Published in final edited form as:

Nat Struct Mol Biol. 2016 June 7; 23(6): 475–480. doi:10.1038/nsmb.3231.

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

Patrick Barth^{1,2,3} and Alessandro Senes⁴

¹Structural and Computational Biology and Molecular Biophysics Graduate Program, Baylor College of Medicine, Houston, Texas 77030

²Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas 77030

³Department of Pharmacology Baylor College of Medicine, Houston, Texas 77030

⁴Department of Biochemistry, University of Wisconsin-Madison, 433 Babcock Dr., Madison WI 53706

Abstract

The computational design of α -helical membrane proteins is still in its infancy but has made important progress. *De novo* design has produced stable, specific and active minimalistic oligomeric systems. Computational re-engineering can improve stability and modulate the function of natural membrane proteins. Currently, the major hurdle for the field is not computational, but the experimental characterization of the designs. The emergence of new structural methods for membrane proteins will accelerate progress

Introduction

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

Although we are still far from predicting and designing structure purely from physical principles – we rely heavily on empirical information derived from analysis of structural and sequence information – great progress has been achieved over the years, especially for 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 energy functions that are effective in ranking the most optimal conformations^{1,2}. Progress has been also promoted by the exponential growth of structural and genome sequence databases. The availability of protein structures in the Protein Data

Bank (PDB) has provided testing ground for structural prediction methods; structural analysis has provided a conformational “alphabet” – favorable side chain, secondary and tertiary and quaternary conformation – to be used as building blocks; structural data also offer templates for homology-based prediction. Similarly, the enormous sequence data available today is rich in information about the relationship between primary sequence and the fold; it also contains evolutionary information that can inform about structurally important contacts between residues that co-evolve. As a result, soluble protein design has progressed far beyond the early days’ focus on the exploration of secondary structure and short sequence assembly³, achieving milestones such as the creation of conformational switches⁴, the creation of folds not yet found in nature⁵ and even the design of enzymatically active proteins^{6–9}.

The α -helical membrane proteins – representing a sizable fraction (20–30%) of entire proteomes^{10–12} – have lagged behind in this trend, in part because of the reduced structural information and the less established understanding of folding in the membrane. Their lag is also a testament of the enormous difficulties encountered during structural characterization of membrane proteins, which prevent the necessary tight integration between computational development and experimental testing. Nevertheless, computational methods are already proving to be important for furthering our biological and biophysical understanding of membrane proteins. In this Perspective, we discuss the current state, the hurdles and the future developments for α -helical membrane protein design. We consider *de novo* design, as well as functional re-engineering aimed to rationally alter the function of natural proteins or enhance their stability. Finally, we briefly summarize how structural prediction is supporting the design and experimental study of membrane proteins.

***De novo* design: success of minimalist functional models**

The most complex *de novo* designs of helical membrane proteins to date have been created by DeGrado and collaborators. PRIME¹³ and Rocker¹⁴ are two designs based on the same basic backbone motif, a D_2 -symmetrical antiparallel helical homo-tetramer (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 anti-parallel homo-dimeric arrangement found in natural transporters that are inserted in opposite orientation with respect to the membrane (or pseudo-antiparallel arrangement, when two ancestral monomers have been 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 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 a stable complex that bound the cofactors tightly. In addition to stability, the complex had remarkable specificity, and mutations predicted to alter the packing or the hydrogen-bonding network severely affected its assembly. The structure of PRIME was not solved experimentally, but the predicted geometry of metal coordination was verified by EPR. Although full assessment of the functional capabilities of PRIME was not performed,

initial characterization showed that PRIME was able to facilitate electron transfer in a phospholipid bilayer.

