

## Review Article

# Designing minimalist membrane proteins

 Paul Curnow<sup>1,2</sup>

<sup>1</sup>School of Biochemistry, University of Bristol, Bristol, U.K.; <sup>2</sup>BrisSynBio, Life Sciences Building, Tyndall Avenue, Bristol, U.K.

**Correspondence:** Paul Curnow (p.curnow@bristol.ac.uk)



The construction of artificial membrane proteins from first principles is of fundamental interest and holds considerable promise for new biotechnologies. This review considers the potential advantages of adopting a strictly minimalist approach to the process of membrane protein design. As well as the practical benefits of miniaturisation and simplicity for understanding sequence-structure-function relationships, minimalism should also support the abstract conceptualisation of membrane proteins as modular components for synthetic biology. These ideas are illustrated with selected examples that focus upon  $\alpha$ -helical membrane proteins, and which demonstrate how such minimalist membrane proteins might be integrated into living biosystems.

## Why design membrane proteins?

It is estimated that up to 30% of all proteins are integral membrane proteins [1], with at least one part of the protein sequence passing through a lipid bilayer membrane. This abundance reflects the universal role of such proteins in essential life processes such as signalling, solute transport, bioenergetics and much more. Developing a basic understanding of membrane protein biosynthesis, trafficking, insertion, folding, and assembly continues to attract considerable attention [2–7]. There is also now a growing interest in using membrane proteins in the emerging field of synthetic biology (e.g.[8–11]). In this minireview, we discuss the prospects of designing artificial membrane proteins from scratch — ‘*de novo*’ — as a particular way to explore the fundamental principles of membrane biology and to realise new applications in synthetic biosystems.

The motivations for membrane protein design are much the same as those put forward for ‘water-soluble’ proteins, which have recently been discussed in several excellent reviews and perspectives [12–19]. We do not intend to reproduce these arguments in full here, but they can be broadly and briefly summed up as follows. Natural proteins are intricate and complex, which can obscure the core physical principles underlying their structure and function. Building simplified model proteins can, therefore, be a useful way to cut through this complexity and understand the fundamental connections between sequence, structure and function. Another motivation for the designer is that natural selection has not had sufficient time to sample every possible combination of amino acids. Hence there are a large number of protein sequences (and by extension, structures and functions) that have never yet occurred in the natural world. *De novo* design allows us to survey the full scope of protein sequence and structural space, and to see whether artificial constructs could replace natural proteins or have useful non-natural functions. Exciting innovations in the field take advantage of new enabling technologies such as gene synthesis, automation and recombinant engineering as well as the improved accessibility, lower cost and greater sophistication of established techniques including peptide synthesis, mass spectrometry, NMR and computation.

In the case of integral membrane proteins, there are at least two additional motivations for the prospective designer. The most obvious stimulus is that our understanding of membrane proteins is now being revolutionised by a dramatic increase in the number of high-resolution structures available — although at the time of writing this still only accounts for <2% of all structures deposited in the protein databank [20]. This progress in structural biology provides a rich new source of information for design; for example, as discussed below, analysing known structures can identify general sequence motifs involved in transmembrane helix packing [21,22]. A second impetus is that the design of

Received: 4 July 2019  
Revised: 5 August 2019  
Accepted: 12 August 2019

Version of Record published:  
11 October 2019

membrane proteins has been less well-explored compared with their water-soluble counterparts and so there remains much to learn. One aspect of this is that membrane proteins have a notorious reputation for being difficult to work with. Low expression levels, misfolding and non-specific aggregation can often frustrate efforts to apply biophysical methods to membrane proteins. This can make it hard to judge the success of the design process. However, both the knowledge base and practical toolkit around expressing, purifying and reconstituting membrane proteins continue to rapidly mature, which now substantially improves the chances of overcoming these barriers for *de novo* membrane proteins [23–29].

Given that there seem to be good reasons to engage in membrane protein design, how should one proceed? A number of different approaches are possible, and the field embraces a swathe of creative computational and knowledge-based strategies that account for some notable recent successes [30–33]. Here, we will discuss ascetic minimalism as one particular tactic for the design of *de novo* integral membrane proteins. We suggest that the stage is set for the development of genetically encoded, low-complexity proteins that can provide new insights into membrane biogenesis and can be harnessed to introduce novel functions into living cells.

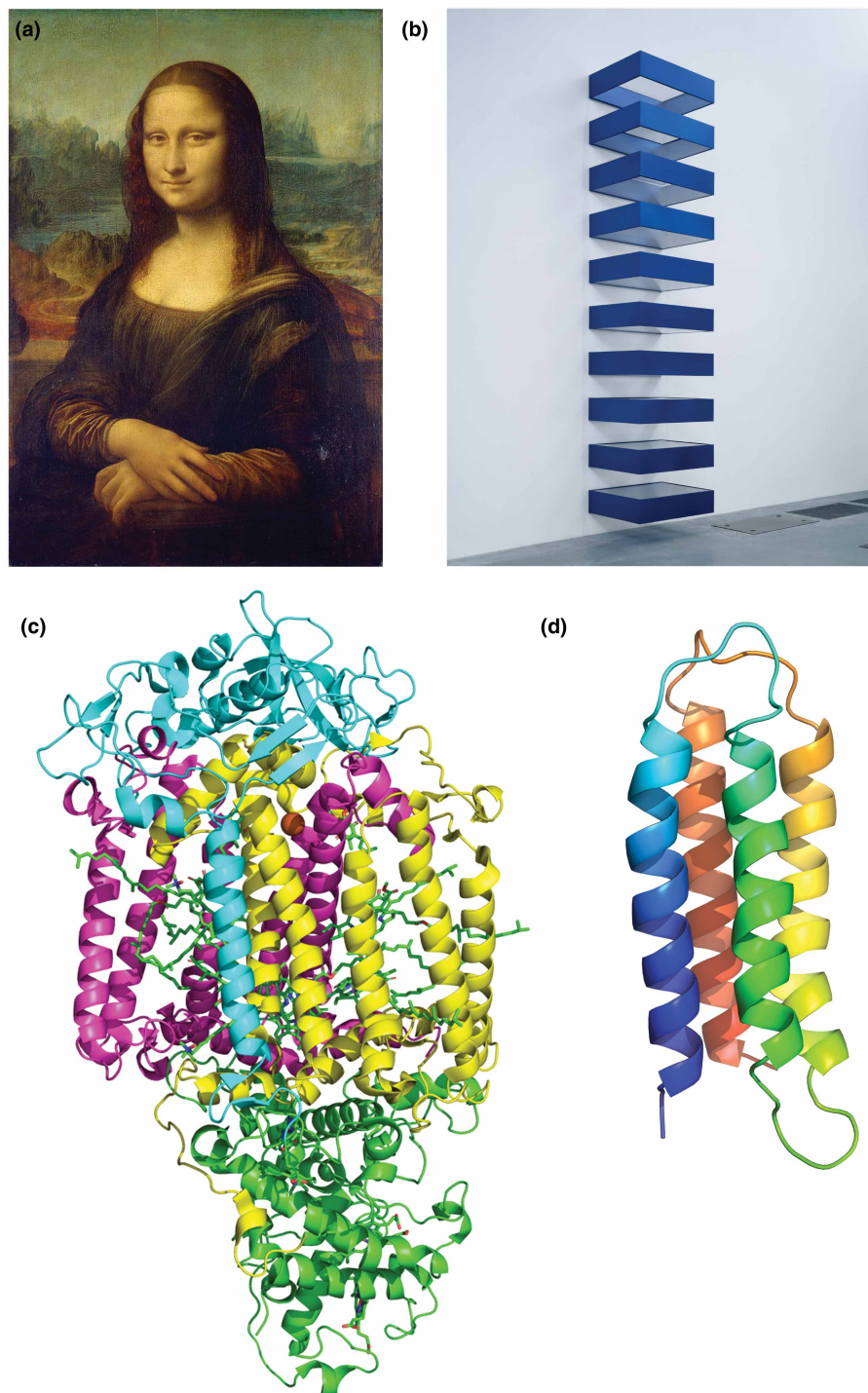
## A manifesto for minimal membrane proteins

The concept of minimalism emerged from the art world in the 1960s and has since been adopted by theatre, choreography, music, architecture, and other spheres. Although minimalism invokes simplicity and reductionism, it is more than just a paring-back of superfluous elements [34]. The original minimalist artworks are three-dimensional objects that lack any ornament or obvious makers' craft. They are modular and repetitive, and the methods and materials of construction are obvious. They are geometric and deliberately eschew any reference to the organic or natural. Fundamental to minimalist art is that the work contains little illusion or allusion, ensuring that the meaning of the piece is constructed by the viewer rather than the originator (Figure 1). By analogy, minimalist *de novo* membrane proteins should have little or no sequence similarity to natural proteins. They should be of the lowest-possible sequence complexity, using the fewest available amino acids to achieve their form and function, and be modular at the level of either sequence or structure. They ought to stand alone and not be fusions to other, natural, proteins. And they ought to ideally be genetically encoded, so that they might eventually become part of cell physiology and be susceptible to laboratory evolution.

