nature structural & molecular biology

Toward high-resolution computational design of the structure and function of helical membrane proteins

Patrick Barth[1–3] & Alessandro Senes[4]

The computational design of α-helical membrane proteins is still in its infancy but has already made great progress. *De novo* design allows stable, specific and active minimal oligomeric systems to be obtained. Computational reengineering can improve the stability and function of naturally occurring membrane proteins. Currently, the major hurdle for the field is the experimental characterization of the designs. The emergence of new structural methods for membrane proteins will accelerate progress.

The fields of protein-structure prediction and protein design started after the elucidation of the principles that govern protein folding. After all, the ability to predict structure *ab initio* stringently validates the understanding of how information is encoded in the primary sequence. Similarly, the ability to design structure and enzymatic activities *de novo* is the ultimate test for understanding of the biophysical principles that govern folding, stability and function.

Although the field is still far from being able to predict and design structure purely on the basis of physical principles—and it currently relies heavily on empirical information derived from analysis of 3D structures and primary sequences—great progress has been achieved, especially regarding soluble proteins. Protein prediction and design have been fostered by the development of computational methods that can efficiently explore the vast space of possible protein conformations and of energy functions that can effectively rank the most optimal conformations[1,2]. Progress has also been promoted by the rapid growth of structural and genome sequence databases. The availability of protein structures in the Protein Data Bank (PDB) has provided a testing ground for structural-prediction methods; structural analysis has provided a conformational 'alphabet' (i.e., favorable side chain, secondary, tertiary and quaternary conformations) to be used as building blocks; and structural data also offer templates for homology-based prediction. Similarly, the enormous amount of primary sequence data available today is highly informative regarding the relationship between amino acid sequence and protein folds; the data also contain evolutionary information on structurally important contacts between residues that coevolve. As a result, the design of soluble proteins has progressed far beyond the early focus on the exploration of secondary structure and short-sequence assembly[3], achieving milestones such as the creation of conformational switches[4] and folds not yet found in nature[5], and even the design of enzymatically active proteins[6–9].

The design of α-helical membrane proteins, which represent ~30% of the entire proteome[10–12], has lagged behind in this trend, in part because less structural information is available for membrane proteins, and understanding of folding in the membrane is not fully established. This lag is also a testament to the enormous difficulties encountered during structural characterization of membrane proteins, which hinder the necessary tight integration between computational development and experimental testing. Nevertheless, computational methods are already proving to be important for advancing the biological and biophysical understanding of membrane proteins. In this Perspective, we discuss the current state of design of α-helical membrane proteins, including hurdles and prospects for future development. We consider *de novo* design as well as functional reengineering aimed at rationally altering the function of natural proteins or enhancing their stability. Finally, we briefly summarize how structural prediction is supporting the design and experimental study of membrane proteins.

De novo design: success of minimalistic functional models

The most complex *de novo* designs of helical membrane proteins to date have been created by DeGrado and collaborators. PRIME[13] and Rocker[14] are two designs based on the same basic backbone motif, a D_2-symmetrical antiparallel helical homotetramer (**Fig. 1**). D_2 symmetry is ideal for transfer and transport proteins because it places an axis of symmetry parallel to the membrane (*x* axis in **Fig. 1a**,**c**). It resembles the antiparallel homodimeric arrangement found in natural transporters that are inserted in an opposite orientation with respect to the membrane (or pseudoantiparallel arrangement, when two ancestral monomers have fused into a single polypeptide chain during evolution)[15].

PRIME is a 24 amino acid–long peptide designed to perform electron transfer in membranes. It forms a tetrameric helical bundle, with two pairs of helices that sandwich two non-natural iron porphyrins (**Fig. 1a**,**b**). The design of PRIME was based on the optimization of a number of interactions that stabilize structure and support function: the coordination of the iron, the hydrogen-bonding network and the packing among the side chains and cofactors. The design produced

[1]Structural and Computational Biology and Molecular Biophysics Graduate Program, Baylor College of Medicine, Houston, Texas, USA. [2]Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas, USA. [3]Department of Pharmacology, Baylor College of Medicine, Houston, Texas, USA. [4]Department of Biochemistry, University of Wisconsin–Madison, Madison, Wisconsin, USA. Correspondence should be addressed to P.B. (patrickb@bcm.edu) or A.S. (senes@wisc.edu).