PRIME was a functional design, but it was static. Recently, Grigoryan, DeGrado and coworkers made a significant step toward the creation of multi-state artificial membrane proteins with the Zn^{2+} transporter Rocker, a 25 amino acid-long peptide (Fig. 1c, d)¹⁴. The design of Rocker is a departure from the rigid tetrameric coiled coil of PRIME, with the introduction of backbone motions that facilitate metal transfer between two binding sites, and thus their release on the opposite side of the membrane. To enable this transition, Rocker was designed to have two types of helix-helix interfaces, an Ala-coil dimer with tight inter-helical contact and a second looser interface separating the two Ala-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 innate preference of the homo-tetramer 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 by means of a negative-design algorithm selecting against the canonical symmetric configuration^{16,17}. The strategy is schematically illustrated in Fig. 1e. The design resembles the mechanism of the natural drug efflux pump EmrE, which also cycles through asymmetrical homo-dimeric states¹⁸.

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

Structural motifs as membrane protein design building blocks

Analysis of sequence-structure relationships has been a key contributor to our 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 rather 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 those that led to the discovery of the GxxxG motif^{19,20}. With the growth of 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 stability of specific helical topologies.

Early analysis of helix-helix interaction motifs showed that ~75% of all helix-helix interactions fall within a few structural motifs²¹. Subsequent analysis 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 to facilitate the formation of hydrogen bonds (both side-chain mediated and backbone mediated hydrogen bonds, i.e. $C\alpha-H\cdots O$)²². This 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 identified helical trimers as the most recurrent minimal packing units in multi-pass membrane proteins²⁵. Remarkably, just six structural classes of trimers cover a large fraction of the topology of interacting trimers. Bioinformatic analysis indicated that the recurrent sequence motifs mediating inter-helical contacts in the different trimer classes had a high-degree of covariation, implying that these contacts are indeed preserved during evolution even in functionally unrelated protein families. Moreover, the motifs anti-correlated with structural flexibility (Fig. 2a), further suggesting that their frequent occurrence is primarily linked 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 the CHAMPs (Computed Helical Anti-Membrane Protein). The CHAMPs are biologically active peptides designed to bind to the transmembrane domain of specific classes of integrins, thus disrupting association with their natural partner by competing for the same binding surface²⁶. The scaffold chosen for the CHAMPs' design was the GAS_{right} motif, one of the most common and stable dimeric motifs, characterized by its signature GxxxG sequence motif and stabilized by extensive van der Waals contacts and networks of $C\alpha-H$ 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

PRIME, Rocker and CHAMPs are designs based on short peptide oligomers. Designing small sequences reduces the space of conformations to be searched, but this simplification is not what has driven the field toward these type of “minimal” functional models – as opposed to longer polytopic monomeric proteins. In fact, working with short sequences imposes severe limitations on the complexity of the design, which, paradoxically, may reduce the number of possible solutions available for stable and active designs.

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

A reduction of the experimental barriers that are 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. Reduction of these barriers will also enable at tighter integration between computational development and experimental testing, which 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 today²⁸. Indeed, recent technical advances in membrane protein structural biology (X-ray crystallography, NMR and cryo electron microscopy)^{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 our 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 allosteric response.

Engineering stability is particularly important for biophysical studies. Until recently, membrane proteins have been mostly solubilized in detergent micelles, which poorly mimic the physico-chemical properties of lipid bilayers, often leading to destabilization or unfolding³⁷. Additionally, since a majority of membrane proteins function by changing conformations^{35,38}, high intrinsic conformational flexibility may hinder crystallization, preventing their structural determination. This has led to empirical experimental efforts to stabilize membrane proteins (large scale scanning of point mutations³⁹, directed evolution and selection⁴⁰ and others⁴¹). Although these approaches have been successful in identifying thermostabilized variants of numerous membrane proteins, most of these variants were trapped in a given conformational state, and thus were often functionally impaired. For example, neurotensin receptor variants evolved for stability could not signal unless functionally important residues were reverted back to their identity in the WT receptor⁴². 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, selectively stabilizing the inactive state conformation of a GPCR⁴³ (Fig. 3a). By combining structure and evolutionary-based sequence analysis, the authors identified non-conserved residues forming suboptimal interactions with their environment (Fig. 3b). They hypothesized that mutations could readily be identified at these sites stabilizing the protein without disrupting its folding and function. Combinations of mutations were designed *in silico* from 20 possible amino-acids at each putative metastable site, using objective physical criteria. Remarkably, designed variants displayed experimentally enhanced thermostability by up to 31 °C from WT and up to 11 °C from a variant selected