What is the case for such austere minimalism in membrane protein design? There should be some straightforward advantages to diminution and simplification, which have been recognised previously [35,36]. Minimal membrane proteins should allow us to identify the most basic sequence requirements for particular behaviours and functions. Such elementary proteins could be more predictable, tractable, pliable, adaptable and robust than many natural membrane proteins. When recombinantly expressed one might expect minimalist proteins to impose a relatively low demand for cellular resources and to be highly orthogonal to the host cell, minimising the metabolic load on the host and reducing the potential for unwanted interference with natural pathways. The lower the complexity of the protein sequence, the easier it should be to understand the impact of individual mutations on the protein as a whole. In many ways, this seems like a natural progression of the widespread and influential use of minimal transmembrane segments to probe multiple aspects of membrane protein biophysics [37–39]. A full consideration of this extensive body of work lies outside the scope of the current short review, but examples include the study of protein–lipid interactions [40,41], protein–protein interactions [42–45] and the biogenesis of transmembrane helices [46,47].

Minimal proteins could provide other benefits to the field of synthetic biology. The unifying vision of synthetic biology is the *artificial cell*; a man-made complex system that diverges to a greater or lesser degree from extant modern life [48–50]. Realising this ambitious goal in its fullest sense requires not just an artificial *genome*, but an artificial *proteome*. Minimal *de novo* proteins could provide modular 'building blocks' for assembling such a proteome, and will be useful in allowing the size of the synthetic genome to be as small as possible. But there are also implications of minimalism that go beyond this.

One important part of synthetic biology is abstraction. With regards to proteins, this requires that they are seen less as the compelling products of natural history and more as engineering components: durable, reliable, consistent and interchangeable parts to be adopted or discarded as the situation demands. But natural proteins come with an origin story — a narrative constructed around their discovery and study. We ought to recognise that these narratives and contexts make it difficult to really see proteins in the same light as a house brick or hose-clip, however much we might think otherwise. For example, one of the cornerstones of synthetic biology research is the green fluorescent protein (GFP). Scientists and schoolchildren alike cherish the story of



**Figure 1. The concept of minimalism in protein design.**

(a and b) Minimalist art rejects traditional representations of the natural world, such as landscapes and portraits, in favour of abstract, serial, and repetitive 3D forms. (c and d) By analogy, minimalist membrane proteins do not attempt to recreate the complex sequences and structures of natural proteins but instead provide modular biocompatible objects with some specific properties. (a) *La Gioconda* ('*The Mona Lisa*') Leonardo Da Vinci (1452–1519) from [https://en.wikipedia.org/wiki/Mona\\_Lisa](https://en.wikipedia.org/wiki/Mona_Lisa). (b) *Untitled*, 1990 Donald Judd (1928–1994) © Judd Foundation/ARS, NY and DACS, London 2019. From the Tate Images Collection © Tate, London 2019. (c) Photosynthetic reaction centre from *R. viridis*, PDB ID 1PRC [102]. (d) Molecular model of a minimal *de novo* membrane protein (not to scale with (c)) adapted from ref. [82] (<https://creativecommons.org/licenses/by/4.0/>).

GFP; how a seemingly unpromising source, the jellyfish, gave rise to a revolutionary technology that has transformed the molecular biosciences. This means that like many natural proteins GFP remains indelibly associated with certain people, places, and times. Truly artificial proteins circumvent this issue since they have no such narrative constructed around them, but instead are deliberately utilitarian. It seems reasonable to suggest that the simpler such *de novo* sequences are, and the less they echo their natural counterparts, the easier it will be to think of them in the abstract. In this spirit, minimalism appears to be an important, productive and liberating area for protein design.

## How to realise minimalism in membrane protein design?

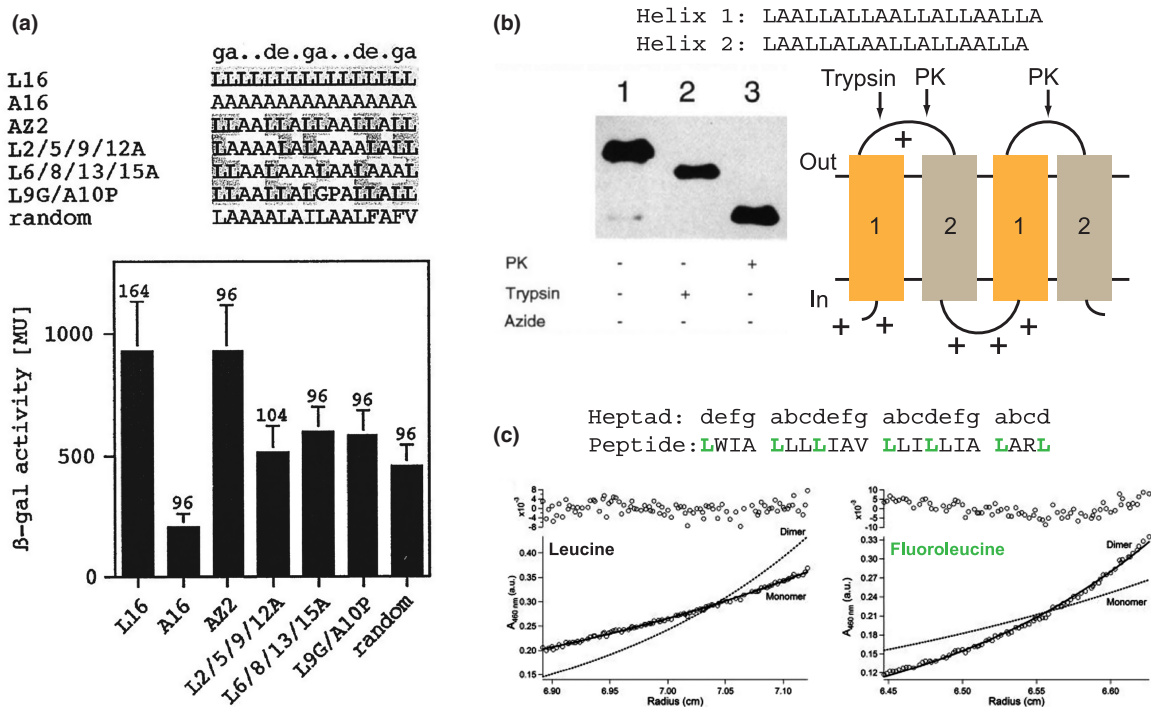
The most obvious targets for the minimal design of membrane proteins are transmembrane  $\alpha$ -helical bundles, and this will be the focus of the discussion below. Nearly all of the natural proteins in cell plasma membranes feature  $\alpha$ -helical secondary structure and these integral membrane proteins are responsible for a wealth of different cellular functions. The remainder of membrane proteins are  $\beta$ -barrels, which occur in the outer membranes of bacteria, mitochondria and chloroplasts and have received less attention in terms of engineering and design [51,52].

