Peptides as transmembrane segments: Decrypting the determinants for helix–helix interactions in membrane proteins

ARTICLE *in* BIOPOLYMERS · JANUARY 2007

Impact Factor: 2.39 · DOI: 10.1002/bip.20668 · Source: PubMed

3 AUTHORS, INCLUDING:

Arianna Rath
SickKids
32 PUBLICATIONS **766** CITATIONS

SEE PROFILE

Peptides as Transmembrane Segments: Decrypting the Determinants for Helix–Helix Interactions in Membrane Proteins

Arianna Rath,[1,2] Rachel M. Johnson,[1,2] Charles M. Deber[1,2]

[1] *Division of Molecular Structure and Function, Research Institute, Hospital for Sick Children, Toronto, Ont., Canada M5G 1X8*

[2] *Department of Biochemistry, University of Toronto, Toronto, Ont., Canada M5S 1A8*

Received 30 October 2006; accepted 19 December 2006
Published online 5 January 2007 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.20668

ABSTRACT:

Although the structural analysis of membrane proteins is advancing, an understanding of the basic principles that underlie their folding and assembly remains limited because of the high insolubility intrinsic to these molecules and concomitant challenges in obtaining crystals. Fortunately, from an experimental standpoint, membrane protein folding can be approximated as the rigid-body docking of pre-formed α-helical transmembrane segments one with another to form the final functional protein structure. Peptides derived from the sequences of native α-helical transmembrane segments and those that mimic their properties are therefore valuable in the experimental evaluation of protein folding within the membrane. Here we present an overview of the progress made in our laboratory and elsewhere in using peptide models toward defining the sequence requirements and forces stabilizing membrane protein folds. © *2007 Wiley Periodicals, Inc. Biopolymers (Pept Sci) 88: 217–232, 2007.*

Keywords: membrane protein folding; transmembrane peptides; helix–helix interactions

Correspondence to: Charles M. Deber, Division of Molecular Structure and Function, Research Institute, Hospital for Sick Children, Toronto, Ontario, Canada M5G 1X8; e-mail: deber@sickkids.ca

Contract grant sponsors: Canadian Institutes of Health Research (CIHR); Natural Sciences and Engineering Research Council of Canada (NSERC)

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

INTRODUCTION

The molecular traffic and information flow across biological membranes is controlled and regulated by membrane proteins. Conduits for ions and small molecules, transporters for drugs and other molecules, as well as components of certain enzymes are all constructed from membrane-embedded proteins. Membrane proteins also underlie fundamental cellular processes such as protein translocation, respiration, and signal transduction. This important class of proteins comprises ~70% of all current pharmaceutical targets,[1] and has been estimated to constitute ~20–30% of all proteins in sequenced genomes.[2]

Stable incorporation of membrane proteins into the nonpolar interior of lipid bilayers requires the existence of one or more contiguous portions of amino acid sequence with largely hydrophobic character, termed transmembrane (TM) segments. The structures of individual TM segments are further constrained in the cytoplasmic membrane to predominantly α-helical structures that simultaneously sequester the peptide backbone away from the low dielectric environment of the membrane and satisfy the amide-carbonyl hydrogen bonds.[3–5] The high hydrophobicity of TM segments, however, presents technical challenges in the overexpression, purification, and structural characterization of membrane proteins. To date, membrane proteins represent only ~0.3% of the >34,000 structures deposited in the Protein Data Bank.[7] Though it has been noted that the structural characterization of membrane proteins is moving forward exponentially,[6] at the present pace of structural determinations it still may take

more than three decades to obtain representative structures of most membrane protein folds.[7]

Gaining insight into the membrane protein folding problem has not relied exclusively on the availability of high-resolution structures. In a series of early folding experiments performed on the α-helical membrane protein bacteriorhodopsin, the native protein fold was regenerated from various protein fragments[8–10]—an indication that TM helix–helix interactions have sufficient specificity to generate tertiary structure.[9,10] Studies of loop-severed or loop-deleted fragments confirmed the assembly of rhodopsin family members via helix–helix contacts[11,12] and extended these findings to other TM proteins [yeast α-factor transporter STE6; lactose permease; rat m3 muscarinic acetylcholine receptor].[13–15] Taken together, these experiments support the division of the basic folding pathway of α-helical membrane proteins into two energetically distinct categories according to the two-stage model.[3] In the first step of folding, individual segments of *ca.* 18–25 residues with sufficient hydrophobic character are inserted into the membrane environment, whereupon they fold into stable TM α-helices in order to minimize the free energy associated with the burial of main-chain polar groups. In the second step, these independently-folded helices dock one with another within the membrane plane to adopt tertiary and/or quaternary structures stabilized by lateral, sequence-dependent, helix–helix interactions.

Because individual TM α-helices can be considered as true protein domains—*i.e.*, independently folded units—in the process of membrane protein assembly, the challenge of overexpressing and purifying full-length membrane proteins can be circumvented by a divide-and-conquer approach; individual peptides derived from single TM sequences can be used to evaluate experimentally the sequence requirements for stable insertion and folding. TM peptides have been used in membrane-mimetic and biological membrane environments, in our laboratory and others to contribute an understanding the general principles that underlie folding in the membrane. In this review, we focus on the advances made using TM peptides in understanding the properties that produce TM segments capable of spanning the bilayer as α-helices, and once inserted, the nature of the TM-TM helix–helix interactions that dictate the tertiary and/or quaternary structures of membrane protein folds.

BACKGROUND

Hydrophobic residues (Ala, Leu, Val, Ile, Phe) dominate the TM sequences of integral membrane proteins, although polar residues (Pro, Ser, Thr, Gln, Asn, Tyr, Trp; and occurrences of Glu, Asp, Lys, and Arg) account for nearly 25% of overall composition.[16,17] Because of their intrinsic hydrophobic character, peptides derived from TM sequences are difficult to produce and purify via chemical synthesis or in recombinant systems; those derived from TM sequences have a high tendency to aggregate nonspecifically during synthesis, leading to the use of multiple couplings of amino acids and/or the use of elevated temperatures to improve yields.[18–22] Synthetic procedures may also be complicated by the binding of the peptide irreversibly to the stationary phase of reverse phase high performance liquid chromatography columns during purification.[18] Subsequent solubilization and avoidance of aggregation may also require the use of harsh conditions or detergent amphiphiles, which must later be separated from the TM peptide.[18] Overexpression of TM proteins in cellular systems also presents challenges to protein insolubility and low protein yield in addition to avoidance of aggregation during purification procedures [reviewed in Ref. 23]. The study of TM peptides as models of membrane protein folding in vitro therefore requires a means to increase peptide solubility—whether produced synthetically or in a recombinant system—while maintaining the hydrophobic character and folding properties native to the TM domain.

The "Tagging" Approach to TM Peptide Synthesis

Placement of charged and/or polar residues at the N- and/or C-terminal ends of hydrophobic sequences—termed polar residue "tagging"—has proven to be useful in the design of aqueous-soluble "TM-mimetic" peptides. In previous works, we synthesized a series of peptides with a 10-residue hydrophobic core [typical sequence: H_2N-SKSKAXAAXAWAX-AKSKSKS-OH, where X = an uncharged amino acid (excluding Trp and Cys)] and found that these sequences were soluble in aqueous buffer and a variety of membrane-mimetic media [sodium dodecyl sulfate (SDS) and lysophosphatidylglycerol micelles; and dimyristoyl-phosphatidylglycerol bilayer vesicles].[24–26] Moreover, these peptides circumvented aggregation in all experimental media, remaining monomeric at concentrations up to 300 μM.[26]

The 10-residue hydrophobic segment of the peptides used in the earlier studies, however, is approximately half the length of the *ca.* 20 nonpolar residues that comprise natural TM domains. Key to studying the energetics of TM helix incorporation into bilayers and of helix–helix interactions is the design of water-soluble peptides with hydrophobic cores closer in length to natural TM segments. To achieve this goal, we designed a series of peptides with a 19-residue hydrophobic core flanked at the N- and C-termini with Lys residues to enhance aqueous solubility, with the prototypical sequence H_2N-KKAAAXAAAAAXAAWAAXAAAKKKK-NH_2, where

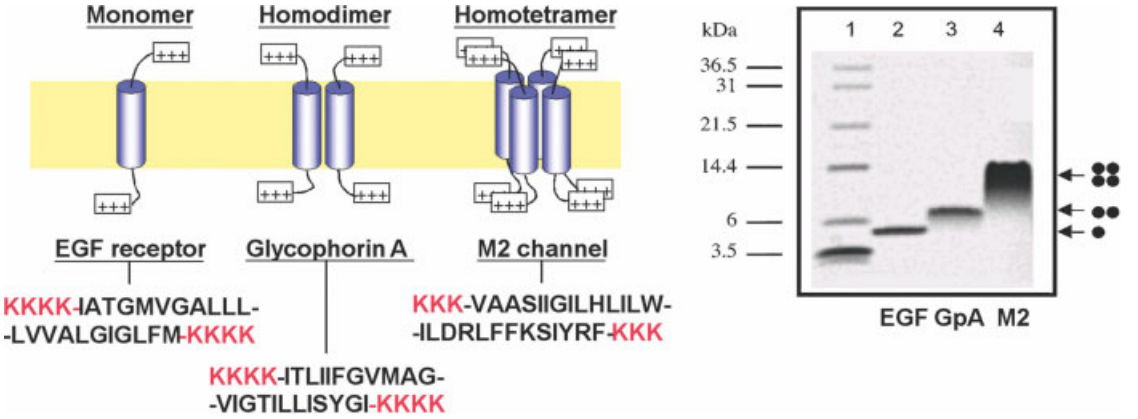

FIGURE 1 Lys-tagged peptide approach to analysis of assembly states of TM segments. Single-spanning TM domain sequence peptides from the epidermal growth factor receptor (EGF), glycophorin A (GpA), and the influenza A virus M2 ion channel (M2) were designed and synthesized with 3- to 4-Lys residues at both N- and C-termini. *Left panel:* Synthetic peptides depicted with the oligomeric states known for these segments when in the corresponding intact proteins (EGF, monomer; GpA, dimer; M2, tetramer). *Right panel:* SDS-PAGE gel showing that each peptide migrates according to its native oligomerization state. Adapted from Melnyk et al., Biochemistry, 2001, 40, 11106–11113.

