol) Tris buffer to quench the reaction. Incubate for 20 min at RT.

124 Transfer the reaction solution to a dialysis cassette (3,000 MWCO) and dialyze it for 3 h against 1 L of OG-label buffer, stirring the solution. Repeat the dialysis two more times to achieve complete removal of the sulfo-NHS-acetate in excess. During the dialysis, add 3 mg of DHPC per h directly to the dialysis cassette.

▲ CRITICAL STEP During the dialysis, the DHPC gradually diffuses outside the cassette while the DMPC remains trapped inside. Replenish the lost DHPC to maintain the bicelle q ~ 0.5–0.7.

125 Transfer the sample to a 4-mL concentrator (3,000 MWCO) and concentrate it to ~50 μM by centrifuging at 6,000*g* for ~1 hour at RT.

126 Mix the TM sample with TriNTA-GB1 (from Step 120) in a molar ratio of 1:1.5. Incubate for 10 min at RT. After mixing, the concentration of the TM protein (monomeric) is ~30 μM. Split the mixture among four tubes. Add BS3 crosslinker to each tube so that its concentration is 1 mM, and then incubate for 30 min at RT. Add glutaraldehyde to the tubes so that its final concentration is 0.1, 0.3, 1, and 3 mM (one condition per tube). Incubate for 5 min at RT.

? TROUBLESHOOTING

127 Prepare a negative control by adding OG-label buffer (use the same volume as for the TM protein in Step 126) to the TriNTA-GB1. Add BS3 crosslinker so that its concentration is 1 mM, then incubate for 30 min at RT. Add glutaraldehyde so that its final concentration is 1 mM and incubate for 5 min at RT.

128 Add 5% (vol/vol) Tris buffer to each tube to quench the reaction. Incubate for 20 min at RT.

129 Add EDTA to each tube to a final concentration of 25 mM to release the TriNTA-GB1 from the TM protein.

130 Quantify the oligomeric state of TriNTA-GB1 via SDS–PAGE analysis[23].

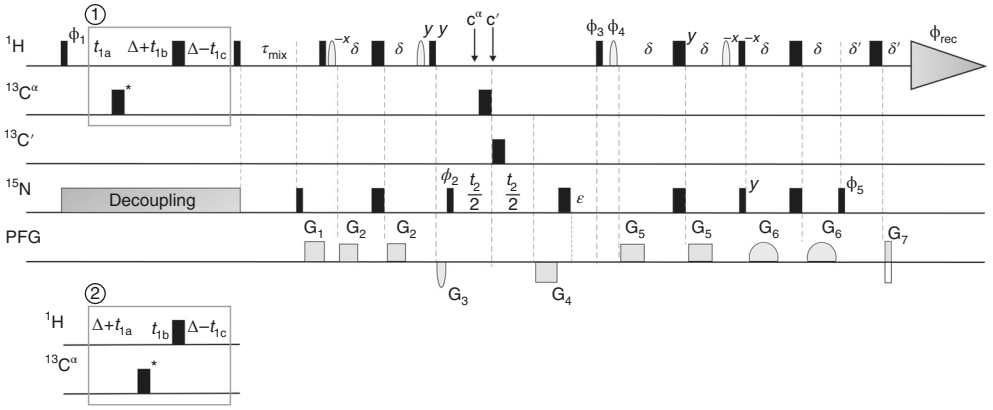


Fig. 4 | Pulse sequence of the 3D J_{CH}-modulated NOE experiment. Narrow and wide black rectangles stand for 90° and 180° pulses, respectively. The pulses are applied along the *x* axis unless noted differently. Empty rectangles represent non-selective pulses; round shapes represent band-selective pulses. The line labeled 'PFG' represents a pulse field gradient applied along the *z* axis. The two different variants necessary to achieve $J(^1H\text{-}^{13}C)$ modulation are shown in the boxes labeled (1) and (2); the two pulse schemes are recorded in an interleaved manner. The arrows indicate switching of the ^{13}C carrier frequency. The ^{13}C pulse marked by an asterisk (23.5 µs @ 800 MHz) is applied at 60 p.p.m. and inverts all aliphatic and aromatic carbons; the other ^{13}C pulses (35.6 µs @ 800 MHz) are applied either on C^α or C′ and selectively invert carbons on the carrier. ^{15}N decoupling is achieved using the WALTZ-16 (1.5 kHz) sequence. Selective ^1H pulses for water suppression have a Gaussian shape and a 1-ms duration. Empty rectangles are used for G_1, G_2, G_4, G_5, and G_7 gradients, whereas a sine shape is used for G_3 and G_6 gradients. The following phase cycle is used: $\Phi_1 = x, -x$; $\Phi_2 = 2(y), 2(-y)$; $\Phi_3 = -y$; $\Phi_4 = y$; $\Phi_5 = -x$; $\Phi_{rec} = x, 2(-x), x$. The delays are set as follows: $\Delta = 4.0$ ms; $\delta = 2.4$ ms; $\varepsilon = t_2(0) + p_4$ (where p_4 is the duration of G_4); $\delta' = 0.5$ ms; τ_{mix} denotes the mixing time. The gradients have the following durations and relative strengths: G_1 (2,000 µs) = 48%; G_2 (1,200 µs) = 35%; G_3 (1,000 µs) = −45%; G_4 (2,500 µs) = −50%; G_5 (1,200 µs) = 33%; G_6 (2,200 µs) = 35%; G_7 (251 µs) = 50% (with 100% corresponding to 0.56 T/m). The semi-constant time mode for the evolution of chemical shift is used in the ^1H evolution period, whereas real-time evolution is used in the ^{15}N evolution period. For the semi-constant time mode, the delays are set as follows: $t_{1a} = t_1/2$; $t_{1b} = t_{1a} - t_{1c}$; $t_{1c} = \Delta/N_1$ (where N_1 is the number of total increments in the t_1 dimension). Quadrature detection in the ^1H indirect dimension is obtained by incrementing phase Φ_1 by 90° in a States manner; for quadrature detection in the ^{15}N indirect dimension, echo-anti-echo data are recorded by inverting the sign of G_7 and increasing phases Φ_3, Φ_4, and Φ_5 by 180°.

Part 3: J_{CH}-modulated NOE experiment for detecting unambiguous interprotomer NOEs

Sample preparation

131 Reconstitute the ~1:1 (wt/wt) mixture of $^2H,^{15}N$-labeled protein (perdeuterated or partially deuterated) and ^{13}C-labeled protein (expressed and purified as described in Steps 1–47) as indicated in Steps 48–53.

NMR experiments

132 Repeat Steps 67–69.

133 Acquire a high-resolution 2D $^1H\text{-}^{15}N$ TROSY-HSQC spectrum as a 2D reference for the subsequent NOESY analysis and process it with standard parameters. Typically, raw data are multiplied by an apodization function, zero-filled, and Fourier-transformed in both dimensions. Proper zero- and first-order phase correction and baseline correction in both dimensions are applied after the Fourier transformation.

134 Acquire a high-resolution 2D $^1H\text{-}^{13}C$ HSQC spectrum as a 2D reference for the NOESY analysis and process it with standard parameters. Typically, raw data are multiplied by an apodization function, zero-filled, and Fourier-transformed in both dimensions. Proper zero- and first-order phase correction and baseline correction in both dimensions are applied after the Fourier transformation.

135 Create a new experiment for acquiring the 3D J_{CH}-modulated NOE spectrum. Set the experimental parameters as described in Fig. 4. On the basis of the previously acquired 2D $^1H\text{-}^{15}N$ TROSY-HSQC spectrum (Step 133), set the proper carrier positions and spectral windows. Set the recovery delay to at least 1.2 s and the NOE mixing time (τ_{mix}) to 200 ms.

136 Acquire the $^{15}N\text{-}^1H$ and $^1H\text{-}^1H$ planes to verify that all the parameters have been set correctly and that the experiment is working as expected.

! CAUTION The two interleaved spectra are stored along the ^{15}N dimension; therefore, to correctly process the ^{15}H-^{1}H plane, it is necessary to first split the free induction decay (FID) into the two sub-spectra.

137 Set the number of ^{1}H and ^{15}N increments to a resolution of ~8–10 and ~30–35 ms, respectively.
! CAUTION The actual number of points along the ^{15}N indirect dimension is half that which is set, because the two interleaved spectra are stored in this dimension.

138 Acquire the 3D J_{CH}-modulated NOE spectrum.

Spectra processing

139 Split the FID along the ^{15}N dimension to obtain the two sub-datasets, which follow pathways (1) and (2) in Fig. 4. Process the two spectra with standard parameters. Typically, raw data are multiplied by an apodization function, zero-filled, and Fourier-transformed in all dimensions. Proper zero- and first-order phase correction and baseline correction in both dimensions are applied after the Fourier transformation.
? TROUBLESHOOTING

140 Combine the two sub-datasets by summing and subtracting their FIDs to obtain the spectra selective for the intra- ((1) + (2)) and inter-NOEs ((1) − (2)), respectively. Process the two spectra with parameters identical to those used in Step 139.

Part 4: Oligomerization solution and structure determination

Generation of the monomer structure

141 Run TALOS+ to predict the protein phi and psi backbone torsion angles from the available assigned chemical shifts (H^N, H^α, C^α, C^β, C′, N), which can be obtained by recording NMR triple-resonance experiments for backbone assignment (e.g., HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB).
? TROUBLESHOOTING

142 Use XPLOR-NIH to build the protein monomer structure, using the dihedral restraints from Step 141.
▲ CRITICAL STEP Include only the dihedral restraints classified as 'Good' by TALOS+.

Exhaustive search to determine the mode of oligomeric assembly

143 To use the ExSSO program to perform an exhaustive search of the symmetric conformational space guided by interprotomer NOEs to identify the mode of assembly of the symmetric oligomer (Steps 143–146), first set up the parameters of the configure.py script by providing (i) the protein name (`$proName`), (ii) the file path of the protomer structure (from Step 142), (iii) the file path of the inter-NOE restraints table (determined using the 3D J_{CH}-modulated NOE experiment, Steps 131–140), and (iv) the number of protomers in the oligomer (as determined by OG-label, Steps 77–130). Instructions on the format to be used for (ii) and (iii) can be seen by typing `$ python check_format.py -p $homeDir/$proName/param -t pdb` and `$ python check_format.py -p $homeDir/$proName/param -t noe`, respectively.

144 Run the script by typing: `$ python configure.py`. The file param, containing all the user defined parameters, is created in $homeDir/$proName/.

145 Perform the exhaustive search of the symmetric conformational space by running the following script: `$ python exhaustive_search.py -p $homeDir/$proName/param`. Representative conformations of the solution(s) found are stored in $homeDir/$proName/$proName.cs.
? TROUBLESHOOTING

146 Generate a representative structure for each solution found in Step 145 by running the following script: `$ python get_structure.py -p $homeDir/$proName/param`. The calculated structures ($proName_*.pdb) are stored in $homeDir/$proName/ and sorted from best to worst according to the ExSSO scoring system.