### Amino acid sequences

The interior of a lipid bilayer membrane is hydrophobic. It is energetically favourable for hydrophobic amino acid sidechains to partition into this environment, but energetically unfavourable for exposed peptide bonds to do the same. These competing effects are resolved by the spontaneous formation of a transmembrane  $\alpha$ -helix, which exposes the hydrophobic sidechains to the bilayer interior and fully satisfies backbone hydrogen bonding. This means that virtually any sufficiently hydrophobic sequence of  $\sim 18$ – $26$  amino acids can form transmembrane  $\alpha$ -helices that are long enough to span a biological lipid bilayer [4]. This suggests that it will be relatively easy to reduce the complexity of transmembrane segments while retaining secondary structure, and indeed the lipophilic domains of natural helical membrane proteins are already built from a somewhat restricted palette of amino acids. Slightly over half of all the amino acids in biological transmembrane helices are the hydrophobic residues Leu, Ile, Val, Phe and Ala, with Leu being most prevalent [53–55]. Others, especially the charged residues Glu, Arg, Lys and Asp, are relatively rare (although certainly can be found, and Lys and Arg in particular can be accommodated by their ‘snorkelling’ into the lipid headgroup region [56,57]). The primary consideration then is not whether a minimal sequence based on one or more of the common hydrophobic amino acids will form a transmembrane helix, since we may assume that it will do so, but how these helices will interact with the membrane and with each other.

### Interactions of transmembrane helices

The membrane milieu impacts the type of interactions that can drive helix association. In the low dielectric of the membrane interior, sidechain hydrogen bonding can be a means of impelling helix association [43,45]. However, it appears that for many natural membrane proteins a key driving force is Van der Waals interactions. These forces are optimised through shape complementarity and sidechain geometry that allow the close approach of transmembrane helices [58]. In some cases, this also allows for the formation of interhelical C $\alpha$ -H...O hydrogen bonds that likely influence both stability and packing specificity [59,60]. Close packing can be successfully accomplished even with a small subset of hydrophobic amino acids, probably supported by a lipophilic effect that favours protein–protein over protein–lipid interactions. For example, transmembrane poly-leucine helices can self-associate via a ‘knobs-into-holes’ leucine zipper interaction also found in natural proteins (Figure 2a), and this is maintained when Leu is replaced at the helix interface by other hydrophobic amino acids [44,61]. It does seem that natural membrane proteins adopt a relatively limited range of helix–helix packing orientations. About 30% of natural antiparallel helical pairs exploit an Ala-coil motif to pack with a slight left-handed crossing angle [22]. In this motif, small amino acid sidechains (typically Gly, Ala or Ser) occupy positions *a* and *d* or *e* of the helical heptad to allow close packing via sidechain interdigitation. A frequent feature of both parallel and antiparallel right-handed pairwise packing interactions is the well-known small-xxx-small motif, where small residues occur at *i* and *i* + 4 sequence positions [22,55,59,62]. These small residues are often flanked by the  $\beta$ -branched sidechains Val and Ile, and it was suggested that the limited rotational freedom of these sidechains reduces the entropic cost of helix association [55]. These analyses were recently extended from helical pairs to consider helical trimers, which Feng & Barth [21] recognised as the most basic structural unit in natural membrane proteins. Just over half of these helical trimers could be sorted



**Figure 2. Some approaches in the minimalist design of membrane sequences.**

(a) A transcriptional activation assay shows that polyleucine sequences self-assemble *in vivo*. Mutational analysis confirms the importance of leucine at heptad positions *a* and *d*, suggesting a leucine zipper interface. Republished with permission of The American Society for Biochemistry and Molecular Biology from Gurezka et al. [44] © 1999. Permission conveyed through Copyright Clearance Center, Inc. (b) Topological control of a recombinant minimal membrane protein. Two slightly different sequences for minimal TM helices are used to assemble a four-helix construct as shown, and charged residues are introduced at the N-terminus and intracellular loop to control topology according to the ‘positive inside’ rule. Topology within the *E. coli* inner membrane is assessed by proteolysis of the exposed periplasmic loops, which leads to the predictable differences in gel migration shown. PK, Proteinase K. Left panel adapted by permission from Springer Nature. Whitley et al. [81] © 1994. (c) Introducing hexafluoroLeucine at heptad positions *a* and *d* encourages dimerisation of a minimal peptide as determined by equilibrium analytical ultracentrifugation. Adapted with permission from Bilgiçer and Kumar [86]. © 2004 National Academy of Sciences, U.S.A.

into one of six major categories. They identified 13 common sequence motifs that are able to optimise side-chain Van der Waals interactions as well as interhelical backbone–backbone and backbone–sidechain hydrogen bonds. Many of these motifs featured large hydrophobic residues at the helical interfaces. Although such simple sequence motifs are only one aspect of membrane protein folding [63], the examples above certainly describe features of helix packing in natural proteins that can be incorporated into minimal designs.

## Topology and biogenesis

Polytopic (or multipass) membrane proteins contain multiple helices that thread back-and-forth across the membrane and are connected by extramembrane loops. Establishing and maintaining the topology of these helices across the membrane is more complex than might have previously been thought [64–66], but a few general principles can be distilled. In bacterial systems, the major pathway for biogenesis of polytopic proteins involves the first TM helix acting as the membrane localisation sequence, also known as a signal-anchor. This engages with the signal recognition particle to target the nascent polypeptide to the cytoplasmic membrane and supports insertion via the Sec translocon. Protein topogenesis then has a close correlation with the character of the extramembrane loops, with cytoplasmic loops being enriched in positively charged amino acids. This is the ‘positive inside’ rule [67,68], which has been experimentally verified many times and can be exploited to deliberately dictate topology [69,70]. Another notable feature of natural proteins is the statistical preference for

Trp and Tyr residues in the lipid headgroup region, which appear to be important as ‘anchors’ that control helical positioning and dynamics [40,71–74]. In the popular recombinant host *Escherichia coli*, about 80% of all membrane proteins have a topology that results in their C-terminus residing in the cytoplasm, and probably 60% have both their N- and C-termini within the cytoplasm [75]. A similar orientation bias was also found in the model eukaryote *Saccharomyces cerevisiae* [76]. Hence when contemplating the transmembrane topology of multipass *de novo* membrane proteins, locating both termini in the cytoplasm seems like a sensible design choice.

### Polyleucine as a minimal scaffold

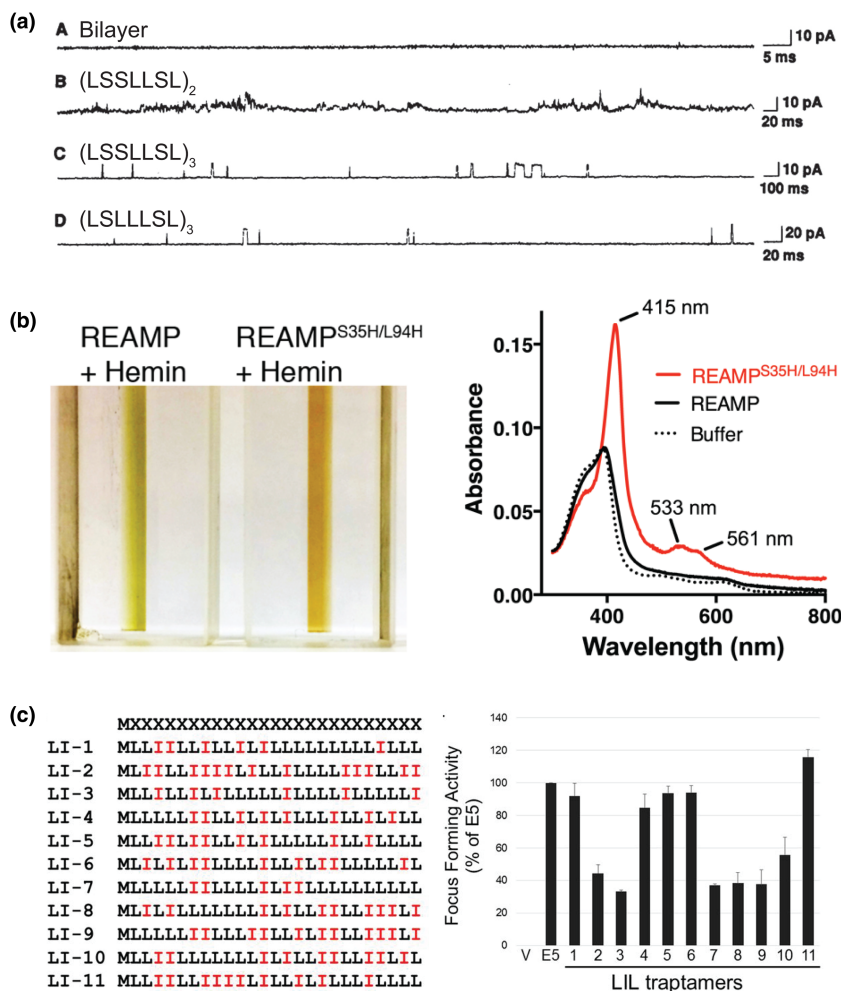
Considering all of the above, polyleucine emerges as an appropriate starting point for the design of minimal  $\alpha$ -helical membrane proteins. Leu has a high helical propensity, is the most common amino acid in natural transmembrane helices, and can engage in both protein–protein and protein–lipid interactions. Polyleucine is already established as a suitable and neutral ‘host’ for known motifs than can drive helical association. PolyLeu helices can be tolerated by biological systems and the topology of these helices can be controlled by the charge distribution in the flanking extramembrane domains.