X = each of the 20 commonly occurring amino acids.[27,28] The "Lys-tags" used in these peptides to enhance aqueous solubility were anticipated to be more effective than Ser-Lys pairs in preventing peptide–peptide aggregation through the resulting high local positive charges, and the peptide C-terminus was amidated to eliminate potential electrostatic attractions that might occur within or between peptide molecules. This "tagging" approach—which, in effect, borrows the feature that most native TM segments in membrane proteins are flanked by Lys (or Arg) residues at membrane entry/exit points—rendered these otherwise highly hydrophobic model peptides soluble in aqueous phases, reduced peptide–peptide aggregation, and greatly facilitated synthesis and purification.[27,28] These "TM-mimetic" peptides further retained the ability to spontaneously partition into—and adopt α-helical conformation in—various membrane-mimetic environments (lipid micelles, small unilamellar vesicles, and organic solvents),[29] and, importantly, remained monomeric in all aqueous and apolar experimental media through the concentration range of 5–250 μM.[29] Subsequent work on a TM-mimetic peptide with a 21-residue hydrophobic core (termed TMX-1) indicated that Ser-Lys pairs could also be used to increase peptide solubility in aqueous media, although high solubility in this case was due to the formation of molecular aggregates that persisted at peptide concentrations down to at least 0.1 μM. Aqueous TMX-1 was nevertheless capable of strongly partitioning into membrane vesicles and showed a very strong preference for TM α-helical conformation.[30] Tagging hydrophobic peptides with polar residues—in particular Lys residues—therefore provides a means to increase peptide solubility that can simultaneously maintain the hydrophobic character and folding properties native to the TM domain in membrane mimetic-media but provide monomeric solubility in aqueous solution.

Expression of TM Sequences as Fusion Proteins

An alternative approach to the solubilization and production of TM segments by synthetic means is fusion of the TM peptide sequence(s) of interest to a soluble protein to enhance solubility and facilitate purification from bacterial systems. In such a chimeric protein system developed by the Engelman group, fusion of the TM α-helical domain of interest to the C-terminus of *Staphylococcal* nuclease, a monomeric soluble protein,[31] allows the resulting chimera to be expressed at high levels in *E. coli*, readily purified, and used to assess the stability and stoichiometry of contacts between the TM domain of interest in vitro.[31] We have also used fusions to the soluble protein thioredoxin—in conjunction with Lys tags—to overexpress in *E. coli* and purify in milligram amounts helix-loop-helix fragments of membrane proteins [reviewed in Ref. 23]. In this case, however, the thioredoxin is cleaved from the TM protein before characterization. The native folds of TM segments are preserved in both these systems, and each has been used to investigate helix–helix associations as will be described in detail later.

Use of Chimeric Reporter Systems to Assay TM Segment Folding

The solubility issues inherent to the production and purification of TM segments in sufficient amounts for biophysical

characterization may be circumvented completely through the use of chimeric systems that report bilayer insertion and/or helix–helix association of a TM sequence of interest. Fusion of a TM sequence of interest to the luminal P2 domain of integral membrane protein leader peptidase (Lep) followed by in vitro translation in rough microsomes and electrophoretic separation, for example, allows for the efficiency of membrane insertion to be quantified.[32] A chimeric protein system for the study of TM helix–helix assembly in the bacterial cytoplasmic membrane has also been developed.[33] Termed TOXCAT, this assay fuses the TM domain of interest at its N-terminus to the DNA binding domain of ToxR (a dimerization-dependent transcription factor) and at its C-terminus to maltose binding protein (a monomeric periplasmic anchor). Upon chimera expression in *E. coli*, self-association of the TM segment of interest results in the ToxR-driven activation of a reporter gene encoding chloramphenicol acetyltransferase (CAT), with the readout level of CAT expression indicating the strength of helix–helix interactions. Several modified TOXCAT procedures have since been developed.[34,35] Hetero-dimerization of TM peptide sequences in the *E. coli* inner membrane, for example, can be monitored by fusing the sequences of interest to two different LexA DNA binding domains in the GALLEX assay,[36] with heterodimeric association reported by repression of β-galactosidase activity.

INVESTIGATING HELIX–HELIX INTERACTIONS WITH TM PEPTIDE SEQUENCES

After insertion, individual TM helices associate one with another within the bilayer plane to build the tertiary and/or quaternary structures of membrane proteins.[3] In order to be useful in the investigation of these helix–helix interactions, water-solubilized TM peptides must not only partition into membrane-like environments and adopt native-like structure but must also recapitulate their native helix–helix contacts. We initially investigated the ability of Lys-tagged TM peptides to adopt native quaternary structures by synthesizing peptides corresponding to the TM domains from a variety of oligomeric, single-spanning membrane proteins [epidermal growth factor receptor (EGFR), the human erythrocyte protein glycophorin A (GpA), and the influenza A virus M2 ion channel (M2)] with 3–4 lysine residues placed at both N- and C-termini.[37] Each peptide was shown to adopt an α-helical conformation consistent with native structure formation upon insertion into SDS micelles, and more importantly migrated in SDS micelles on SDS polyacrylamide gel electrophoresis (SDS-PAGE) gels according to their respective native-like oligomeric states: EGFR was monomeric, GpA was dimeric, and M2 was tetrameric (see Figure 1). The Lys-tag approach to TM peptide solubilization proved therefore to be compatible with the formation of the secondary structure and the sequence-specific helix–helix contacts of native TM domains.

Identification of TM Helices That Stabilize Membrane Protein Oligomers

We and others have subsequently used Lys-tagged TM peptides to pinpoint the helix–helix contacts that build and stabilize several different oligomeric folds (Figure 2). Lys-tagged TM peptides were used to examine helix–helix interactions in the *E. coli* aspartate receptor (Tar), a double-spanning membrane protein known to form homodimers. To determine if one or both of its TM segments (termed Tar-1 and Tar-2) were responsible for the stability of the oligomer, pep-

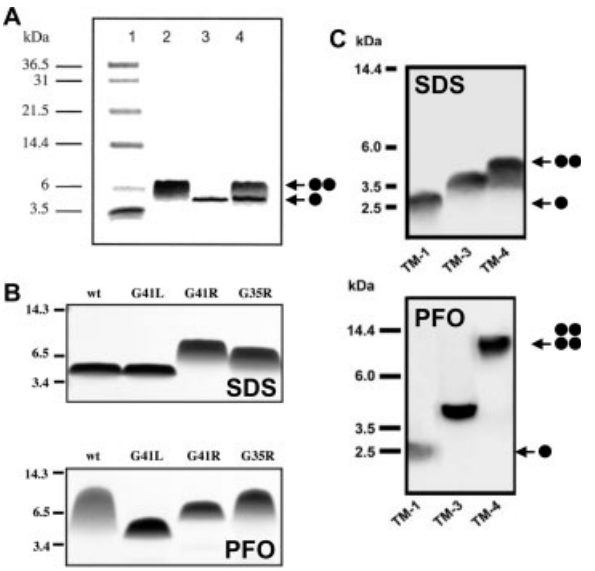

FIGURE 2 Determination of helices that stabilize oligomeric folds using Lys-tagged peptides. (A) Oligomeric states of Tar-1 and Tar-2 Lys-tagged TM peptides on SDS-PAGE. 6K-Tar-1 is dimeric (lane 2), while 6K-Tar-2 is monomeric (lane 3). Mixing of 6K-Tar-1 and 6K-Tar-2 prior to equilibration in SDS buffer resulted in no change in migration of either band (lane 4). Oligomeric states are indicated at the left of the panel. Adapted from Melnyk et al., Biochemistry, 2001, 40, 11106–11113. (B) Self-assembly of wild-type (wt) and mutant peptides in corresponding to the gamma/FYXD2 TM segment in SDS (*top panel*) and PFO (*bottom panel*) micelles. In PFO, the G41L and G41R substitutions alter peptide oligomerization, while the G35R peptide migrates similarly to wild-type (wt). Adapted from Therien and Deber, J Mol Biol, 2002, 322, 583–550. Oligomeric states of Lys-tagged TM peptides corresponding to TM helices 1, 3, and 4 of Hsmr in SDS (*top panel*) and PFO (*bottom panel*) micelles. The oligomeric state of TM-4 is indicated at the left of the panels. Adapted from Rath and Melnyk, J Biol Chem 2006, 281, 15546–15553.

tides corresponding to the Tar-1 and Tar-2 sequences were synthesized with 3 Lys residues at their N- and C-termini, and were found to be α-helical in SDS micelles.[37] While Tar-1 formed SDS-resistant homodimers, Tar-2 was monomeric, implicating the Tar-1 helix as the oligomeric determinant for the Tar protein.[37] Subsequent investigation of Lys-tagged peptides encompassing the TM region of the gamma/FYXD2 subunit of the Na^+, K^+ -ATPase pump showed that while the wild-type sequence is capable of tetramerization, a mutant peptide with a disease-causing Gly to Arg substitution could no longer oligomerize, suggesting a connection between the quaternary structure of the gamma/FYXD2 subunit and familial renal hypomagnesaemia.[38]

We have also used Lys-tagged peptides to investigate the assembly of hSMR, a member of the bacterial small multidrug resistance (SMR) family. These proteins are drug/proton antiporters embedded in the bacterial cytoplasmic membrane that can confer clinically significant bacterial resistance to drugs and cytotoxic compounds. SMRs must oligomerize in order to function, a process that requires the assembly of at least two inactive monomers by intermolecular association of their four TM helices. Lys-tagged peptides corresponding to each of the four wild type Hsmr TM helices (denoted TM-1 to TM-4) and a corresponding library of mutant peptides were examined to determine the interactive surfaces that likely contribute to protein oligomerization.[39] TM-1, TM-3, and TM-4 were found to adopt α-helical conformation in SDS and perfluorooctanoate (PFO) micelles. While TM-4 formed SDS-resistant dimers and PFO-resistant tetramers, TM-1 and TM-3 were found to be monomeric, implicating TM-4 as the predominant stabilizer of hSMR assembly and providing a model for SMR dimer assembly.

