

## Sequence Context Strongly Modulates Association of Polar Residues in Transmembrane Helices

Jessica P. Dawson<sup>1†</sup>, Roman A. Melnyk<sup>2,3†</sup>, Charles M. Deber<sup>2,3\*</sup> and Donald M. Engelman<sup>1\*</sup>

<sup>1</sup>Department of Molecular Biophysics and Biochemistry, Yale University, Bass Centre Room 420A, 266 Whitney Avenue, New Haven, CT 06520-8114, USA

<sup>2</sup>Division of Structural Biology and Biochemistry, Research Institute, Hospital for Sick Children, Toronto, Ont. M5G 1X8, Canada

<sup>3</sup>Department of Biochemistry, University of Toronto, Toronto Ont. M5S 1A8, Canada

Polar residues are capable of mediating the association of membrane-embedded helices through the formation of side-chain/side-chain interhelical hydrogen bonds. However, the extent to which native van der Waals packing of the residues surrounding the polar locus can enhance, or interfere with, the interaction of polar residues has not yet been studied. We examined the propensities of four polar residues (aspartic acid, asparagine, glutamic acid, and glutamine) to promote self-association of transmembrane (TM) domains in several biologically derived sequence environments, including (i) four naturally occurring TM domains that contain a Glu or Gln residue (Tnf5/CD40 ligand, C79a/Ig- $\alpha$ , C79b/Ig- $\beta$ , and Fut3/ $\alpha$ -fucosyltransferase); and (ii) variants of bacteriophage M13 major coat protein TM segment with Asp and Asn at interfacial and non-interfacial positions. Self-association was quantified by the TOXCAT assay, which measures TM helix self-oligomerization in the *Escherichia coli* inner membrane. While an appropriately placed polar residue was found in several cases to significantly stabilize TM helix–helix interactions through the formation of an interhelical hydrogen bond, in other cases the strongly polar residues did not enhance the association of the two helices. Overall, these results suggest that an innate structural mechanism may operate to control non-specific association of membrane-embedded polar residues.

© 2003 Elsevier Ltd. All rights reserved

**Keywords:** transmembrane helix; helix–helix interactions; M13 major coat protein; TOXCAT assay; polar residues in membranes

\*Corresponding authors

### Introduction

The folding of  $\alpha$ -helical membrane proteins can be conceptualized as a process including two thermodynamically distinct steps, involving (i) the formation of independently stable transmembrane (TM) helices; and (ii) the specific TM helix–helix interactions that give rise to higher-order structures.<sup>1</sup> In the study of helix–helix interactions, the two forces considered in step (ii) to be the major contributors are specific packing interactions (van der Waals contacts) and hydrogen bonding between polar residues. While earlier studies focused on the contribution of van der Waals interactions to the second stage of membrane protein

folding, much recent research has been devoted to determining the role of H-bonding between TM helices. For example, using model TM segments, it was shown by SDS-PAGE gel assays that sequences containing a single strongly polar residue (e.g. Gln, Asn, Glu, Asp) self-associate, while those having a single Ser or Thr do not.<sup>2–5</sup> However, in a separate study, it was found that multiple Ser and Thr residues interacting cooperatively across the helical interface can provide sufficient interaction energy to drive inter-helical association.<sup>6</sup>

Since single strongly polar residues, or multiple weakly polar residues in a TM helix, can be sufficient to drive association in some simplified sequences, polar residues in TM domains could form strong, and potentially deleterious interhelical H-bonds. Thus, a polar substitution within a TM helix of a functional helical bundle could give rise to the formation of a non-native interhelical H-bond when partnered with a wild-type polar

† These authors contributed equally to this work.

Abbreviations used: TM, transmembrane; MCP, major coat protein; CAT, chloramphenicol acetyl transferase.

E-mail addresses of the corresponding authors: [deber@sickkids.ca](mailto:deber@sickkids.ca); [donald.engelman@yale.edu](mailto:donald.engelman@yale.edu)

residue in an adjacent helix, yielding a potentially misfolded and non-functional protein. Indeed, mutations from non-polar-to-polar residues in TM domains have been implicated in a number of diseases involving proteins such as CFTR (cystic fibrosis)<sup>7</sup> and neu-ErbB (cancer).<sup>8</sup> In CFTR, a Val-to-Asp mutation in TM4 results in the formation of a non-native H-bond with an endogenous Gln in TM3, which is thought to hinder protein function.<sup>7</sup> In a conceptually related investigation, the interaction between Gln and Asp both present in a single CFTR TM4 helix forms only in the context of the wild-type TM4 sequence, but not in a randomized hydrophobic sequence.<sup>9</sup> In the case of neu-erbB-2, a mutation from Val-to-Glu constitutively activates the receptor, and as with CFTR, the location of the Glu within the sequence was found to be critical for receptor activation.<sup>10</sup> Subsequent research showed that although Glu is involved in strong H-bonding,<sup>11</sup> it does not promote homo-oligomerization of the Glu-containing TM domain.<sup>12</sup> Interestingly, the Glu residue activates the receptor only in the presence of a GG4 motif,<sup>10,13</sup> suggesting that the interface must be pre-disposed to interaction.