### Minimal design: from concept to reality

The ideas outlined above can now be illustrated with some key examples of success in producing minimal membrane proteins. The first example seems to be from Goodall & Urry, who showed in 1973 that synthetic peptides built around an AAG repeat motif (e.g. (AAG)<sub>4</sub>) could form conductive ion channels across artificial membranes [77]. Subsequently, Kennedy et al. [78] extended this observation to polypeptides of various lengths containing the tetrad repeat motif LSLG. Peptides (LSLG)<sub>12</sub> or an N-formyl derivative of (LSLG)<sub>6</sub> also formed conductive channels that were proposed to arise from individual  $\beta$ -helices similar to gramicidin. However, a detailed molecular understanding was obviously difficult at that time [79]. A major leap forward was taken in a landmark paper from DeGrado and colleagues in 1988 [35,80]. This described the rational design of minimal peptides composed only of Ser and Leu that could recapitulate the dimensions and properties of individual  $\alpha$ -helices found within natural ion channels. Two heptad repeat motifs were chosen for this work. The first of these motifs, LSSLLSL, was repeated in triplicate to give peptide (LSSLLSL)<sub>3</sub>. The expectation was that these peptides would self-assemble, with the Ser residues generating a polar helical face that would be sequestered from the bilayer interior to form an aqueous pore. The (LSSLLSL)<sub>3</sub> peptide formed relatively permissive cation channels in synthetic lipid bilayers, with a pore size of  $\sim 8$  Å and conductance properties reminiscent of natural channel proteins (Figure 3a). Substituting Leu for Ser in the repeat motif to give peptide (LSLLLSL)<sub>3</sub> produced channels with markedly different properties. These were apparently highly selective with much smaller pore size, being only permeable to protons. Computational models suggested that both peptides could assemble into parallel helical bundles with tightly packed helical interfaces driven by interdigitation of the Leu sidechains. However, (LSSLLSL)<sub>3</sub> was likely to be a hexamer and (LSLLLSL)<sub>3</sub> probably a trimer or tetramer. Both of these different tertiary structures allowed the Ser residues to be accommodated in the polar channel interior.

The minimalist approach was also exemplified by work from von Heijne and colleagues [81]. The sequence LAALLALLAALLALLAALLA was designed with the expectation that this would form ‘Janus’-like  $\alpha$ -helices, with one face of the helix being Leu and the other face being Ala. Constructs comprising one, two or four successive TM helices were successfully biosynthesised from a recombinant plasmid (Figure 2b). These *de novo* proteins were localised to the *E. coli* membrane and the transmembrane topology was controlled by adjusting the charge in the interconnecting loops. These highly orthogonal constructs were easily accommodated by the cell — for example their expression did not inhibit cell growth.

Our own work [82] has built upon this by expressing a different minimal sequence into cellular membranes and then exploring the purification and characterisation of this *de novo* protein. Our design is an antiparallel four-helix bundle in which helices of sequence LLLSGLGLLLLSLGLLLLS are connected by short loops to give a N<sub>in</sub>/C<sub>in</sub> topology. This bioinspired design was an abstracted version of the consensus sequence of the SMR group of small transporters. It could be expressed with the expected topology in the *E. coli* membrane and was surprisingly tractable to further study. The synthetic protein was purified in the gentle maltoside detergents that are widely used in membrane biochemistry and shown to be stable and monodisperse *in vitro*. The sequence was only designed for expression and did not specify any tertiary packing interactions, and so unsurprisingly appeared to be a dynamic molten-globule [82]. Goparaju and colleagues also showed that a minimal



**Figure 3. Examples of functionality in minimal membrane proteins.**

(a) Electrophysiology of synthetic lipid bilayers containing minimal ion channels. Panel A shows a negative control without any peptide; Panel B is control using a short peptide that cannot span the bilayer; Panel C is peptide with sequence (LSSLLSL)<sub>3</sub>; and Panel D is (LSLLLSL)<sub>3</sub>. From Lear et al. [80]. Reprinted with permission from AAAS, adapted with permission of the authors. (b) A minimal 4-helix bundle ('REAMP'), in which each helix has the sequence LLLSGLGLLLLSLLGLLLLS, can be induced to bind a haem cofactor via a bis-Histidine site. The reddish-brown colour and the spectroscopic fingerprint of the oxidised hemoprotein are both reminiscent of natural cytochromes. Adapted from ref. [82] <https://creativecommons.org/licenses/by/4.0/>. (c) A series of minimal single TM helices — termed 'LIL proteins' — interact with the PDGF-β receptor and transform mammalian cells. The activity of each sequence LI1–11 is compared with a natural activator, the small oncoprotein E5. Adapted with permission from Heim et al. [99].

four-helix transmembrane bundle, with each helix built around subtle variations of polyLeu/Ala, could be recovered in detergents for *in vitro* analysis after being deliberately expressed into cellular inclusion bodies [83].

How easy is it to obtain well-defined structures from minimal sequences? A recent study demonstrated that Van der Waals packing interactions are sufficient to strongly constrain the structure of minimal membrane proteins, although the packing requirements are rather precise [33]. Inspired by the repeating LxxLxxx heptad packing motif (register *abcdefg*) observed in phospholamban, the authors computationally designed and chemically synthesised transmembrane peptides that assembled into tightly packed homopentamers. This was driven entirely by pairwise nonpolar packing interactions between neighbouring helices. These interactions could be very sensitive to mutation. For example, even a conservative Leu to Ile mutation at one *g* position was sufficient to disrupt packing and abolish oligomerisation. This provides substantial encouragement that low-complexity

sequences could pack well enough to form well-defined structures, but that this will require judicious placement of residues at key positions.

An interesting alternative is to exploit the properties of non-natural amino acids. With regards to membrane proteins, fluorinated amino acid sidechains should favour protein–protein interactions over protein–lipid interactions [84,85]. This idea was validated by a study showing that low-complexity hydrophobic peptides containing hexafluoroleucine at the *a* and *d* heptad positions could assemble into either stable helical dimers or probable tetramers (Figure 2c), with the oligomerisation state depending on the presence of a single Asn in the membrane core [86,87].

Overall then, two key themes emerge. The first is that minimal membrane proteins are biosynthetically accessible — that they can be produced in cell systems and targeted to cellular membranes. The second is that minimal sequences could potentially form robust and well-defined structures by exploiting or mimicking the helical packing ‘rules’ observed in natural proteins.

## Functional minimal membrane proteins

Once minimal sequences are identified, how can these be turned to some useful function? As discussed above, some minimal proteins can clearly form channels and pores (Figure 3a). In the rest of this review, we look at other means of functionalising the inert scaffold of a minimal membrane protein.

### Cofactor binding

The binding of porphyrins and related tetrapyrroles is an obvious target for introducing function into minimal membrane proteins. This is because these compounds are hydrophobic, so will spontaneously partition into lipid membranes and are good interaction partners for transmembrane segments. Natural porphoproteins are widespread and very well-studied, and the principles governing their structure and function are established. Metallated porphoproteins, particularly those containing haem, have a number of potentially interesting applications in electron transfer, gas binding and redox (bio)chemistry. Those containing chlorophylls can be photo-excited. Cofactor binding can rigidify otherwise dynamic *de novo* proteins [88,89] and indeed some structural flexibility in the apoprotein is probably desirable to support the interaction [90]. To this end, most studies with redesigned or *de novo* proteins have focussed upon the binding of haem. A popular strategy is to co-ordinate the haem iron with the imidazole sidechain of histidine, a binding mode that predominates in natural hemoproteins.