Self-association of the anthrax toxin receptor ANTXR1/2 has also been probed using Lys-tagged peptides corresponding to its single TM segment. ANTXR TM domain peptides were found adopt α-helical conformation and to oligomerize in SDS micelles, and a series of mutant ANTXR1 peptides was generated to validate the self-association of the ANTXR1 TM domain predicted by computer modeling.[40] The mutations identified in TM peptides to disrupt self-association were subsequently shown in vivo to reduce the rate of PA63 heptamer formation on the mammalian cell surface, indicating that ANTXR1 TM domains self-associate and that these interactions may stabilize intermediate oligomerization states of ANTXR1-PA63 complexes.

Investigation of Sequence Motifs That Dictate Close-Packed Helix–Helix Interactions

As described already, TM peptides have proved useful in identification of helices that stabilize membrane protein qua-

Table I Examples of TM Segments Known to Form Homo-Oligomers

Motif/protein[a]	Transmembrane Segment	Reference
Right-handed crossing (GG4 motif)		
GpA 72–98	EITLIIFGVMAGVIGTILLISYGIRRL	42
M13 MCP 25 – 45	AWAMVVVIVGATIGIKLFKKF	43, 44
Left-handed crossing (heptad motif)		
Influenza M2 26–44	LVVAASIIGILHLILWILD	45
Phospholamban 31–52	LFINFCLILICLLLICIIVMLL	46, 47
ANTXR1 318-337	ILAIALLILFLLLALALLWW	40
Ser zipper	QLFAASLLLFAVSLILFIASARLRYLAG	48
Polar motifs		
Strongly polar residue	VEELLSKNYHLENEVARLKKLVGLLLLLLLN/D/Q/E/HLLLLLLLLL LLLLLLQLLIAVLLLIAVN/D/Q/ELILLIAVARLYLVG	49–52
Multiple Ser/Thr	SxxSSxxT, SxxxSSxxT	53
Tar-1	VVTLLVMVLGVFALLQLISGSLFF	37, 54
Motif(s) undetermined		
DGK-2 50-75	VDAITRVLLISSVMLVMIVEILNSAI	55
FYXD2/GAMMA	TVRKGGLIFAGLAFVVGLLILLSKRF	38
D2 dopamine receptor TM7	LYSAFTWLGYVNSAVNPIIY	56
MPZ	GVVLGAVIGGVLGVVLLLLLLFYVV	[unpub.]
TNF	IFMYLLTVFLITQMIGSALFAVYLH	57
Hsmr-4 85-105	VAGVVGLALIVAGVVVLNVAS	39

[a] GpA: human erythrocyte protein glycophorin A; MCP: bacteriophage M13 major coat protein; ANTXR1: TM segment from human anthrax toxin receptor; DGK-2: TM2 from diacylglycerol kinase; Tar-1: TM1 from *E. coli* aspartate receptor; GAMMA: Na^+/K^+ ATPse γ-subunit TM segment; MPZ: TM segment from myelin protein zero; TNF: TM segment from tumor necrosis factor; Hsmr-4: TM4 from *H. salinarum* small multidrug resistance protein Hsmr.

ternary structures. Examining the folding behavior of TM peptides has also provided insight into the precise residue motifs within a TM helix sequence—and by extension the forces—that stabilize helix–helix packing and assembly. Much of our current knowledge concerning the molecular details of this process has been obtained from a few well-characterized homo-oligomeric single spanning TM domains and systems used because of their symmetry and relative simplicity.[41] Van der Waals interactions and electrostatic interactions, including hydrogen bonding, and solvophobic effects have been identified in these systems as factors that determine membrane protein stability, but the magnitude and relative importance of each likely vary from one system to another. A list of the residue motifs found to mediate helix–helix contacts in several of these systems is given as Table I.

The Gly-xxx-Gly Helix–Helix Interaction Motif

The vast majority of studies examining the sequence-dependence of helix–helix interactions have relied on the production of TM domains as individual peptides and/or fusion proteins. Early work examining dimerization of the TM domain from the single-spanning α-helical human erythrocyte glycoprotein A (GpA)[31,58] relied on fusion of the TM peptide to staphylococcal nuclease for expression and purification. Dimerization of the wild-type mutant chimeras on SDS-PAGE established that apolar residues defined much of the GpA interface—an early indication that precise packing interactions and van der Waals forces could stabilize membrane protein assembly. Also noted in early characterizations of the GpA dimerization site was the involvement of two Gly residues separated by three intervening positions—later termed the Gly-xxx-Gly or GG4 motif—at interfacial positions.[58] Subsequent fusion of the GpA TM sequence in the TOXCAT chimera and study of its homodimerization revealed the entire dimerization motif of GpA [LIxxGVxxGVxxT], and it was found that the GG4 motif by itself was capable of inducing strong self-assembly of host sequences, and that the three-residue spacing between both glycines to be optimal for the interaction.[59]

We have systematically examined the role of GG4 motifs in stabilizing TM helix dimerization of GpA and of other TM dimers such as the Ff bacteriophage M13 major coat protein (MCP) using a combination of Lys-tagged peptides and TOXCAT chimeras. MCP exists as a single-spanning membrane protein in the *E. coli* host inner membrane prior to assembly of lipid-free virions. The 50-residue coat protein sequence, ^{1}AEGDDPAKAA^{11}FFSLQASATE^{21}YIGYAWAMVV^{31}VIVGATIGIK^{41}LFKKFTSKAS, contains a mid-hydrophobic segment (the effective TM segment) as highlighted by the underline (adapted from Ref. 60). Through mutagenic analysis, observation of SDS-resistant MCP dimers, and melting curves obtained from CD spectra, we had previously mapped a "two in–two out" helix–helix association motif in MCP that contained a GG4 motif (VVxxGAxxGI).[43] Molecular models of the M13 coat protein V31A dimer generated from these results indicated an interfacial location for the GG4 motif (see Figure 3). Comparisons between intact wild type and mutant (V29A, V31A) bacteriophage M13 coat proteins with their corresponding Lys-tagged TM peptide constructs (M13 residues 24–42 in KKK-TM-KKK peptides)—using thermally-induced α-helix to β-sheet transitions monitored by CD experiments in lipid micelles—demonstrated that the relevant helix–helix tertiary contacts found in the intact proteins indeed persist in the peptide mimics.[61]

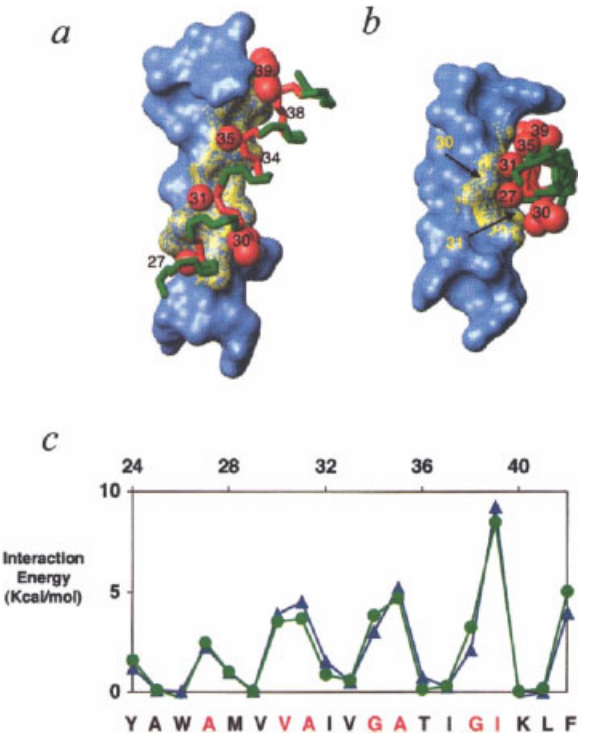

FIGURE 3 Model for M13 V31A dimer as generated by a global conformation search. (A) The dimeric interface is depicted from a side view. One helix is rendered as a surface projection (blue), with the interaction face highlighted in yellow. In the second helix, the residues within the dimer interface are in red, with side chains shown as space-filled spheres. (B) Dimer interface depicted as an end view from N-to-C terminus. (C) Residue interaction energy of the two helices calculated for the V31A dimer. Navy and green lines indicate the sites and extents of interaction energy contributed by each of the two helices. Residues within the interface are in red on the sequence axis. Adapted from Wang and Deber, J Biol Chem 2000, 275, 16155–16159.

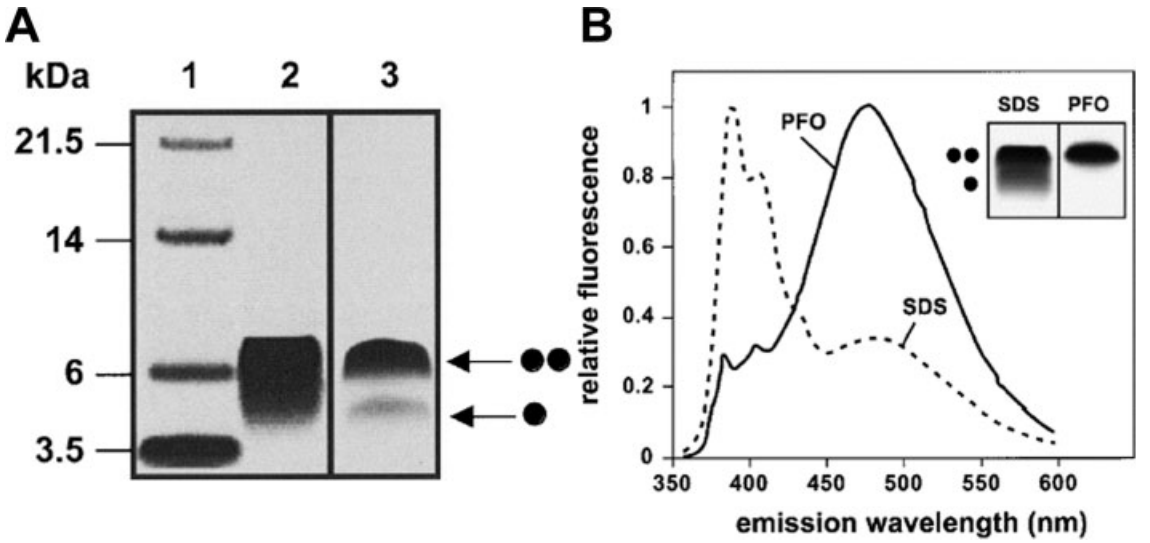

FIGURE 4 Association of MCP TM domain peptides. (A) SDS-PAGE analysis of wild-type MCP peptide using either 4–12% NuPAGE (lane 2) or 16% Tricine (lane 3) gels (Invitrogen). (B) Excimer fluorescence of pyrene-labeled wild-type MCP peptide in SDS or PFO micelles. *Inset*: comparison of the extent of dimer formation on SDS- and PFO-PAGE. Adapted from Melnyk et al. J Mol Biol 2002, 315, 63–72.