Received 21 March; accepted 20 April; published online 7 June 2016; doi:10.1038/nsmb.3231

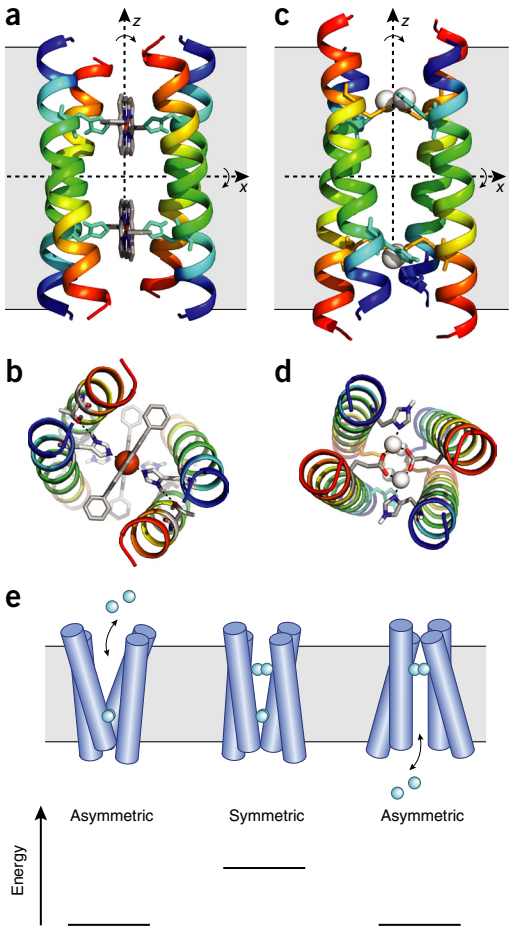

Figure 1 Minimalistic active membrane designs. (**a**–**d**) Side and top views of PRIME (**a**,**b**) and Rocker (**c**,**d**). Both designs are D_2-symmetrical antiparallel helical bundles of four identical subunits. The helices are oriented in alternating directions with respect to the membrane (gray). D_2 symmetry has two axes of symmetry: one perpendicular (*z*) and one parallel (*x*) to the membrane. This symmetry is ideal for transfer or transport proteins because it resembles the antiparallel homodimeric arrangement found in many natural transporters. The PRIME peptide (**a**,**b**) is 24 amino acids long and was designed for electron transfer. It binds two non-natural iron porphyrins (shown as gray sticks, with iron as red sphere), coordinated by histidine residues. The Rocker peptide (**c**,**d**) is 25 amino acids long and was designed to bind Zn^{2+} ions (shown as white spheres). Rocker was designed to have two types of helix-helix interfaces: an alanine-coil dimer with close interhelical contacts and a second looser interface between two alanine-coil dimers (**d**). The zinc is coordinated by histidine and glutamate amino acids. (**e**) The symmetric conformation of Rocker, which would inhibit transport, has been preferentially destabilized through a negative design protocol performed by assessing the thermodynamic preference for symmetric and nonsymmetric states of the candidate sequences with the VALOCIDY algorithm[16]. The resulting preference for two degenerate asymmetric configurations enables the complex to alternate between states that expose a binding site to either face of the membrane.

a stable complex that binds cofactors tightly. Beyond stability, the complex has remarkable specificity, and mutations predicted to alter the packing or the hydrogen-bonding network severely affect its assembly. The structure of PRIME has not been solved experimentally, but the predicted geometry of metal coordination was verified by EPR. Although full assessment of the functional capabilities of PRIME has not been performed, initial characterization showed that PRIME facilitates electron transfer in a phospholipid bilayer.

PRIME is a functional design, but it is static. Recently, Grigoryan, DeGrado and co-workers have made a major step toward the creation of multistate artificial membrane proteins with the Zn^{2+} transporter Rocker, a 25 amino acid–long peptide[14] (**Fig. 1c**,**d**). The design of Rocker represents a departure from the rigid tetrameric coiled coil of PRIME and includes introduced backbone motions that facilitate metal transfer between two binding sites and subsequent metal release on the opposite side of the membrane. To enable this transition, Rocker was designed to have two types of helix-helix interfaces: an alanine-coil dimer with tight interhelical contacts and a looser second interface separating the two alanine-coil dimers (**Fig. 1d**). The dimer of dimers was designed to adopt two degenerate asymmetric configurations that open a Zn^{2+}-binding site to either face of the membrane, thus breaking the intrinsic preference of the homotetramer for the symmetric configuration, which would inhibit transport by enabling both Zn^{2+}-binding sites to be simultaneously occupied. The necessary conformational frustration was achieved through a negative-design algorithm selecting against the canonical symmetric configuration[16,17] (schematic illustration in **Fig. 1e**). The design resembles the mechanism of the natural drug-efflux pump EmrE, which also cycles through asymmetrical homodimeric states[18].