An emerging paradigm from these studies is that polar residues differ in their effects depending upon where they occur within the sequence, which suggests that not all mutations to polar residues necessarily potentiate the formation of aberrant H-bonds. It is evident that further studies on the interplay between packing and H-bonding in the context of the membrane are required in order to understand membrane protein folding and the role of polar residues in manifesting disease phenotypes. Though it is established<sup>2-5</sup> that strongly polar residues do not necessarily require specific packing at the interface for strong association, the extent to which the surrounding residues can interfere with, or enhance, the interaction of these residues has not yet been addressed. Here, we sought to examine both van der Waals packing interactions and H-bonding in concert to determine when and how they influence each other. Using four distinct, biologically derived TM segments, we show here that polar residues in TM domains display a wide spectrum of effects, demonstrating that the participation of polar residues on TM helix-helix interactions can be attenuated by their local sequence and packing context.

## Results

### Analysis of polar residue distribution in TM helices

The likelihood of two polar residues from two neighboring helices in a bilayer engaging in a side-chain/side-chain hydrogen bond is dependent upon both the frequency and location of the polar groups in these TM domains. Initially, we sought to examine these parameters by analyzing

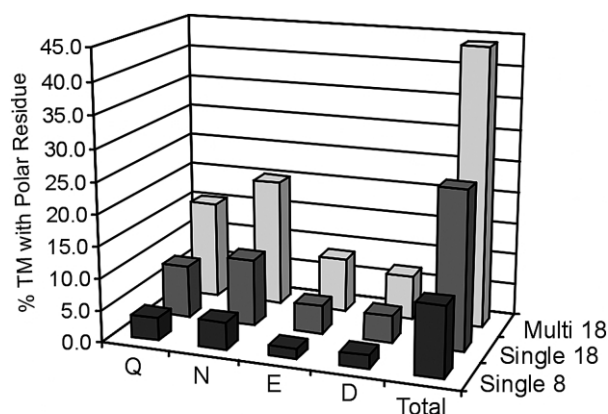
the composition of TM segments to determine how often they contain strongly polar residues, and furthermore, where they occur within each TM helix. The TM helices that would be at the greatest risk of undergoing such non-native associations would be those from single-span membrane proteins, which are not tethered to their interacting partners during biosynthesis as are helices in multi-spanning membrane proteins.

To determine the potential for such interactions, previously compiled databases<sup>14</sup> were segregated into single and multi-spanning TM helices and the percentage of TM domains that contain the strongly polar residues Gln, Asn, Glu, and/or Asp were determined. These two databases were homology-purged, and each TM was reduced to its most hydrophobic 18 residues, resulting in 12,743 TMs from multi-spanning membrane proteins and 1410 TMs from single-spanning membrane proteins. A third database of single-spanning TM helices was then created, wherein the 18-residue window was further reduced to the central eight residues. These three databases (viz. all multi-spanning TMs, all single-spanning TMs, and the central eight residues of single-spanning TMs) were probed for the number(s) of TMs containing Gln, Asn, Glu, and Asp. Unlike previous work, which looked for the percentage of each residue in TM domains, this study calculated the percent of TMs that contain a specific polar residue. The total percent of TMs containing Gln, Asn, Glu, and Asp was also determined.

As seen in [Figure 1](#), approximately twice as many TMs from multi-spanning membrane proteins contain polar residues when compared to single-spanning TMs. This is to be expected because many multi-spanning membrane proteins require polar residues in the performance of their specific functions as channels and transporters. Interestingly, ca 25% of all TMs from single-spanning membrane proteins contain a strongly polar residue. To determine if these polar residues are clustered near the ends of the single-spanning TMs, the central eight residues were searched for strongly polar residues. The results in [Figure 1](#) indicate that the strongly polar residues are evenly distributed across the helix because the occurrences within the central eight residues are almost exactly 50% of the total occurrences in the corresponding 18-residue TMs. These findings confirm that strongly polar residues are present in significant quantities in the central core of TM domains where helix-helix interactions may be maximal, and thus provide a biologically relevant context to the question: do the residues surrounding a strongly polar residue modulate its ability to engage in H-bonds?

### Glutamine/glutamate-containing sequences in TM domains

To determine initially whether strongly polar residues necessarily drive TM helix-helix association



	Q	N	E	D	Total
Multi 18	15.5	20.2	8.5	7.0	44.0
Single 18	8.3	10.6	4.5	4.3	25.0
Single 8	3.6	4.3	1.6	2.1	11.0

**Figure 1.** Percentage of TM domains containing Gln, Glu, Asp, and/or Asn residues. The bar graph denotes the percentage of TM domains containing each of the four polar residues (Gln, Glu, Asp, Asn). Multi 18 and Single 18 are the results from a database of the 18 most hydrophobic residues in the TMs of multi-spanning and single-spanning TM helices. Single 8 are the results from the central eight residues derived directly from the Single 18 database. Total denotes the number of TM domains containing at least one strongly polar residue. The numerical data are shown below the graph.