Given that the MCP TM peptides could recapitulate native contexts, we went on to synthesize additional Lys-tagged wild-type and mutant TM peptides corresponding to the effective α-helical TM segment of MCP,[44] and examined the specificity and stoichiometry of MCP helix–helix interactions in detergent micelles. Fluorescence resonance energy transfer (FRET) experiments on the dansyl and dabcyl-labeled MCP TM domain peptides in detergent micelles and SDS-PAGE analysis demonstrated that the peptides specifically associated into noncovalent homodimers, as postulated for the biologically relevant membrane-embedded MCP oligomer. MCP peptides labeled with short-range pyrene fluorophores at the N terminus also displayed excimer fluorescence consistent with homodimerization occurring in a parallel fashion (Figure 4).

When placed in the TOXCAT chimera, however, the MCP TM sequence exhibited dimer stability significantly lower than that of the GpA sequence.[57,62] We thus undertook to understand the mechanism by which the affinity of the GG4 motif in the MCP and GpA peptide sequences was controlled at the sequence level. Molecular modeling experiments indicated that the GpA and MCP dimers possess the same overall fold (see Figure 5), though the interfacial residues in GpA (LIxxGVxxGVxxT) differ from those in MCP (VVxxGAxx-GIxxF) apart from the GG4 motif. To help decrypt the determinants of dimer affinity, we introduced GpA interfacial residues, alone and in combination, into the MCP sequence.[62] Using both in vivo TOXCAT assays and SDS-PAGE gel migration rates of synthetic peptides derived from TM regions of the proteins, we found that the most distal interfacial sites, 12 residues apart (and ∼18 Å in structural space), work in concert to control TM-TM affinity synergistically (see Figure 6).

The residue context surrounding the GpA GG4 motif, however, does not completely account for the higher stability of GpA compared to MCP; in the case where every non-Gly

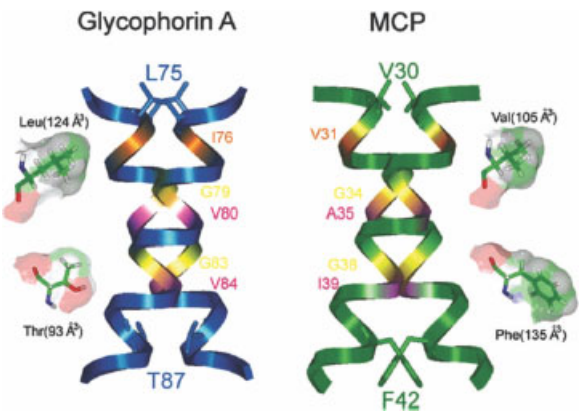

FIGURE 5 Ribbon drawings of the solid-state NMR structure of GpA dimer, residues 73–90 (left panel) and the predicted structures for the MCP dimer, residues 28–45 (right panel). Both the TM helices contain the GG4 motif at the packing contacts, whereas the other wild-type residues at interfacial positions vary. The interfacial residues at the *a* and *g* positions of GpA and MCP structures are shown along with an illustration of the van der Waals volumes of the amino acids. Adapted from Melnyk et al. J Biol Chem 2004, 279, 16591–16597.

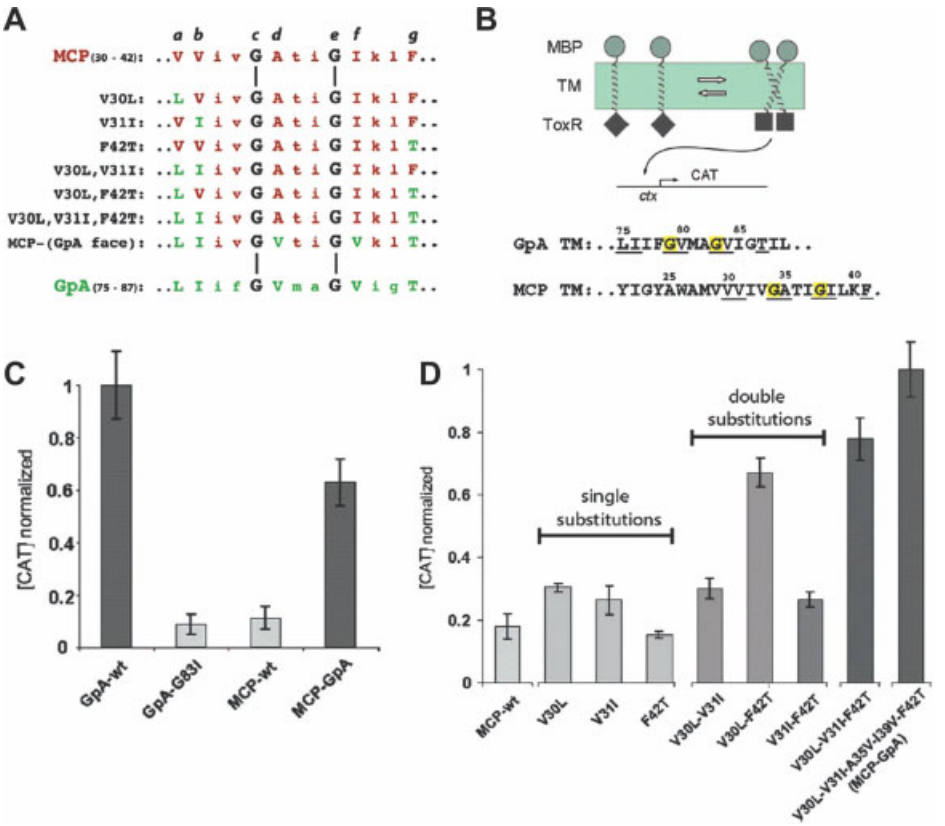

FIGURE 6 Affinity of MCP versus GpA dimerization as measured by the TOXCAT assay. (A) Sequence alignment of wild-type and mutant MCP and GpA constructs studied. Singly, double, and triply substituted MCP constructs are shown. For MCP-GpA, all of the dimer interfacial residues in MCP were changed simultaneously to those of GpA. (B) TOXCAT assay to measure TM helix–helix affinity. The TM domain of interest is fused between an N-terminal cytosolic ToxR transcriptional activator and the C-terminal, periplasmic maltose-binding protein (MBP). When induced to form dimers by the attached TM domains, ToxR becomes functional, resulting in the expression of a chloramphenicol acetyltransferase (CAT) reporter gene. The peptide sequences fused in the TOXCAT chimera for GpA and MCP are shown. The GG4 motifs are not in equivalent positions relative to the ends of the TM helix sequences. (C) CAT levels normalized to wild-type GpA. On the y-axis, the expression of GpA wild type dimer is normalized to 1.0. G83I is a control mutant in GpA that is known to destabilize dimerization. Upon changing all of the interfacial residues in MCP to that of GpA (MCP-GpA), CAT levels are increased, albeit to levels lower than wild-type GpA. (D) CAT levels normalized such that MCP-GpA was assigned a value of 1.0. Adapted from Melnyk et al. J Biol Chem 2004, 279, 16591–16597.

interfacial residue in the MCP helix was replaced with its corresponding residue from GpA, the resulting mutant (termed MCP-GpA) retained reduced helix–helix interaction strength[62] (see Figure 6). We therefore suspected that factors other than the identity of distal interfacial residues might also be involved in modulating the affinity of GG4 motif-mediated dimerization in the GpA and MCP systems. Upon comparison of the GpA and MCP sequences, we noticed that the two proteins differ not only in their distal interfacial residues but also in the position of the GG4 motif residues with respect to the ends of each TM helix sequence; the GpA motif is located near the center, while the MCP motif is positioned towards the C-terminal end (Figure 7). To test if the location of the GG4 motif in the GpA versus MCP sequence is an additional factor that modulates helix–helix affinity, we re-positioned the GpA and/or MCP interaction motifs using Leu residues at either the midpoint or C-terminal end of *de novo* designed TM peptide sequences inserted in the TOXCAT fusion protein.[63] We found that—in the context of either the GpA, MCP, or MCP-GpA interaction interfaces—centrally located GG4 motifs are capable of stronger helix–helix interactions than those proximal to TM helix ends (Figure 7).