Rocker transports Zn^{2+} down a gradient and functions as a Zn^{2+}/H^+ antiporter. Rocker's rate of transport is slow compared with that of natural transporters, but its level of function is notable for a protein obtained by pure computational design, without the aid of further screening of directed evolution. The structure of Rocker was solved by X-ray crystallography, and although the structure is partial (a single alanine-coil dimer with no metal bound), it is in good agreement with the designed architecture of the complex. This work represents a milestone because it is the first time that the structure of a *de novo*–designed membrane protein has been solved at high resolution.

Structural motifs as building blocks in membrane-protein design

Analysis of sequence-structure relationships has been a key contributor to understanding of membrane-protein folding. Unlike soluble proteins, which combine different secondary elements to display a wide variety of folds, the helical membrane proteins adopt a fairly homogeneous helical-bundle architecture. Therefore, a great deal of analysis has been performed to understand the most favored helix-helix interaction motifs.

The search for recurrent sequence motifs at interacting transmembrane helices began with experimental screening and statistical analysis of the sequence database, such as the work that led to the discovery of the GxxxG motif[19,20]. With the growth of the number of structures in the PDB, it has become possible to analyze structures for recurring themes. These recurring elements, found in a variety of unrelated proteins, are relevant for design because they are likely to be useful for engineering the stability of specific helical topologies.

Early analysis of helix-helix interaction motifs revealed that ~75% of all such interactions fall within a few structural motifs[21]. Subsequent studies based on a much-expanded structural database indicated that helical dimer motifs of soluble and membrane proteins share similar geometries, but the membrane motifs are enriched in small amino acids, thus facilitating the formation of hydrogen bonds (both side chain–mediated and backbone-mediated hydrogen bonds, i.e., Cα-H···O)[22]. This finding is consistent with the particular importance of hydrogen-bonding interactions in establishing tertiary and quaternary structure in membrane proteins[23,24]. Most recently, Feng and Barth have identified helical trimers as being

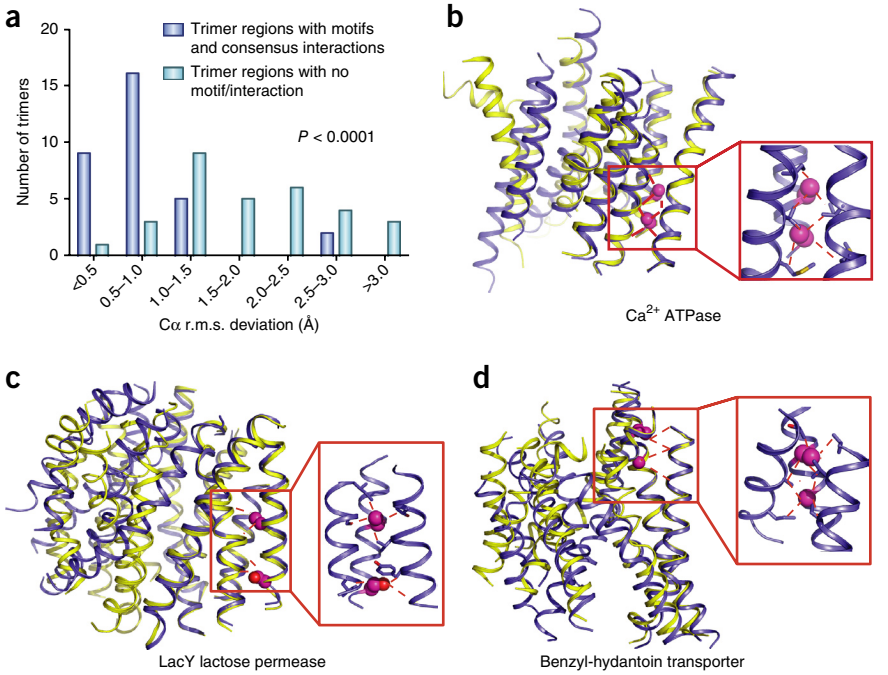

Figure 2 Sequence and 3D contact motifs are strong predictors of local conformational stability. (**a**) Distribution of trimer-unit structural changes (measured by Cα r.m.s. deviation in angstroms) in multipass membrane proteins crystallized in distinct conformations. Data for trimers containing sequence or contact motifs are in blue, and others are in teal. (**b**–**d**) Examples of multipass-membrane-protein X-ray structures are Ca^{2+} ATPase (**b**), LacY (**c**) and benzyl-hydantoin transporter (**d**), crystallized in two distinct conformations (superimposed backbone representations in blue and yellow). The trimer units containing a sequence or contact motif (red window) do not change conformations. Adapted from ref. 25, Nature Publishing Group.

the most recurrent minimal packing units in multipass membrane proteins[25]. Remarkably, only six structural classes of trimers cover a large fraction of the topology of interacting trimers. Bioinformatic analysis has indicated that the recurrent sequence motifs mediating interhelical contacts in the different trimer classes show a high degree of covariation, thus suggesting that these contacts have indeed been preserved during evolution, even in functionally unrelated protein families. Moreover, the motifs anticorrelate with structural flexibility (**Fig. 2a**), thus linking their frequent occurrence to a role in local conformational stability (**Fig. 2b–d**).