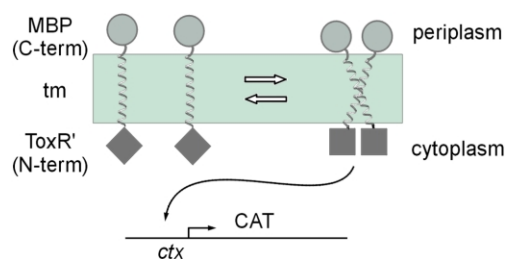
in bilayers, the homo-oligomerization of four TM domains containing a wild-type Glu or Gln (Figure 2) was determined using the TOXCAT system, which measures TM helix homo-oligomerization in the *Escherichia coli* inner membrane (Figure 3). Interactions between the TM domains drive expression of chloramphenicol acetyl transferase (CAT), which renders the bacteria resistant to

### TOXCAT Constructs

```

MCPwt ...nrrarYIGYAWAMVVVIIVGATIGKLF1i1inps...
C79a      IITAEGIILLFCAVPPGTLFF
C79b      GIIMIQTLLIILFIIIVPIFLL
C79b QA   GIIMIATLLIILFIIIVPIFLL
Fut3      LAALLFQLLVAVCFFSYL
Fut3 QA   LAALLFALLVAVCFFSYL
Tnf5      YLLTVFLITQMIGSLFAVYL
Tnf5 QA   YLLTVFLITAMIGSALFAVYL
Tnf5 NP   YLLTVFLIVQMIGALFAVYL
  
```

**Figure 2.** The wild-type sequence of M13 major coat protein. The highlighted residues were each mutated to Asp and Asn to yield a total of eight variants. The lower-case flanking residues are present in all of the TOXCAT constructs. The remaining TOXCAT constructs contain a naturally occurring polar residue (Gln or Glu). Mutation of the Gln to Ala was made in three sequences, and in Tnf5 a mutation (termed NP) of two neighboring polar residues (Thr and Ser) to two non-polar residues (Val and Ala) was also created. Sequences are aligned by their Gln or Glu positions.

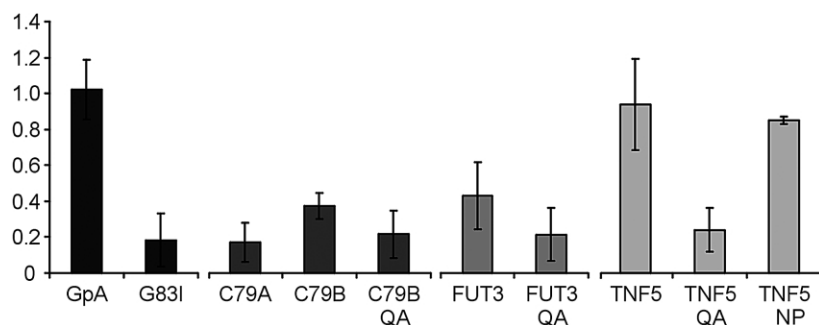


**Figure 3.** TOXCAT assay. Oligomerization of the TM domains drives dimerization of the cytoplasmic ToxR domains, which are then able to activate transcription at the *ctx* promoter that controls expression of chloramphenicol acetyl transferase (CAT).

chloramphenicol (CAM). An increase in TM helix association is manifested as an increase in CAT activity. Extents of TM helix association in the TOXCAT assay were normalized to the standard of the tightly-associated glycoprotein A (GpA) wild-type dimer.<sup>15,16</sup> Also, the equivalent expression of all TOXCAT constructs was verified each time CAT assays were performed. Of the four proteins chosen, one (Tnf5/CD40 ligand) is a homo-trimer,<sup>17</sup> two (C79a/Ig- $\alpha$  and C79b/Ig- $\beta$ ) form a hetero-dimer pair,<sup>18-20</sup> while for the last (Fut3/ $\alpha$ -fucosyltransferase), there is no evidence for oligomerization. The specific role of the TM domain in each of these interactions has not been previously examined. In the case of Tnf5, a truncated soluble form was found to be a homo-trimer, and for C79a and b the interaction occurs through a disulfide bond.

Despite the fact that all four constructs contained either a native Glu or Gln residue, C79a, C79b and Fut3 gave weak-to-moderate dimer signals, whereas Tnf5 interacted as strongly as GpA wild-type (Figure 4). One possible explanation for the higher association of Tnf5 relative to the other sequences might be that in addition to the Gln, the two other polar residues, Ser and Thr, near the Gln in Tnf5, are increasing the number of H-bonds between the helices. We tested the role of the Ser and Thr in the association by mutating them to Ala and Val, respectively. There was no significant change in association for the double mutant (Tnf5 NP in Figure 4), indicating that it is the sequence context of the Gln that produces the strong association, rather than additional polar residues. To determine if the association seen for Tnf5 requires the native Gln residue, it was mutated to Ala; as seen in Figure 4, mutation of this Gln caused a significant decrease in interaction.