Fusion protein constructs have been used by other groups in vivo to examine the role of the GG4 motif—and the

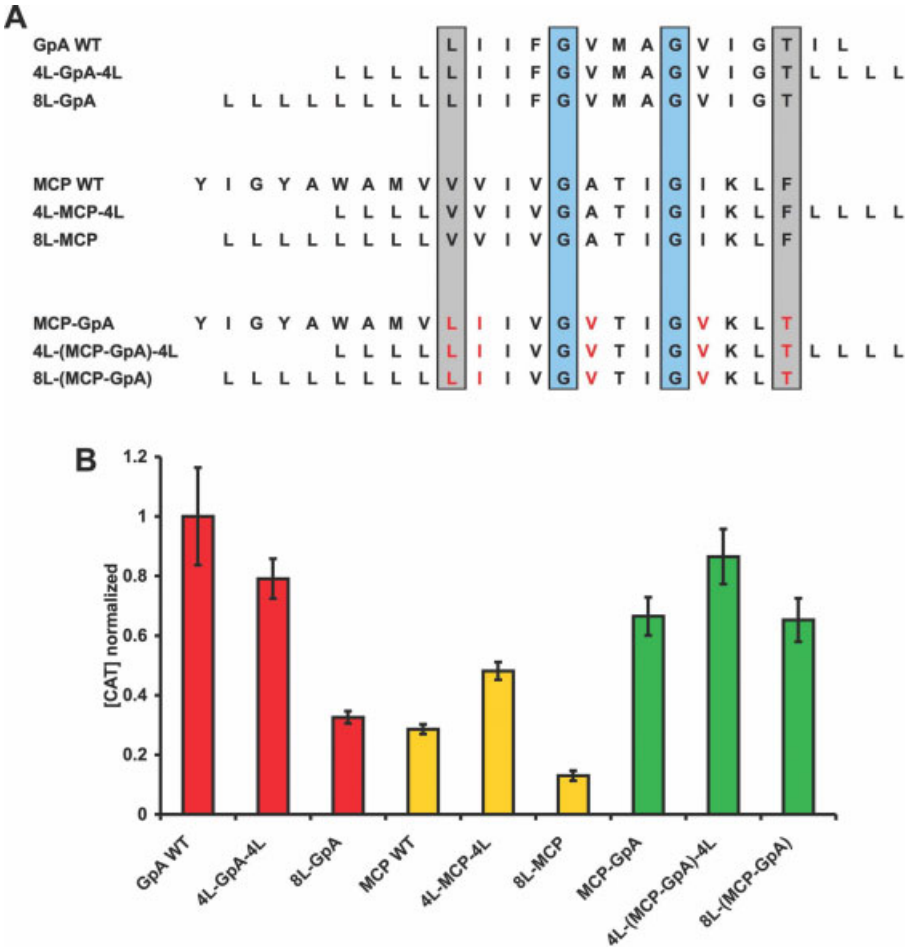

FIGURE 7 Position-dependent affinity of GpA and MCP interfaces. (A) Sequence alignment of wild-type and motif-shifted constructs of GpA, MCP, and MCP-GpA, with wild-type GpA interfacial residues shown in red. Blue boxes represent the position of the Gly-xxx-Gly motif; grey boxes represent the distal residues responsible for modulating the affinity of the GG4 motif. (B) CAT expression normalized to GpA WT for GpA (red), MCP (gold), and MCP-GpA (green) mutants. Adapted from Johnson et al. Biochem Cell Biol 2000, 84, 1006–1012.

related AA4 and SS4 motifs—in modulating helix–helix affinity in GpA,[64] with the conclusion that the presence small-xxx-small motifs can serve as a useful clue to the interaction site of TM helices, though the context of adjacent side-chains is essential for stability.

Heptad Repeat Motifs

TM peptides have proven useful in the investigation of other close-packed helix–helix interactions. Characterization of the heptad repeat motif that allows helices to pack together via a left-handed packing angle was accomplished via examining the assembly of phospholamban (PL), a 52-residue homo-pentameric calcium regulatory protein, and of the influenza A M2 protein, a 97-residue homo-tetrameric proton channel. Mutagenesis of the PL TM domain, produced either as a fusion to staphylococcal nuclease,[46] or by in vitro translation,[47] identified a 3- to 4-residue periodicity of residues essential for helix–helix assembly, and suggested that the TM helices of PL adopt a pentameric coiled-coil structure stabilized by a leucine zipper motif (see Table I). Both in vivo expression of the full-length M2 sequence,[45] and synthetic production of the full sequence as well as a peptide corresponding to its TM segment[65] have proved useful in investigating its assembly into a homo-tetramer. Mutagenesis of the M2 TM segment in the context of the full-length protein[45] identified an ~3.5-residue periodicity of severely disruptive mutations consistent with a heptad repeat and left-handed crossing angle. Residues forming rapid cross-links displayed a similar heptad repeat pattern.[66] The M2 TM peptide was

shown to recapitulate the assembly of the full-length protein produced by chemical synthesis,[65] and M2 TM peptides have been subsequently used to examine the precise residue determinants of oligomer assembly and conformational stability of the TM tetramer[67,68] (see Table I), and to investigate the role of a His residue in the M2 TM segment in pH-dependent association.[69]

Investigation of Polar Residues in Helix–Helix Interactions

Involvement of polar and/or charged residues in the association of TM segments has been evaluated in several other TM peptide model systems, using both *de novo* designed and natural TM sequences. One such peptide, termed MS1, was derived from a dimeric membrane-soluble coiled-coil engineered from the water-soluble yeast transcription factor GCN4.[49] Produced by chemical synthesis, MS1 is entirely hydrophobic throughout a 20-residue segment with the exception of a single Asn side chain, and associates in a monomer-dimer-trimer equilibrium. Mutation of this Asn to Val essentially eliminated oligomerization of MS1, providing evidence that interactions involving a polar Asn side chain can provide a strong driving force for helix–helix association within the membrane. The oligomerization of MS1 variants with a variety of substitutions at its single Asn residue (Gln, Asp, Glu, Lys, Ala, Val, Leu, Ser, Thr) were also examined.[50] MS1 peptides with two polar atoms at the guest site—Asn, Gln, Asp, and Glu—formed stable trimers, whereas residues with one or fewer polar atoms showed a much weaker tendency to associate.

Another *de novo* designed model TM helix based on the GCN4 Leu zipper was also found to be driven to self-associate by a single Asn residue, both in vitro as a staphylococcal nuclease fusion, and within the *E. coli* cytoplasmic membrane as part of a TOXCAT chimera.[51] Here, hydrogen bonding between TM helices, rather than exposure of strongly hydrogen bonding groups to lipids, was proposed as the means by which Asn could drive helix–helix association. Host-guest studies of poly-Leu sequences with various polar residues expressed in the same two fusion systems confirmed that polyleucine sequences with Asn, Asp, Gln, Glu, and His—residues capable of being simultaneously hydrogen bond donors and acceptors—form homo- or heterooligomers.[52] In contrast, polyleucine sequences with Ser, Thr, and Tyr do not associate more than the polyleucine sequence alone.

Ser and Thr, however, can mediate the self-association of model TM peptides in the correct context. TOXCAT chimera-based studies, for example, found that a short polar motif QXXS found in the *E. coli* Tar protein is sufficient to induce stable TM-TM interactions.[54] Also, when high-affinity homo-oligomerizing sequences were selected from a randomized library of TM peptide interfaces generated in the TOXCAT chimera, the two most frequently occurring motifs were SxxSSxxT and SxxxSSxxT.[53] Isosteric mutations of any one of the Ser and Thr residues to non-polar residues abolished oligomerization, indicating that the interaction between these positions is specific and requires an extended motif of Ser and Thr hydroxyl groups to drive strong and specific association through a cooperative network of interhelical hydrogen bonds. Synthetic Ser-Leu model peptides containing a heptad repeat pattern of Ser residues were also shown to self-associate as TM α-helices into ion-conducting channels,[70] and co-operative interaction of two or more appropriately placed Ser side chains per helix to allow formation of a "Ser zipper" that can stabilize TM helix association has been demonstrated in a peptide containing three copies of a Ser(a)Leu(d) heptad motif.[48] In this latter case, however, replacement of the Ser residues with Ala resulted in a small decrease in the stability of the dimer, indicating that packing interactions rather than hydrogen bonds provided the primary driving force for helix–helix interaction.

The interplay between van der Waals and electrostatic interactions in stabilizing TM helix–helix assemblies is an area actively being explored with TM peptides. In collaborative work, we examined the role of polar residues in promoting the self-association of natural TM domains using four naturally occurring TM domain sequences that each contained a Glu or Gln residue (Tnf5/CD40 ligand, C79a/Ig-α, C79b/Ig-β, and Fut3/α-fucosyltransferase—see Table I); and variants of bacteriophage M13 major coat protein TM segment with Asp and Asn at interfacial and non-interfacial positions.[57] The TM peptides were expressed in the TOXCAT chimera, and self-association monitored in the *E. coli* inner membrane. In some cases, a polar mutation significantly stabilized TM helix–helix interactions through the formation of an interhelical H-bond; other instances arose where the strongly polar residues did not enhance, or actually destabilized, the association of the two helices. These findings suggested that a balance prevails between the energetic contributions of van der Waals and electrostatic interactions in controlling nonspecific association of membrane-embedded polar residues.

We went on to examine the relative contributions and interplay between van der Waals contacts and electrostatic interactions in a *de novo* designed highly hydrophobic helix–loop–helix construct termed AI.[71] This construct, with the prototypic sequence K_1KKKKKKFAIAIAIIAWAX$_{19}$AIIAIAIAI-KSPGSKIAIAIAIIAZ$_{44}$AWAIIAIAIAFKKKKKKK$_{62}$, was designed

with "small" (Ala) and "large" (Ile) residues in order to maximize the area of potential van der Waals contacts available for helix–helix contacts within the hairpin sequence, and two guest positions X and Z for incorporation of polar and/or nonpolar residues.[71] With apolar residues at the two "guest" positions, synthetic peptides corresponding to the single TM derived from the AI hairpin formed SDS-resistant dimers on PAGE gels, and formed a strong dimer when expressed in the TOXCAT chimera—an indication that the TM segments in the AI hairpin construct formed helix–helix contacts based on van der Waals packing alone (see Figure 8). The van der Waals-packed AI hairpin construct was subsequently mutagenized to all 20 residues at the X position and expressed in *E. coli*. Analysis of AI hairpin mutants for oligomerization on SDS-PAGE showed that those mutants with Z = Ile and X = Glu, Asp, Gln, Asn, Arg, His, and Lys formed SDS-resistant dimers that corresponded to H-bonded four-helix bundles, while the remaining (e.g., X = Phe, Trp, Leu, Ile, Met, Val, Cys, Tyr, Ala, Thr, Ser, Gly, and Pro) remained monomeric (see Figure 10). Systematic studies of X/Z double mutants indicated that hairpin dimerization to four-helix bundles was the result of the disruption of interactions between the two helices within the AI construct. Overall, our work on the AI hairpin indicated that in situations where hydrophobic van der Waals packing energy between helices is sufficient to prevent significant rotation about the major axes of interacting helices, intrahairpin side chain-side chain H-bond formation will occur mainly when pairs of polar residues are interfacially located and proximal.