A clear example of how structural motifs provide opportunities for engineering stability is the creation of computed helical anti-membrane proteins (CHAMPs). CHAMPs are biologically active peptides designed to bind to the transmembrane domain of specific classes of integrins, thereby disrupting association with their natural partner by competing for the same binding surface[26]. The scaffold chosen for the CHAMPs' design was the GAS_{right} motif, one of the most common and stable dimeric motifs, which is characterized by its signature GxxxG sequence motif and stabilized by extensive van der Waals contacts and networks of Cα-H hydrogen bonds[21,24,27]. As membrane-protein design continues to target more complex and ambitious goals, common and well-understood structural motifs such as GAS_{right} will provide the components for engineering function around structurally defined and stable frameworks.

Beyond minimalism: the need for better experimental methods

The designs of PRIME, Rocker and CHAMPs are based on short peptide oligomers. Designing small sequences decreases the space of conformations to be searched, but this simplification is not what has driven the field toward this type of 'minimal' functional model rather than toward longer polytopic monomeric proteins. In fact, working with short sequences imposes severe limitations on the complexity of the design and thus paradoxically may decrease the number of possible solutions available for stable and active designs.

In principle, it would be interesting to test the ability to design larger polytopic proteins *de novo*, but experimental validation of these types of designs is the main challenge. Even with natural membrane proteins, each step in their biophysical characterization—expression, purification, determination of their stability and structure determination—can be difficult. The minimalist strategy is relatively approachable: the chemical synthesis of synthetic peptides is robust, and measuring oligomerization is easier than determining the folding of a large polytopic protein. Furthermore, the binding of cofactors and the measurement of activity offer an avenue for evaluating designs at an early stage, allowing for validation even if structural characterization fails.

A lowering of the experimental barriers faced today in the structural characterization of polytopic membrane proteins will certainly open the door for expanding these initial successes to a broader variety of designs. This reduced barrier would also enable a closer integration of computational development and experimental testing, as is necessary, for example, to improve the energy functions and to foster the creation of more realistic lipid-bilayer models, beyond the implicit solvent approximations commonly used today[28]. Indeed, recent technical advances in membrane-protein structural biology (X-ray crystallography, NMR and cryo-EM)[29–36] suggest that more systematic high-resolution characterization of designed membrane proteins may be achieved in the near future.

Functional design and redesign of helical membrane proteins

Beyond *de novo* design, which stringently tests understanding of the physical principles governing protein structure, protein engineering has also proven to be very effective at manipulating naturally evolved proteins. This strategy is important for better understanding the biophysical and functional properties of biological systems. Other important protein-design goals are improving stability to enable *in vitro* and structural studies, and functional reprogramming by modulating protein state occupancy, altering binding specificities or influencing the allosteric response.

Engineering stability is particularly important in biophysical studies. Until recently, membrane proteins have primarily been solubilized in detergent micelles, which poorly mimic the physicochemical properties of lipid bilayers and often lead to destabilization or unfolding[37]. Additionally, because a majority of membrane proteins function by changing conformations[35,38], high intrinsic conformational flexibility may hinder crystallization, thereby preventing structural determination. This phenomenon has prompted empirical experimental efforts to stabilize membrane proteins (through large-scale scanning of point mutations[39], directed evolution and selection[40], and other techniques[41]). Although these approaches have successfully identified

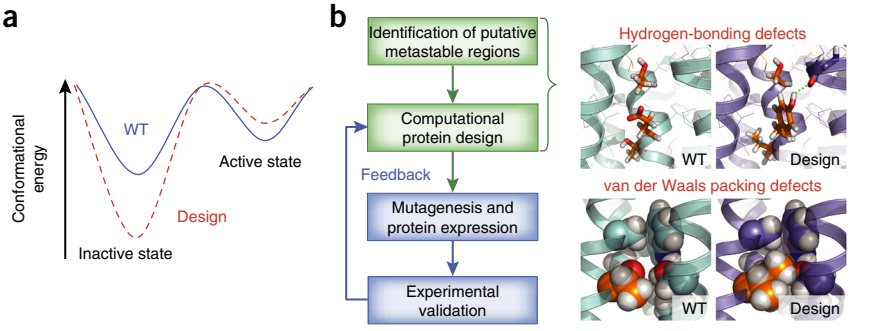

Figure 3 Conformational membrane-protein thermostabilization by computational design. (**a**) Conformational energy landscape depicting the effects of designed mutations selectively stabilizing the inactive state of a membrane receptor. (**b**) Integrated bioinformatics, computational design and experiments identifying metastable sites and designing thermostabilizing mutations.

