Biochimica et Biophysica Acta 1848 (2015) 1248–1257

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamem

# De novo design of transmembrane helix–helix interactions and measurement of stability in a biological membrane

Anthony Nash [a,1], Rebecca Notman [b], Ann M. Dixon [c,*]

[a] MOAC Doctoral Training Centre, University of Warwick, Coventry CV4 7AL, UK
[b] Centre for Scientific Computing, University of Warwick, Coventry CV4 7AL, UK
[c] Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK

ARTICLE INFO

*Article history:*
Received 28 November 2014
Received in revised form 29 January 2015
Accepted 18 February 2015
Available online 27 February 2015

*Keywords:*
Membrane protein
Rational protein design
GALLEX
Gibbs free energy
Dissociation constant

ABSTRACT

Membrane proteins regulate a large number of cellular functions, and have great potential as tools for manipulation of biological systems. Developing these tools requires a robust and quantitative understanding of membrane protein folding and interactions within the bilayer. With this in mind, we have designed a series of proteins to probe the net thermodynamic contribution of well-known sequence motifs to transmembrane helix-helix association in a biological membrane. The proteins were designed from first principles (*de novo*) using current knowledge about membrane insertion and stabilizing interaction motifs. A simple poly-Leu "scaffold" was decorated with individual helix interaction motifs (G-*XXX*-G, polar residues, heptad repeat) to create transmembrane helix–helix interactions of increasing strength. The GALLEX assay, an *in vivo* assay for measurement of transmembrane helix self-association, was combined with computational methods to characterize the relative strength and mode of interaction for each sequence. In addition, the apparent free energy contribution ($\Delta\Delta G^{app}$) of each motif to transmembrane helix self-association was measured in a biological membrane, results that are the first of their kind for these *de novo* designed sequences, and suggest that the free energy barrier to overcoming weak association is quite small (<1.4 kcal $mol^{-1}$) in a natural membrane. By quantifying and rationalizing the contribution of key motifs to transmembrane helix association, our work offers a route to direct the design of novel sequences for use in biotechnology or synthetic biology (e.g. molecular switches) and to predict the effects of sequence modification in known transmembrane domains (for control of cellular processes).



Membrane proteins control the flow of critical information across membranes, and thus regulate a large number of cellular functions. Understanding the mechanisms that govern the folding and interactions of membrane proteins, including their ability to interact with lipids, is the key to unlocking their huge potential as tools for manipulation of biological systems. Many investigators have used a rational protein design approach to reveal these mechanisms while making non-natural membrane proteins with unique functions. For example, Lear and coworkers designed a model ion channel, which was selectively permeable to proton ions, using model amphipathic peptides containing only Leu and Ser residues [1]. Similar approaches have been applied to design of electron-transfer systems that bind novel cofactors [2] and cell-penetrating peptides that target tumour cells [3].

Artificial protein systems are typically developed in one of two ways, either by *protein engineering* or by *de novo protein design* [4]. The majority of membrane protein design studies in the literature fall under the heading of protein engineering, where investigators made small changes (mutations) to a native protein sequence in order to improve or alter the properties of the protein (e.g. lipophilicity, oligomeric state [5], fusogenicity [6]). These studies can reveal regions of structural importance in a protein sequence and may impart new properties to the protein, but results are often difficult to interpret in the context of the native sequence (which has been shown to exert significant influence on interactions [7]).