Final structural refinement

147 Use the best scorer from Step 146 as starting model for structural refinement with XPLOR-NIH.

148 Identify more self-consistent interchain NOEs from the regular ^{15}N-edited and ^{13}C-edited NOESY spectra by checking whether tentative assignments are compatible with the starting model. Run XPLOR-NIH, including also the newly assigned NOE restraints, to improve the structural ensemble.

149 Update the starting model with the average structure of the new structural ensemble calculated in Step 148.

150 Repeat Steps 148 and 149 iteratively until no new NOEs are found or until the desired RMSD of the structural ensemble (15 lowest-energy structures out of 150 calculated) is reached.

Part 5: Characterization of the protein TM partition

Sample preparation

151 Express, purify, and reconstitute the ^{15}N-labeled protein sample as indicated in Steps 1–53.

NMR experiments

152 Repeat Steps 67–69.

153 Acquire a high-resolution 2D ^{1}H-^{15}N TROSY-HSQC spectrum that will serve as reference for the titration.

▲ **CRITICAL STEP** The recovery delay should be set to at least 3.5 s to prevent signal loss due to incomplete Boltzmann recovery.

154 Process the spectrum with standard parameters. Typically, raw data are multiplied by an apodization function, zero-filled, and Fourier-transformed in both dimensions. Proper zero- and first-order phase correction and baseline correction in both dimensions are applied after the Fourier transformation.

155 Remove the sample from the magnet. Add the desired amount of paramagnetic probe (Gd-DOTA or 16-DSA) to the sample (see Step 157 for quantities) and mix well.

156 Repeat Steps 153 and 154 to acquire a 2D ^{1}H-^{15}N TROSY-HSQC spectrum for the actual paramagnetic probe concentration. Use acquisition and processing parameters identical to those used for the reference spectrum.

▲ **CRITICAL STEP** When performing the titration with Gd-DOTA, a high concentration of the latter (~10–20 mM) broadens the water signal and hinders the shimming of the magnet. Make sure a good shimming is achieved before recording the spectra.

? TROUBLESHOOTING

157 Repeat Steps 155 and 156 to acquire all the points of the titration. After the reference spectrum, the following points are typically measured:

- *[Gd-DOTA]*. 2, 4, 6, 8, 10, 15, and 20 mM (e.g., for a 360-μL sample, add 3.6 μL of 200 mM Gd-DOTA solution for the first five points, and then add 3.0 μL of 600 mM Gd-DOTA solution for the last two. Dilution can be neglected in the calculation because the volumes added are much less than that of the sample).
- *[16-DSA]*. 0.6, 1.2, 1.8, 2.4, 3.0, 3.6, and 4.2 mM (e.g., for a 360-μL sample, add for each point 10.8 μL of 20 mM 16-DSA solution. Dilution can be neglected in the calculation).

? TROUBLESHOOTING

Data analysis

158 Load the assignment peak list on top of the 2D ^{1}H-^{15}N TROSY-HSQC spectra. Finely adjust the peak positions to exactly match the top of the NMR cross-peaks. Export the NMR peak intensities of all the spectra. Discard overlapping peaks from the analysis.

159 Normalize the peak intensities of all the spectra with respect to those of the reference spectrum (I/I_0).

160 For each residue, generate a (I/I_0) versus [paramagnetic probe] plot and fit it to the following equation with two unknown parameters (PRE_{amp} and τ):

$$\frac{I}{I_0} = 1 - PRE_{amp}\left(1 - e^{-\frac{[\text{paramagnetic probe}]}{\tau}}\right), \quad (1)$$

where I and I_0 are the peak intensities in the presence and absence of the paramagnetic agent, respectively, [paramagnetic probe] is the concentration of the paramagnetic agent (Gd-DOTA or 16-DSA), τ is the decay constant, and PRE_{amp} is the PRE amplitude.

161 Plot PRE_{amp} versus (residue number) to obtain information on the topology of the protein (e.g., how many times the protein crosses the bilayer center) or for obtaining structural information for structure validation/refinement (the slopes of the plot reflect the relative orientation of the different protein segments with respect to the bicelle normal axis).

162 Determine the protein symmetry axis. Rotate the protein structure so that its symmetry axis is parallel to the bicelle bilayer normal (r_Z axis). Arbitrarily place the protein along the axis, assigning $r_Z = 0$ to the position of a chosen protein amide proton. Calculate the positions of all the other amide protons along the axis (r_Z, Å) with respect to the assigned $r_Z = 0$.

163 Apply an offset to the previously assigned $r_Z = 0$ (e.g., ±1, ±2, ±3 Å, and so on) and recalculate the positions of all the amide protons along the r_Z axis on the basis of the new $r_Z = 0$. Repeat the procedure for several offsets, placing them at a distance of 2–3 Å to probe the entire protein (i.e., $r_Z = 0$ is assigned at regular intervals from the N terminus to the C terminus of the protein). For each offset, calculate the positions of all the amide protons (r_Z dataset).

164 For each r_Z dataset, calculate the absolute values of the measured r_Z ($|r_Z|$) and plot PRE_{amp} versus $|r_Z|$.

165 Fit each PRE_{amp} versus $|r_Z|$ curve to the following sigmoidal function with four unknown parameters (PRE_{amp}^{min}, PRE_{amp}^{max}, r_Z^I, and SLOPE):

$$PRE_{amp} = PRE_{amp}^{min} + \frac{\left(PRE_{amp}^{max} - PRE_{amp}^{min}\right)}{1 + e^{\left(r_Z^I - |r_Z|\right)/SLOPE}}, \quad (2)$$

where PRE_{amp}^{min} and PRE_{amp}^{max} are the limits within which PRE_{amp} can vary for a given protein system, r_Z^I is the inflection point (the distance from the bilayer center at which PRE_{amp} is halfway between PRE_{amp}^{min} and PRE_{amp}^{max}), and SLOPE is a parameter that reports the steepness of the curve at the inflection point.

166 Compare the adjusted coefficient of determination (R^2_{adj}) of each fitting and identify which r_Z dataset (or which offset applied to the initially assigned $r_Z = 0$ (Step 162)) yields the best fitting (highest R^2_{adj}). The corresponding protein position will be very close to the actual protein partition in the bicelle.

167 Starting from the best r_Z dataset identified in Step 166, repeat Step 163 but apply smaller offsets to achieve better accuracy (e.g., ±0.1, ±0.2, ±0.3 Å, and so on).

168 Repeat Steps 164–166. The best fitting of PRE_{amp} versus $|r_Z|$ to Eq. 2 (highest R^2_{adj}) provides the accurate membrane partition of the protein (r_Z dataset), with the bilayer center corresponding to the assigned $r_Z = 0$.

Troubleshooting

Troubleshooting advice can be found in Table 1.

Table 1 | Troubleshooting table

Step	Problem	Possible reason	Possible solution
21	Weak or absent protein expression	The amount of IPTG used for induction and/or the expression temperature may not be optimal, or other issues (such as codon usage and leaky expression) could occur	Induce cells at an OD_{600} value within the range 0.6–0.8; screen for the best protein expression conditions, varying the temperature in the range of 16–35 °C and the [IPTG] from 0.1 to 0.5 mM. Codon-optimize the gene for *E. coli* or change cell lines, if necessary
35	CNBr cleavage is not efficient	Too much oxygen is present in the reaction system	Before the reaction starts, pretreat the solution by blowing a gentle nitrogen gas stream for 5 min; maintain the gas stream during the entire reaction
43	The column is not clean, even after several washes	90% FA may be not strong enough to completely clean the column	Run the 'cleaning' HPLC method, loading 1 mL of TFE, instead of 5 mL of 90% FA. Repeat the column washing two to three times, if needed
45	Elution peaks are not resolved	The HPLC method used to purify the protein was not optimized	Change the composition of HPLC buffers A and B, varying the amount of IPA and ACN; change the HPLC method by using slower gradients
46	Excessive protein formylation occurred	The protein was dissolved in 90% FA for too long	Reduce the CNBr-cleavage reaction time (Step 35) or the FA concentration (Steps 31 and 32)
51	Excessive protein aggregation or precipitation	Denaturant removed too quickly	Check the bicelle q[19] (Step 54) to ensure it does not become too large due to loss of DHPC. Use a small-MWCO dialysis cassette (3,500 MWCO) and slower stirring speed (<200 r.p.m.)
54	The bicelle q is >0.5	An insufficient amount of DHPC is present	Calculate the amount of DHPC present in the sample on the basis of the actual q and the amount of lipids used in Step 48. Calculate the amount of DHPC that would be needed to obtain $q = 0.5$ for

Table continued

Table 1 (continued)