thermostabilized variants of numerous membrane proteins, most of these variants have been trapped in a given conformational state and consequently have often been functionally impaired. For example, neurotensin-receptor variants evolved for stability cannot signal unless their functionally important residues are reverted back to their identity in the wild-type receptor[42]. This issue highlights the need for rational computational engineering approaches that can deconstruct mutational effects on protein stability and function.

Chen *et al.* combined structure-based computational protein design techniques and bioinformatics analysis to select highly thermostabilized variants with a minimal number of designed mutations, as demonstrated by their selective stabilization of the inactive-state conformation of a GPCR[43] (**Fig. 3a**). By combining structure and evolution-based sequence analysis, the authors have identified nonconserved residues forming suboptimal interactions with their environment (**Fig. 3b**). The authors initially hypothesized that mutations could readily be identified at the sites stabilizing the protein without disrupting protein folding and function. They then designed combinations of mutations *in silico* from 20 possible amino acids at each putative metastable site, by using objective physical criteria. Remarkably, the designed variants displayed experimentally enhanced thermostability by up to 31 °C from the wild type and up to 11 °C from a variant selected by scanning mutagenesis. As intended by the calculations, the designed mutations induced up to a 396-fold decrease in agonist affinity, thus indicating that computational design can modulate receptor pharmacology. The method was also able to predict the thermostabilization effects of 70% of empirically selected mutations for diverse GPCRs. Interestingly, whereas packing defects and unsatisfied polar residues were repeatedly found in GPCR structures, their exact locations appeared to be specific to each receptor member or subfamily. These findings suggest that, by promoting specific local conformational flexibility, metastable motifs might encode functional selectivity in the conformational changes governing GPCR signaling.

The origin of the stabilizing effects of empirically selected mutations in GPCRs has been rationalized through molecular dynamics[44]. Overall, the analyzed mutations were found to contribute both enthalpically and entropically to receptor thermostabilization. The combination of effects includes increased favorable tertiary interactions and increased receptor rigidity associated with decreased collective motions and the presence of ordered water molecules. These findings highlight the diversity of potential stabilizing determinants that can be computationally targeted through use of rational amino acid substitutions. However, modulating many of these properties—such as conformational dynamics, correlated motions and protein core hydration—by computational design remains a major challenge, owing to the high associated computational costs.

Reprogramming membrane-protein function represents another major area of computational protein engineering. Because the binding of extracellular molecules regulates intracellular signals or transport in numerous classes of membrane receptors, developing protein variants with altered ligand binding or sensing properties is an active area of research. Empirical approaches have been used to engineer GPCR variants activated solely by synthetic drugs not recognized by native receptors[45,46], through mimicking the lock-and-key strategy used to design selective protein kinase inhibitors[47]. GPCRs with reprogrammed ligand binding selectivity should provide the next generation of molecules to allow deconstruction of the role of native GPCRs in complex signaling responses and to design of new therapies. Recent advances in computational techniques to design protein ligand-binding sites[48] have suggested that such approaches may be feasible to manipulate membrane-protein functions.

A second avenue for reprogramming the responses of receptors to extracellular stimuli is modulating the allosteric properties encoded by the receptor sequence and structure. Recent approaches interpreting conformational dynamic fluctuations extracted from molecular dynamic simulations have identified networks of highly dynamically correlated residues in membrane proteins[49–51]. These residue networks can in principle propagate changes in structure and dynamics across the receptor by providing allosteric pathways connecting extracellular to intracellular binding sites. Preferentially targeting these residues in protein design may provide an effective approach to engineer membrane proteins with altered allosteric and signaling responses to native ligands.

All the above rational engineering strategies for thermostabilization and functional reprogramming hold promise for furthering the structural and functional understanding of membrane proteins. A major complication, however, is that these design approaches rely on high-resolution structural information. In the absence of experimental structures, the application of these methods can be extended to structurally uncharacterized membrane proteins, depending on the accuracy of models obtained by structure-prediction techniques.