Glycophorin A is a natural single-pass membrane protein in which the membrane domain forms a homodimer. This was used as the basis for a *de novo* helical dimer that could bind haem via bis-Histidine ligation with apparent  $K_d$  in the low  $\mu\text{M}$  range [91]. The subsequent hemoprotein was catalytically active as a peroxidase, with a redox potential of  $-128$  mV. A single amino acid substitution was sufficient to both improve the binding affinity and adjust the haem redox potential by  $-44$  mV [92]. Similar results emerged from the redesign of a natural helical peptide to adopt the Ala-coil motif [93]. This work generated a tetrameric dihaem membrane protein, which is of interest since two haems are needed to transfer electrons across a lipid bilayer. Korendovych and colleagues used computational design to develop a *de novo* four-helix bundle that also exploited the Ala-coil motif, and showed that this chemically synthesised peptide could co-ordinate two iron diphenylporphyrins via bis-His geometry in both detergent micelles and artificial membranes [94]. The redox potential of the two bound porphyrins was separated by 71 mV, being  $-97$  and  $-168$  mV. This splitting of the redox potential is typical for porphyrins that are in close proximity and could potentially facilitate biological electron transfer. The membrane segments of designs that blended water-soluble and hydrophobic regions were also found to bind haem and bacteriochlorophyll [95,96]. With regards to minimal proteins, Goparaju et al. [83] showed that four-helix bundles built from low-complexity Leu/Ala sequences could bind multiple haems as well as the photoactive complex zinc protoporphyrin IX via Histidine ligation. Bound haems had distinct redox potentials, and simultaneously binding both Fe- and Zn-substituted porphyrins allowed electron transfer between the two cofactors. Our own work has also shown that genetically encoded minimal proteins could be induced to bind a single haem *in vitro* (Figure 3b) with little modification other than introducing coordinating histidines [82]. These proteins had a redox potential of  $-101$  mV, a shift in the haem potential of  $+32$  mV from a hydrophobic non-protein environment, and nascent peroxidase activity that demonstrated the potential for developing these sequences for redox biochemistry and electron transfer.

It will be fascinating now to try and translate this understanding into living systems to influence cell physiology. This will require cofactor binding *in vivo* as well as the deliberate ‘tuning’ of redox potentials for specific



applications. To our knowledge cofactor binding within a cellular membrane has not yet been achieved for *de novo* integral membrane proteins, although there is precedent for this from some soluble designs [97]. The biochemical pathways underlying cofactor synthesis are well-characterised, and so metabolic engineering to encourage such cofactor loading is plausible.

## Protein–protein interactions

Another potential application is to use minimal proteins to disrupt natural protein–protein contacts within the membrane, which are essential for the assembly of functional membrane complexes. This has already been achieved with more complex *de novo* peptides [98]. A recent report showed that small proteins comprising a single transmembrane domain consisting of only Leu and Ile (termed LIL proteins; Figure 3c) could be recombinantly expressed in mammalian cell lines [99]. Through library screening, specific sequence variants were identified that could influence cell biology by interacting with, and so activating, the TM domain of platelet-derived growth factor  $\beta$  receptor (PDGF $\beta$ R) to initiate a signalling cascade. This remarkable result suggests that even proteins with elementary amino acid sequences can specifically interact with natural proteins with functional consequences.

## Conclusions

The examples above provide substantial encouragement for the membrane protein designer. Sequences with minimal chemical diversity are chemically and biosynthetically accessible and introducing activity, for example by cofactor binding, appears to be achievable. Perhaps the major challenge with minimal proteins (just as for other *de novo* proteins) lies in understanding the particular sequence features required to arrive at a compact folded state. As put by Richardson: ‘the hardest part of protein folding, or protein design, is the last little bit’ [100]. Transporters, receptors, ligand-gated channels and others exert their function through specific conformational changes; so defining explicit conformations, and the means to transition between them, will be important in any effort to recreate these functions [31]. Well-structured *de novo* membrane proteins — albeit with greater sequence complexity — have already been achieved by exploiting the helix packing properties apparent in natural proteins [30–33] and this will be supported by the continuing development of bioinformatic and computational methods [15,101]. Minimalism is a logical extension of prior work on model membrane systems that develops our essential understanding of membrane biology and could generate novel synthetic components. This makes the minimalist approach an interesting and useful stitch in the broader tapestry of protein design.

## Perspectives

- Integral membrane proteins are very important in biological systems, but are underrepresented in *de novo* design. A design strategy that embraces strict minimalism will offer fundamental insights and could be useful for synthetic biology applications.
- Much of our current knowledge comes from studies of chemically-synthesised peptides in model bilayers. The use of naturally occurring sequence motifs to control helical packing has been a productive approach.
- Future directions should include biosynthesis and functionalisation *in vivo*, in order to integrate artificial proteins into living systems.

## Abbreviations

GFP, green fluorescent protein; PDGF $\beta$ R, platelet-derived growth factor  $\beta$  receptor.

## Acknowledgements

We thank Ross Anderson for helpful discussions. There are no competing financial interests. BrisSynBio is a BBSRC/EPSRC Synthetic Biology Research Centre (BB/L01386X/1).

## Competing Interests

The Author declares that there are no competing interests associated with this manuscript.