Step	Problem	Possible reason	Possible solution
			the lipids used. Add DHPC (100 mg/mL) to cover the difference. Measure a 1D ^{1}H NMR spectrum (Step 54) to ensure the q is now 0.5
	The bicelle q is <0.5	Too much DHPC is present	Calculate the amount of DHPC present in the sample on the basis of the actual q and the amount of lipids used in Step 48. Calculate the amount of DHPC that would be needed to obtain $q = 0.5$ for the lipids used. Dilute the sample with the amount of NMR buffer (pH 6.7) required to retain the excess of detergent, knowing that in solution the free [DHPC] is ~5 mM. Repeat Steps 53 and 54
63	MTSL labeling is not efficient	The sample pH is not sufficiently high	Increase the sample pH to ~7.5 before adding MTSL
		The cysteine residue of the Cys mutant is partially or completely oxidized	Increase the amount of DTT in Steps 59 and 60 or use a stronger reducing agent (e.g., tris-(2-carboxyethyl)phosphine (TCEP)); lower the sample pH to ~6.0 while removing the DTT by dialysis (Step 60)
		The amount of MTSL used was not sufficient	Increase the excess of MTSL to ~20–30 times
65	Free MTSL is not removed completely	The DHPC removal rate is too slow	Increase the stirring speed to aid DHPC removal; increase the number of dialyses
91	Only a small fraction of GB1 binds to the StrepTrap HP column	The sample injection is too fast	Inject the sample into the column slowly (<1 mL/min); connect a few columns together; clean the columns after each use
96	The efficiency of the reaction between TriNTA and SMCC is low	Incorrect buffer conditions	Use fresh sodium carbonate buffer (pH 8.0). The pH of the buffer can increase to 10 after ~1 week
103	The efficiency of the reaction between SMCC-TriNTA and GB1 is low	The GB1 is not completely reduced; the GB1 is oxidized by air during the reaction; the amount of SMCC-TriNTA is not sufficient	Make sure that the GB1 is completely reduced before starting the reaction; degas the reaction buffer; use a higher excess of SMCC-TriNTA
117	The binding between the TriNTA-GB1 and Foldon resin is weak	The resin is not saturated by Foldon; the incubation time is too short; residual imidazole is present in the buffer	Use at least 20 mg of Foldon per 150 mg of dry resin; incubate the TriNTA-GB1 and the Foldon resin for at least 30 min; make sure that no imidazole is present on the resin or in the TriNTA-GB1 buffer; repeat the purification procedure a few times to achieve high TriNTA-GB1 purity
126	The efficiency of the crosslinking reaction is low	The His_6-tags of the TM protein are too closely spaced; the protein concentration is low; the crosslinker is too short	When multiple His_6-tags are too closely spaced, the binding stoichiometry between the His_6-tag and TriNTA may no longer be 1:1. Introduce the His_6-tag on the basis of the protein structure so that the His_6-tags are fairly distant; add charged residues in the proximity of the His_6-tag to introduce electrostatic repulsion near the latter; use concentrated protein; use longer crosslinkers
139	No inter-NOEs are observed	The ^{2}H-,^{15}N- and ^{13}C-labeled proteins have not been mixed adequately	Mix the two preparations before performing CNBr cleavage
	Positive inter-NOEs are present in sub-spectrum (1), but negative inter-NOEs are weak or not present in sub-spectrum (2)	NMR relaxation is more pronounced in pathway (2) of the experiment	Perform the analysis as described. Inter-NOEs in the inter-NOE selective spectra ((1) − (2)) may appear weaker, but intra-NOEs will still be completely canceled
141	TALOS+ predictions are not classified as 'Good'	Chemical shift may be not calibrated; chemical shift assignment may be too incomplete	Calibrate the chemical shifts against 4,4-dimethyl-4-silapentane-1-sulfonic acid; acquire additional NMR spectra to improve the chemical shift assignments
145	More than one solution is found	The number of inter-NOEs may be insufficient	Increase the number of assigned interchain NOEs
156	Gradient shim fails at high concentrations of Gd-DOTA	Gd-DOTA generates a strong PRE effect on the D_2O signal	Adjust the field to center the deuterium signal; lock and load the shim map used for the previous point of the titration. Perform the gradient shim. If it still fails, shim manually on z, x, y, xz, yz, and z.
157	The 20 mM 16-DSA solution turns unclear after a few hours	The [16-DSA] is close to its maximum solubility in bicelles	Sonicate the solution for 5 min right before use; prepare a new 16-DSA solution at lower concentration (~15 mM)

Timing

Part 1

Steps 1–21, protein expression: 4 d
Steps 22–32, Ni-NTA purification: 1 d
Steps 33–39, CNBr cleavage: 1 d
Steps 40–47, reverse-phase HPLC: 1 d
Steps 48–54, protein reconstitution in DMPC/DHPC bicelles: 2 d

Part 2

- Steps 55–66, sample preparation: 3 d
- Steps 67–73, NMR experiments: 3–6 d, depending on the sample concentration and spectrometer sensitivity
- Steps 74–76, data analysis: 1 d
- Steps 77–94, GB1 expression and purification: 4 d
- Steps 95–105, TriNTA-GB1 conjugation: 3 d
- Steps 106–120, conjugated TriNTA purification: 2 weeks
- Steps 121–130, OG-label: 3 d

Part 3

- Step 131, sample preparation: 2 d
- Steps 132–138, NMR experiments: 10–12 d, depending on the sample concentration and spectrometer sensitivity
- Steps 139 and 140, spectra processing: 1 h

Part 4

- Steps 141 and 142, generation of the monomer structure: 1 d
- Steps 143–146, exhaustive search to determine the mode of oligomeric assembly: 1 d
- Steps 147–150, final structural refinement: days to weeks, depending on the quality of the NMR spectra and the user's experience

Part 5

- Step 151, sample preparation: 2 d
- Steps 152–157, NMR experiments: 3–6 d, depending on the sample concentration and spectrometer sensitivity
- Steps 158–168, data analysis: 1 d

Measurement

- Box 1, quantification of the sample deuteration level: 1 day

Anticipated results

Part 1: Expression, purification, and reconstitution of TMDs

The protocol describes the steps needed to prepare bicelle-reconstituted TMD samples suitable for structural studies by NMR; it is demonstrated below for the TMD of Fas (Fas TMD), a member of the TNF receptor superfamily. Although only shown here for the Fas TMD as an example, the applicability of the protocol is very broad, as has been demonstrated in numerous studies, including ones that involve MP regions[1,8,12,24,53–57].

The Fas TMD was expressed in *E. coli* as a fusion to the C terminus of the TrpLE sequence with a methionine residue added in between (Figs. 1 and 2)[55]. The expression level of the fusion protein was ~2 mg/L, which is comparable to that of most other hydrophobic protein fragments expressed using the same system (Fig. 3b)[8,56]. The methionine residue was used as a cleavage site for CNBr, separating the TMD from the TrpLE (Figs. 1 and 2b). Importantly, the fusion protein is cleaved right after the methionine such that the TMD consists of only its native amino acid sequence. After cleavage, the Fas TMD was separated by HPLC, using the Zorbax SB-C3 column (HPLC, Fig. 3c). The HPLC profile and SDS–PAGE analysis both indicated high protein purity (Fig. 3c). The expression protocol above can be used to achieve an extremely high level of protein deuteration. Measurements from analytical liquid chromatography–mass spectrometry (LC–MS) of our ^{2}H-labeled Fas TMD sample show that

the protein deuteration percentage was ~98.3% (Box 1). This high level of deuteration is ideal for the interchain NOE experiment (described in part 3) because the intra-chain NOEs from aliphatic groups would be invisible, allowing exclusive detection of interchain contacts. Before reconstitution in bicelles, the SDS–PAGE analysis of the Fas TMD powder resolved the protein as a monomer (~4 kDa). However, after reconstitution in DMPC/DHPC bicelles with q ~ 0.55, which yields lipid discs with diameter of ~50 Å (Fig. 3d), the Fas TMD migrated on SDS–PAGE as a trimer (~13 kDa), indicating that the Fas TMD spontaneously trimerizes in lipid bilayer and the trimers apparently resisted the denaturing power of SDS–PAGE (Fig. 3d). After reconstitution into bicelles, the sample quality was further evaluated using a 2D ^{1}H-^{15}N TROSY-HSQC NMR spectrum. The NMR spectrum has good chemical shift dispersion and one peak per residue (Fig. 3e), indicating that the protein is well folded and adopts a single trimeric conformation. We emphasize that some TMD oligomers do not survive the denaturing environment of SDS–PAGE. It is therefore critical that TMD oligomerization be examined using the methods in part 2. We note that formation of homogeneous oligomers in bicelles does not guarantee that the oligomeric form represents the native state on the cell surface. Hence, it is important to perform functional mutagenesis assays[5,24] to test the functional relevance of the TM oligomers.

Part 2: Determination of the oligomeric state of TMDs in bicelles

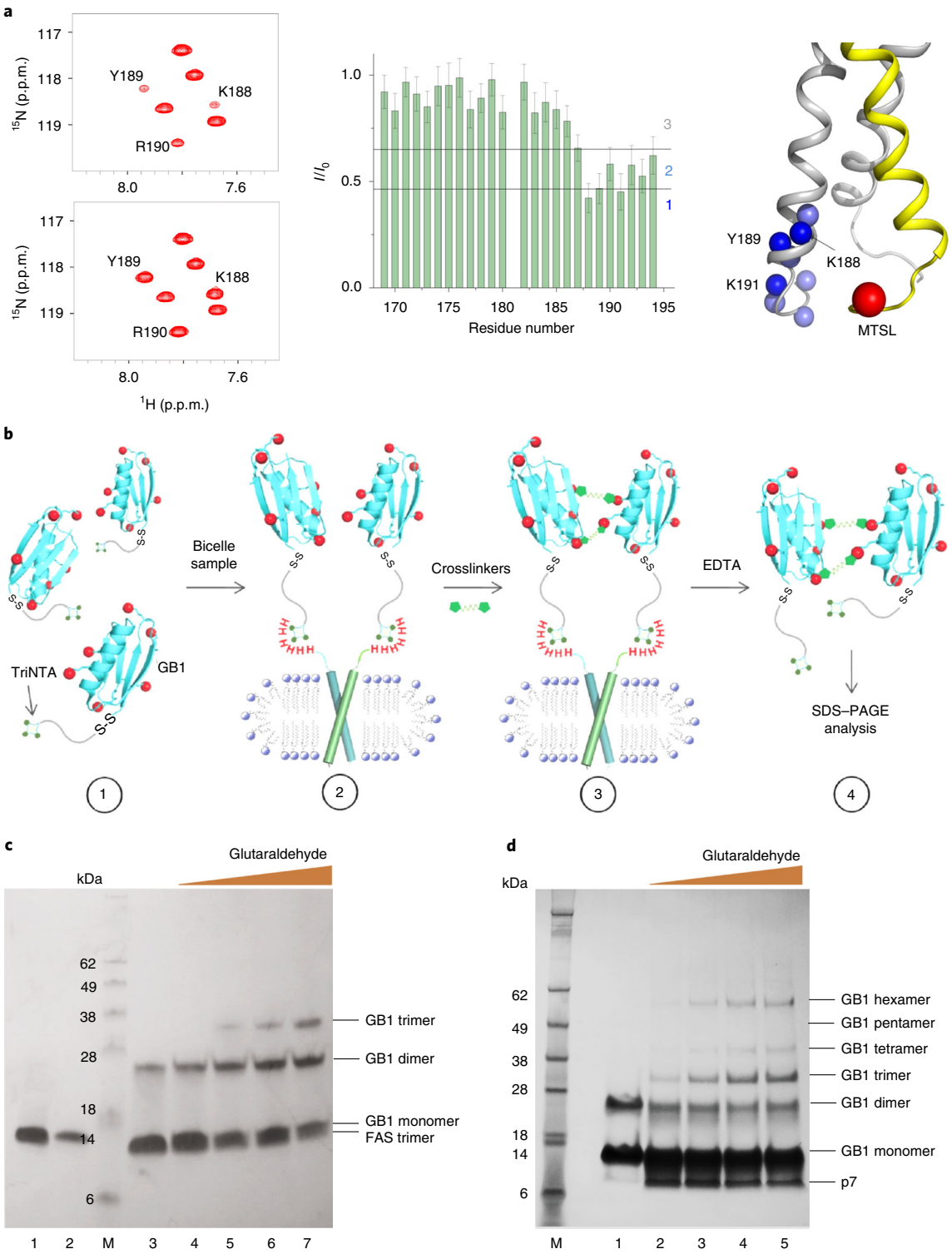
Intermolecular PRE analysis

The protocol describes an efficient strategy for site-specific paramagnetic labeling of TMD to detect interchain PREs if the TMD forms an oligomer. For this experiment, two protein preparations are needed: one for ^{15}N-labeled protein, which serves as the NMR readout, and the other for spin-labeled protein that would influence the NMR signals of the ^{15}N-labeled protein if the two mix and oligomerize. For the spin-labeled protein, a cysteine mutation is introduced (usually in regions predicted to be accessible, e.g., one of the protein termini) for labeling with MTSL. The two preparations are then mixed at a ratio of 1:1 for optimal measurement of interchain PREs. The main purpose of the experiment is to address whether the TMD is monomeric or multimeric in bicelles, although it can also be used later to validate NOE-derived oligomeric structures.