*De novo* protein design involves the creation of entirely new, non-natural protein sequences. These new sequences can be generated in large numbers using a combinatorial approach, such as a genetic library of thousands of mutants. This approach has been used in the past to select transmembrane (TM) domain sequences that strongly self-associate in a bacterial membrane [8], and has revealed the enrichment

*Abbreviations:* TM, transmembrane; MHC, major histocompatibility complex; G-XXX-G, Gly-XXX-Gly, where X is any residue; ΔG, Gibbs free energy; MBP, maltose-binding protein; β-gal, β-galactosidase; GpA, glycophorin A; CHI, CNS searching of helix interactions; IPTG, isopropyl β-D-1-thiogalactopyranoside

* Corresponding author. Tel.: +44 2476 150037; fax: +44 2476 524112.
*E-mail address:* ann.dixon@warwick.ac.uk (A.M. Dixon).
[1] Current address: Department of Chemistry, University College London, London, WC1E 6BT, UK.

http://dx.doi.org/10.1016/j.bbamem.2015.02.020
0005-2736/

of specific sequence motifs and residues, such as Trp [9], Phe [10], Ser and Thr [11], in interacting sequences. Likewise, computational approaches have been used to generate large numbers of new membrane protein sequences and identify those with desirable properties [12,13].

New protein sequences have also been created from first principles, using current understanding of hydrophobicity, membrane insertion, and helix–helix interactions to create model systems that illustrate a sequence/structure/function relationship. These protein sequences are often composed of a very simple "scaffold" or background sequence (poly-Leu [14–19], poly-Ala [20], poly-Met and poly-Val [21] have been used in the past, as have combinations of Leu/Ala [14,15,22–25], Leu/Val [26], Leu/Ser [1] or Ala/Ile [27]) upon which a sequence motif of interest is grafted. Using this approach, investigators have studied the role of amino acid sequence in TM helix interactions [14–21,27], membrane insertion [25], topology [24], membrane fusion [26], membrane anchoring and hydrophobic mismatch [22,23]. Less abundant in the literature are quantitative thermodynamic parameters that describe the interactions of these non-native systems. While some values have been published for model systems using analytical ultracentrifugation [13,28] and Forster resonance energy transfer [25] methods, there is a much richer literature for naturally-occurring TM domains including those in glycophorin A [7,29–33], fibroblast growth factor receptor 3 [34], the rat Neu oncoprotein [35], the bovine papillomavirus E5 protein [36] and class II MHC invariant chain [37]. These values help us to understand the driving forces behind biochemical processes, act as the foundation for new mechanistic models, and have been fundamental in our understanding of protein folding thus far. Therefore, in the case of model proteins, understanding the net thermodynamic contribution of small changes in sequence to overall interaction strength could direct the design of sequences for use in biotechnology or synthetic biology.

With this in mind, we have designed a series of TM domain proteins that allow us to estimate the net thermodynamic contribution of well-known sequence motifs to helix-helix association. TM domain sequences were designed *de novo* (from first principles) using current knowledge about membrane insertion and interaction motifs, and were not derived from a genetic library or a native protein sequence. Because of its propensity to spontaneously insert into bilayers and form a highly α-helical structure [26], the poly-Leu backbone was used as a simple scaffold. Poly-Leu has been reported to exist as a monomer [16,18,19] as well as an oligomer [14]. We then attempted to create TM helical interactions of increasing strength by "decorating" the surface of the low-complexity poly-Leu TM sequence with well-known helix interaction motifs. In a 2006 survey of 445 helical pairs obtained from high-resolution membrane protein structures, investigators found that 75% of helix interactions could be described by simple principles of helix-helix packing and highlighted the importance of two motifs in particular [38], namely the G-*XXX*-G motif and the heptad repeat motif. Extensive manipulation of the heptad repeat motif has been used in the *de novo* design of soluble coiled-coil structures [4,39–41], and similar packing has been observed in membrane proteins [42]. Therefore, both the G-*XXX*-G and heptad repeat motifs were used in this investigation. The contribution of polar residues was also studied, as these have been shown to contribute significantly to oligomerization of TM domains [15,19,43] and are an evolutionarily conserved feature of many native TM proteins [44]. Using an *in vivo* reporter assay called GALLEX, helix interaction strength was measured here both qualitatively and quantitatively for a series of TM domains of increasing sequence complexity. The individual contributions of a range of well-known helix interaction motifs to the overall Gibbs free energy (ΔG) of association were measured in a natural membrane bilayer, the inner membrane of *Escherichia coli,* allowing us to directly compare the efficiency of these motifs for stimulating helix interactions. We supplement this biochemical study with a computational search for favorable dimer conformations over a selected number of *de novo* designs. The results were used to design a bi-functional TM domain that could potentially interact via two competing mechanisms.