Although C79a, C79b and Fut3 gave weaker signals than Tnf5, it is possible that the Gln or Glu in these sequences is providing some stabilization energy. To determine if the weak association for these sequences could be disrupted still further, Gln-to-Ala mutations of C79b and Fut3 were constructed. C79b was chosen over C79a because its signal is slightly higher and therefore disruptions should be easier to detect. However, neither the



**Figure 4.** CAT activity measurements of the single-spanning TM domains with a naturally occurring Gln or Glu residue. The bars indicate the amount of CAT activity normalized to wild-type GpA. Error bars represent the standard deviation of at least four measurements for each construct, except TNF NP (three measurements). QA corresponds to a Gln to Ala mutation; NP is the double mutation of Ser and Thr to Ala and Val, respectively.

mutation in C79b nor in Fut3 produced a statistically significant drop in signal, indicating that the Gln in these sequences is not a major factor in stabilizing the association.

#### Effect of Asp and Asn on M13 major coat protein TM helix association

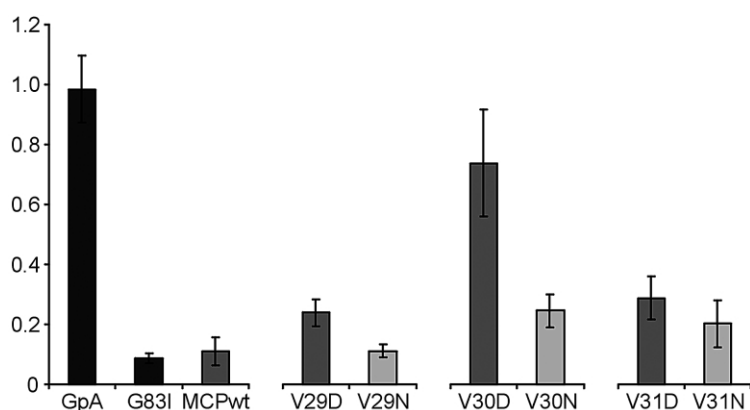
The major coat protein (MCP) TM segment is an attractive model system to study the effects of point mutations on TM domain association since its wild-type MCP exists as a metastable homodimer, both in detergent micelles<sup>21,22</sup> and within the *E. coli* inner membrane.<sup>23</sup> As a result, in randomized mutagenesis experiments, point mutations that stabilize association occur with a comparable frequency to those that destabilize association, as has been seen from extensive mutagenesis of both the full-length protein and synthetic TM peptides.<sup>21–23</sup> In this framework, polar substitutions that result in an interhelical H-bond are expected to produce an increase in the dimer population. If structures featuring interhelical H-bonds represent the lowest energy conformation(s) of polar residue-containing TM helices, as previous studies have suggested,<sup>2–5</sup> then “walking” polar residues around one full turn of the MCP helix should result in an equivalent stabilization of the dimeric conformation in all cases. If, however, there are positional or context-specific considerations that affect the strength of interhelical H-bonds, then such a walk should give rise to a range of interaction energies.

Each position in the central Val-rich core (V<sub>29</sub>V<sub>30</sub>V<sub>31</sub>I<sub>32</sub>) of the MCP TM segment was substituted with Asp and Asn residues to generate eight variant TM peptides along with the MCP wild-type segment. For these experiments, Asp and Asn were chosen over Glu and Gln because previous reports showed that an Asp mutation in the Val-rich core of MCP prevented homooligomerization.<sup>22</sup> To identify whether mutations in the TM domain of MCP cause an increase or decrease in association, migration properties of MCP peptides were analyzed on SDS-PAGE. However, instead of producing distinct monomer and dimer species, the MCP peptides formed smeared

or shifted bands that could not be reliably interpreted. This result was observed previously for polar mutations of GpA. During saturation mutagenesis of the TM domain of GpA, it was found that all polar residues disrupted helix association when using SDS-PAGE to measure the monomer/dimer equilibrium.<sup>16</sup> Several of these polar mutations that were later measured in the *in vivo* TOXCAT system were not disruptive.<sup>24</sup> These results suggested that SDS gels can be an unreliable reporter of helix oligomerization when polar amino acids are present within the TM domain.

Consequently, we tested for the homo-oligomerization of the same TM domains in a biological membrane using the TOXCAT assay. To guide the correct orientation of TOXCAT chimeras within the *E. coli* inner membrane, the “positive inside” rule<sup>25</sup> was considered when choosing which residues to include from the TM segment. Inspection of the MCP TM sequence used in previous studies<sup>22</sup> revealed that several charged residues present at the C-terminal end could misorient the chimera or abrogate insertion if included. Thus, the C-terminal residues (KKFTSK) of the peptides were omitted in the design of the MCP TOXCAT chimera (Figure 2). The membrane insertion and correct orientation of each construct was tested by protease sensitivity in a spheroplast assay, as described<sup>24</sup> (data not shown). All constructs were judged to be inserted correctly into the membrane except for I32D and I32N. These two constructs were either inserted backwards, or not inserted at all.