by scanning mutagenesis. As intended by the calculations, designed mutations induced an up-to 396-fold decrease in agonist affinity, 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, while packing defects and unsatisfied polar residues were recurrently found in GPCR structures, their exact location appeared to be specific to each receptor member or subfamily. These findings suggested 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 using molecular dynamics⁴⁴. Overall, the analyzed mutations were found to contribute both enthalpically and entropically to the receptor thermostabilization. The combination of effects included increased favorable tertiary interactions, increased receptor rigidity associated with decreased collective motions and presence of ordered water molecules. These findings outline the diversity of potential stabilizing determinants that could be computationally targeted using 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 large associated computational costs.

Reprogramming membrane protein function represents another major area of impact for computational protein engineering. Since 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 represents 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}, mimicking the lock and key strategy used to design selective protein kinase inhibitors⁴⁷. GPCRs with reprogrammed ligand binding selectivity would provide the next generation of molecules to deconstruct the role of native GPCRs in complex signaling responses and to design novel therapies. Recent advances in computational techniques to design protein ligand binding sites⁴⁸ suggest that such approaches may be feasible to manipulate membrane protein functions.

A second avenue for reprogramming the responses of receptors to extracellular stimuli could be 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 our structural and functional understanding of membrane proteins. A major challenge, however, is that these design approaches rely on

high-resolution structural information. In the absence of experimental structure, 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 methods. Therefore, progress in one area may lead to progress in the other area. However, structure prediction's most important role in relation to design is to provide structural templates. Since 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). While folding small protein domains has become rapidly feasible^{52,53}, predicting the structure of larger proteins with complex topologies remains challenging⁵⁴. Recent breakthroughs in predicting inter-residue structural contacts from genome sequence coevolutionary signals 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 accurate enough for designing transmembrane protein regions⁵⁸. 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 Modeller⁵⁹, I-TASSER⁶⁰ and Medeller⁶¹ 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, 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 between homologs. Since these regions contribute to the recognition of substrate, allosteric modulators and cytoplasmic effectors, our ability to completely redesign membrane proteins will largely depend on future improvements in loop modeling techniques^{64,65}.

Conclusions

The computational design of α -helical membrane proteins is still emerging compared to soluble protein design. Nevertheless, the progress made in recent years – adopting minimalistic oligomeric systems to create active assemblies, and succeeding in re-engineering of natural membrane proteins for modulating their stability and function – is encouraging. 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 the structural and biophysical characterization of membrane proteins will likely support a tighter integration of computational development and experimental testing, which will accelerate technological progress. With these advances, computational design holds promise to become a key tool for investigating structure and function of membrane proteins and an integral component in biotechnology, synthetic biology and therapeutic applications.

Acknowledgments

PB was supported by National Institutes of Health Grant R01GM097207, by a Lilly Research Award Program and by a supercomputer allocation from XSEDE (MCB120101). AS was supported by National Institutes of Health Grant R01GM0997522 and National Science Foundation Grant CHE-1415910. We are grateful to Samson Condon, Samantha Anderson, Xiang Feng and Hongnan Cao for critical reading of the manuscript and to Gevorg Grigoryan for providing the model of Rocker.