The results expected from this type of analysis are exemplified here in the case of the Fas TMD. The ^{15}N-labeled Fas TMD and the unlabeled mutant (with a cysteine introduced at the C terminus) were mixed at a 1:1 ratio right after HPLC purification. The mixed protein was reconstituted into DMPC/DHPC bicelles (q = 0.5), followed by linking the cysteine and MTSL via disulfide bonding. After recording the two NMR spectra, one before (paramagnetic) and one after (diamagnetic)

Fig. 5 | Characterization of the oligomeric state of TMDs in bicelles. a, Interchain PRE analysis demonstrated here for the Fas TMD reconstituted in bicelles with q = 0.55. The sample consisted of an ~1:1 mixture of ^{15}N-labeled Fas TMD and ^{14}N Fas TMD spin-labeled with MTSL ((1-oxyl-2,2,5,5-tetramethyl pyrroline-3-methyl) methanethiosulfonate) at the C terminus (Cys194). Left: 2D ^{1}H-^{15}N TROSY-HSQC spectra before (top) and after (bottom) reducing the nitroxide with ascorbic acid, acquired at 600 MHz and 303 K. Middle: residue-specific PRE, defined as the ratio of the peak intensity before (I) and after (I_0) ascorbic acid addition. Right: mapping of the PREs onto the Fas TMD trimer structure, showing the backbone amide protons (blue spheres) with strong PREs and the spin label at the Cβ position (red sphere). Amide protons are colored on the basis of the PRE regions in the middle panel: 1, dark blue; 2, light blue. ^{15}N-labeled and MTSL-labeled chains are shown as gray and yellow ribbons, respectively. **b**, Schematic illustration of the OG-label procedure for characterizing the protein oligomerization state. (1) TriNTA-GB1 is sufficiently dilute (<100 μM) for minimizing nonspecific crosslinking. (2) In the presence of bicelle-reconstituted TM oligomers (containing one His_6-tag per protomer), the TriNTA-GB1 is recruited to the TM oligomer in stoichiometric amounts via strong affinity between the His_6-tag and the TriNTA. Before incubation with the TriNTA-GB1, the primary amines of the TM protein are blocked. (3) Addition of crosslinkers (first BS3 and then glutaraldehyde) to crosslink the GB1s. (4) TriNTA-GB1 is released from the TM protein by stripping Ni^{2+} with EDTA, followed by SDS–PAGE analysis. **c**, OG-label application to the Fas TMD reconstituted in DMPC/DHPC bicelles with q = 0.5. As a control, 45 μM TriNTA-GB1 was treated with 1 mM BS3 and then with 1 mM glutaraldehyde for 5 min (lane 3). In the presence of 30 μM bicelle-reconstituted Fas TMD, 45 μM TriNTA-GB1 was treated first with 1 mM BS3 and then with increasing amounts of glutaraldehyde for 5 min: 0.1 mM (lane 4), 0.3 mM (lane 5), 1 mM (lane 6), and 3 mM (lane 7). The bands of TriNTA-GB1 and bicelle-reconstituted Fas TMD (trimer) alone are shown in lane 1 and 2, respectively. The SDS–PAGE gel used was a 12% bis-acrylamide (Bis)–Tris protein gel. M, marker. **d**, OG-label application to the HCV p7 reconstituted in DMPC/DHPC bicelles with q = 0.6. As a control, 30 μM TriNTA-GB1 was treated first with 0.6 mM 3,3′-dithiobis(sulfosuccinimidyl propionate) (DTSSP) and then with 0.5 mM glutaraldehyde for 5 min (lane 1). In the presence of 20 μM bicelle-reconstituted p7, 30 μM TriNTA-GB1 was treated first with 0.6 mM DTSSP and then with increasing amounts of glutaraldehyde for 5 min: 0.1 mM (lane 2), 0.5 mM (lane 3), 1 mM (lane 4), and 2.5 mM (lane 5). The SDS–PAGE gel used was a 4–12% Bis–Tris protein gel. **b** adapted with permission from ref. [23], Elsevier.

ascorbic acid addition, the peak intensity ratios of the paramagnetic (I) to the diamagnetic (I_0) state (defined here as the PRE) were measured for all residues. The PRE versus residue plot (Fig. 5a) shows strong PREs for the C-terminal region of the ^{15}N-labeled Fas TMD but not for the N-terminal region, indicating that the PREs are due to specific, parallel oligomerization of the TMD. Moreover, average signal loss of the affected region is ~50%, which is expected from a 1:1 mixed sample. The interchain PRE results indicate that the Fas TMD is oligomeric.



Several potential pitfalls deserve special attention. The mixed PRE analysis is meaningful only when two differently labeled chains are adequately mixed within oligomers, but this can be problematic for certain TMDs that oligomerize very strongly. For example, the TMD of the HIV-1 gp41 fusion protein trimerizes so strongly that, once the trimer is formed, the monomers can no longer be mixed between the trimers[12,24]. In that case, the two differently labeled species were mixed before CNBr cleavage, as the TrpLE TMD fusion protein is monomeric in FA. Another potential pitfall is the incomplete removal of MTSL labels. MTSL is hydrophobic and partitions weakly in micelles or bicelles. We found that the sample must be dialyzed extensively (e.g., four times) to completely remove the free MTSL. Otherwise, even tiny residual amounts of MTSL could generate substantial but nonspecific PREs. A simple way to test the robustness of the MTSL removal protocol is to test it on a control sample containing only the ^{15}N-labeled protein (without cysteine) and the same amount of MTSL used for labeling. If free MTSLs are completely removed, no PRE should be observed. Finally, caution must be taken when reducing the spin label with ascorbic acid. As ascorbic acid is acidic, it is important to ensure that the ascorbic acid solution is buffered to the same pH as the that of protein sample, as reducing the sample pH could change NMR peak intensity by changing solvent exchange. A final pH check after ascorbic acid addition should be performed before recording the TROSY-HSQC spectrum.

OG-label for characterizing the oligomerization state

The mixed PRE experiment can address only whether the TMD is an oligomer; it cannot determine the oligomerization number, which is an essential parameter for NMR structure determination. The OG-label method (Fig. 5b) was developed as a standard tool for addressing TMD oligomeric state in bicelles. It can be considered a standard because the use of an SCP as the crosslinking readout of the TMD oligomer makes the method much less sensitive to variations in the ability of the target TMDs to be crosslinked. This method is demonstrated here for the trimeric Fas TMD.

The GB1 protein was the SCP used; it was linked to a TriNTA molecule via a PEG-2-SMCC; the efficiency of conjugating GB1 and TriNTA was ~90%. The TriNTA-GB1 was purified using the His_6-tag column, and the purified TriNTA-GB1 migrated as a 10-kDa band on SDS–PAGE. For the OG-label experiment, the Fas TMD with a C-terminal His_6-tag was expressed, purified, and reconstituted as described in part 1. To prevent unwanted crosslinking between Fas TMD and GB1, the active amine groups of Fas TMD were first blocked by the addition of a 100-fold molar excess of sulfo-NHS-acetate to the bicelle sample. Excessive sulfo-NHS-acetate was removed by dialysis while controlling the bicelle q. After dialysis, 30 μM Fas TMD (monomer concentration) was mixed with 45 μM TriNTA-GB1 to ensure saturation of the protein His_6-tags by TriNTA-GB1. The mixture was then treated with 1 mM BS3 crosslinkers for 30 min, followed by incubation with various concentrations of glutaraldehyde (0.1, 0.3, 1.0, and 3 mM) for 5 min. The crosslinking reaction was quenched with a 20 mM Tris buffer (pH 7.5). As a negative control, 1 mM of BS3 and 1 mM of glutaraldehyde were added sequentially to 45 μM TriNTA-GB1 in the absence of His_6-tagged Fas TMD. SDS–PAGE of the crosslinked GB1 showed oligomeric species up to trimers and that the trimer band intensified with increased amounts of glutaraldehyde. The results indicate that the Fas TMD in bicelles forms trimers (Fig. 5c). In addition to the Fas TMD, use of the OG-label has been demonstrated for higher-order oligomers such as the p7 protein from the Hepatitis C virus (HCV)[23]. In that case, the OG-label managed to detect hexamers (Fig. 5d).

The major potential pitfall of the OG-label method is a situation in which multiple His_6-tags in the TMD oligomer are very close, as this could result in the binding of one TriNTA-GB1 to multiple His_6-tags and thereby underestimation of the oligomeric state. An effective solution to this problem is to insert a linker between the His_6-tag and the protein to spread the His_6-tags from different protomers. If the 1:1 binding between TriNTA-GB1 and the His_6-tag still cannot be achieved for extremely tight TMD oligomers, charged residues such as arginines can be added before and/or after the His_6-tag sequence to separate the His_6-tags by electrostatic repulsion. Another potential pitfall is that, for higher oligomeric states (e.g., $n > 4$), the crosslinked GB1 species may not migrate in SDS–PAGE exactly according to their true MW. We thus recommend performing independent validation of the MW by mass spectrometry (Box 1).

Part 3: J_{CH}-modulated NOE experiment for detecting unambiguous interprotomer NOEs

As the regular ^{15}N- or ^{13}C-edited NOESY spectrum is usually very complex, it is critical to first detect exclusively interchain NOEs using mixed samples consisting of differently isotopically labeled chains

as described in the protocol. We again use the Fas TMD as an example to illustrate the expected results of this experiment.

We used a mixed sample in which half of the Fas TMDs were (^{15}N/^{2}H)-labeled and the other half 15% ^{13}C-labeled to measure exclusively NOEs between the ^{15}N-attached protons of one subunit and aliphatic protons of the neighboring subunits (Fig. 6). The non-deuterated protein was 15% ^{13}C-labeled for recording the ^{1}H-^{13}C HSQC spectrum as an internal aliphatic proton chemical shift reference while providing stereospecific assignment of leucine and valine methyl groups[58]. As discussed in part 2, adequate mixing of two differently labeled proteins is critical for this type of experiment. We found that the HPLC profile can provide useful information on the miscibility of the TMDs. For example, the HPLC elution peaks of deuterated and protonated Fas TMDs are separated (Fig. 7), indicating that the Fas TMD is monomeric in the HPLC buffer. Therefore, mixing separately purified Fas TMDs immediately after HPLC elution should achieve thorough mixing. By contrast, the deuterated and protonated TMDs of HIV-1 gp41 eluted as one peak (Fig. 7b), suggesting that the protein oligomerized even in the HPLC organic solvent. For gp41 TMD, it is thus important to mix the differently labeled protein before CNBr cleavage. For the mixed Fas TMD sample, the simple ^{15}N-edited NOESY-TROSY-HSQC spectrum was recorded and showed obvious interchain NOE peaks between backbone amide and methyl groups (Fig. 6c). These visible NOEs must be interchain because the ^{15}N labeled Fas TMD is >98% deuterated, as indicated by the mass spectrometer analysis (Fig. 6b).