Structure prediction in support of membrane-protein design

Computational design and structure prediction are closely related technologies. Prediction seeks to find the most favorable structure for a given sequence within a large conformational space. Design aims at identifying the sequence that would best stabilize a given backbone structure among a large number of possible amino acid combinations. Although the objectives are different, the two procedures share many underlying methodologies. Therefore, progress in one area may lead to progress in the other area. However, the most important role of structure prediction in relation to design is to provide structural templates. Because structural information remains limited for membrane proteins, accurate structure prediction can in principle extend membrane-protein engineering efforts beyond the minority of proteins of known structure.

Until recently, most prediction efforts focused on *de novo* structure prediction (i.e., from sequence). Whereas folding small protein

domains has become feasible[52,53], predicting the structure of larger proteins with complex topologies remains challenging[54]. Recent breakthroughs in predicting inter-residue structural contacts from genome-sequence coevolutionary signals have improved the accuracy of large predicted membrane-protein structures and protein binding interfaces[55–57].

Whereas *de novo* prediction has not yet reached the level of accuracy necessary to guide protein-design applications, homology-based models are becoming sufficiently accurate to allow transmembrane protein regions to be designed[58]. As the membrane-protein structural database continues to grow, an increasing number of proteins can now be modeled from homologs with solved structures. Methods such as Modeller[59], I-TASSER[60] and Medeller[61] regularly achieve high structural accuracy when multiple close-homolog structures are available. Approaches combining *de novo* structure reconstruction with homology-based techniques have also emerged, thereby extending the high-accuracy regime to proteins whose closest solved homologs have lower sequence identity (i.e., down to 15%)[58,62,63]. The major remaining challenge concerns the accuracy of loop regions, which often diverge among homologs. Because these regions contribute to the recognition of substrate, allosteric modulators and cytoplasmic effectors, the ability to completely redesign membrane proteins will largely depend on future improvements in loop modeling techniques[64,65].

Concluding remarks

The computational design of α-helical membrane proteins, compared with soluble proteins, is still emerging. Nevertheless, encouraging progress has been made in recent years, including adopting minimal oligomeric systems to create active assemblies and succeeding in reengineering natural membrane proteins to modulate their stability and function. Currently, the major hurdles for membrane-protein design reside in the difficult experimental structure determination of both natural and artificial proteins. The emergence of new methods for structural and biophysical characterization of membrane proteins will probably support a tighter integration of computational development and experimental testing and consequently accelerate technological progress. With these advances, computational design holds promise to become a key tool for investigating the structure and function of membrane proteins and an integral component in biotechnology, synthetic biology and therapeutic applications.