## References

- Wallin, E. and von Heijne, G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* **7**, 1029–1038 <https://doi.org/10.1002/pro.5560070420>
- Cymer, F., von Heijne, G. and White, S.H. (2015) Mechanisms of integral membrane protein insertion and folding. *J. Mol. Biol.* **427**, 999–1022 <https://doi.org/10.1016/j.jmb.2014.09.014>
- Harris, N.J., Charalambous, K., Findlay, H.E. and Booth, P.J. (2018) Lipids modulate the insertion and folding of the nascent chains of  $\alpha$  helical membrane proteins. *Biochem. Soc. Trans.* **46**, 1355–1366 <https://doi.org/10.1042/BST20170424>
- McKay, M.J., Afrose, F., Koeppe, R.E. and Greathouse, D.V. (2018) Helix formation and stability in membranes. *Biochim. Biophys. Acta* **1860**, 2108–2117 <https://doi.org/10.1016/j.bbamem.2018.02.010>
- Chaturvedi, D. and Mahalakshmi, R. (2017) Transmembrane  $\beta$ -barrels: evolution, folding and energetics. *Biochim. Biophys. Acta* **1859**, 2467–2482 <https://doi.org/10.1016/j.bbamem.2017.09.020>
- Konavalova, A., Kahne, D.E. and Silhavy, T.J. (2017) Outer membrane biogenesis. *Annu. Rev. Microbiol.* **71**, 539–556 <https://doi.org/10.1146/annurev-micro-090816-093754>
- Popot, J.-L. and Engelman, D.M. (2016) Membranes do not tell proteins how to fold. *Biochemistry* **55**, 5–18 <https://doi.org/10.1021/acs.biochem.5b01134>
- Thomik, T., Wittig, I., Choe, J.-Y., Boles, E. and Oreb, M. (2017) An artificial transport metabolon facilitates improved substrate utilization in yeast. *Nat. Chem. Biol.* **13**, 1158–1163 <https://doi.org/10.1038/nchembio.2457>
- Bayoumi, M., Bayley, H., Maglia, G. and Sapra, K.T. (2017) Multi-compartment encapsulation of communicating droplets and droplet networks in hydrogel as a model for artificial cells. *Sci. Rep.* **7**, 45167 <https://doi.org/10.1038/srep45167>
- Booth, M.J., Restrepo Schild, V., Graham, A.D., Olof, S.N. and Bayley, H. (2016) Light-activated communication in synthetic tissues. *Sci. Adv.* **2**, e1600056 <https://doi.org/10.1126/sciadv.1600056>
- Leister, D. (2019) Genetic engineering, synthetic biology and the light reactions of photosynthesis. *Plant Physiol.* **179**, 778–793 <https://doi.org/10.1104/pp.18.00360>
- Woolfson, D.N., Bartlett, G.J., Burton, A.J., Heal, J.W., Niitsu, A., Thomson, A.R. et al. (2015) De novo protein design: how do we expand into the universe of possible protein structures? *Curr. Opin. Struct. Biol.* **33**, 16–26 <https://doi.org/10.1016/j.sbi.2015.05.009>
- Huang, P.-S., Boyken, S.E. and Baker, D. (2016) The coming of age of de novo protein design. *Nature* **537**, 320–327 <https://doi.org/10.1038/nature19946>
- Brisendine, J.M. and Koder, R.L. (2016) Fast, cheap and out of control - insights into thermodynamic and informatic constraints on natural protein sequences from de novo protein design. *Biochim. Biophys. Acta* **1857**, 485–492 <https://doi.org/10.1016/j.bbabi.2015.10.002>
- Barth, P. and Senes, A. (2016) Towards high-resolution computational design of the structure and function of helical membrane proteins. *Nat. Struct. Mol. Biol.* **23**, 475–480 <https://doi.org/10.1038/nsmb.3231>
- Watkins, D.W., Armstrong, C.T. and Anderson, J.L.R. (2014) De novo protein components for oxidoreductase assembly and biological integration. *Curr. Opin. Chem. Biol.* **19**, 90–98 <https://doi.org/10.1016/j.cbpa.2014.01.016>
- Baker, D. (2019) What has de novo design taught us about protein folding and biophysics? *Protein Sci.* **28**, 678–683 <https://doi.org/10.1002/pro.3588>
- Nastri, F., Chino, M., Maglio, O., Bhagi-Damodaran, A., Lu, Y. and Lombardi, A. (2016) Design and engineering of artificial oxygen-activating metalloenzymes. *Chem. Soc. Rev.* **45**, 5020–5054 <https://doi.org/10.1039/C5CS00923E>
- Hecht, M.H., Zarchitsky, S., Karas, C. and Chari, S. (2018) Are natural proteins special? Can we do that? *Curr. Opin. Struct. Biol.* **48**, 124–132 <https://doi.org/10.1016/j.sbi.2017.11.009>
- Membrane proteins of known structure. Available from: <https://blanco.biomol.uci.edu/mpstruc/>
- Feng, X. and Barth, P. (2016) A topological and conformational stability alphabet for multipass membrane proteins. *Nat. Chem. Biol.* **12**, 167–173 <https://doi.org/10.1038/nchembio.2001>
- Walters, R. and DeGrado, W.F. (2006) Helix-packing motifs in membrane proteins. *Proc. Natl Acad. Sci. U.S.A.* **103**, 13658–13663 <https://doi.org/10.1073/pnas.0605878103>
- Stroud, Z., Hall, S. and Dafforn, T.R. (2018) Purification of membrane proteins free from conventional detergents: SMA, new polymers, new opportunities and new insights. *Methods* **147**, 106–117 <https://doi.org/10.1016/j.ymeth.2018.03.011>
- Skrzypek, R., Iqbal, S. and Callaghan, R. (2018) Methods of reconstitution to investigate membrane protein function. *Methods* **147**, 126–141 <https://doi.org/10.1016/j.ymeth.2018.02.012>
- Dilworth, M.V., Piel, M., Bettaney, K.E., Ma, P., Luo, J., Sharples, D. et al. (2018) Microbial expression systems for membrane proteins. *Methods* **147**, 3–39 <https://doi.org/10.1016/j.ymeth.2018.04.009>
- Le Bon, C., Marconnet, A., Masscheleyn, S., Popot, J.-L. and Zoonens, M. (2018) Folding and stabilising membrane proteins in amphipol A8-35. *Methods* **147**, 95–105 <https://doi.org/10.1016/j.ymeth.2018.04.012>
- Lyons, J.A., Shahsavari, A., Paulsen, P.A., Pedersen, B.P. and Nissen, P. (2016) Expression strategies for structural studies of eukaryotic membrane proteins. *Curr. Opin. Struct. Biol.* **38**, 137–144 <https://doi.org/10.1016/j.sbi.2016.06.011>
- Denisov, I.G. and Sligar, S.G. (2016) Nanodiscs for structural and functional studies of membrane proteins. *Nat. Struct. Mol. Biol.* **23**, 481–486 <https://doi.org/10.1038/nsmb.3195>
- Moraes, I., Evans, G., Sanchez-Weatherby, J., Newstead, S. and Shaw Stewart, P.D. (2014) Membrane protein structure determination - the next generation. *Biochim. Biophys. Acta* **1838**, 78–87 <https://doi.org/10.1016/j.bbamem.2013.07.010>
- Lu, P., Min, D., DiMaio, F., Wei, K.Y., Vahey, M.D., Boyken, S.E. et al. (2018) Accurate computational design of multipass transmembrane proteins. *Science* **359**, 1042–1046 <https://doi.org/10.1126/science.aag1739>
- Joh, N.H., Wang, T., Bhate, M.P., Acharya, R., Wu, Y., Grabe, M. et al. (2014) De novo design of a transmembrane Zn<sup>2+</sup>-transporting four-helix bundle. *Science* **346**, 1520–1524 <https://doi.org/10.1126/science.1261172>