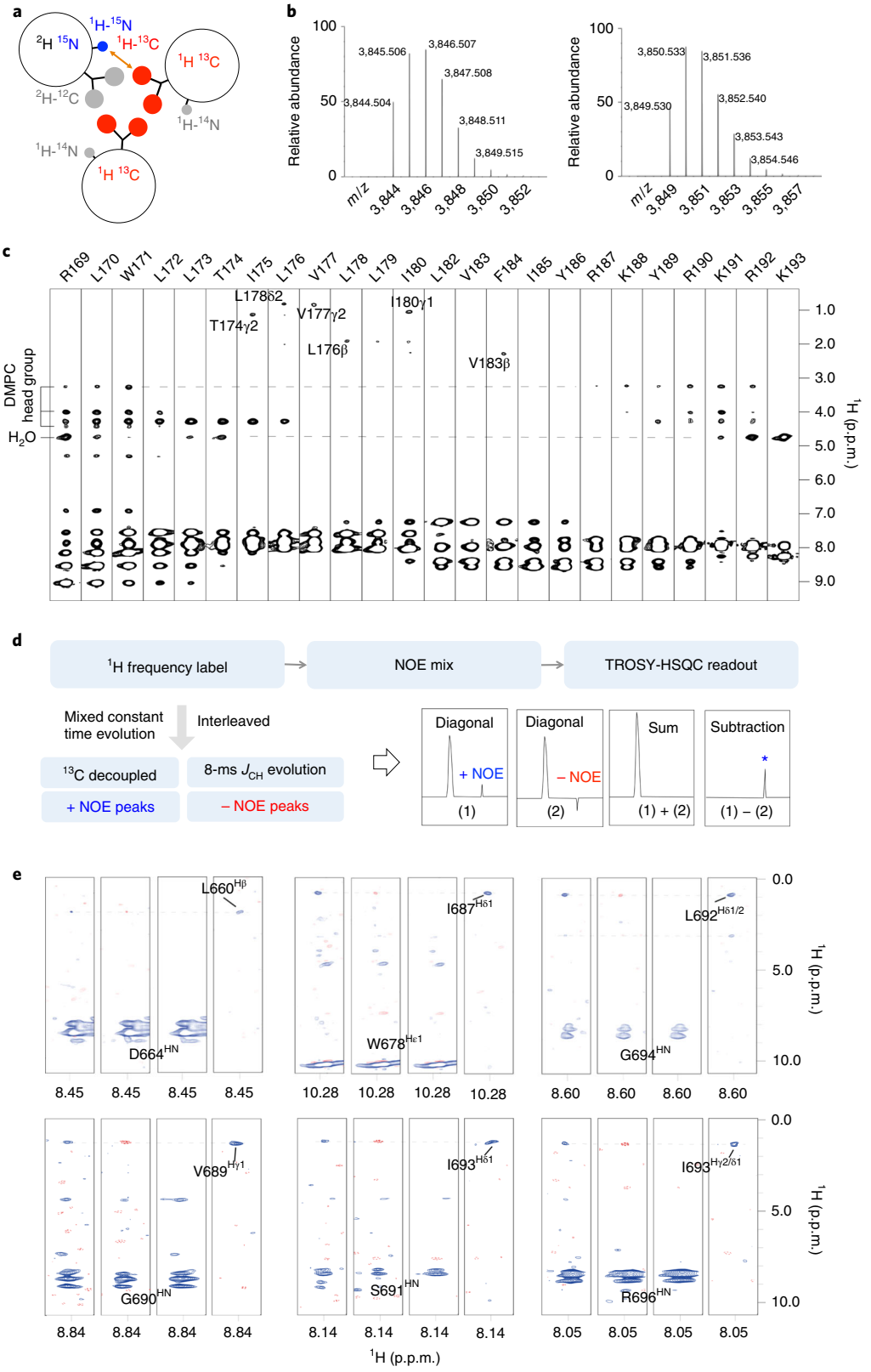
For certain applications in which the protein cannot be perdeuterated due to high cost, the J_{CH}-modulated NOE experiment[24] can be used to selectively detect NOEs between ^{15}N-attached ^{1}H and ^{13}C-attached ^{1}H. For this experiment, the mixed sample should contain 50% (^{15}N, ~85% ^{2}H)-labeled and 50% (^{13}C,^{1}H)-labeled chains (Fig. 6a). The J_{CH}-modulated NOE experiment is based on the regular 3D ^{15}N-edited NOESY-TROSY-HSQC, in which the ^{1}H evolution period before the NOE mixing is changed to a 'mixed constant-time' evolution to introduce ^{1}H-^{13}C J evolutions (Figs. 4 and 6d). Two interleaved 3D spectra are recorded. In one case, the total J evolution is 0 ms, and the inter- and intra-molecular NOE peaks are both positive (Fig. 6d). In the other case, the total J evolution is 8 ms, making intermolecular NOE peaks negative and intramolecular NOE peaks positive (Fig. 6d). Collecting the two datasets interleaved with each other allow us to add the two spectra to see only intramolecular NOEs or to subtract the two spectra to see only intermolecular NOEs. This method was applied to a fragment of the HIV-1 gp41 protein containing the MPER and TMD, designated MPER-TMD. As shown in Fig. 6e, the inter- and intra-molecular NOEs could be distinguished by simple addition and subtraction of the interleaved spectra[24].

Part 4: Oligomerization solution and structure determination

The main purpose of using the ExSSO program is to examine whether the interchain NOE restraints from part 3 above are sufficient to derive a unique solution of oligomeric assembly, and if so, to provide a starting model of the oligomer for further structural refinement. As explained in the

Fig. 6 | Detection of interchain NOEs. a, Isotopic-labeling scheme for detecting interchain NOEs. For regular NOESY experiments, the (^{15}N,^{2}H)-labeled chain is perdeuterated (>98%, with ^{1}H^{N} exchanged back), whereas the other half is (^{1}H,^{13}C) labeled. When using the J_{CH}-modulated NOE experiment, the (^{15}N,^{2}H)-labeled chain may have a lower deuteration level (e.g., ~80%), whereas the other half is uniformly ^{13}C-labeled (>99%). **b**, LC/MS analysis of a perdeuterated Fas TMD sample (Box 1), showing the calculated mass distribution assuming 100% deuteration at non-labile positions (right) and experimental mass distribution (left). The two dominant peaks in the theoretical spectrum are at 3,850 and 3,851 Da, whereas the corresponding peaks in the experimental spectrum are at 3,845 and 3,846 Da. The 5-Da difference means that, on average, only 5 out of the 288 possible deuteration positions of the Fas TMD are protonated, corresponding to a deuteration level of 98.3%. **c**, Residue-specific strips from the 3D ^{15}N-edited NOESY-TROSY-HSQC spectrum (NOE mixing time = 200 ms) recorded at 800 MHz and 303 K using a Fas TMD sample containing 50% (^{15}N,^{2}H)-labeled and 50% (^{1}H, 15% ^{13}C)-labeled chains. The labeled cross-peaks are interchain NOEs between the backbone amide and side-chain aliphatic protons. **d**, Flow diagram of the interleaved J_{CH}-modulated NOE experiment (Fig. 4). Interchain NOEs are positive in (1) (J_{CH} evolution = 0) and negative in (2) (J_{CH} evolution = 8 ms). (1) + (2) selects for the intra-chain NOEs, whereas (1) − (2) selects for the interchain NOEs. **e**, Residue-specific strips from the J_{CH}-modulated NOESY (NOE mixing time = 200 ms) recorded at 800 MHz and 308 K using an HIV-1 gp41 MPER TMD sample containing 50% (^{15}N,^{2}H)-labeled and 50% (^{1}H,^{13}C)-labeled chains. For each selected residue, four strips are shown from left to right: (1) (positive inter-NOEs, blue), (2) (negative inter-NOEs, red), (1) + (2) (inter-NOEs are canceled), and (1) − (2) (inter-NOEs are selected). In the example, the interchain NOEs selected are D664^{HN}-L660$^{H\beta}$, W678$^{H\varepsilon1}$-I687$^{H\delta1}$, G694^{HN}-L692$^{H\delta1/2}$, G690^{HN}-V689$^{H\gamma1}$, S691^{HN}-I693$^{H\delta1}$, and R696^{HN}-I693$^{H\gamma2/\delta1}$. **c** adapted with permission from ref. [5], Elsevier.

original ExSSO paper[43], for TMD oligomers that are higher than dimers, the interchain NOE restraints all have twofold directional ambiguity. The ExSSO program performs exhaustive search of the symmetric conformational space to resolve such ambiguity; it is demonstrated below for the