References

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.* 2013; 42:315–335. [PubMed: 23451890]
2. Pantazes RJ, Grisewood MJ, Maranas CD. Recent advances in computational protein design. *Current Opinion in Structural Biology.* 2011; 21:467–472. [PubMed: 21600758]
3. Richardson JS, Richardson DC. The de novo design of protein structures. *Trends Biochem Sci.* 1989; 14:304–309. [PubMed: 2672455]
4. Ambroggio XI, Kuhlman B. Computational design of a single amino acid sequence that can switch between two distinct protein folds. *J Am Chem Soc.* 2006; 128:1154–1161. [PubMed: 16433531]
5. Kuhlman B, et al. Design of a novel globular protein fold with atomic-level accuracy. *Science.* 2003; 302:1364–1368. [PubMed: 14631033]
6. Jiang L, et al. De novo computational design of retro-aldol enzymes. *Science.* 2008; 319:1387–1391. [PubMed: 18323453]
7. Röthlisberger D, et al. Kemp elimination catalysts by computational enzyme design. *Nature.* 2008; 453:190–195. [PubMed: 18354394]
8. Siegel JB, et al. Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science.* 2010; 329:309–313. [PubMed: 20647463]
9. Khare SD, et al. Computational redesign of a mononuclear zinc metalloenzyme for organophosphate hydrolysis. *Nat Chem Biol.* 2012; 8:294–300. [PubMed: 22306579]
10. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 2001; 305:567–580. [PubMed: 11152613]
11. Almén MS, Nordström KJV, Fredriksson R, Schiöth HB. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol.* 2009; 7:50. [PubMed: 19678920]
12. Liu J, Rost B. Comparing function and structure between entire proteomes. *Protein Sci.* 2001; 10:1970–1979. [PubMed: 11567088]
13. Korendovych IV, et al. De novo design and molecular assembly of a transmembrane diporphyrin-binding protein complex. *J Am Chem Soc.* 2010; 132:15516–15518. [PubMed: 20945900]
14. Joh NH, et al. De novo design of a transmembrane Zn²⁺-transporting four-helix bundle. *Science.* 2014; 346:1520–1524. [PubMed: 25525248]
15. Forrest LR. Structural Symmetry in Membrane Proteins. *Annu Rev Biophys.* 2015; 44:311–337. [PubMed: 26098517]
16. Grigoryan G. Absolute free energies of biomolecules from unperturbed ensembles. *J Comput Chem.* 2013; 34:2726–2741. [PubMed: 24132787]

17. Hallen MA, Keedy DA, Donald BR. Dead-end elimination with perturbations (DEEPer): a provable protein design algorithm with continuous sidechain and backbone flexibility. *Proteins*. 2013; 81:18–39. [PubMed: 22821798]
18. Morrison EA, et al. Antiparallel EmrE exports drugs by exchanging between asymmetric structures. *Nature*. 2012; 481:45–50. [PubMed: 22178925]
19. Senes A, Gerstein M, Engelman DM. 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*. 2000; 296:921–36. [PubMed: 10677292]
20. Russ WP, Engelman DM. The GxxxG motif: a framework for transmembrane helix-helix association. *J Mol Biol*. 2000; 296:911–9. [PubMed: 10677291]
21. Walters RFS, DeGrado WF. Helix-packing motifs in membrane proteins. *Proc Natl Acad Sci USA*. 2006; 103:13658–63. [PubMed: 16954199]
22. Zhang SQ, et al. The membrane- and soluble-protein helix-helix interactome: similar geometry via different interactions. *Structure*. 2015; 23:527–541. [PubMed: 25703378]
23. Senes A, Engel DE, DeGrado WF. Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol*. 2004; 14:465–479. [PubMed: 15313242]
24. Senes A, Ubarretxena-Belandia I, Engelman DM. The Calpha ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc Natl Acad Sci USA*. 2001; 98:9056–61. [PubMed: 11481472]
25. Feng X, Barth P. A topological and conformational stability alphabet for multipass membrane proteins. *Nat Chem Biol*. 2016; 12:167–173. [PubMed: 26780406]
26. Yin H, et al. Computational design of peptides that target transmembrane helices. *Science*. 2007; 315:1817–22. [PubMed: 17395823]
27. Mueller BK, Subramaniam S, Senes A. A frequent, GxxxG-mediated, transmembrane association motif is optimized for the formation of interhelical C α -H hydrogen bonds. *Proc Natl Acad Sci USA*. 2014; 111:E888–895. [PubMed: 24569864]
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*. 2016; doi: 10.1016/j.bbamem.2015.12.032
29. Cheng Y, Grigorieff N, Penczek PA, Walz T. A primer to single-particle cryo-electron microscopy. *Cell*. 2015; 161:438–449. [PubMed: 25910204]
30. Bai X, McMullan G, Scheres SHW. How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci*. 2015; 40:49–57. [PubMed: 25544475]
31. Poulos S, Morgan JLW, Zimmer J, Faham S. Bicyclics coming of age: an empirical approach to bicelle crystallization. *Meth Enzymol*. 2015; 557:393–416. [PubMed: 25950975]
32. Williamson JA, et al. Structure and multistate function of the transmembrane electron transporter CcdA. *Nat Struct Mol Biol*. 2015; 22:809–814. [PubMed: 26389738]
33. Wang S, et al. Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. *Nat Methods*. 2013; 10:1007–1012. [PubMed: 24013819]
34. Isogai S, et al. Backbone NMR reveals allosteric signal transduction networks in the β 1-adrenergic receptor. *Nature*. 2016; 530:237–241. [PubMed: 26840483]
35. Manglik A, et al. Structural Insights into the Dynamic Process of β 2-Adrenergic Receptor Signaling. *Cell*. 2015; 161:1101–1111. [PubMed: 25981665]
36. Liu JJ, Horst R, Katritch V, Stevens RC, Wüthrich K. Biased signaling pathways in β 2-adrenergic receptor characterized by 19F-NMR. *Science*. 2012; 335:1106–1110. [PubMed: 22267580]
37. Zhou Y, Bowie JU. Building a thermostable membrane protein. *J Biol Chem*. 2000; 275:6975–6979. [PubMed: 10702260]
38. Katritch V, Cherezov V, Stevens RC. Structure-function of the G protein-coupled receptor superfamily. *Annu Rev Pharmacol Toxicol*. 2013; 53:531–556. [PubMed: 23140243]
39. Magnani F, Shibata Y, Serrano-Vega MJ, Tate CG. Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. *Proc Natl Acad Sci USA*. 2008; 105:10744–10749. [PubMed: 18664584]