- 32 Mahendran, K.R., Niitsu, A., Kong, L., Thomson, A.R., Sessions, R.B., Woolfson, D.N. et al. (2017) A monodisperse transmembrane  $\alpha$ -helical peptide barrel. *Nat. Chem.* **9**, 411–419 <https://doi.org/10.1038/nchem.2647>
- 33 Mravic, M., Thomaston, J.L., Tucker, M., Solomon, P.E., Liu, L. and DeGrado, W.F. (2019) Packing of apolar side chains enables accurate design of highly stable membrane proteins. *Science* **363**, 1418–1423 <https://doi.org/10.1126/science.aav7541>
- 34 Meyer, J. (2010) *Minimalism*, Phaidon Press, London, U.K.
- 35 DeGrado, W.F., Wasserman, Z.R. and Lear, J.D. (1989) Protein design, a minimalist approach. *Science* **243**, 622–628 <https://doi.org/10.1126/science.2464850>
- 36 Robertson, D.E., Farid, R.S., Moser, C.C., Urbauer, J.L., Mulholland, S.E., Pidikiti, R. et al. (1994) Design and synthesis of multi-haem proteins. *Nature* **368**, 425–432 <https://doi.org/10.1038/368425a0>
- 37 Killian, J.A. (2003) Synthetic peptides as models for intrinsic membrane proteins. *FEBS Lett.* **555**, 134–138 [https://doi.org/10.1016/S0014-5793\(03\)01154-2](https://doi.org/10.1016/S0014-5793(03)01154-2)
- 38 Killian, J.A. and Nyholm, T.K. (2006) Peptides in lipid bilayers: the power of simple models. *Curr. Opin. Struct. Biol.* **16**, 473–479 <https://doi.org/10.1016/j.sbi.2006.06.007>
- 39 London, E. (2007) Using model membrane-inserted hydrophobic helices to study the equilibrium between transmembrane and nontransmembrane states. *J. Gen. Physiol.* **130**, 229–232 <https://doi.org/10.1085/jgp.200709842>
- 40 de Planque, M.R.R., Kruijzer, J.A.W., Liskamp, R.M.J., Marsh, D., Greathouse, D.V., Koeppe, R.E. et al. (1999) Different membrane anchoring positions of tryptophan and lysine in synthetic transmembrane  $\alpha$ -helical peptides. *J. Biol. Chem.* **274**, 20839–20846 <https://doi.org/10.1074/jbc.274.30.20839>
- 41 Weiss, T.M., van der Wel, P.C.A., Killian, J.A., Koeppe, R.E. and Huang, H.W. (2003) Hydrophobic mismatch between helices and lipid bilayers. *Biophys. J.* **84**, 379–385 [https://doi.org/10.1016/S0006-3495\(03\)74858-9](https://doi.org/10.1016/S0006-3495(03)74858-9)
- 42 Ruan, W., Linder, E. and Langosch, D. (2004) The interface of a membrane-spanning leucine zipper mapped by asparagine-scanning mutagenesis. *Protein Sci.* **13**, 555–559 <https://doi.org/10.1110/ps.03357404>
- 43 Zhou, F.X., Cocco, M.J., Russ, W.P., Brunger, A.T. and Engelman, D.M. (2000) Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat. Struct. Biol.* **7**, 154–160 <https://doi.org/10.1038/72430>
- 44 Gurezka, R., Laage, R., Brosig, B. and Langosch, D. (1999) A heptad motif of leucine residues found in membrane proteins can drive self-assembly of artificial transmembrane segments. *J. Biol. Chem.* **274**, 9265–9270 <https://doi.org/10.1074/jbc.274.14.9265>
- 45 Zhou, F.X., Merianos, H.J., Brunger, A.T. and Engelman, D.M. (2001) Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl Acad. Sci. U.S.A.* **98**, 2250–2255 <https://doi.org/10.1073/pnas.041593698>
- 46 Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H. et al. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* **433**, 377–381 <https://doi.org/10.1038/nature03216>
- 47 Hessa, T., Meindl-Beinker, N.M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M. et al. (2007) Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature* **450**, 1026–1030 <https://doi.org/10.1038/nature06387>
- 48 Powell, K. (2018) How biologists are creating life-like cells from scratch. *Nature* **563**, 172–175 <https://doi.org/10.1038/d41586-018-07289-x>
- 49 Göpfrich, K., Platzman, I. and Spatz, J.P. (2018) Mastering complexity: towards bottom-up construction of multifunctional eukaryotic synthetic cells. *Trends Biotechnol.* **36**, 938–951 <https://doi.org/10.1016/j.tibtech.2018.03.008>
- 50 Forlin, M., Lentini, R. and Mansy, S.S. (2012) Cellular imitations. *Curr. Opin. Chem. Biol.* **16**, 586–592 <https://doi.org/10.1016/j.cbpa.2012.10.020>
- 51 Slusky, J.S.G. (2017) Outer membrane protein design. *Curr. Opin. Struct. Biol.* **45**, 45–52 <https://doi.org/10.1016/j.sbi.2016.11.003>
- 52 Stapleton, J.A., Whitehead, T.A. and Nanda, V. (2015) Computational redesign of the lipid-facing surface of the outer membrane protein OmpA. *Proc. Natl Acad. Sci. U.S.A.* **112**, 9632–9637 <https://doi.org/10.1073/pnas.1501836112>
- 53 Baeza-Delgado, C., Marti-Renom, M.A. and Mingarro, I. (2013) Structure-based statistical analysis of transmembrane helices. *Eur. Biophys. J.* **42**, 199–207 <https://doi.org/10.1007/s00249-012-0813-9>
- 54 Saidijam, M., Azizpour, S. and Patching, S.G. (2018) Comprehensive analysis of the numbers, lengths and amino acid compositions of transmembrane helices in prokaryotic, eukaryotic and viral integral membrane proteins of high-resolution structure. *J. Biomol. Struct. Dyn.* **36**, 443–464 <https://doi.org/10.1080/07391102.2017.1285725>
- 55 Senes, A., Gerstein, M. and Engelman, D.M. (2000) Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighbouring positions. *J. Mol. Biol.* **296**, 921–936 <https://doi.org/10.1006/jmbi.1999.3488>
- 56 Öjemalm, K., Higuchi, T., Lara, P., Lindahl, E., Suga, H. and von Heijne, G. (2016) Energetics of side-chain snorkeling in transmembrane helices probed by nonproteinogenic amino acids. *Proc. Natl Acad. Sci. U.S.A.* **113**, 10559–10564 <https://doi.org/10.1073/pnas.1606776113>
- 57 Strandberg, E. and Killian, J.A. (2003) Snorkeling of lysine side chains in transmembrane helices: how easy can it get? *FEBS Lett.* **544**, 69–73 [https://doi.org/10.1016/S0014-5793\(03\)00475-7](https://doi.org/10.1016/S0014-5793(03)00475-7)
- 58 Eilers, M., Shekar, S.C., Shieh, T., Smith, S.O. and Fleming, P.J. (2000) Internal packing of helical membrane proteins. *Proc. Natl Acad. Sci. U.S.A.* **97**, 5796–5801 <https://doi.org/10.1073/pnas.97.11.5796>
- 59 Mueller, B.K., Subramaniam, S. and Senes, A. (2014) A frequent, GxxxG-mediated, transmembrane association motif is optimized for the formation of interhelical C $\alpha$ -H hydrogen bonds. *Proc. Natl Acad. Sci. U.S.A.* **111**, E888–E895 <https://doi.org/10.1073/pnas.1319944111>
- 60 Senes, A., Ubarretxena-Belandia, I. and Engelman, D.M. (2001) The C $\alpha$ -H $\cdots$ O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl Acad. Sci. U.S.A.* **98**, 9056–9061 <https://doi.org/10.1073/pnas.161280798>
- 61 Gurezka, R. and Langosch, D. (2001) In vitro selection of membrane-spanning leucine zipper protein-protein interaction motifs using POSSYCCAT. *J. Biol. Chem.* **276**, 45580–45587 <https://doi.org/10.1074/jbc.M105362200>
- 62 Russ, W.P. and Engelman, D.M. (2000) The GxxxG motif: a framework for transmembrane helix-helix association. *J. Mol. Biol.* **296**, 911–919 <https://doi.org/10.1006/jmbi.1999.3489>
- 63 Li, E., Wimley, W.C. and Hristova, K. (2012) Transmembrane helix dimerisation: beyond the search for sequence motifs. *Biochim. Biophys. Acta* **1818**, 183–193 <https://doi.org/10.1016/j.bbamem.2011.08.031>
- 64 Xie, K. and Dalbey, R.E. (2008) Inserting proteins into the bacterial cytoplasmic membrane using the Sec and YidC translocases. *Nat. Rev. Microbiol.* **6**, 234–244 <https://doi.org/10.1038/nrmicro3595>