## 1. Materials and methods

### 1.1. GALLEX assay

The self-association of the rationally-designed TM proteins in the *Escherichia coli* inner membrane was studied using the GALLEX assay, the details of which have been described previously [45]. DNA encoding the designed TM domain of interest was ligated into the pBLM100 plasmid (provided by Prof. D. Schneider) according to the reported protocol [45] to produce a fusion protein containing maltose binding protein (MalE) at the C-terminus and residues 1–87 of the LexA protein from *E. coli* at the N-terminus. During the original validation of the GALLEX protocol [45], Schneider demonstrated that a GpA TM segment of 17 residues in length yielded a ten-fold increase in association compared with a TM segment of 19 residues. On a similar note, the establishment of the ToxR *in vivo* TM reporter assays was optimized to yield significant signal strength using only 13 residues from the GpA TM segment [46]. The vast majority of left-handed dimers are thought to pack with a crossing-angle at around + 20° whereas the crossing-angle of right-handed dimer is around - 50° [47–49]. In order for a TM segment to span the hydrophobic region of the bilayer, short sequences may refrain from adopting an otherwise favorable right-handed orientation. However, it is clear from previous work [45] that a 17 residue long GpA TM segment results in a stable right-handed helical dimer. In light of this, all of our constructs contained 17 amino acid TM domains.

The resulting GALLEX chimeras were expressed in *E. coli* strain SU101 (provided by Prof. D. Schneider) after induction with 0.01 mM IPTG at 37 °C. Self-association of the TM domains leads to dimerization of LexA domains, which then bind to the wild-type *lacZ* promoter in SU101 and repress expression of β-galactosidase (β-gal). Therefore, the degree of repression of β-gal is proportional to the strength of TM domain homo-oligomerisation. Self-association of the LexA-TM-MalE fusion proteins was assessed via measurement of β-gal activity as described previously [50]. The β-gal Miller units were calculated using Eq. (1):

$$\frac{1000(OD_{420}-1.75\times OD_{550})}{t\times v\times OD_{600}} \quad (1)$$

where $t$ is time (min) and $v$ is volume (mL). Prior to performing the GALLEX assay, membrane insertion of all constructs was confirmed using the *malE* complementation assay [45], where cells were grown on M9 agar plates containing 0.4% maltose. A sodium hydroxide wash was also performed to give a qualitative indication of membrane insertion according to the reported protocol [51]. Expression levels for all constructs were confirmed via Western analysis against the MBP domain. The resulting band intensities were determined using the ImageJ software [52], and used to normalize all β-gal activities to total fusion protein expression levels. For all GALLEX measurements described above, a minimum of three independent outgrowths were prepared and measured in order to assess the magnitude of error. The final result was taken from the average of the three samples and the error was given by the standard error of the mean. Statistical analysis of all data was carried out using ANOVA with a probability (P) of 0.05 (95% confidence interval) to establish whether a given change in the "scaffold" sequence resulted in a significant change in the strength of interaction.

In order to estimate the thermodynamic contributions of selected motifs on the low complexity background sequence, quantitative measurements of TM interactions in the *E. coli* inner membrane were also performed using the GALLEX assay according to the reported protocol [30]. Briefly, β-gal activity measurements were collected from cell samples induced with a range of IPTG concentrations ($1\times10^{-4}-1$ mM). Expression of the chimera at each IPTG concentration was assessed by Western blotting against MBP and quantification of the band intensities using ImageJ [52], and these values were plotted against [IPTG]. The plot