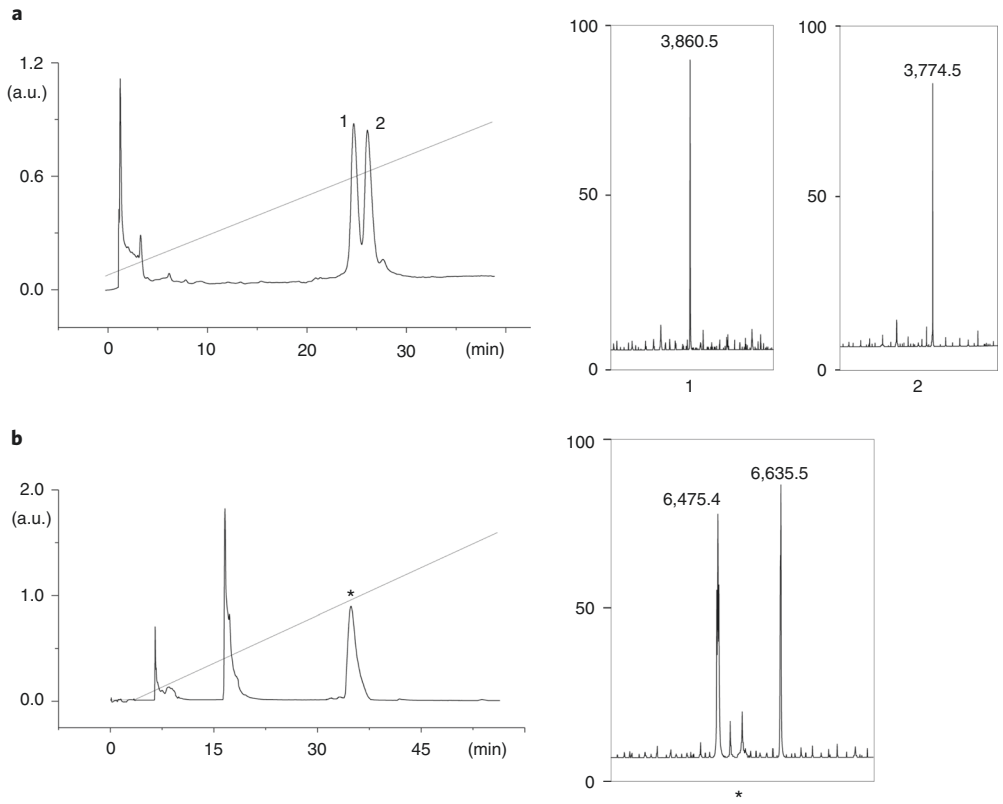


Fig. 7 | Sample mixing profiles for TMDs. a, HPLC chromatogram (left) of a mixed sample of Fas TMD half (^{15}N,^{2}H)-labeled (1) and the other half (^{1}H, 15% ^{13}C)-labeled (2), showing that the deuterated and the protonated proteins are eluted separately. MALDI-TOF mass spectroscopy analysis of the two peaks (right) confirmed that they indeed contain the pure deuterated and protonated protein, respectively. **b**, HPLC chromatogram (left) of a mixed sample of HIV-1 gp41 TMD half (^{15}N,^{2}H)-labeled and the other half (^{1}H,^{13}C)-labeled. The HPLC profile shows that the two different species cannot be resolved (the asterisk indicates the mixed peak), as confirmed by MALDI-TOF mass spectroscopy analysis (right, the asterisk indicates that the corresponding HPLC peak contains both mixed species).

Fas-TMD. In the ExSSO algorithm, each Fas TMD monomer is treated as a rigid body whose orientation and position relative to the symmetry axis were evaluated in the context of the symmetric oligomer and against the interchain NOE restraints. The monomer structure (Fig. 8) was initially constructed with the backbone dihedral angles derived from chemical shifts (using the TALOS+ program[59]). The oligomerization number was determined to be 3 by SDS–PAGE (Fig. 3d) and OG-label (Fig. 5c). The algorithm assigns the symmetry axis (z axis) and samples the orientation of each protomer by performing a Euler rotation around its center of mass. The reoriented protomers are placed at distance r between the z axis and their center of mass. Then the oligomer structure is evaluated against the interchain restraints (from Fig. 6c) using the ExSSO scoring system.

The procedure was repeated for the Fas TMD trimer to cover the entire symmetry space and showed that the top conformation clusters converged to the same mode of trimeric assembly (Fig. 8). The best scorer was used as the starting model for further structural refinement, which involved (i) identifying more self-consistent interchain NOEs from the regular ^{15}N-edited and ^{13}C-edited NOESY spectra, and (ii) updating the structure in XPLOR-NIH[44] with the new NOE restraints. The above two steps were performed iteratively until the structural ensemble had an RMSD of ~0.829 and ~1.392 Å for backbone and heavy atoms, respectively (Fig. 8).

Part 5: Characterization of the protein TM partition

The PPT protocol can be used to generate quantitative information about protein partition in the bilayer region of the bicelles, and it is demonstrated here for the Fas TMD trimer in bicelles (Fig. 9). For this application, (^{15}N,^{2}H)-labeled Fas TMD was reconstituted in DMPC/DHPC bicelles with $q = 0.5$ (Fig. 9a), and Gd-DOTA was used to generate solvent PRE. We emphasize that, for the

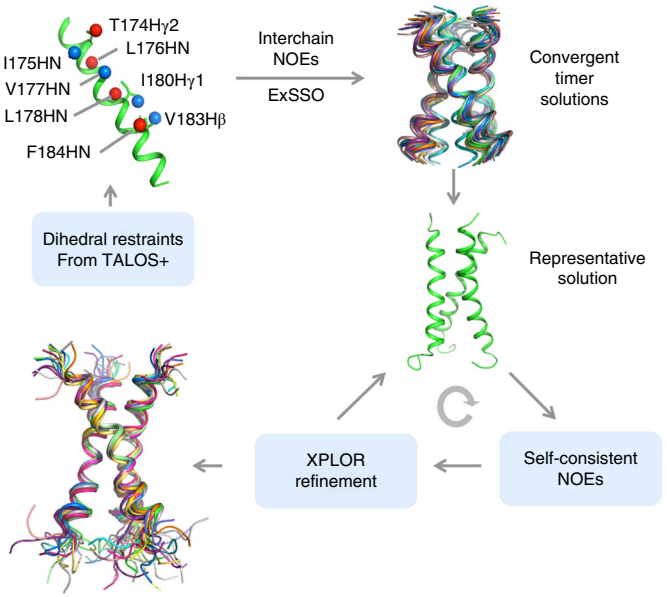


Fig. 8 | Structure determination of TMDs. The protocol used to calculate the structure of oligomeric TMDs is demonstrated for the Fas TMD. The monomer structure was generated using backbone dihedral angle restraints derived from chemical shifts using TALOS+. Backbone amide and side-chain protons exhibiting interchain NOEs are highlighted as spheres (red and blue spheres represent interprotomer pairs). An exhaustive search is then performed (by the ExSSO program) to resolve the directional ambiguity of interchain NOE restraints and to derive convergent assembly solutions. The trimeric state was assigned to the Fas TMD on the basis of the OG-label result in Fig. 5c. A representative trimer structure was finally selected and used as a starting model for further structural refinement using the XPLOR-NIH program. Additional self-consistent NOEs were added, and the structure was updated. The process was repeated iteratively. The final structural ensemble of 15 low-energy structures has backbone and heavy-atom RMSDs of 0.829 and 1.392 Å, respectively.

PPT method to be applicable, bicelles with $q \geq 0.5$ should be used, because when bicelles are sufficiently wide, the lateral contribution to the measurable PRE becomes negligible, so that the observed PREs can be used directly to probe residue-specific depth immersion of the protein along the bicelle normal axis. Another critical aspect of PPT is to add the paramagnetic agent via serial titration, rather than a single addition, as a wide range of paramagnetic field strengths (from very weak to strong) are required to probe both solvent-exposed and buried protein regions.

At each titration point, a 2D ^{1}H-^{15}N TROSY-HSQC spectrum was recorded to measure residue-specific PRE, defined here as I/I_0, in which I and I_0 are the intensities of a peak in the presence and absence of the paramagnetic agent, respectively (Fig. 9b). For each of the residues, the PRE titration curve was then fitted to the exponential decay function (Eq. 1) to derive the residue-specific PRE_{amp} (Fig. 9c). To determine the Fas TMD position relative to the bilayer center, the PPT method exploits the knowledge of the protein structure and of its symmetry axis, which is parallel to the bicelle normal and defines the protein orientation. The only degree of freedom, the protein position along the bicelle normal, is determined by translating the protein as a rigid body until its position agrees with the measured PRE_{amp}. Specifically, the PRE_{amp} versus residue number plot was converted to PRE_{amp} versus bilayer immersion depth by calculating, for each residue i, the distance along the bilayer normal (r_z) from the amide proton to an arbitrary reference point on that basis of the protein structure (Fig. 9d). The PRE_{amp} versus r_z plot showed three main PRE regimes: (i) a PRE-saturated regime near the bicelle surface; (ii) a PRE-insensitive regime in the lipid core; (iii) a PRE-sensitive regime connecting the previous two. Owing to the symmetric nature of the quickly tumbling bicelles, the PRE_{amp} versus $|r_z|$ can be described by the symmetric sigmoidal function (Eq. 2). Likewise, if the protein position in the bicelle is correctly assigned, the PRE_{amp} should fit well to the symmetric function. On the basis of the above rationale, the Fas TMD trimer was systematically slid along the bicelle normal by moving the reference point ($r_z = 0$) until the best fit of PRE_{amp} to Eq. 2 was reached (Fig. 9e). The Fas TMD position yielding the highest adjusted coefficient of determination (R^2_{adj}) represents the membrane partition of the protein (Fig. 9f).

The PPT is not limited to solvent PRE. A lipophilic probe such as 16-DSA is an equally powerful probe that provides a reciprocal PRE profile to that of the Gd-DOTA. Although 16-DSA can also