All TOXCAT constructs were equivalently expressed and the reported CAT activity of each construct is an average of at least four separate measurements, each normalized to GpA wild-type (Figure 5). While the CAT activity signal for the moderately stable MCP wild-type was expected to be lower than the highly stable GpA wild-type, the observed value was significantly lower than expected. The signal for MCP wild-type fell in a range comparable to that of GpA G83I, a mutation known to disrupt helix association when evaluated as a Staphylococcal nuclease fusion protein in SDS-PAGE assays.<sup>15,16</sup> This “lower-than-expected” association of MCP in the TOXCAT system may be



**Figure 5.** CAT activity measurements used to assess the degree of association of each construct. The bars indicate the amount of CAT activity normalized to the wild-type GpA dimer. Both wild-type GpA and its disruptive mutant G83I are shown for comparison. Error bars represent the standard deviation of at least four measurements for each construct.

a result of deleting the C-terminal residues FTSK, which were previously shown to play a role in stabilizing the MCP dimer (unpublished data). The observed metastability of the MCP dimer has been reasoned to be a necessary requirement for phage viability, since mutants that significantly enhanced or decreased the MCP wild-type TM dimer were not viable.<sup>21,22</sup>

Despite the low TOXCAT values for MCP, the mutant constructs exhibited a range of helix association strengths with Asp mutations producing a higher association than Asn at the same position (Figure 5). Surprisingly, most of the mutations produced no (V29N) to small increases (V29D, V30N, V31D and V31N) in helix association. Only the V30D mutation produced a helix interaction close to that of GpA wild-type.

## Discussion

Compared to soluble proteins, the hydrophobic bilayer-embedded regions of membrane proteins contain fewer polar residues. However, despite the overwhelming predominance of hydrophobic residues within the lipid bilayer, analyses of TM segments reveal that the polar residues Asp, Asn, Glu, and Gln, constitute 4–6% of the total amino acid composition in TM helices.<sup>14,26</sup> If these residues are evenly distributed, this translates to an average of ~1 such hydrophilic residue per typical 20–25 residue TM  $\alpha$ -helix. Since these polar residues are able to drive strong helix associations in the absence of detailed packing,<sup>2–5</sup> this observation indicates that there are many polar residues available to engage in strong, potentially non-native, interactions with other TM helices. Here, we sought to examine hydrogen bonding interactions in the presence of the specific packing (van der Waals) interactions responsible for native helix–helix association to determine (i) the situations in which polar associations occur; and (ii) whether the associations between polar residues in TM helices are a function of their context in the presence of native packing effects.

## Analysis of polar residue distribution in TM segments

Since a TM helix from a single-spanning membrane protein may have a greater freedom of association than a helix within a multi-spanning membrane protein, it is more likely to engage in interactions with other TM domains. Therefore, the occurrence of, and interaction between, strongly polar residues in single-spanning membrane proteins, is an area of particular interest. As shown in Figure 1, nearly half of all TM domains from multi-spanning membrane proteins contain a strongly polar residue. This number only drops to 25% for single-spanning membrane proteins, perhaps because many if not most of such proteins participate in oligomeric complexes. Our results further show that 11% of all single-spanning membrane proteins contain either a Gln, Glu, Asp or Asn in the central region of their TM segment. If these polar residues were prone to interactions with other polar residues that frequently proved detrimental to the organism, then it is unlikely that they would be present in so many TM helices. One must also consider the possibility that these TM domains could be expressed across a range of organisms and cell types, such that several different TM helices with a strongly polar residue would rarely be present in the same membrane at the same time. If the opportunity for improper hetero-oligomerization is rare, then unregulated homo-oligomerization would be a greater concern. However, as shown here, homo-oligomerization of TM domains containing a strongly polar residue is indeed regulated by the sequence environment.

## Effects of native polar residues in TM segment assembly

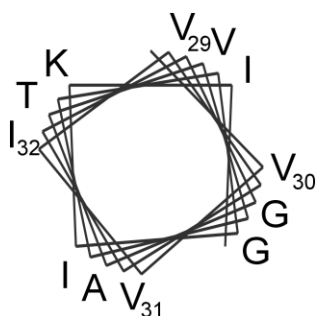
Of the three Gln-containing sequences that we examined experimentally, C79b and Fut3 have similar interaction strengths, while Tnf5 is approximately twice that, near the value of GpA wild-type. In Tnf5, the mutation of Glu-to-Ala shows that the polar residue is primarily responsible for mediating the observed association. Why does the Gln in

Tnf5 cause a substantial increase in association over the Gln-to-Ala mutation, but no increase in C79b and Fut3? Examination of the Tnf5 TM domain reveals the presence of two other polar residues (Ser and Thr) that could be further H-bonded with the Gln and thereby strengthen the association. However, mutation of these two residues (Tnf5 NP) produced no detectable change in the association, indicating that it is indeed the surrounding non-polar residues that are creating the difference in association between the sequences.