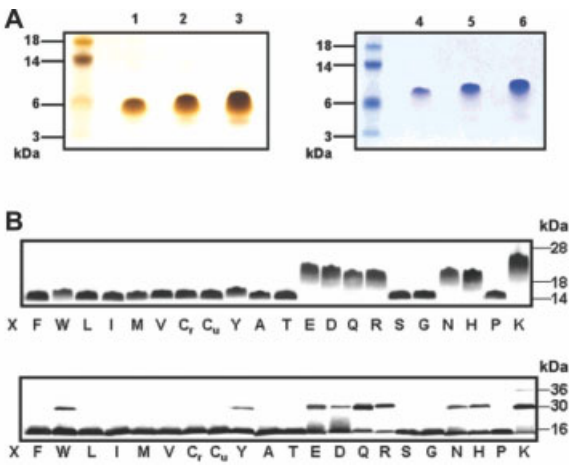

FIGURE 8 Van der Waals interactions pack together the TM segments of the AI hairpin. (A) Self-association of the single TM of the AI hairpin. Purified AI peptide [sequence KKKK-FAIAIAIAI-WAIAIIAIAIAI-KKKKK; 3404.4 Da) analyzed on SDS-PAGE and visualized by silver stain (*left panel*) or Coomassie blue stain (*right panel*). The amounts of peptide loaded are 37.5 ng (lane 1), 75 ng (lane 2), 187.5 ng (lane 3), 1500 ng (lane 4), 3000 ng (lane 5), and 7500 ng (lane 6). The peptide is dimeric at all concentrations. (B) Oligomerization assays for each of the 20 commonly occurring acids at position X (Ile 19) when Z = Ile (Ile 44) is kept constant on the NuPAGE (*top panel*) and Tris-glycine (*bottom panel*) gel systems. The X = C mutant was run under both reducing and non-reducing conditions (C_r and C_u, respectively). Adapted from Johnson et al. Biochemistry 2004, 43, 14361–14369.

The stabilizing effect of interhelical hydrogen bond formation has additionally been demonstrated to have a position-specificity with respect to TM helix ends. In the synthetic MS1 peptide, Asn stabilized self-association to a trimer when placed near the middle of the TM helix, but was neither stabilizing nor destabilizing when located at the interface between the hydrophobic and polar regions of the peptide.[72] The magnitude of stabilization of helix–helix interactions by polar interactions thus depends on the location of the polar residue within the hydrophobic TM segment.

INVESTIGATING INSERTION WITH TM PEPTIDE SEQUENCES

Following the two-stage model,[3] the first step of membrane protein folding is the stable incorporation of TM segments into the apolar environment of the lipid bilayer, with concomitant α-helical structure formation. The precise relationship between TM segment sequence, secondary structure, and bilayer insertion is therefore a central question concerning the biosynthesis of integral membrane proteins. The water solubility of Lys-tagged TM-mimetic peptides in aqueous buffer as well as a wide variety of membrane-mimetics [lipid micelles, small unilamellar vesicles, and organic solvents] was a key development that permitted us to investigate and compare specifically the secondary structural preferences of TM segments in aqueous versus membrane-mimetic environments.[29] Spontaneous partitioning of such peptides into membrane-like environments[29] also allowed for a quantitative determination of the minimum hydrophobicity required of a given peptide for spontaneous integration into a membrane.[29] This "threshold hydrophobicity" (Figure 9) was found to be satisfied by >96% of over 5000 TM segments derived from single- and multi-spanning intrinsic membrane proteins.[29] In related research, the α-helical propensity of each residue was determined in the isotropic nonpolar environment of *n*-butanol, with the finding that the helical propensity of individual side chains scales roughly with their hydrophobicity - and helped to explain, in part, why *β*-promoting residues in soluble proteins can become helix promoters in membranes.[73] From these combined experimentally determined parameters, we developed the automated prediction program *TM Finder* that can successfully locate

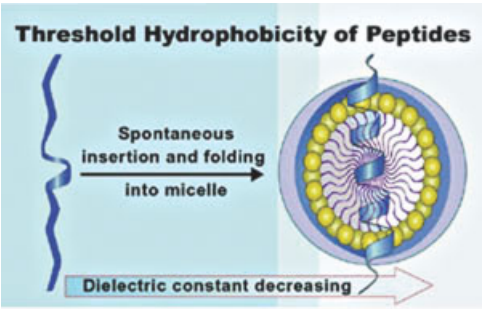

FIGURE 9 "Threshold hydrophobicity" of TM segments. When the average hydrophobic character of a peptide segment is equivalent to just above that of a poly-Ala segment, the segment transfers spontaneously from an aqueous phase to a membrane phase[73,74] and assumes a helical conformation. In the process, aqueous solvation of peptide bonds is replaced by intrahelical H-bonds. This phenomenon is essentially a manifestation of stage 1—the insertion step—of the two-stage model for membrane protein folding.[3]

most TM segments in proteins[76] (a web version of *TM Finder* is available at http://www.bioinformatics-canada.org/TM/).

Applied in vivo, where protein insertion is mediated by other proteins, the phenomenon of "threshold hydrophobicity" is consistent with the selective incorporation of TM segments into the lipid bilayer during the biosynthetic translocation process without additional expenditure of energy.[74] However, work with model TM peptide sequences inserted into the chimeric Lep protein system has revealed that the efficiency of membrane incorporation may depend on factors in addition to the hydrophobicity threshold.[32] Specifically, insertion into the membrane by the translocation machinery depends strongly on the position of polar residues within TM segments,[32] and direct protein–lipid interactions appear to be critical during translocon-mediated membrane insertion.

Lipid Acyl Chain Length and Helix Insertion

The length of lipid acyl chains—and by extension bilayer thickness—has also been shown to influence the insertion behavior of TM peptides in vitro. The TM orientation of a Lys-flanked α-helical model peptide with a 19-residue mixed Leu/Ala hydrophobic sequence with a single Trp residue was found by fluorescence measurements to be destabilized by differences between the width of the bilayer and the length of the hydrophobic sequence (termed hydrophobic mismatch). When the bilayer width exceeded the length of the hydrophobic segment, mismatch induced formation of a non-TM peptide orientation close to the interface of the aqueous and apolar phases,[77] suggesting that variations in TM helix length vs. membrane thickness influence TM insertion. Subsequent work on a variety TM-mimetic peptides consisting of a poly-Leu or poly-Leu/Ala hydrophobic sequence, flanked on both sides by either Trp or Lys residues, has demonstrated that the nature of the flanking residues modulates the peptide response to hydrophobic mismatch in terms of TM segment tilt or adoption of non-TM orientation (reviewed in Ref. 78).

The TM segment of a natural membrane protein that inserts into bilayers—the single-spanning Pf3 coat protein—has also been used to investigate the effect of hydrophobic mismatch on peptide behavior.[79] The insertion efficiency of the protein into membrane-mimetic large unilamellar vesicles (LUVs) was shown to depend on the bilayer thickness, with most efficient insertion under hydrophobic matching conditions. Pf3 mutants with different TM segments were also constructed and generally inserted into LUVs in a mismatch-dependent manner, though longer and less hydrophobic TM segments were most efficiently inserted into thinner bilayers than was expected on the basis of hydrophobic matching.

The effects of hydrophobic mismatch and insertion into model membrane vesicles was further investigated with synthetic Lys-tagged poly-Leu-rich peptides where one or two consecutive Leu residues in the hydrophobic core of the helix were substituted with Ala, Phe, Gly, Ser, Asp, Lys, His, Pro, Gly-Gly, Ser-Ser, Pro-Gly, Pro-Pro, Lys-Lys, or Asp-Asp residues.[80] Peptide behavior was assessed via fluorescence measurements of a Trp residue placed at the center of the sequence. All of the peptides with single substitutions inserted into dioleoylphosphatidylcholine (DOPC) bilayers and adopted a TM orientation. However, in thicker bilayers composed of dierucoylphosphatidylcholine (DEuPC), TM states were destabilized by mismatch between helix length and bilayer thickness. Asp-Asp and Lys-Lys replacements abolished the TM orientation of the peptides in both DOPC and DEuPC vesicles, while Gly-Gly and Ser-Ser substitutions had little effect. In even thicker bilayers, the Ser-Ser substituted peptide was compromised in its ability to form a TM orientation. These studies highlighted the complex relationship between TM peptide sequence, bilayer thickness, and the stability of TM helix insertion, and showed that the effects of polar residue substitutions on hydrophobic helices can be modulated by properties such as mismatch.

Lipid Solvation and Helix Insertion

The importance of protein-lipid interactions in translocon-mediated membrane insertion[32] also raises the possibility of a relationship between biological insertion and the ability of bilayer lipids to solvate individual TM helices. We have begun to investigate the role of lipid solvation in mediating membrane protein folding using the lipid accessible surface

area (LASA) as a measure of the ability of lipid to solvate the surfaces of TM α-helical sequences. In this method, TM α-helical sequences are modeled as energy-minimized monomers, and the surface area accessible to a methylene-sized probe used as a measure of solvation by membrane components. Thus, larger LASAs represent more contact between the lipid-sized probe and the surface of the helix monomer (better helix–lipid packing and solvation), and smaller LASAs represent less contact (poorer packing and solvation). Working on the hypothesis that better lipid solvation of completely hydrophobic TM helices (*viz.* larger LASA) should facilitate insertion compared to hydrophobic TM helices with smaller LASA, we undertook to examine the relationship between LASA and the reported apparent free energy of membrane insertion (ΔG_{app}) of 28 individual 19-residue TM sequences composed solely of Ala and Leu residues in various combinations.[32] Each of these individual model TM peptides was modeled as an energy-minimized monomer and their total surface area exposed to a methylene-sized probe calculated as described.[75] We noted a strong ($R = 0.68$) inverse correlation between (ΔG_{app}) and the LASA of the individual α-helices (see Figure 10), such that helices with smaller LASAs—hence those more poorly solvated by lipid—are less favorable for insertion, whereas insertion may be promoted by larger LASAs.