40. Sarkar CA, et al. Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci USA*. 2008; 105:14808–14813. [PubMed: 18812512]
41. Chun E, et al. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure*. 2012; 20:967–976. [PubMed: 22681902]
42. Egloff P, et al. Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in *Escherichia coli*. *Proc Natl Acad Sci USA*. 2014; 111:E655–662. [PubMed: 24453215]
43. Chen KYM, Zhou F, Fryszczyn BG, Barth P. Naturally evolved G protein-coupled receptors adopt metastable conformations. *Proc Natl Acad Sci USA*. 2012; 109:13284–13289. [PubMed: 22847407]
44. Vaidehi N, Grisshammer R, Tate CG. How Can Mutations Thermostabilize G-Protein-Coupled Receptors? *Trends Pharmacol Sci*. 2016; 37:37–46. [PubMed: 26547284]
45. Conklin BR, et al. Engineering GPCR signaling pathways with RASSLs. *Nat Methods*. 2008; 5:673–678. [PubMed: 18668035]
46. Roth BL. DREADDs for Neuroscientists. *Neuron*. 2016; 89:683–694. [PubMed: 26889809]
47. Knight ZA, Shokat KM. Features of selective kinase inhibitors. *Chem Biol*. 2005; 12:621–637. [PubMed: 15975507]
48. Tinberg CE, et al. Computational design of ligand-binding proteins with high affinity and selectivity. *Nature*. 2013; 501:212–216. [PubMed: 24005320]
49. Bhattacharya S, Vaidehi N. Differences in allosteric communication pipelines in the inactive and active states of a GPCR. *Biophys J*. 2014; 107:422–434. [PubMed: 25028884]
50. Miao Y, Nichols SE, Gasper PM, Metzger VT, McCammon JA. Activation and dynamic network of the M2 muscarinic receptor. *Proc Natl Acad Sci USA*. 2013; 110:10982–10987. [PubMed: 23781107]
51. LeVine MV, 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*. 2014; 10:e1003603. [PubMed: 24785005]
52. Yarov-Yarovoy V, Schonbrun J, Baker D. Multipass membrane protein structure prediction using Rosetta. *Proteins*. 2006; 62:1010–25. [PubMed: 16372357]
53. Barth P, Schonbrun J, Baker D. Toward high-resolution prediction and design of transmembrane helical protein structures. *Proc Natl Acad Sci US A*. 2007; 104:15682–15687.
54. Barth P, Wallner B, Baker D. Prediction of membrane protein structures with complex topologies using limited constraints. *Proc Natl Acad Sci US A*. 2009; 106:1409–1414.
55. Hopf TA, et al. Three-dimensional structures of membrane proteins from genomic sequencing. *Cell*. 2012; 149:1607–1621. [PubMed: 22579045]
56. Ovchinnikov S, et al. Large-scale determination of previously unsolved protein structures using evolutionary information. *Elife*. 2015; 4:e09248. [PubMed: 26335199]
57. Wang Y, Barth P. Evolutionary-guided de novo structure prediction of self-associated transmembrane helical proteins with near-atomic accuracy. *Nat Commun*. 2015; 6:7196. [PubMed: 25995083]
58. Chen KYM, Sun J, Salvo JS, Baker D, Barth P. High-resolution modeling of transmembrane helical protein structures from distant homologues. *PLoS Comput Biol*. 2014; 10:e1003636. [PubMed: 24854015]
59. Eswar N, Eramian D, Webb B, Shen MY, Sali A. Protein structure modeling with MODELLER. *Methods Mol Biol*. 2008; 426:145–159. [PubMed: 18542861]
60. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*. 2010; 5:725–738. [PubMed: 20360767]
61. Kelm S, Shi J, Deane CM. MEDELLER: homology-based coordinate generation for membrane proteins. *Bioinformatics*. 2010; 26:2833–2840. [PubMed: 20926421]
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*. 2015; 23:1538–1549. [PubMed: 26190572]