- 65 von Heijne, G. (2006) Membrane protein topology. *Nat. Rev. Mol. Cell Biol.* **7**, 909–918 <https://doi.org/10.1038/nrm2063>
- 66 Bogdanov, M., Dowhan, W. and Vitrac, H. (2014) Lipids and topological rules governing membrane protein assembly. *Biochim. Biophys. Acta* **1843**, 1475–1488 <https://doi.org/10.1016/j.bbamcr.2013.12.007>
- 67 von Heijne, G. (1986) The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J.* **5**, 3021–3027 <https://doi.org/10.1002/j.1460-2075.1986.tb04601.x>
- 68 Nilsson, J., Persson, B. A.R. and von Heijne, G. (2005) Comparative analysis of amino acid distributions in integral membrane proteins from 107 genomes. *Proteins* **60**, 606–616 <https://doi.org/10.1002/prot.20583>
- 69 von Heijne, G. (1989) Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature* **341**, 456–458 <https://doi.org/10.1038/341456a0>
- 70 von Heijne, G. and Manoil, C. (1990) Membrane proteins: from sequence to structure. *Prot. Eng.* **4**, 109–112 <https://doi.org/10.1093/protein/4.2.109>
- 71 Sparks, K.A., Gleason, N.J., Gist, R., Langston, R., Greathouse, D.V. and Koeppe, R.E. (2014) Comparisons of interfacial Phe, Tyr, and Trp residues as determinants of orientation and dynamics for GWALP transmembrane peptides. *Biochemistry* **53**, 3637–3645 <https://doi.org/10.1021/bi500439x>
- 72 Yuen, C.T.K., Davidson, A.R. and Deber, C.M. (2000) Role of aromatic residues at the lipid-water interface in micelle-bound bacteriophage M13 major coat protein. *Biochemistry* **39**, 16155–16162 <https://doi.org/10.1021/bi0016117>
- 73 Killian, J.A. and von Heijne, G. (2000) How proteins adapt to a membrane-water interface. *Trends Biochem. Sci.* **25**, 429–434 [https://doi.org/10.1016/S0968-0004\(00\)01626-1](https://doi.org/10.1016/S0968-0004(00)01626-1)
- 74 Norimatsu, Y., Hasegawa, K., Shimizu, N. and Toyoshima, C. (2017) Protein-phospholipid interplay revealed with crystals of a calcium pump. *Nature* **545**, 193–198 <https://doi.org/10.1038/nature22357>
- 75 Daley, D.O., Rapp, M., Granseth, E., Melén, K., Drew, D. and von Heijne, G. (2005) Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* **308**, 1321–1323 <https://doi.org/10.1126/science.1109730>
- 76 Kim, H., Melén, K., Österberg, M. and von Heijne, G. (2006) A global topology map of the *Saccharomyces cerevisiae* membrane proteome. *Proc. Natl Acad. Sci. U.S.A.* **103**, 11142–11147 <https://doi.org/10.1073/pnas.0604075103>
- 77 Goodall, M.C. and Urry, D.W. (1973) A synthetic transmembrane channel. *Biochim. Biophys. Acta* **291**, 317–320 [https://doi.org/10.1016/0005-2736\(73\)90425-2](https://doi.org/10.1016/0005-2736(73)90425-2)
- 78 Kennedy, S.J., Roeske, R.W., Freeman, A.R., Watanabe, A.M. and Besche, Jr, H.R. (1977) Synthetic peptides form ion channels in artificial lipid bilayer membranes. *Science* **196**, 1341–1342 <https://doi.org/10.1126/science.867034>
- 79 Spach, G., Trudelle, Y. and Heitz, F. (1983) Peptides as channel-making ionophores: conformational aspects. *Biopolymers* **22**, 403–407 <https://doi.org/10.1002/bip.360220153>
- 80 Lear, J.D., Wasserman, Z.R. and DeGrado, W.F. (1988) Synthetic amphiphilic peptide models for protein ion channels. *Science* **240**, 1177–1181 <https://doi.org/10.1126/science.2453923>
- 81 Whitley, P., Nilsson, I. and von Heijne, G. (1994) De novo design of integral membrane proteins. *Nat. Struct. Biol.* **1**, 858–862 <https://doi.org/10.1038/nsb1294-858>
- 82 Lalaurie, C.J., Dufour, V., Meletiou, A., Ratcliffe, S., Harland, A., Wilson, O. et al. (2018) The de novo design of a biocompatible and functional integral membrane protein using minimal sequence complexity. *Sci. Rep.* **8**, 14564 <https://doi.org/10.1038/s41598-018-31964-8>
- 83 Goparaju, G., Fry, B.A., Chobot, S.E., Wiedman, G., Moser, C.C., Dutton, P.L. et al. (2016) First principles design of a core bioenergetic transmembrane electron-transfer protein. *Biochim. Biophys. Acta* **1857**, 503–512 <https://doi.org/10.1016/j.bbabi.2015.12.002>
- 84 Yoder, N.C. and Kumar, K. (2002) Fluorinated amino acids in protein design and engineering. *Chem. Soc. Rev.* **31**, 335–341 <https://doi.org/10.1039/b201097f>
- 85 Marsh, E.N.G. (2016) Designing fluorinated proteins. *Methods Enzymol.* **580**, 251–278 <https://doi.org/10.1016/bs.mie.2016.05.006>
- 86 Bilgiçer, B. and Kumar, K. (2004) De novo design of defined helical bundles in membrane environments. *Proc. Natl Acad. Sci. U.S.A.* **101**, 15324–15329 <https://doi.org/10.1073/pnas.0403314101>
- 87 Naarmann, N., Bilgiçer, B., Meng, H., Kumar, K. and Steinem, C. (2006) Fluorinated interfaces drive self-association of transmembrane  $\alpha$  helices in lipid bilayers. *Angew. Chem. Int. Ed. Engl.* **45**, 2588–2591 <https://doi.org/10.1002/anie.200503567>
- 88 Farid, T.A., Kodali, G., Solomon, L.A., Lichtenstein, B.R., Sheehan, M.M., Fry, B.A. et al. (2013) Elementary tetrahelical protein design for diverse oxidoreductase functions. *Nat. Chem. Biol.* **9**, 826–833 <https://doi.org/10.1038/nchembio.1362>
- 89 Choma, C.T., Lear, J.D., Nelson, M.J., Dutton, P.L., Robertson, D.E. and DeGrado, W.F. (1994) Design of a heme-binding four-helix bundle. *J. Am. Chem. Soc.* **116**, 856–865 <https://doi.org/10.1021/ja00082a005>
- 90 Solomon, L.A., Kodali, G., Moser, C.C. and Dutton, P.L. (2014) Engineering the assembly of heme cofactors in man-made proteins. *J. Am. Chem. Soc.* **136**, 3192–3199 <https://doi.org/10.1021/ja411845f>
- 91 Cordova, J.M., Noack, P.L., Hilcove, S.A., Lear, J.D. and Ghirlanda, G. (2007) Design of a functional membrane protein by engineering a heme-binding site in Glycophorin A. *J. Am. Chem. Soc.* **129**, 512–518 <https://doi.org/10.1021/ja057495i>
- 92 Shinde, S., Cordova, J.M., Woodrum, B.W. and Ghirlanda, G. (2012) Modulation of function in a minimalist heme-binding membrane protein. *J. Biol. Inorg. Chem.* **17**, 557–564 <https://doi.org/10.1007/s00775-012-0876-1>
- 93 Mahajan, M. and Bhattacharjya, S. (2014) Designed di-heme binding helical transmembrane protein. *Chembiochem* **15**, 1257–1262 <https://doi.org/10.1002/cbic.201402142>
- 94 Korendowych, I.V., Senes, A., Kim, Y.H., Lear, J.D., Fry, H.C., Therien, M.J. et al. (2010) De novo design and molecular assembly of a transmembrane diporphyrin-binding protein complex. *J. Am. Chem. Soc.* **132**, 15516–15518 <https://doi.org/10.1021/ja107487b>
- 95 Discher, B.M., Noy, D., Strzalka, J., Ye, S., Moser, C.C., Lear, J.D. et al. (2005) Design of amphiphilic protein maquettes: controlling assembly, membrane insertion, and cofactor interactions. *Biochemistry* **44**, 12329–12343 <https://doi.org/10.1021/bi050695m>
- 96 Noy, D., Discher, B.M., Rubtsov, I.V., Hochstrasser, R.M. and Dutton, P.L. (2005) Design of amphiphilic protein maquettes: enhancing maquette functionality through binding of extremely hydrophobic cofactors to lipophilic domains. *Biochemistry* **44**, 12344–12354 <https://doi.org/10.1021/bi050696e>
- 97 Anderson, J.L.R., Armstrong, C.T., Kodali, G., Lichtenstein, B.R., Watkins, D.W., Mancini, J.A. et al. (2014) Constructing a man-made c-type cytochrome maquette in vivo: electron transfer, oxygen transport and conversion to a photoactive light harvesting maquette. *Chem. Sci.* **5**, 507–514 <https://doi.org/10.1039/C3SC52019F>

- 98 Yin, H., Slusky, J.S., Berger, B.W., Walters, R., Vilaire, G., Litvinov, R.I. et al. (2007) Computational design of peptides that target transmembrane helices. *Science* **315**, 1817–1822 <https://doi.org/10.1126/science.1136782>
- 99 Heim, E.N., Marston, J.L., Federman, R.S., Edwards, A.P.B., Karabadzhak, A.G., Petti, L.M. et al. (2015) Biologically active LIL proteins built with minimal chemical diversity. *Proc. Natl Acad. Sci. U.S.A.* **112**, E4717–E4725 <https://doi.org/10.1073/pnas.1514230112>
- 100 Richardson, J.S., Richardson, D.C., Tweedy, N.B., Gernert, K.M., Quinn, T.P., Hecht, M.H. et al. (1992) Looking at proteins: representations, folding, packing and design. *Biophys. J* **63**, 1186–1209 [https://doi.org/10.1016/S0006-3495\(92\)81696-X](https://doi.org/10.1016/S0006-3495(92)81696-X)
- 101 Perez-Aguilar, J.M. and Saven, J.G. (2012) Computational design of membrane proteins. *Structure* **20**, 5–14 <https://doi.org/10.1016/j.str.2011.12.003>
- 102 Deisenhofer, J., Epp, O., Sinning, I. and Michel, H. (1995) Crystallographic refinement at 2.3 Å resolution and refined model of the photosynthetic reaction centre from *Rhodospseudomonas viridis*. *J. Mol. Biol.* **246**, 429–457 <https://doi.org/10.1006/jmbi.1994.0097>