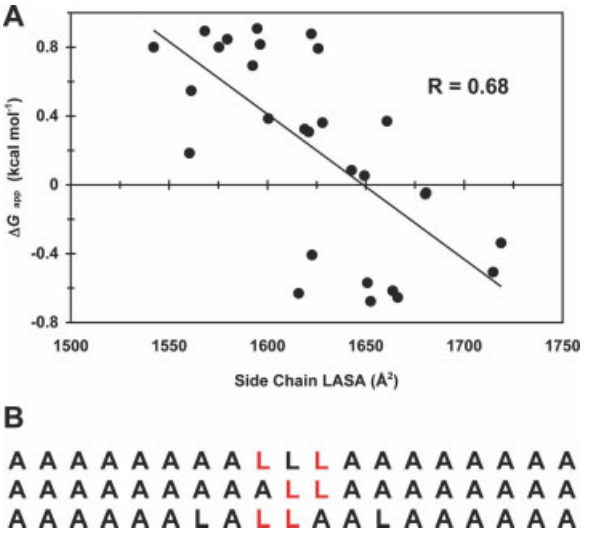

FIGURE 10 (A) Correlation of apparent free energy of insertion (ΔG_{app}) reported in vivo[32] with monomeric helix lipid accessible surface area (LASA). (B) Sequences of hydrophobic segments used to generate energy-minimized α-helical structures for LASA analysis, adapted from Hessa et al. Nature 2005, 433, 377–381. The two red Leu residues are moved symmetrically outwards from their initial positions to generate the 28 sequences analyzed.

Lipid Acyl Chain Length and Helix–Helix Interactions

Helix–helix association also appears to be controlled by hydrophobic mismatch effects in several model peptide systems. Lys-tagged poly-Leu peptides with 11- to 23-residue hydrophobic cores studied by fluorescence spectroscopy in various model membranes showed that mismatches between helix length and bilayer width can control not only membrane location and orientation, but also helix–helix interactions.[81] The self-assembly of poly-Leu/Ala TM-mimetic peptides was also dependent on the lipid environment. In dioleoylphosphatidylcholine (DOPC), where the peptide hydrophobic length approximately matched the bilayer thickness, helix–helix association was found to require peptide:lipid ratios exceeding 1:25. However, self-association of the model TM peptides was promoted by either increasing or decreasing the bilayer thickness, and by adding cholesterol.[82]

However, it may be favorable helix–helix interactions—rather than helix–lipid effects—that modulate the oligomerization response to hydrophobic mismatch. Upon determination of the complete set of thermodynamic parameters for dimer formation by a TM peptide [hydrophobic core $(AALALAA)_3$] in phosphocholine lipid bilayers with different hydrophobic thicknesses (C14-C22) at 5–55°C,[83] stronger dimerization in thicker membranes and at lower temperatures was found to be driven by favorable interactions between helix macrodipoles. The concept of dipole–dipole interaction as a basic driving force of helix dimerization may help to understand the structural and functional modifications of membrane proteins in response to hydrophobic mismatch.

Lipid Solvation and Helix–Helix Interactions

Although favorable helix–helix interactions, rather than unfavorable helix–lipid interactions, may modulate self-assembly in cases of hydrophobic mismatch, experiments on synthetic TM peptides with both natural and *de novo* designed sequences in vitro suggest that unfavorable helix–lipid interactions nevertheless can influence TM peptide self-association. The release of detergent molecules during oligomerization of a peptide corresponding to residues 69–101 of GpA provides a small entropic contribution to stability,[84] implying that the decreased motional freedom of lipid molecules packed against the relatively rigid surfaces of TM helices are entropically unfavorable helix–lipid interactions that favor helix assembly. Poor packing between the lipid fatty acyl chains and irregular TM helix surfaces has also been proposed to drive helix–helix association[85]; helix–helix packing has been demonstrated to be more efficient than helix–lipid

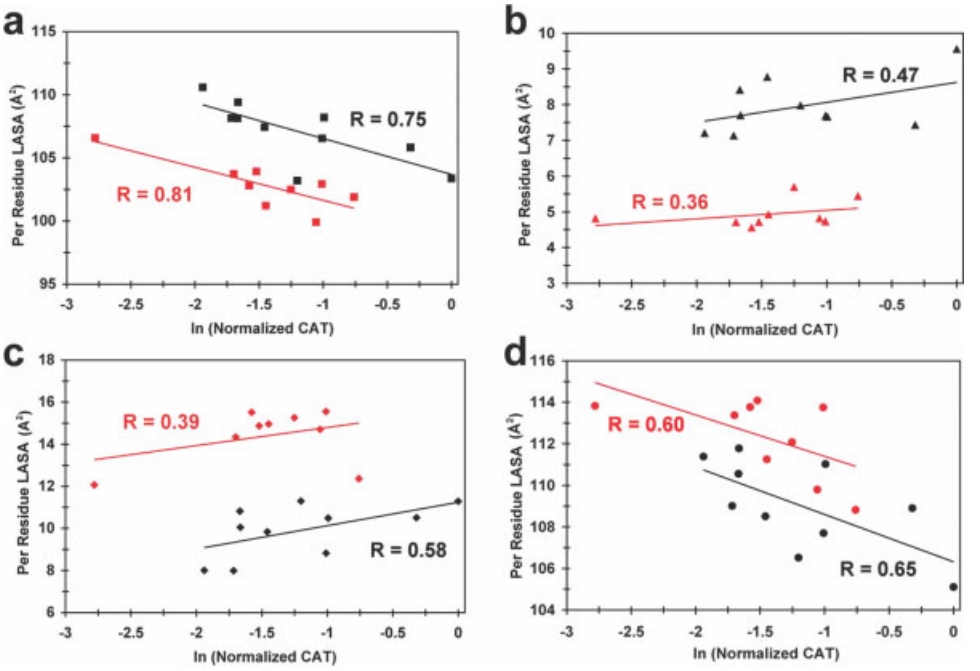

FIGURE 11 Relationship between dimer affinity and helix lipid accessibility. *Per residue* LASA of (A) non-polar, (B) backbone, (C) polar, and (D) all side chain, atoms for WT and mutant GpA (black) and MCP (red) model helices, plotted versus the natural logarithm of TOXCAT signal normalized to wild-type GpA. The *y*-axis scales of the graphs in (A)–(D) have been adjusted to allow for display of both the GpA and MCP data sets. The G83I mutant of GpA is included in the data set, as are all the GG4 motif mutants. The correlation coefficient of each linear fit determined by linear regression is indicated. The correlations observed for per residue non-polar LASA and GpA or MCP affinity shown in (A) have $P = 0.013$, and $P = 0.004$, respectively. Adapted from Johnson et al. Biochemistry 2006, 45, 8507–8515.

packing in synthetic poly-Leu peptides by approximately the volume of a methyl group.[86] Long-trajectory molecular dynamics simulations of a peptide corresponding to the TM region of GpA[87] find that the first step of helix–helix association is driven primarily by the solvent. The TM helices of GpA are thus "pushed" together by the solvent in this simulation—even in the absence of strongly polar groups (see Table I)—allowing for initial helix–helix contacts to form.

Using the surface area of monomeric helices accessible to a methylene-sized probe—LASA—as a measure of lipid solvation, we have investigated the role of lipid solvation in mediating the stability of helix–helix contacts within the bacterial cytoplasmic membrane using peptide fusions to the TOXCAT chimera. We compared the relative in vivo dimer affinities of 18 combinations of Gly, Ala, and Ser "small-xxx-small" mutations in glycophorin A (GpA) and bacteriophage M13 major coat protein (MCP) homodimers (Figure 11) to their monomeric LASA values.[75] A strong inverse correlation was found between non-polar group lipid accessibility and dimer affinity ($R = 0.75$ for GpA, $P = 0.013$; and $R = 0.81$ for MCP, $P = 0.004$, see Figure 11). This observation implies that poor packing/solvation of helix monomers by membrane components may also contribute to the stability of helix–helix association—particularly for those TM segments that form strong dimers in the absence of polar residues.

CONCLUSIONS

The large amount of data concerning the properties of TM segments required for stable insertion into biological membranes—and the motifs and sequence-specific forces that underlie the self-assembly of TM α-helices—underscore the utility of synthetic TM peptides, peptide fusions, and fusion protein systems monitoring self-association in furthering our understanding of the membrane protein folding problem. And, although progress is being made in the structural characterization of membrane proteins, the ease with which TM peptides and peptide fusions can be prepared synthetically and in bacterial expression systems in quantities suitable for biophysical studies suggests that TM peptides will continue be valuable tools in uncovering the principles that dictate the three-dimensional folds of membrane proteins from their amino acid sequences.

R.M.J. held an Ontario Graduate Scholarship and a Research Training Committee award from the Hospital for Sick Children. A.R. holds a post-doctoral award from the CIHR Training Program in "Protein Folding: Principles and Diseases."