### Effect of Asp and Asn substitutions on Ff phage major coat protein association

To further examine the interplay between packing and H-bonding, a series of Asp and Asn mutations were introduced into the Val-rich core (V<sub>29</sub>V<sub>30</sub>V<sub>31</sub>I<sub>32</sub>) of the MCP TM, for which there is extensive information regarding its association both *in vivo* and *in vitro*.<sup>21–23</sup> Questions this study addressed were: (i) do H-bonds alone dictate the lowest energy conformation for TM helices containing polar residues? and (ii) do carboxamides and carboxyl side-chains produce comparable effects on interhelical assembly?

In membranes (TOXCAT assay), polar substitutions in the central Val-rich core of the MCP TM gave rise to a wide spectrum of effects, further supporting a context-specific effect of polar residues in TM segments (Figure 5). In agreement with previous studies, there were several examples where a polar substitution resulted in an increase in dimer stability. In particular, the mutation V30D resulted in a ~fivefold increase in CAT activity over MCP wt, indicating the presence of a particularly strong interhelical H-bond. This mutation, which places a polar residue at the proposed dimeric interface, above a GG4 motif (Figure 6), may result in a situation where H-bond formation can occur ideally with no disruption of the existing packing interactions. If this is the case, the result-



**Figure 6.** Helical wheel diagram of the MCP TM segment assuming a periodicity of 3.9 residues per turn. In the case of helices interacting with a right-handed crossing-angle (as with the GpA homodimer), the apparent number of residues per turn is elevated by projection along the helix axis from 3.6 to 3.9 residues per turn.<sup>17</sup> In the representation shown, V30 resides directly above the interfacial GG4 motif.

ing dimer will benefit synergistically from contributions of both van der Waals interactions and H-bonding. However, at other positions (V29) there was no, or little, change in MCP dimer stability seen upon the introduction of a polar residue, indicating either the absence of any H-bond formation, only very weak bonding, or that an alternative structure is formed with an association energy approximating the native interaction.

An unexpected discovery from the Asp and Asn walk of the MCP TM segment was that in certain cases a polar substitution could actually prevent TM helix insertion into the bilayer. In the context of the TOXCAT chimera, spheroplast assays revealed that neither of the I32 mutants inserted into the bacterial inner membrane. A possible explanation for this, is that the polar (Asp, Asn) residue may interact intramolecularly with either another residue in the helix (e.g. T36), or with the helix backbone. Such interactions, if present, could give rise to a significant distortion or bending in the helix backbone that might hinder insertion. Mutation of Thr36–Ala in the presence of I32D, however, did not restore self-association or insertion (data not shown), suggesting that T36 is not interacting with D32 in this context.

Ultimately, the strength of an H-bond will depend on both the orientation and angle of the polar groups. For the C=O···H–N hydrogen bond, most bond angles fall between 120–160° with an average distance for the H-bond of ~2 Å.<sup>27,28</sup> Once the H-bond moves beyond these ranges, it loses its H-bonding character and becomes a weaker electrostatic interaction. The formation of a strong H-bond may be inhibited by steric clashes and/or aberrant packing interactions, which may arise through attempts to bring the polar partners into H-bonding range.

Although the mutations at I32 in TOXCAT did not provide information about helix association, their failure to insert into the bilayer indicates another interesting context effect of polar residues. These results suggest that the ability of a polar residue to insert into the bilayer is not simply a function of the overall hydrophobicity of the helix, but involves the sequence environment of the polar residue within the helix. In this instance moving the polar residue one position from V31 to I32 was sufficient to prevent insertion. Therefore, polar residues may prevent correct membrane protein folding by altering helix association or hindering helix insertion and both cases show apparent context dependence.

Interestingly, examination of results shown here for Asn and Asp substitutions at a given residue within MCP reveals that Asp residues generally cause a greater effect on TM helix–helix association in membranes than the corresponding Asn residue. At V29 and V30 for instance, replacements to Asn are relatively benign when compared with those to Asp at these positions (Figure 5). This difference may be due, in part, to the fact that carboxyl groups are known to form stronger H-bonds

than carboxamides. However, since the difference between Asp and Asn is not consistent, there are obviously other factors (possibly steric) contributing to the ability of Asp and Asn to form H-bonds of varying strength.

## Conclusion

Helix–helix association *via* side-chain/side-chain hydrogen bond interactions among strongly polar residues (Gln, Glu, Asp and Asn) is observed to be greatly influenced by their sequence environment. In their native TM sequences, we found that Glu and Gln can either have no observable effect on helix association (C79b and Fut3) or produce a strong association (Tnf5) (Figure 4). When Asp and Asn are introduced into the moderately stable MCP dimer, there are a variety of stabilizing/non-stabilizing outcomes (Figure 5), which again point to the importance of the sequence context in polar residue-mediated helix association. The results with MCP suggest that within a given TM domain, there may be “hot spots” where polar mutations create “spurious” associations, emphasizing the complexity of the task of dissecting the individual sequence-dependent roles of H-bonding and van der Waals packing in a given TM–TM association.