63. Bhattacharya S, et al. Critical analysis of the successes and failures of homology models of G protein-coupled receptors. *Proteins*. 2013; 81:729–739. [PubMed: 23042299]
64. Mandell DJ, Coutsias EA, Kortemme T. Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat Methods*. 2009; 6:551–552. [PubMed: 19644455]
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*. 2014; 10:e1003539. [PubMed: 24763317]

Figure 1. Minimalistic active membrane designs

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 orientation of the helices goes in alternating direction with respect to the membrane (depicted in grey). 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 its symmetry resembles the antiparallel homo-dimeric arrangement found in many natural transporters. PRIME (a, b) is a 24 amino-acid peptide long and was designed for electron transfer. It binds two non-natural iron porphyrins, coordinated by His residues. Rocker (c, d) is a 25 amino-acid long peptide

designed to bind to Zn^{2+} ions. Rocker was designed to have two types of helix-helix interfaces: an Ala-coil dimer with close inter-helical contact and a second looser interface between two Ala-coil dimers (d). The zinc is coordinated by His and Glu amino-acids. e. The symmetric conformation of Rocker, which would inhibit transport, was preferentially destabilized by means of a negative design protocol, performed by assessing the thermodynamic preference for symmetric and non-symmetric states of the candidate sequences with the VALOCIDY algorithm¹⁶. The resulting preference for two degenerate asymmetric configurations enabled the complex to alternate between states that expose a binding site to either face of the membrane.

Figure 2. Sequence/3D contact motifs are strong predictors of local conformational stability
a. Distribution of trimer unit structural changes (measured by C α rmsd in Å) in multi-pass membrane proteins crystallized in distinct conformations. Data for trimers containing sequence/contact motifs are in black, others are in grey. b–d. Three examples of multi-pass membrane protein X-ray structures crystallized in two distinct conformations (superimposed backbone representations in blue and yellow). The trimer units containing a sequence/contact motif (red window) do not change conformations.

Figure 3. Conformational membrane protein thermostabilization by computational design

a. Conformational energy landscape depicting the effects of designed mutations stabilizing selectively the inactive state of a membrane receptor. b. integrated bioinformatics, computational design and experiments identifying metastable sites and designing thermostabilizing mutations.