ACKNOWLEDGMENTS

P.B. is supported by National Institutes of Health grant R01GM097207, by a Lilly Research Award Program and by a supercomputer allocation from XSEDE (MCB120101). A.S. is supported by National Institutes of Health grant R01GM0997522 and National Science Foundation grant CHE-1415910. We are grateful to S. Condon, S. Anderson, X. Feng and H. Cao for critical reading of the manuscript and to G. Grigoryan for providing the model of Rocker.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Li, Z., Yang, Y., Zhan, J., Dai, L. & Zhou, Y. Energy functions in de novo protein design: current challenges and future prospects. *Annu. Rev. Biophys.* **42**, 315–335 (2013).
2. Pantazes, R.J., Grisewood, M.J. & Maranas, C.D. Recent advances in computational protein design. *Curr. Opin. Struct. Biol.* **21**, 467–472 (2011).
3. Richardson, J.S. & Richardson, D.C. The de novo design of protein structures. *Trends Biochem. Sci.* **14**, 304–309 (1989).
4. Ambroggio, X.I. & Kuhlman, B. Computational design of a single amino acid sequence that can switch between two distinct protein folds. *J. Am. Chem. Soc.* **128**, 1154–1161 (2006).
5. Kuhlman, B. *et al.* Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**, 1364–1368 (2003).
6. Jiang, L. *et al.* De novo computational design of retro-aldol enzymes. *Science* **319**, 1387–1391 (2008).
7. Röthlisberger, D. *et al.* Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195 (2008).
8. Siegel, J.B. *et al.* Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **329**, 309–313 (2010).
9. Khare, S.D. *et al.* Computational redesign of a mononuclear zinc metalloenzyme for organophosphate hydrolysis. *Nat. Chem. Biol.* **8**, 294–300 (2012).
10. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E.L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
11. Almén, M.S., Nordström, K.J.V., Fredriksson, R. & Schiöth, H.B. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol.* **7**, 50 (2009).
12. Liu, J. & Rost, B. Comparing function and structure between entire proteomes. *Protein Sci.* **10**, 1970–1979 (2001).
13. Korendovych, I.V. *et al.* De novo design and molecular assembly of a transmembrane diporphyrin-binding protein complex. *J. Am. Chem. Soc.* **132**, 15516–15518 (2010).
14. Joh, N.H. *et al.* De novo design of a transmembrane Zn^{2+}-transporting four-helix bundle. *Science* **346**, 1520–1524 (2014).
15. Forrest, L.R. Structural symmetry in membrane proteins. *Annu. Rev. Biophys.* **44**, 311–337 (2015).
16. Grigoryan, G. Absolute free energies of biomolecules from unperturbed ensembles. *J. Comput. Chem.* **34**, 2726–2741 (2013).
17. Hallen, M.A., Keedy, D.A. & Donald, B.R. Dead-end elimination with perturbations (DEEPer): a provable protein design algorithm with continuous sidechain and backbone flexibility. *Proteins* **81**, 18–39 (2013).
18. Morrison, E.A. *et al.* Antiparallel EmrE exports drugs by exchanging between asymmetric structures. *Nature* **481**, 45–50 (2012).
19. Senes, A., Gerstein, M. & Engelman, D.M. Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J. Mol. Biol.* **296**, 921–936 (2000).
20. Russ, W.P. & Engelman, D.M. The GxxxG motif: a framework for transmembrane helix-helix association. *J. Mol. Biol.* **296**, 911–919 (2000).
21. Walters, R.F.S. & DeGrado, W.F. Helix-packing motifs in membrane proteins. *Proc. Natl. Acad. Sci. USA* **103**, 13658–13663 (2006).
22. Zhang, S.-Q. *et al.* The membrane- and soluble-protein helix-helix interactome: similar geometry via different interactions. *Structure* **23**, 527–541 (2015).
23. Senes, A., Engel, D.E. & DeGrado, W.F. Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr. Opin. Struct. Biol.* **14**, 465–479 (2004).
24. Senes, A., Ubarretxena-Belandia, I. & Engelman, D.M. The Cα—H···O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. USA* **98**, 9056–9061 (2001).
25. Feng, X. & Barth, P. A topological and conformational stability alphabet for multipass membrane proteins. *Nat. Chem. Biol.* **12**, 167–173 (2016).
26. Yin, H. *et al.* Computational design of peptides that target transmembrane helices. *Science* **315**, 1817–1822 (2007).
27. Mueller, B.K., Subramaniam, S. & Senes, A. A frequent, GxxxG-mediated, transmembrane association motif is optimized for the formation of interhelical Cα-H hydrogen bonds. *Proc. Natl. Acad. Sci. USA* **111**, E888–E895 (2014).
28. Mori, T., Miyashita, N., Im, W., Feig, M. & Sugita, Y. Molecular dynamics simulations of biological membranes and membrane proteins using enhanced conformational sampling algorithms. *Biochim. Biophys. Acta* http://dx.doi.org/10.1016/j.bbamem.2015.12.032 (2016).
29. Cheng, Y., Grigorieff, N., Penczek, P.A. & Walz, T. A primer to single-particle cryo-electron microscopy. *Cell* **161**, 438–449 (2015).
30. Bai, X.C., McMullan, G. & Scheres, S.H.W. How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* **40**, 49–57 (2015).
31. Poulos, S., Morgan, J.L.W., Zimmer, J. & Faham, S. Bicelles coming of age: an empirical approach to bicelle crystallization. *Methods Enzymol.* **557**, 393–416 (2015).
32. Williamson, J.A. *et al.* Structure and multistate function of the transmembrane electron transporter CcdA. *Nat. Struct. Mol. Biol.* **22**, 809–814 (2015).
33. Wang, S. *et al.* Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. *Nat. Methods* **10**, 1007–1012 (2013).
34. Isogai, S. *et al.* Backbone NMR reveals allosteric signal transduction networks in the β_1-adrenergic receptor. *Nature* **530**, 237–241 (2016).
35. Manglik, A. *et al.* Structural insights into the dynamic process of β2-adrenergic receptor signaling. *Cell* **161**, 1101–1111 (2015).
36. Liu, J.J., Horst, R., Katritch, V., Stevens, R.C. & Wüthrich, K. Biased signaling pathways in β_2-adrenergic receptor characterized by ^{19}F-NMR. *Science* **335**, 1106–1110 (2012).
37. Zhou, Y. & Bowie, J.U. Building a thermostable membrane protein. *J. Biol. Chem.* **275**, 6975–6979 (2000).
38. Katritch, V., Cherezov, V. & Stevens, R.C. Structure-function of the G protein-coupled receptor superfamily. *Annu. Rev. Pharmacol. Toxicol.* **53**, 531–556 (2013).