The overall results indicate that strongly polar residues interact with a wide range of affinities that are modulated by the packing interactions provided by their surrounding residues. These findings demonstrate that potentially strong inter-helical H-bonding interactions can be attenuated by their local environment, suggesting that an innate structural mechanism may operate to select against strong and non-specific association of membrane-embedded polar residues. In addition, recent work indicates that non-specific association might be avoided by assembly pathways that favor a specific oligomer *in vivo*, as appears to be the case in the T-cell receptor.<sup>29</sup>

## Materials and Methods

### TOXCAT chimera construction

Oligonucleotides encoding the transmembrane domain of each construct were restriction digested with *StyI* and *DpnII* and ligated into the *NheI* and *BamHI* sites of the restriction-digested plasmid pccKAN. Each construct was tested for membrane insertion using 0.1 M NaOH washes to remove all soluble and peripheral membrane proteins from the membrane.<sup>30</sup> Integral membrane proteins remain in the pelleted membrane fraction. Additionally, membrane insertion and correct orientation were verified for the MCP constructs using the protease sensitivity in the spheroplast assay described,<sup>24</sup> except that the LexA antibody (Invitrogen) was used to verify the intact state of the spheroplasts. Briefly, spheroplasts were created by digesting bacteria with lysozyme to remove the peptidoglycan layer. Spheroplasts were then digested with proteinase K to remove those proteins present in the periplasm. Only if the TOX-

CAT chimera is in the membrane and correctly oriented will the maltose binding protein (MBP) portion of the fusion be digested away from the remaining protein. Following digestion the samples were run on SDS-PAGE and then Western analysis was carried out using antibodies against MBP (NEB) and LexA (1:5000 dilution). Blots were developed using goat anti-rabbit alkaline phosphatase secondary antibody (Pierce).

### CAT assays

Cell-free extracts were made by pelleting 200  $\mu$ l of cells at an  $A_{600}$  of 0.6, resuspending in 500  $\mu$ l of 0.1 M Tris (pH 8.0), then lysing with 20  $\mu$ l of 100 mM EDTA, 100 mM DTT, 50 mM Tris (pH 8.0) and one drop of toluene from a drawn-out Pasteur pipette, at 30 °C for 30 minutes. The cell-free extract was then diluted 1:100 before being used in the CAT assays. The CAT assays were performed using the Quant-T-CAT kit (Amersham). All measurements were performed at least four times.

### Searching of Swiss Prot

All databases used in this study were created previously<sup>14</sup> from the SWISS PROT, version 37.<sup>31</sup> The databases of single and multi-spanning transmembrane helices used here were homology purged and each TM was reduced to an 18-residue window of maximum hydrophobicity. For the single-spanning TMs, an additional database was created, containing the central eight residues of the 18-residue window. See the text for a further discussion.

---

---

## Acknowledgements

This work was supported, in part, by grants to C.M.D. from the Canadian Institutes for Health Research (CIHR) and to D.M.E. from NSF (2MCB-990567) and NIH (GM54160). J.P.D. held a National Science Foundation Graduate Fellowship. R.A.M. held a CIHR Doctoral Research Award.

## References

1. Popot, J. L. & Engelman, D. M. (1990). Membrane protein folding and oligomerization: the two-stage model. *Biochemistry*, **29**, 4031–4037.
2. Choma, C., Gratkowski, H., Lear, J. D. & DeGrado, W. F. (2000). Asparagine-mediated self-association of a model transmembrane helix. *Nature Struct. Biol.* **7**, 161–166.
3. Gratkowski, H., Lear, J. D. & DeGrado, W. F. (2001). Polar side chains drive the association of model transmembrane peptides. *Proc. Natl Acad. Sci. USA*, **98**, 880–885.
4. Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T. & Engelman, D. M. (2000). Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nature Struct. Biol.* **7**, 154–160.
5. Zhou, F. X., Merianos, H. J., Brunger, A. T. & Engelman, D. M. (2001). Polar residues drive association of poly-leucine transmembrane helices. *Proc. Natl Acad. Sci. USA*, **98**, 2250–2255.
6. Dawson, J. P., Weinger, J. S. & Engelman, D. M.