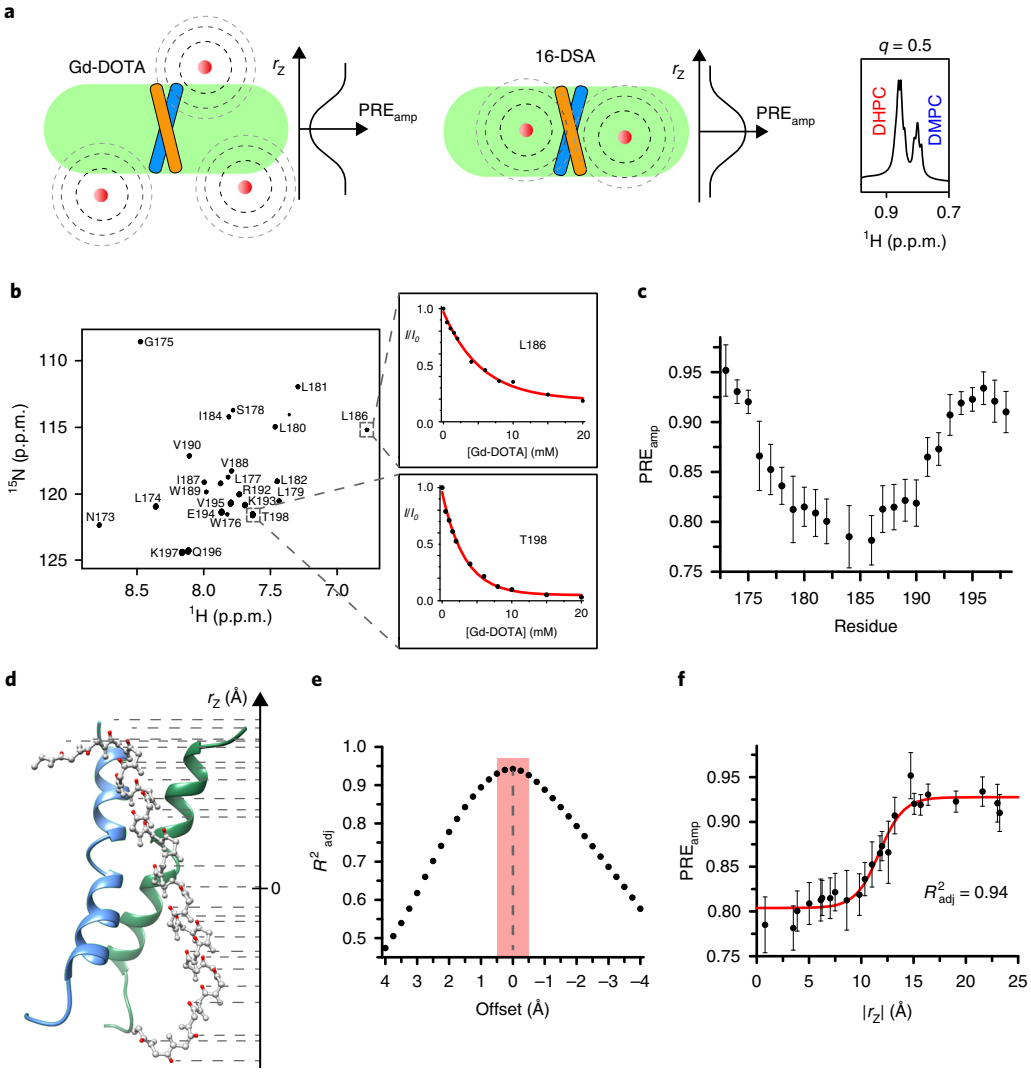


Fig. 9 | The PPT method for determination of the protein membrane partition. An overview of the method is reported for the Gd-DOTA case, using the Fas TMD as an example. The NMR spectra were acquired at 600 MHz at 303 K. **a**, Schematic illustration of a bicelle-reconstituted TMD titrated with either a water-soluble paramagnetic probe such as Gd-DOTA (left) or a lipophilic paramagnetic probe such as 16-DSA (middle). Next to each bicelle, the expected PRE profile along the bicelle normal is shown for both cases (reciprocal). On the right, the zoomed region of the 1D ^{1}H NMR spectrum shows that the bicelle q was 0.5. **b**, The 2D ^{1}H-^{15}N TROSY-HSQC spectrum of the Fas TMD in bicelles. As examples, the I/I_0 versus [Gd-DOTA] plots of L186 (buried) and T198 (exposed) are shown. The data fitting (red line) to the exponential decay function (Eq. 1) yielded PRE$_{amp}$ for the two residues. **c**, Residue-specific PRE$_{amp}$ plotted versus the residue number. **d**, The NMR structure of the Fas TMD homotrimer with the symmetry axis aligned with the bicelle normal (r_Z), showing the amide protons (red spheres) for which the PRE$_{amp}$ has been determined. The dashed lines point to the projected positions (r_Z) of the amide protons on the bicelle normal. **e**, The adjusted coefficient of determination (R^2_{adj}) is used to evaluate the agreement between the assigned protein position and the PRE data. The position yielding the best fit (highest R^2_{adj}) represents the actual membrane partition of the protein. The error in the measure is within ±0.5 Å, as shown in the red area **f**, The optimized fit of the PRE$_{amp}$ versus r_Z data from **e** to Eq. 2, yielding the membrane partition of the Fas TMD. Image adapted with permission from ref. [19], John Wiley & Sons.

weakly partition in the detergent rim of the bicelles, it preferentially resides in the bilayer region due to better compatibility of the acyl chain length. In fact, PRE$_{amp}$ profiles derived from GD-DOTA and 16-DSA titrations are reciprocal of each other (Fig. 10a)[22,23], meaning that the 16-DSA in the bilayer region of the bicelles dominates the membrane partition analysis. Given their opposite solubility, Gd-DOTA results are better suited to investigating residues well-buried in the bicelle, whereas 16-DSA is better at probing residues in the bicelle head-group region (Fig. 10a).

The PPT method can also be useful for studying the topology of a TM protein with unknown structure, or for obtaining structural information for structure validation/refinement. In fact, the slopes of the PRE$_{amp}$ versus residue number plots reflect the relative orientation of the different protein segments with respect to the bicelle normal axis. For example, the PRE$_{amp}$ versus residue

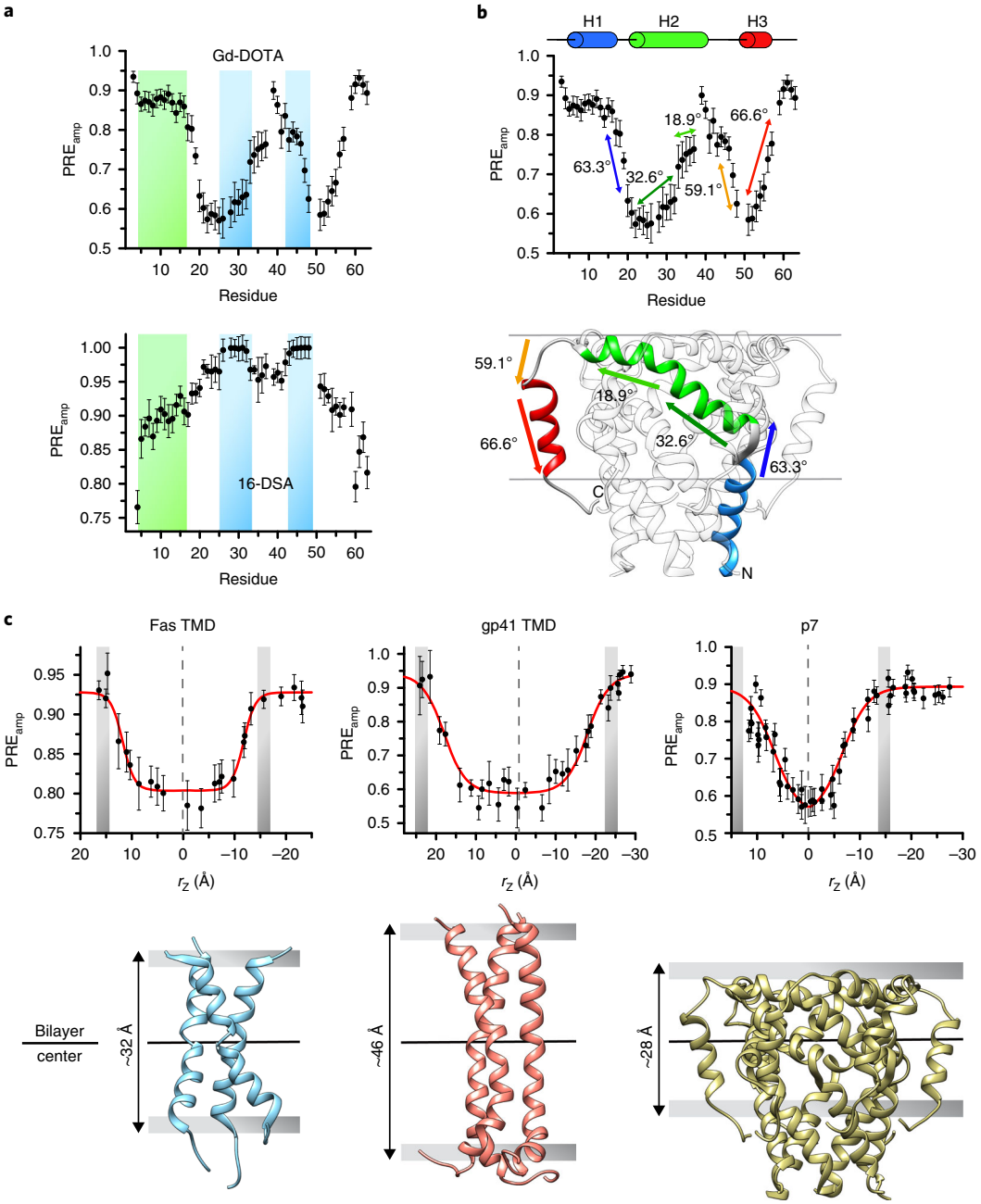


Fig. 10 | Analysis of the protein TM partition. a, Comparison between residue-specific PRE_{amp} derived from titrating Gd-DOTA (top) and 16-DSA (bottom), shown for the HCV p7 reconstituted in bicelles with $q = 0.6$. The two plots are reciprocal and their analysis (performed as described in Fig. 9) yielded identical membrane partitions. As can be seen, Gd-DOTA is better suited to investigation of residues well-buried in the bicelle (shaded in blue), showing steeper slope in the Gd-DOTA plot, whereas 16-DSA is more sensitive at probing residues in the bicelle head-group region (shaded in green). **b**, The PRE_{amp} versus residue plot from Gd-DOTA titration in **a**, showing its correlation with the p7 structure. The arrows (in different colors) indicate the fragments of the p7 structure (bottom) for which the slope of PRE_{amp} correlates well with the steepness of the fragment in the hexamer structure. The steepness is reported as the angle of the fragment relative to the bilayer plane. **c**, Membrane partition calibration curves obtained for the Fas TMD (left), HIV-1 gp41 TMD (middle), and HCV p7 (right), describing the partition of the proteins in the bicelle. The gray bars represent the bicelle boundaries, identified as the positions at which the PRE_{amp} values begin reaching saturation. At the bottom, the membrane partitions of the three proteins show different lipid bilayer thickness around the proteins (~32, ~46, and ~28 Å for Fas TMD, HIV-1 gp41 TMD, and HCV p7, respectively). Image adapted with permission from ref. [23], Elsevier.

number plot for the HCV p7 shows a 'W' or 'M' shape (for Gd-DOTA or 16-DSA, respectively), indicating that the protein crosses the bilayer center two times. Importantly, the slopes of the plot correlate very well with the orientation of the protein helices, providing useful information for structure validation or refinement (Fig. 10b)[23].

Finally, the PPT method can also be used to estimate the lipid bilayer thickness around the TMD. In fact, the r_z values in Fig. 10c at which the PRE_{amp} starts to become saturated should represent the bicelle boundaries. We note that for accurate determination of the membrane boundaries, the use of solvent PRE is preferred over that of lipophilic PRE, because soluble paramagnetic probes can usually be applied at higher concentration than the lipophilic probes, making it easier to reach PRE saturation. Using this simple PPT analysis, we found that different TMDs can indeed modulate different lipid bilayer thicknesses around them. For example, the bilayer thickness around the Fas TMD is 32 Å (Fig. 10c, left)[5]. The bilayer thickness around the HIV-1 gp41 MPER TMD is substantially larger, at 46 Å (Fig. 10c, middle)[22]. However, the HCV p7 caused drastic thinning of the bilayer to 28 Å (Fig. 10c, right)[23], which is consistent with the fact that the endoplasmic reticulum membrane, where the p7 resides, is much thinner than other membranes.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability

The code and instructions for the ExSSO program are freely accessible from the website: http://www.csbio.sjtu.edu.cn/bioinf/ExSSO/. The software is also provided as Supplementary Software 1 and 2.