39. Magnani, F., Shibata, Y., Serrano-Vega, M.J. & Tate, C.G. Co-evolving stability and conformational homogeneity of the human adenosine A_{2a} receptor. *Proc. Natl. Acad. Sci. USA* **105**, 10744–10749 (2008).
40. Sarkar, C.A. *et al.* Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc. Natl. Acad. Sci. USA* **105**, 14808–14813 (2008).
41. Chun, E. *et al.* Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* **20**, 967–976 (2012).
42. Egloff, P. *et al.* Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **111**, E655–E662 (2014).
43. Chen, K.-Y.M., Zhou, F., Fryszczyn, B.G. & Barth, P. Naturally evolved G protein-coupled receptors adopt metastable conformations. *Proc. Natl. Acad. Sci. USA* **109**, 13284–13289 (2012).
44. Vaidehi, N., Grisshammer, R. & Tate, C.G. How can mutations thermostabilize G-protein-coupled receptors? *Trends Pharmacol. Sci.* **37**, 37–46 (2016).
45. Conklin, B.R. *et al.* Engineering GPCR signaling pathways with RASSLs. *Nat. Methods* **5**, 673–678 (2008).
46. Roth, B.L. DREADDs for neuroscientists. *Neuron* **89**, 683–694 (2016).
47. Knight, Z.A. & Shokat, K.M. Features of selective kinase inhibitors. *Chem. Biol.* **12**, 621–637 (2005).
48. Tinberg, C.E. *et al.* Computational design of ligand-binding proteins with high affinity and selectivity. *Nature* **501**, 212–216 (2013).
49. Bhattacharya, S. & Vaidehi, N. Differences in allosteric communication pipelines in the inactive and active states of a GPCR. *Biophys. J.* **107**, 422–434 (2014).
50. Miao, Y., Nichols, S.E., Gasper, P.M., Metzger, V.T. & McCammon, J.A. Activation and dynamic network of the M2 muscarinic receptor. *Proc. Natl. Acad. Sci. USA* **110**, 10982–10987 (2013).
51. LeVine, M.V. & Weinstein, H. NbIT: a new information theory-based analysis of allosteric mechanisms reveals residues that underlie function in the leucine transporter LeuT. *PLoS Comput. Biol.* **10**, e1003603 (2014).
52. Yarov-Yarovoy, V., Schonbrun, J. & Baker, D. Multipass membrane protein structure prediction using Rosetta. *Proteins* **62**, 1010–1025 (2006).
53. Barth, P., Schonbrun, J. & Baker, D. Toward high-resolution prediction and design of transmembrane helical protein structures. *Proc. Natl. Acad. Sci. USA* **104**, 15682–15687 (2007).
54. Barth, P., Wallner, B. & Baker, D. Prediction of membrane protein structures with complex topologies using limited constraints. *Proc. Natl. Acad. Sci. USA* **106**, 1409–1414 (2009).
55. Hopf, T.A. *et al.* Three-dimensional structures of membrane proteins from genomic sequencing. *Cell* **149**, 1607–1621 (2012).
56. Ovchinnikov, S. *et al.* Large-scale determination of previously unsolved protein structures using evolutionary information. *eLife* **4**, e09248 (2015).
57. Wang, Y. & Barth, P. Evolutionary-guided *de novo* structure prediction of self-associated transmembrane helical proteins with near-atomic accuracy. *Nat. Commun.* **6**, 7196 (2015).
58. Chen, K.-Y.M., Sun, J., Salvo, J.S., Baker, D. & Barth, P. High-resolution modeling of transmembrane helical protein structures from distant homologues. *PLoS Comput. Biol.* **10**, e1003636 (2014).
59. Eswar, N., Eramian, D., Webb, B., Shen, M.-Y. & Sali, A. Protein structure modeling with MODELLER. *Methods Mol. Biol.* **426**, 145–159 (2008).
60. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
61. Kelm, S., Shi, J. & Deane, C.M. MEDELLER: homology-based coordinate generation for membrane proteins. *Bioinformatics* **26**, 2833–2840 (2010).
62. Zhang, J., Yang, J., Jang, R. & Zhang, Y. GPCR-I-TASSER: a hybrid approach to G protein-coupled receptor structure modeling and the application to the human genome. *Structure* **23**, 1538–1549 (2015).
63. Bhattacharya, S. *et al.* Critical analysis of the successes and failures of homology models of G protein-coupled receptors. *Proteins* **81**, 729–739 (2013).
64. Mandell, D.J., Coutsias, E.A. & Kortemme, T. Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat. Methods* **6**, 551–552 (2009).
65. Tang, K., Zhang, J. & Liang, J. Fast protein loop sampling and structure prediction using distance-guided sequential chain-growth Monte Carlo method. *PLoS Comput. Biol.* **10**, e1003539 (2014).