- (2002). Motifs of serine and threonine can drive association of transmembrane helices. *J. Mol. Biol.* **316**, 799–805.
7. Therien, A. G., Grant, F. E. & Deber, C. M. (2001). Interhelical hydrogen bonds in the CFTR membrane domain. *Nature Struct. Biol.* **8**, 597–601.
  8. Bargmann, C. I., Hung, M. C. & Weinberg, R. A. (1986). Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. *Cell*, **45**, 649–657.
  9. Partridge, A. W., Melnyk, R. A. & Deber, C. M. (2002). Polar residues in membrane domains of proteins: molecular basis for helix–helix association in a mutant CFTR transmembrane segment. *Biochemistry*, **41**, 3647–3653.
  10. Cao, H., Bangalore, L., Bormann, B. J. & Stern, D. F. (1992). A subdomain in the transmembrane domain is necessary for p185neu<sup>\*</sup> activation. *EMBO J.* **11**, 923–932.
  11. Smith, S. O., Smith, C. S. & Bormann, B. J. (1996). Strong hydrogen bonding interactions involving a buried glutamic acid in the transmembrane sequence of the neu/erbB-2 receptor. *Nature Struct. Biol.* **3**, 252–258.
  12. Mendrola, J. M., Berger, M. B., King, M. C. & Lemmon, M. A. (2002). The single transmembrane domains of ErbB receptors self-associate in cell membranes. *J. Biol. Chem.* **277**, 4704–4712.
  13. Chen, L. I., Webster, M. K., Meyer, A. N. & Donoghue, D. J. (1997). Transmembrane domain sequence requirements for activation of the p185c-neu receptor tyrosine kinase. *J. Cell. Biol.* **137**, 619–631.
  14. Senes, A., Gerstein, M. & 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 neighboring positions. *J. Mol. Biol.* **296**, 921–936.
  15. Lemmon, M. A., Flanagan, J. M., Hunt, J. F., Adair, B. D., Bormann, B. J., Dempsey, C. E. & Engelman, D. M. (1992). Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices. *J. Biol. Chem.* **267**, 7683–7689.
  16. Lemmon, M. A., Flanagan, J. M., Treutlein, H. R., Zhang, J. & Engelman, D. M. (1992). Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry*, **31**, 12719–12725.
  17. Mazzei, G. J., Edgerton, M. D., Losberger, C., Lecoanet-Henchoz, S., Graber, P., Durandy, A. *et al.* (1995). Recombinant soluble trimeric CD40 ligand is biologically active. *J. Biol. Chem.* **270**, 7025–7028.
  18. Campbell, K. S., Hager, E. J., Friedrich, R. J. & Cambier, J. C. (1991). IgM antigen receptor complex contains phosphoprotein products of B29 and mb-1 genes. *Proc. Natl Acad. Sci. USA*, **88**, 3982–3986.
  19. Reth, M., Hombach, J., Wienands, J., Campbell, K. S., Chien, N., Justement, L. B. & Cambier, J. C. (1991). The B-cell antigen receptor complex. *Immunol. Today*, **12**, 196–201.
  20. Venkitaraman, A. R., Williams, G. T., Dariavach, P. & Neuberger, M. S. (1991). The B-cell antigen receptor of the five immunoglobulin classes. *Nature*, **352**, 777–781.
  21. Deber, C. M., Khan, A. R., Li, Z., Joensson, C., Glibowicka, M. & Wang, J. (1993). Val → Ala mutations selectively alter helix–helix packing in the transmembrane segment of phage M13 coat protein. *Proc. Natl Acad. Sci. USA*, **90**, 11648–11652.
  22. Melnyk, R. A., Partridge, A. W. & Deber, C. M. (2002). Transmembrane domain mediated self-assembly of major coat protein subunits from Ff bacteriophage. *J. Mol. Biol.* **315**, 63–72.
  23. Haigh, N. G. & Webster, R. E. (1998). The major coat protein of filamentous bacteriophage f1 specifically pairs in the bacterial cytoplasmic membrane. *J. Mol. Biol.* **279**, 19–29.
  24. Russ, W. P. & Engelman, D. M. (1999). TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc. Natl Acad. Sci. USA*, **96**, 863–868.
  25. von Heijne, G. (1992). Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* **225**, 487–494.
  26. Eilers, M., Shekar, S. C., Shieh, T., Smith, S. O. & Fleming, P. J. (2000). Internal packing of helical membrane proteins. *Proc. Natl Acad. Sci. USA*, **97**, 5796–5801.
  27. Taylor, R., Kennard, O. & Versichel, W. (1983). Geometry of the imino-carbonyl (NH–O=C) hydrogen bond. 1. Lone-pair directionality. *J. Am. Chem. Soc.* **105**, 5761–5766.
  28. Taylor, R., Kennard, O. & Versichel, W. (1984). Geometry of the NH–O=C hydrogen bond. 3. Hydrogen bond distances and angles. *Acta Crystallog. sect. B*, **40**, 280–288.
  29. Call, M. E., Pyrdol, J., Wiedmann, M. & Wucherpfennig, K. W. (2002). The organizing principle in the formation of the T-cell receptor-CD3 complex. *Cell*, **111**, 967–979.
  30. Chen, H. & Kendall, D. A. (1995). Artificial transmembrane segments. Requirements for stop transfer and polypeptide orientation. *J. Biol. Chem.* **270**, 14115–14122.
  31. Bairoch, A. & Apweiler, R. (1999). The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1999. *Nucl. Acids Res.* **27**, 49–54.

Edited by G. von Heijne

(Received 18 March 2003; received in revised form 28 May 2003; accepted 28 May 2003)