REFERENCES

1. Gurrath, M. Curr Med Chem 2001, 8, 1605–1648.
2. Boyd, D.; Schierle, C.; Beckwith, J. Protein Sci 1998, 7, 201–205.
3. Popot, J. L.; Engelman, D. M. Biochemistry 1990, 29, 4031–4037.
4. Popot, J. L.; Engelman, D. M. Annu Rev Biochem 2000, 69, 881–922.
5. White, S. H.; Wimley, W. C. Annu Rev Biophys Biomol Struct 1999, 28, 319–365.
6. White, S. H. Protein Sci 2004, 13, 1948–1949.
7. Oberai, A.; Ihm, Y.; Kim, S.; Bowie, J. U. Protein Sci 2006, 15, 1723–1734.
8. Liao, M. J.; London, E.; Khorana, H. G. J Biol Chem 1983, 258, 9949–9955.
9. Popot, J. L.; Gerchman, S. E.; Engelman, D. M. J Mol Biol 1987, 198, 655–676.
10. Kahn, T. W.; Engelman, D. M. Biochemistry 1992, 31, 6144–6151.
11. Yu, H.; Kono, M.; McKee, T. D.; Oprian, D. D. Biochemistry 1995, 34, 14963–14969.
12. Marti, T. J Biol Chem 1998, 273, 9312–9322.
13. Berkower, C.; Michaelis, S. EMBO J 1991, 10, 3777–3785.
14. Zen, K. H.; McKenna, E.; Bibi, E.; Hardy, D.; Kaback, H. R. Biochemistry 1994, 33, 8198–8206.
15. Schoneberg, T.; Liu, J.; Wess, J. J Biol Chem 1995, 270, 18000–18006.
16. Liu, Y.; Engelman, D. M.; Gerstein, M. Genome Biol 2002, 3, research0054. See: http://genomebiology.com/2002/3/10/research/0054.1.
17. Senes, A.; Gerstein, M.; Engelman, D. M. J Mol Biol 2000, 296, 921–936.
18. Tomich, J. M.; Carson, L. W.; Kanes, K. J.; Vogelaar, N. J.; Emerling, M. R.; Richards, J. H. Anal Biochem 1988, 174, 197–203.
19. Deber, C. M.; Lutek, M. K.; Heimer, E. P.; Felix, A. M. Pept Res 1989, 2, 184–188.
20. Glover, K. J.; Martini, P. M.; Vold, R. R.; Komives, E. A. Anal Biochem 1999, 272, 270–274.
21. Fischer, W. B.; Pitkeathly, M.; Wallace, B. A.; Forrest, L. R.; Smith, G. R.; Sansom, M. S. Biochemistry 2000, 39, 12708–12716.
22. Frank, S.; Kammerer, R. A.; Hellstern, S.; Pegoraro, S.; Stetefeld, J.; Lustig, A.; Moroder, L.; Engel, J. Biochemistry 2000, 39, 6825–6831.
23. Cunningham, F.; Deber, C. M. Methods, in press.
24. Li, S. C.; Deber, C. M. Nat Struct Biol 1994, 1, 368–373.
25. Li, S. C.; Kim, P. K.; Deber, C. M. Biopolymers 1995, 35, 667–675.
26. Deber, C. M.; Li, S. C. Biopolymers 1995, 37, 295–318.
27. Liu, L. P.; Li, S. C.; Goto, N. K.; Deber, C. M. Biopolymers 1996, 39, 465–470.
28. Liu, L. P.; Deber, C. M. Biochemistry 1997, 36, 5476–5482.
29. Liu, L. P.; Deber, C. M. J Biol Chem 1998, 273, 23645–23648.
30. Wimley, W. C.; White, S. H. Biochemistry 2000, 39, 4432–4442.
31. Lemmon, M. A.; Flanagan, J. M.; Hunt, J. F.; Adair, B. D.; Bormann, B. J.; Dempsey, C. E.; Engelman, D. M. J Biol Chem 1992, 267, 7683–7689.
32. Hessa, T.; Kim, H.; Bihlmaier, K.; Lundin, C.; Boekel, J.; Andersson, H.; Nilsson, I.; White, S. H.; von Heijne, G. Nature 2005, 433, 377–381.
33. Russ, W. P.; Engelman, D. M. Proc Natl Acad Sci USA 1999, 96, 863–868.
34. Gurezka, R.; Langosch, D. J Biol Chem 2001, 276, 45580–45587.
35. Lis, M.; Blumenthal, K. Biochem Biophys Res Commun 2006, 339, 321–324.
36. Schneider, D.; Engelman, D. M. J Biol Chem 2003, 278, 3105–3111.
37. Melnyk, R. A.; Partridge, A. W.; Deber, C. M. Biochemistry 2001, 40, 11106–11113.
38. Therien, A. G.; Deber, C. M. J Mol Biol 2002, 322, 583–550.
39. Rath, A.; Melnyk, R. A.; Deber, C. M. J Biol Chem 2006, 281, 15546–15553.
40. Go, M. Y.; Kim, S.; Partridge, A. W.; Melnyk, R. A.; Rath, A.; Deber, C. M.; Mogridge, J. J Mol Biol 2006, 360, 145–156.
41. DeGrado, W. F.; Gratkowski, H.; Lear, J. D. Protein Sci 2003, 12, 647–665.
42. MacKenzie, K. R.; Prestegard, J. H.; Engelman, D. M. Science 1997, 276, 131–133.
43. Deber, C. M.; Khan, A. R.; Li, Z.; Joensson, C.; Glibowicka, M.; Wang, J. Proc Natl Acad Sci USA 1993, 90, 11648–11652.
44. Melnyk, R. A.; Partridge, A. W.; Deber, C. M. J Mol Biol 2002, 315, 63–72.
45. Pinto, L. H.; Dieckmann, G. R.; Gandhi, C. S.; Papworth, C. G.; Braman, J.; Shaughnessy, M. A.; Lear, J. D.; Lamb, R. A.; DeGrado, W. F. Proc Natl Acad Sci USA 1997, 94, 11301–11306.
46. Arkin, I. T.; Adams, P. D.; MacKenzie, K. R.; Lemmon, M. A.; Brunger, A. T.; Engelman, D. M. EMBO J 1994, 13, 4757–4764.
47. Simmerman, H. K.; Kobayashi, Y. M.; Autry, J. M.; Jones, L. R. J Biol Chem 1996, 271, 5941–5946.
48. North, B.; Cristian, L.; Fu Stowell, X.; Lear, J. D.; Saven, J. G.; Degrado, W. F. J Mol Biol 2006, 359, 930–939.
49. Choma, C.; Gratkowski, H.; Lear, J. D.; DeGrado, W. F. Nat Struct Biol 2000, 7, 161–166.
50. Gratkowski, H.; Lear, J. D.; DeGrado, W. F. Proc Natl Acad Sci USA 2001, 98, 880–885.
51. Zhou, F. X.; Cocco, M. J.; Russ, W. P.; Brunger, A. T.; Engelman, D. M. Nat Struct Biol 2000, 7, 154–160.
52. Zhou, F. X.; Merianos, H. J.; Brunger, A. T.; Engelman, D. M. Proc Natl Acad Sci USA 2001, 98, 2250–2255.
53. Dawson, J. P.; Weinger, J. S.; Engelman, D. M. J Mol Biol 2002, 316, 799–805.
54. Sal-Man, N.; Gerber, D.; Shai, Y. J Biol Chem 2005, 280, 27449–27457.
55. Partridge, A. W.; Melnyk, R. A.; Yang, D.; Bowie, J. U.; Deber, C. M. J Biol Chem 2003, 278, 22056–22060.
56. George, S. R.; Ng, G. Y.; Lee, S. P.; Fan, T.; Varghese, G.; Wang, C.; Deber, C. M.; Seeman, P.; O'Dowd, B. F. J Pharmacol Exp Ther 2003, 307, 481–489.
57. Dawson, J. P.; Melnyk, R. A.; Deber, C. M.; Engelman, D. M. J Mol Biol 2003, 331, 255–262.

58. Lemmon, M. A.; Flanagan, J. M.; Treutlein, H. R.; Zhang, J.; Engelman, D. M. Biochemistry 1992, 31, 12719–12725.
59. Brosig, B.; Langosch, D. Protein Sci 1998, 7, 1052–1056.
60. Wickner, W. Biochemistry 1988, 27, 1081–1086.
61. Wang, C.; Deber, C. M. J Biol Chem 2000, 275, 16155–16159.
62. Melnyk, R. A.; Kim, S.; Curran, A. R.; Engelman, D. M.; Bowie, J. U.; Deber, C. M. J Biol Chem 2004, 279, 16591–16597.
63. Johnson, R. M.; Rath, A.; Deber, C. M. Biochem Cell Biol 2006, 84, 1006–1012.
64. Schneider, D.; Engelman, D. M. J Mol Biol 2004, 343, 799–804.
65. Kochendoerfer, G. G.; Salom, D.; Lear, J. D.; Wilk-Orescan, R.; Kent, S. B.; DeGrado, W. F. Biochemistry 1999, 38, 11905–11913.
66. Bauer, C. M.; Pinto, L. H.; Cross, T. A.; Lamb, R. A. Virology 1999, 254, 196–209.
67. Howard, K. P.; Lear, J. D.; DeGrado, W. F. Proc Natl Acad Sci USA 2002, 99, 8568–8572.
68. Stouffer, A. L.; Nanda, V.; Lear, J. D.; DeGrado, W. F. J Mol Biol 2005, 347, 169–179.
69. Salom, D.; Hill, B. R.; Lear, J. D.; DeGrado, W. F. Biochemistry 2000, 39, 14160–14170.
70. Lear, J. D.; Wasserman, Z. R.; DeGrado, W. F. Science 1988, 240, 1177–1181.
71. Johnson, R. M.; Heslop, C. L.; Deber, C. M. Biochemistry 2004, 43, 14361–14369.
72. Lear, J. D.; Gratkowski, H.; Adamian, L.; Liang, J.; DeGrado, W. F. Biochemistry 2003, 42, 6400–6407.
73. Liu, L. P.; Deber, C. M. Biopolymers 1998, 47, 41–62.
74. Deber, C. M.; Liu, L.-P.; Wang, C.; Goto, N. K.; Reithmeier, R. A. In Peptide–Lipid Interactions, Current Topics in Membranes; Simon, S. A.; McIntosh, T. J., Eds.; Academic Press: San Diego; 2002; pp 459–473.
75. Johnson, R. M.; Rath, A.; Melnyk, R. A.; Deber, C. M. Biochemistry 2006, 45, 8507–8515.
76. Deber, C. M.; Wang, C.; Liu, L. P.; Prior, A. S.; Agrawal, S.; Muskat, B. L.; Cuticchia, A. J. Protein Sci 2001, 10, 212–219.
77. Ren, J.; Lew, S.; Wang, Z.; London, E. Biochemistry 1997, 36, 10213–10220.
78. de Planque, M. R.; Killian, J. A. Mol Membr Biol 2003, 20, 271–284.
79. Ridder, A. N.; van de Hoef, W.; Stam, J.; Kuhn, A.; de Kruijff, B.; Killian, J. A. Biochemistry 2002, 41, 4946–4952.
80. Caputo, G. A.; London, E. Biochemistry 2003, 42, 3275–3285.
81. Ren, J.; Lew, S.; Wang, J.; London, E. Biochemistry 1999, 38, 5905–5912.
82. Sparr, E.; Ash, W. L.; Nazarov, P. V.; Rijkers, D. T.; Hemminga, M. A.; Tieleman, D. P.; Killian, J. A. J Biol Chem 2005, 280, 39324–39331.
83. Yano, Y.; Matsuzaki, K. Biochemistry 2006, 45, 3370–3378.
84. Fisher, L. E.; Engelman, D. M.; Sturgis, J. N. J Mol Biol 1999, 293, 639–651.
85. Lee, A. G. Biochim Biophys Acta 2004, 1666, 62–87.
86. Mall, S.; Broadbridge, R.; Sharma, R. P.; East, J. M.; Lee, A. G. Biochemistry 2001, 40, 12379–12386.
87. Henin, J.; Pohorille, A.; Chipot, C. J Am Chem Soc 2005, 127, 8478–8484.