References

1. Call, M. E., Wucherpfennig, K. W. & Chou, J. J. The structural basis for intramembrane assembly of an activating immunoreceptor complex. *Nat. Immunol.* **11**, 1023–1029 (2010).
2. Call, M. E. & Wucherpfennig, K. W. Common themes in the assembly and architecture of activating immune receptors. *Nat. Rev. Immunol.* **7**, 841–850 (2007).
3. Endres, N. F. et al. Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **152**, 543–556 (2013).
4. Arkhipov, A. et al. Architecture and membrane interactions of the EGF receptor. *Cell* **152**, 557–569 (2013).
5. Fu, Q. et al. Structural basis and functional role of intramembrane trimerization of the Fas/CD95 death receptor. *Mol. Cell* **61**, 602–613 (2016).
6. MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. A transmembrane helix dimer: structure and implications. *Science* **276**, 131–133 (1997).
7. Trenker, R., Call, M. E. & Call, M. J. Crystal structure of the glycophorin a transmembrane dimer in lipidic cubic phase. *J. Am. Chem. Soc.* **137**, 15676–15679 (2015).
8. Call, M. E. et al. The structure of the zetazeta transmembrane dimer reveals features essential for its assembly with the T cell receptor. *Cell* **127**, 355–368 (2006).
9. Bocharov, E. V. et al. Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state. *J. Biol. Chem.* **283**, 6950–6956 (2008).
10. Lau, T. L., Kim, C., Ginsberg, M. H. & Ulmer, T. S. The structure of the integrin alphaIIbbeta3 transmembrane complex explains integrin transmembrane signalling. *EMBO J.* **28**, 1351–1361 (2009).
11. Barrett, P. J. et al. The amyloid precursor protein has a flexible transmembrane domain and binds cholesterol. *Science* **336**, 1168–1171 (2012).
12. Dev, J. et al. Structural basis for membrane anchoring of HIV-1 envelope spike. *Science* **353**, 172–175 (2016).
13. Chen, W. et al. Familial Alzheimer's mutations within APPTM increase Abeta42 production by enhancing accessibility of epsilon-cleavage site. *Nat. Commun.* **5**, 3037 (2014).
14. Lee, J. et al. Structure of the Ebola virus envelope protein MPER/TM domain and its interaction with the fusion loop explains their fusion activity. *Proc. Natl. Acad. Sci. USA* **114**, E7987–E7996 (2017).
15. Klammt, C. et al. Facile backbone structure determination of human membrane proteins by NMR spectroscopy. *Nat. Methods* **9**, 834 (2012).
16. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
17. Gasteiger, E. et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784–3788 (2003).
18. Glover, K. J. et al. Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules. *Biophys. J.* **81**, 2163–2171 (2001).
19. Piai, A., Fu, Q., Dev, J. & Chou, J. J. Optimal bicelle size q for solution NMR studies of the protein transmembrane partition. *Chemistry* **23**, 1361–1367 (2017).
20. Sanders, C. R., Hare, B. J., Howard, K. P. & Prestegard, J. H. Magnetically-oriented phospholipid micelles as a tool for the study of membrane-associated molecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **26**, 421–444 (1994).
21. Caldwell, T. A. et al. Low-q bicelles are mixed micelles. *J. Phys. Chem. Lett.* **9**, 4469–4473 (2018).
22. Piai, A., Dev, J., Fu, Q. & Chou, J. J. Stability and water accessibility of the trimeric membrane anchors of the HIV-1 envelope spikes. *J. Am. Chem. Soc.* **139**, 18432–18435 (2017).

23. Chen, W. et al. The unusual transmembrane partition of the hexameric channel of the hepatitis C virus. *Structure* **26**, 627–634.e4 (2018).
24. Fu, Q. et al. Structure of the membrane proximal external region of HIV-1 envelope glycoprotein. *Proc. Natl. Acad. Sci. USA* **115**, E8892–E8899 (2018).
25. Knoblich, K. et al. Transmembrane complexes of DAP12 crystallized in lipid membranes provide insights into control of oligomerization in immunoreceptor assembly. *Cell Rep.* **11**, 1184–1192 (2015).
26. Hofer, N., Aragao, D. & Caffrey, M. Crystallizing transmembrane peptides in lipidic mesophases. *Biophys. J.* **99**, L23–L25 (2010).
27. Thomaston, J. L. & DeGrado, W. F. Crystal structure of the drug-resistant S31N influenza M2 proton channel. *Protein Sci.* **25**, 1551–1554 (2016).
28. Cooper, R. S., Georgieva, E. R., Borbat, P. P., Freed, J. H. & Heldwein, E. E. Structural basis for membrane anchoring and fusion regulation of the herpes simplex virus fusogen gB. *Nat. Struct. Mol. Biol.* **25**, 416–424 (2018).
29. Barbet-Massin, E. et al. Rapid proton-detected NMR assignment for proteins with fast magic angle spinning. *J. Am. Chem. Soc.* **136**, 12489–12497 (2014).
30. Andreas, L. B. et al. Structure and mechanism of the influenza A M218-60 dimer of dimers. *J. Am. Chem. Soc.* **137**, 14877–14886 (2015).
31. van Dam, L., Karlsson, G. & Edwards, K. Direct observation and characterization of DMPC/DHPC aggregates under conditions relevant for biological solution NMR. *Biochim. Biophys. Acta* **1664**, 241–256 (2004).
32. Galperin, M. Y. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J. Bacteriol.* **188**, 4169–4182 (2006).
33. Parkinson, J. S. & Kofoid, E. C. Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.* **26**, 71–112 (1992).
34. Stock, A. M., Robinson, V. L. & Goudreau, P. N. Two-component signal transduction. *Annu. Rev. Biochem.* **69**, 183–215 (2000).
35. Wolanin, P. M., Thomason, P. A. & Stock, J. B. Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biol.* **3**, REVIEWS3013 (2002).
36. Blacklow, S. C. & Kim, P. S. Protein folding and calcium binding defects arising from familial hypercholesterolemia mutations of the LDL receptor. *Nat. Struct. Biol.* **3**, 758–762 (1996).
37. North, C. L. & Blacklow, S. C. Solution structure of the sixth LDL-A module of the LDL receptor. *Biochemistry* **39**, 2564–2571 (2000).
38. Crimmins, D. L., Mische, S. M. & Denslow, N. D. Chemical cleavage of proteins in solution. *Curr. Protoc. Protein Sci.* Chapter 11, Unit 11.4 (2005).
39. Ni, J. & Kanai, M. Site-selective peptide/protein cleavage. *Top Curr. Chem.* **372**, 103–123 (2016).
40. Hwang, P. M., Pan, J. S. & Sykes, B. D. Targeted expression, purification, and cleavage of fusion proteins from inclusion bodies in *Escherichia coli*. *FEBS Lett.* **588**, 247–252 (2014).
41. Caldwell, T. A. et al. Low- q bicelles are mixed micelles. *J. Phys. Chem. Lett.* **9**, 4469–4473 (2018).
42. Lata, S., Reichel, A., Brock, R., Tampe, R. & Piehler, J. High-affinity adaptors for switchable recognition of histidine-tagged proteins. *J. Am. Chem. Soc.* **127**, 10205–10215 (2005).
43. Yang, J., Piai, A., Shen, H. B. & Chou, J. J. An exhaustive search algorithm to aid NMR-based structure determination of rotationally symmetric transmembrane oligomers. *Sci. Rep.* **7**, 17373 (2017).
44. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).
45. Mitra, K., Ubarretxena-Belandia, I., Taguchi, T., Warren, G. & Engelman, D. M. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proc. Natl. Acad. Sci. USA* **101**, 4083–4088 (2004).
46. Sharpe, H. J., Stevens, T. J. & Munro, S. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell* **142**, 158–169 (2010).
47. McMahon, H. T. & Boucrot, E. Membrane curvature at a glance. *J. Cell. Sci.* **128**, 1065–1070 (2015).
48. Rossman, J. S., Jing, X., Leser, G. P. & Lamb, R. A. Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell* **142**, 902–913 (2010).
49. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
50. Guidotti, G. The composition of biological membranes. *Arch. Intern. Med.* **129**, 194–201 (1972).
51. Trenker, R., Call, M. J. & Call, M. E. Progress and prospects for structural studies of transmembrane interactions in single-spanning receptors. *Curr. Opin. Struct. Biol.* **39**, 115–123 (2016).
52. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
53. Xu, C. et al. Regulation of T cell receptor activation by dynamic membrane binding of the CD3epsilon cytoplasmic tyrosine-based motif. *Cell* **135**, 702–713 (2008).
54. Pielak, R. M., Oxenoid, K. & Chou, J. J. Structural investigation of rimantadine inhibition of the AM2-BM2 chimera channel of influenza viruses. *Structure* **19**, 1655–1663 (2011).
55. Schnell, J. R. & Chou, J. J. Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* **451**, 591–595 (2008).
56. OuYang, B. et al. Unusual architecture of the p7 channel from hepatitis C virus. *Nature* **498**, 521–525 (2013).

57. Wang, J., Pielak, R. M., McClintock, M. A. & Chou, J. J. Solution structure and functional analysis of the influenza B proton channel. *Nat. Struct. Mol. Biol.* **16**, 1267–1271 (2009).
58. Szyperski, T., Neri, D., Leiting, B., Otting, G. & Wuthrich, K. Support of 1H NMR assignments in proteins by biosynthetically directed fractional 13C-labeling. *J. Biomol. NMR* **2**, 323–334 (1992).
59. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223 (2009).

Acknowledgements

This work was supported by US National Institutes of Health grants GM116898 and AI127193 to J.J.C.

Author contributions

Q.F., A.P., W.C., and J.J.C. conceived the study. K.X. performed the mass spectrometry analysis. Q.F., A.P., W.C., and J.J.C. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41596-019-0188-9.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to J.J.C.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 11 November 2018; Accepted: 30 April 2019;
Published online: 3 July 2019

Related links

Key references using this protocol

Fu, Q. et al. *Mol. Cell* **61**, 602–613 (2016): https://doi.org/10.1016/j.molcel.2016.01.009
Piai, A., Devi, J., Fu, Q., & Chou, J. J. *J. Am. Chem. Soc.* **139**, 18432–18435 (2017): https://doi.org/10.1021/jacs.7b09352
Chen, W. et al. *Structure* **26**, 627–634.e4 (2018): https://doi.org/10.1016/j.str.2018.02.011
Fu, Q. et al. *Proc. Natl. Acad. Sci. USA* **115**, E8892–E8899 (2018): https://doi.org/10.1073/pnas.1807259115

nature research

Corresponding author(s): James J. Chou

Last updated by author(s): Apr 8, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	N/A
Data analysis	N/A

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Provide your data availability statement here.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	N/A
Data exclusions	No
Replication	The protocol was used on 6 TM domains and all cases are successfully done.
Randomization	N/A
Blinding	N/A

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging