Specificity in Transmembrane Helix-Helix Interactions Mediated by Aromatic Residues*

Neta Sal-Man, Doron Gerber, Itai Bloch, and Yechiel Shai

From the Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel

Aromatic residues have been previously shown to mediate the self-assembly of different soluble proteins through π-π interactions (McGaughey, G. B., Gagne, M., and Rappe, A. K. (1998) J. Biol. Chem. 273, 15458–15463). However, their role in transmembrane (TM) assembly is not yet clear. In this study, we performed statistical analysis of the frequency of occurrence of aromatic pairs in a bacterial TM data base that provided an initial indication that the appearance of a specific aromatic pattern, Aromatic-XX-Aromatic, is not coincidental, similar to the well characterized QXXS motif. The QXXS motif was previously shown to be both critical and sufficient for stabilizing TM self-assembly. Using the ToxR system, we monitored the dimerization propensities of TM domains that contain mutations of interacting residues to aromatic amino acids and demonstrated that aromatic residues can adequately stabilize self-association. Importantly, we have provided an example of a natural TM domain, the cholera toxin secretion protein EpsM, whose TM self-assembly is mediated by an aromatic motif (WXXXW). This is, in fact, the first evidence that aromatic residues are involved in the dimerization of a wild type TM domain. The association mediated by aromatic residues was found to be sensitive to the TM sequence, suggesting that aromatic residue motifs can provide a general means for specificity in TM assembly. Molecular dynamics provided a structural explanation for this backbone sequence sensitivity.

Receptor self-assembly is a central process in a variety of signal transduction cascades. This assembly is mainly mediated by the extracellular or the intracellular domains. However, considerable data have been accumulated concerning the causal involvement of the transmembrane (TM) domains in this process as well (1–5). In contrast to the soluble regions of membrane proteins, our knowledge of the factors that control protein-protein interactions and recognition of the membrane-embedded domains is still limited.

To date, the non-covalent association of native TM domains was reported to be mediated by (i) a heptad motif of leucines through their side chain residues packing interaction (6); (ii) a GXXXG motif, which was first found in the glycophorin A TM domain (4, 8, 9); or (iii) polar residues through the formation of hydrogen bonds (10–14). However, the involvement of additional motifs or key factors that may mediate protein-protein interactions within the membrane merit further investigation.

Examination of the assembly of soluble proteins reveals that aromatic residues serve as key structural elements that mediate the molecular recognition and the self-assembly of amyloid polypeptides as well as bacterial toxins and several proteins such as acetylcholinesterase (15–20). The interactions are formed between the planar aromatic rings and are referred to as π-π interactions (21–23). Even a single mutation of aromatic amino acid in the sequence of the short amyloid peptide abolishes the ability of the peptide to form amyloid fibrils (17). Sequence analysis of membrane immunoglobulin and the T-cell receptor reveals highly conserved aromatic amino acids within the TM domains of these proteins, which may indicate the involvement of these residues in TM-TM interactions (24). An additional piece of evidence that supports this assumption is the recent finding of Langosch and co-workers (25) that demonstrated, using a randomized TM library, tryptophan is enriched in self-interacting TM domains containing a heptad repeat backbone sequence.

In this study, we have investigated the role of aromatic residues in TM self-assembly. We performed a statistical analysis that revealed that an Aromatic-XX-Aromatic pattern (which was chosen based on the QXXS motif) is over-represented in a bacterial TM data base. Using the ToxR system, we measured the ability of aromatic residues to stabilize oligomerization when inserted into the well defined interaction interface of the N-terminal TM domain of the Escherichia coli aspartate receptor (Tar-1). This TM domain, which is part of the bacterial chemotaxis receptor, was previously shown to form a homodimer through its two polar residues (22QXXS25) by forming hydrogen bonds (14). Our results indicate that aromatic residues can enable self-assembly when located at the positions of the polar QXXS motif. Importantly, we have demonstrated that the EpsM TM sequence, which had been found in the statistical analysis search to contain a native double aromatic sequence (WXXXW), was substantially affected in terms of dimerization by mutations in these specific positions. This protein was chosen because it was previously suggested to form dimers (26), data that are not available for most membrane proteins.

**EXPERIMENTAL PROCEDURES**

**Statistical Analysis**—The analysis of the frequencies of occurrence of the QXXW, QXXY, QXXF, WXXS, YXXS, FXXS,
Aromatic Residues Mediate TM Interactions

WXXW, YXXY, and FXFX patterns were performed on a bacterial TM domain data base. The source of the TM sequences was the annotated non-redundant data base Swiss-Prot updated to February 2005. The data base contains 41,916 bacterial TM domains with lengths ranging between 15 and 30 amino acids. The occurrences of the different amino acid combinations in the data base were counted. The averaged expected frequency of occurrences was calculated by counting the number of specific pattern occurrences in 100 randomized data bases. Randomization of the data base was achieved by combining the sequences of the TM domain of a particular length into one long string of characters, which was then shuffled randomly and recut into the original length.

Construction of the ToxR Chimeras—A Nhel-BamHI TM-DNA cassette encoding 16 residues of the Tar-1 wild type (WT) TM domain (13MVGLVFALLQ1SGL28) or the EpsM WT TM domain (26MGALTVALAYWGIWQ41) were inserted between the ToxR transcription activator and the E. coli maltose binding protein MalE within the ToxR-MalE plasmid. The TM domain of interest was inserted into a 6-hydrophobic-amino-acid-sequence, thus creating a hydrophobic sequence of 22 residues, which is a typical length of TM domains. Point mutations of Tar-1 were done in the Gln and/or the Ser at positions 22 and 25, respectively. Point mutations of EpsM were done in one or both Trp residues at positions 37 and 40. The QXXS motif + backbone and WXXW motif + backbone constructs contain the 22QXXS or the 22WXXW motifs within a stretch of leucine and alanine, respectively (see Table 2). The WXXW + Ala16 construct contains the WXXW motif with the sequence of 16 alanines (see Table 2). The sequences of all of the constructs were confirmed by DNA sequencing. The nomenclature of the TM domains represents the two amino acids occupying positions 22 and 25 of the Tar-1 TM domain or positions 37 and 40 of the EpsM TM domain.

In Vivo Detection of Homodimerization of TM Domains within the Membrane—The ToxR transcription activator can be used successfully to assess weak protein-protein interactions within the E. coli membrane. A Tar-1 TM domain encoding the DNA cassette was grafted between the ToxR transcription activator and the maltose binding protein in the ToxR-MalE plasmid. The plasmid was then transformed into E. coli FHK12 cells, which contain β-galactosidase, under the control of a ctx promoter. Dimerization of the TM domains in this system results in association of the ToxR transcription activator, which then becomes active and is able to bind the ctx promoter (27). Quantification of the level of homodimerization was done by measuring the activity of the β-galactosidase reporter gene and by normalizing it to the cell protein content (A950) (Miller units). The base-line activity of a negative control ToxR′Ala16, which remains a monomer, was subtracted from all of the results and referred to as 0% dimerization. The transformed cells were grown in the presence of chloramphenicol for 18 h at 37 °C. β-Galactosidase activities were quantified in crude cell lysates after adding o-nitrophenylgalactosidase and monitoring the reaction at 405 nm for 20 min at intervals of 30 s at 28 °C by a Molecular Devices kinetic reader (27, 28). Specific β-galactosidase activities were calculated from the Vmax of the reaction.

RESULTS

In this study, we have investigated the involvement of aromatic residues in TM self-assembly. To this end, we used the well characterized interaction interface of the dimer formed by the N-terminal TM domain of the E. coli aspartate receptor (QXXS). Here, we have examined whether replacement of one or both polar residues (which are localized to the interaction interface) by aromatic residues will support dimerization.

Statistical Analysis—we analyzed the frequency of occurrence of aromatic residue pairs to determine whether these motifs may be involved in the association of actual TM domains. The pattern investigated, Aromatic-XX-Aromatic,
Aromatic Residues Mediate TM Interactions

To determine the ability of aromatic residues to mediate TM-TM interactions, we utilized the ToxR assembly system, which can detect self-association within the inner membrane of E. coli (27). The two wild type residues, located at the interaction interface of the Tar-1 dimer (at positions 22 and 25, corresponding to Gln and Ser, respectively), were replaced by Trp and were examined for their level of dimerization. Tar W/W, in which both positions 22 and 25 were replaced by Trp, preserved ~75% of the dimerization activity of the WT Tar-1 construct (Fig. 2). Substitution of only one residue of the original motif for Trp yielded high activity for replacement of Ser (Tar Q/W ~125% activity) and moderate activity for replacement of Gln (Tar W/S ~50% activity). These results are expected, considering the different contribution of each residue of the original motif to the dimerization propensity of Tar-1 (~70% was stabilized by Gln and only ~30% by Ser (32)). The localization of the Trp residues relative to the membrane can serve as an additional factor that affects the dimerization ability of the constructs and will be discussed later. Tar-1 G/S, I/S, and M/S mutants showed low dimerization propensities compared with Tar-1 W/S mutant, thus indicating that the association mediated by aromatic residues does not relate to their bulky nature.

To exclude the possibility that the difference between the dimerization activities of the constructs resulted from different expression levels of the chimera proteins, or alternatively, from a failure of the constructs to properly insert into the membrane, we performed Western blotting and maltose complementation assays (Fig. 2, B and C, respectively). The expression levels of the chimera proteins were similar to the Tar-1 WT (Fig. 2B). Note that all activities were normalized to the protein expression level, using the Western blotting results. Correct integration of the ToxR-TM-MalE chimera proteins into the inner membrane of E. coli was assessed by examining the ability of the mutants to functionally complement a MalE-deficient E. coli strain (PD28) (29). Because PD28 cells are unable to grow on minimal medium containing maltose as the only carbon source, only cells that express the chimera protein in the correct orientation (MalE pointing toward the periplasm) would be able to utilize maltose and thus allow cell growth. All constructs exhibited similar rates of cell growth, indicating proper membrane integration (Fig. 2C). A construct with a deleted TM domain (ΔTM) served as a negative control, because it was expected to

![Figure 1](https://example.com/figure1.png)

**Figure 1.** The WXXW motif is over-represented in a bacterial TM database compared with its expected frequency. A, the distribution of the frequency of occurrence of the WXXW motif in 100 randomized TM databases. The arrow marks the actual occurrence of the WXXW motif in the bacterial TM database (with lengths of 15–30 amino acids). B, the actual and the average expected number of occurrences of aromatic residue motifs. STD is the standard deviation of the expected distribution curves. Significance is indicated by the p values calculated as ERFC(x). The actual:expected occurrence ratios of WXXW and YXXY are >1, indicating over-representation in bacterial TM domains. The actual:expected occurrence ratios of QXXW and QXXY are <1, indicating under-representation in bacterial TM domains.

was chosen based on the QXXS composition. A correlation between over- or under-representation of a specific pattern in TM domains and helix-helix interactions was previously found (30). The analysis was done in a broad set of bacterial TM domains. Our results indicate that the aromatic pairs appear to be significantly over-represented compared with their theoretical expectancy. The WXXW pattern was found 419 times within a data base of 41,916 sequences of TM domains, whereas its average random expected occurrence was 303 ± 16 times (Fig. 1), and the YXXY pattern appeared 721 times, whereas its average random expected occurrence was 597 ± 22 times. The ratio of actual:expected occurrence of the WXXW and YXXY sequences is much higher than 1, whereas the ratio of QXXW and QXXY is significantly lower than 1 (Fig. 1B). On the other hand, the WXXS and YXXS motifs approach 1. Taken together, these results suggest that the WXXW, YXXY, QXXW, and QXXY patterns may function as self-association motifs. The frequencies of occurrence of the Phe motifs (XXFX, QXXF, or QXXS) were within the range of the random distribution. These results may indicate that Phe pairs do not serve as association motifs, probably because Phe is much more frequent in TM domains than Trp and Tyr. Trp and Tyr residues within TM domains are not equally distributed along the TM sequence and tend to be clustered at the lipid-water interface (31). Therefore, we statistically analyzed the frequency of occurrence of the aromatic patterns separately in the N-terminal, core, and C-terminal parts of the TM domains. The ratio of actual:expected occurrence of the WXXW sequence was 1.26, 1.88, and 1.47 for the N terminus, core, and C terminus, respectively. The ratios of the YXXY were 1.04, 1.41, and 1.18 and for the QXXF, 1.05, 1.12, and 1.02. These results are consistent with the data acquired from the full TM sequence and imply that the overrepresentation of the WXXW and YXXY motifs is not related to the tendency of W and Y residues to cluster close to the lipid-water interface.

Homodimerization of Different Tar-1 Mutants Containing Aromatic Residues—To determine the ability of aromatic residues to mediate TM-TM interactions, we utilized the ToxR assembly system, which can detect self-association within the inner membrane of E. coli (27). The two wild type residues, located at the interaction interface of the Tar-1 dimer (at positions 22 and 25, corresponding to Gln and Ser, respectively), were replaced by Trp and were examined for their level of dimerization. Tar W/W, in which both positions 22 and 25 were replaced by Trp, preserved ~75% of the dimerization activity of the WT Tar-1 construct (Fig. 2). Substitution of only one residue of the original motif for Trp yielded high activity for replacement of Ser (Tar Q/W ~125% activity) and moderate activity for replacement of Gln (Tar W/S ~50% activity). These results are expected, considering the different contribution of each residue of the original motif to the dimerization propensity of Tar-1 (~70% was stabilized by Gln and only ~30% by Ser (32)). The localization of the Trp residues relative to the membrane can serve as an additional factor that affects the dimerization ability of the constructs and will be discussed later. Tar-1 G/S, I/S, and M/S mutants showed low dimerization propensities compared with Tar-1 W/S mutant, thus indicating that the association mediated by aromatic residues does not relate to their bulky nature.

To exclude the possibility that the difference between the dimerization activities of the constructs resulted from different expression levels of the chimera proteins, or alternatively, from a failure of the constructs to properly insert into the membrane, we performed Western blotting and maltose complementation assays (Fig. 2, B and C, respectively). The expression levels of the chimera proteins were similar to the Tar-1 WT (Fig. 2B). Note that all activities were normalized to the protein expression level, using the Western blotting results. Correct integration of the ToxR-TM-MalE chimera proteins into the inner membrane of E. coli was assessed by examining the ability of the mutants to functionally complement a MalE-deficient E. coli strain (PD28) (29). Because PD28 cells are unable to grow on minimal medium containing maltose as the only carbon source, only cells that express the chimera protein in the correct orientation (MalE pointing toward the periplasm) would be able to utilize maltose and thus allow cell growth. All constructs exhibited similar rates of cell growth, indicating proper membrane integration (Fig. 2C). A construct with a deleted TM domain (ΔTM) served as a negative control, because it was expected to
Aromatic Residues Mediate TM Interactions

Table 1. Different mutants of the Tar-1 TM domain used to determine the involvement of aromatic residues in TM assembly

| Designation of TM domain | Mutations | Sequence^{a,b,c,d,e} |
|--------------------------|-----------|---------------------|
| Tar-1 WT                 | Q/Q/S     | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 W/W                | Q22W S25W | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 Q/W                | S25W      | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 W/S                | Q22W      | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 Y/Y                | Q22Y      | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 Q/Y                | S25Y      | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 Y/S                | Q22Y      | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 F/F                | Q22F S25F | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 Q/F                | S25F      | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 F/S                | Q22F      | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 G/S                | Q22G      | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 M/S                | Q22M      | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 I/S                | Q22I      | ^1 sharply VLGVALLQ/LSL2^2 |
| Tar-1 S/S                | Q22S      | ^1 sharply VLGVALLQ/LSL2^2 |

^{a} Amino acids are numbered according to their position in the WT protein (Swissprot p07017).
^{b} The amino acids in the positions of the original dimerization motif are in bold.
^{c} Mutations in the Tar-1 TM domain are underlined.
^{d} The nomenclature of the TM domains represents the two amino acids at positions 22 and 25 of original polar residues motif, respectively.
^{e} The TM domains were inserted into a 6-hydrophobic amino acid sequence, thus creating a hydrophobic sequence of 22 residues that is of typical length of TM domains.

tate and stabilize association within the membrane environment when localized to the oligomer interface, probably through stacking interactions.

Dimerization of the TM Domain of the Cholera Toxin Secretion Protein EpsM—To examine the involvement of aromatic motif in TM-TM interactions of an actual TM domain, we chose the cholera toxin secretion protein EpsM, which had been found to contain a WXXW sequence in the statistical analysis search. This protein contains a single TM domain and was previously suggested to form dimers (26) (such information is usually missing for many membrane proteins). The aromatic residues of the EpsM TM domain are conserved in related proteins of the bacterial secretion pathway (data not shown). This further suggests that these residues have a role in the function/structure of the protein.

A sequence of 16 residues of the EpsM WT TM domain (^26MGALTVAIYWGIIWQ^31) was inserted within the ToxR-MalE plasmid, and its dimerization propensity was measured. Replacement of the Trp residue by Ala at position 37 had only a small effect on the dimerization propensity (~25% decrease, EpsM AW) and almost no effect on the dimerization for replacement at position 40 (EpsM WA). Replacement of both Trp residues by Ala, on the other hand, completely abolished dimerization (EpsM AA) (Fig. 4). These results suggest that the aromatic residues of the EpsM TM domain are critical for self-association. Nevertheless, the presence of a single aromatic residue is sufficient to preserve stable assembly. This is the first evidence that aromatic residues direct the dimerization of biological TM domains. Replacement of the Trp residues by a different aromatic residue, Tyr, had only a small effect on the association level, thus implying that the aromatic nature of the residues is the critical feature required for stable dimerization.

Backbone-dependent Dimer Association of the WXXW Pattern—To determine whether an aromatic residue pair is sufficient for TM-TM dimerization, we examined the association level of constructs that contain the WXXW pattern within the context of the Alaa_{16} sequence, which serves as a marker for base-line activity for all of the experiments or within a stretch of residues in the cytoplasm and therefore was unable to complement the MalE deficiency.

To determine whether this pattern of dimerization propensity is specific for Trp or represents a general trend among aromatic residues, we prepared six additional constructs that contain aromatic residue replacement at one or both positions of the polar dimerization motif (Table 1). The constructs exhibited an activity pattern similar to the Trp constructs (Fig. 3). Similar expression levels of the chimera proteins (Fig. 3, inset panels) and their correct integration were obtained (data not shown). Thus, we concluded that aromatic residues can facilitate and stabilize association within the membrane environment when localized to the oligomer interface, probably through stacking interactions.

**FIGURE 2.** Tryptophan mutations in the Tar-1 wild type TM domain. A, cells expressing ToxR-TM-MalE chimeras were examined for dimerization activity (normalized relative to the WT Tar-1 TM domain activity and to expression levels). All values represent the average of at least three independent repeats. Error bars represent the estimated S.D. The exact sequences are indicated in Table 1. B, comparison of the expression levels of the ToxR-TM-MalE chimeric proteins (65 kDa). Samples of FHK12 cells containing different sequences of the Tar-1 within the ToxR-MalE chimera protein were lysed in sodium dodecyl sulfate sample buffer, separated on 12% SDS-PAGE, and immunoblotted using anti-maltose-binding protein antibody. The chimera protein mutants exhibited expression levels similar to the WT TM domain. C, correct integration of the ToxR-TM-MalE chimeric proteins was tested by their ability to functionally complement the MalE deficiency of PD28 cells. PD28 cells were transformed with glycophorin A (●), Tar-1 WT (○), Tar W/W (+), Tar W/S (□), Tar Q/W (▲), Tar G/S (◇), Tar M/S (●), Tar I/S (◇), and ΔTM (◊) plasmids and were grown in minimal medium containing maltose. All constructs showed growth curves similar to Tar-1 WT, indicating proper membrane integration. The negative control with a deleted TM domain (ΔTM) showed no growth.
leucines and alanines (the sequences are shown in Table 3). Both backbones were previously shown to have a low dimerization propensity and were properly inserted within the membrane (14). The WXXW motif/H11001 backbone exhibited very low dimerization activity compared with the Tar-1 WT TM domain, whereas the WXXW/H11001 Ala16 demonstrated moderated dimerization propensity. In contrast, the QXXS motif, which is able to stabilize dimerization independently of its TM sequence, showed a dimerization activity similar to the Tar-1 WT TM domain when inserted within the backbone construct (Fig. 5). Unfortunately, the QXXS motif within the Ala16 sequence was unable to insert properly into the membrane (14) and therefore is not shown. Overall, the results indicate that, although the QXXS motif functions as a general dimerization motif, the aromatic pair WXXW can stabilize association in a sequence-specific manner. All constructs exhibited expression levels similar to the Tar-1 WT TM domain and were correctly integrated into the membrane (data not shown).

FIGURE 3. Tyrosine and phenylalanine mutations of the Tar-1 wild type TM domain. A, dimerization of ToxR-TM-MalE chimera proteins with tyrosine mutations in the Tar-1 dimer interface. B, dimerization activity of ToxR-TM-MalE chimera proteins with phenylalanine mutations. The sequences are shown in Table 1. The expression levels of the different TM mutants are shown in the inset panels.

FIGURE 4. The dimerization levels of EpsM WT and mutant constructs. A, dimerization of ToxR-TM-MalE chimera proteins of wild type and several mutants of the EpsM TM domain. The exact sequences are indicated in Table 2. The details are as described in the legend to Fig. 2. B, comparison of the expression levels of ToxR-TM-MalE chimera proteins. C, correct integration of the ToxR-TM-MalE chimera proteins. PD28 cells were transformed with EpsM WT ( ), EpsM AW ( ), EpsM WA ( ), EpsM AA ( ), EpsM YW ( ), EpsM WY ( ), EpsM YY ( ), EpsM YY ( ), and ∆TM ( ) plasmids and were grown in minimal medium containing maltose. All constructs showed similar growth rates, indicating proper membrane integration. The negative control with a deleted TM domain (∆TM) showed no growth.

| TABLE 2 | TM sequences of the EpsM wild type mutants |
|-----------------|---------------------------------|
| Designation of TM domain | Mutations | Sequence<sup>a</sup> |
| EpsM WT (WW) | W37A, W40A | W37Y, W40Y |
| EpsM AW | W37A, W40A | W37A, W40A |
| EpsM WA | W37A, W40A | W37A, W40A |
| EpsM AA | W37A, W40A | W37A, W40A |
| EpsM YW | W37A, W40A | W37A, W40A |
| EpsM YY | W37A, W40A | W37A, W40A |

<sup>a</sup> Amino acids are numbered according to their position in the WT protein (Swissprot Q9KNK8).

<sup>b</sup> The amino acids in the positions of the original dimerization motif are in bold.

<sup>c</sup> Mutations in the Tar-1 TM domain are underlined.

<sup>d</sup> The nomenclature of the TM domains represents the two amino acids at positions 37 and 40 of original polar residues motif, respectively.

<sup>e</sup> The TM domains were inserted into a 6-hydrophobic amino acid sequence, thus creating a hydrophobic sequence of 22 residues that is of typical length of TM domains.

leucines and alanines (the sequences are shown in Table 3). Both backbones were previously shown to have a low dimerization propensity and were properly inserted within the membrane (14). The WXXW motif + backbone exhibited very low dimerization activity compared with the Tar-1 WT TM domain, whereas the WXXW + Ala<sub>16</sub> demonstrated moderated dimerization propensity. In contrast, the QXXS motif, which is able to stabilize dimerization independently of its TM sequence, showed a dimerization activity similar to the Tar-1 WT TM domain when inserted within the backbone construct (Fig. 5). Unfortunately, the QXXS motif within the Ala<sub>16</sub> sequence was unable to insert properly into the membrane (14) and therefore is not shown. Overall, the results indicate that, although the QXXS motif functions as a general dimerization motif, the aromatic pair WXXW can stabilize association in a sequence-specific manner. All constructs exhibited expression levels similar to the Tar-1 WT TM domain and were correctly integrated into the membrane (data not shown).
Aromatic Residues Mediate TM Interactions

TABLE 3
TM sequences used to determine the sensitivity of the aromatic residue motif to the backbone sequence

| Name of inserted TM domain | Sequence |
|---------------------------|----------|
| Leu + Ala backbone        |          |
| QXXS motif + backbone     | 13LLLLLLLAAAAA28 |
| WXXW motif + backbone     | 13LLLLLLLAAAAA28 |
| Ala16                     | 13AAAAAAAAAAA28 |
| WXXW motif + Ala16        | 13AAAAAAAAAAA28 |

A Structural Model for the Backbone-dependent Dimer Association of the Aromatic WXXW Motif—To provide a possible structural explanation for the backbone-dependent activity of the double aromatic residue pattern (WXXW), we conducted a series of parallel docking simulations of homodimerization for four different TM sequences. These sequences contain the QXXS or the WXXW motifs either within the Tar-1 WT sequence or within the Leu + Ala backbone. The resulting structures, obtained from the docking procedure, showed a strong correlation between the packing of the monomers in each dimer and the dimer activity (Fig. 6). Interestingly, the bulky Trp residues prevent tight helical association when the WXXW motif was located within the Leu + Ala backbone sequence. However, in the context of the Tar-1 WT sequence, this pattern displayed a packing ability similar to the Tar-1 WT sequence. We obtained an increase (~2 Å) in the distance between the monomers of the WXXW motif within the Leu + Ala backbone sequence, compared with the other TM sequences (measured between the α-carbons of the residues at position 22). The contact matrix, which presents the mean distance between the center of mass of each pair of residues, clearly demonstrates an overall change in dimer packing. The WXXW motif exhibited closer packing interaction in the Tar-1 WT TM sequence, manifested by a red/yellow tendency (Fig. 6E). A more loosely packed dimer was observed for the WXXW motif in the Leu + Ala backbone, as indicated by the higher frequency of blue (Fig. 6F). Overall, the structural models obtained for the different TM dimers correlate well with the dimerization results, which indicate that the association of the WXXW-containing dimer is sensitive to backbone changes, whereas the wild type QXXS motif is not.

Position Dependence—Aromatic residues are preferentially located at the lipid-water interface (33). To determine whether the dimerization propensity of a TM domain that contains an aromatic residue is affected by the proximity of the aromatic residue to the membrane-water interface, we examined the association levels of three sets of TM domains. Two sets contained a single Trp and one wild type polar residue at either one of the two positions of the Tar-1 dimerization motif (positions 22 and 25). An additional set of TM domains contained pairs of Trp and Ala residues at these interacting positions (the exact sequences are shown in Table 4). Previously, we observed similar dimerization activities for the Tar-1 WT motif (22QXXS25) and the swapped TM dimerization motif (22SXXQ25), suggesting that the contribution of Gln and Ser to TM domain self-association is position-independent (32). The TM domains containing aromatic residues, on the other hand, showed a higher association ability for TM domains that contained Trp closer to the lipid-water interface (Tar Q/W, Tar S/W, and Tar A/W) (Fig. 7) compared with TM domains that contain Trp deeper in the lipid core (Tar W/Q, Tar W/S, and Tar W/A). These results suggest that the association mediated via aromatic residues is position-dependent. A replacement of Ala to Trp at position 19 within the TM domain of Tar W/S mutant created an Aro-XX-Aro motif within the core of the membrane environment (named “shifted motif”). This mutant showed higher dimerization activity relative to the Tar W/S mutant (Fig. 7), thus excluding the possibility that the aromatic residues stabilize TM dimerization only by interaction with the lipid polar head groups or the water molecules. The higher dimerization activity observed for the shifted motif relative to the Tar W/W TM domain attributed to the contribution of the polar serine residue.

DISCUSSION

The first challenge in characterizing helix-helix interactions is to define the interaction interface and the specific residues that are involved in mediating TM-TM associations. In this study we employed the well-characterized dimerization interface of the Tar-1 to investigate the involvement of aromatic residues in the dimerization of TM helices. Our data revealed that mutation of one or both polar residues of this dimerization motif (QXXS) to aromatic residues provides the necessary driving force for establishing an association. Importantly, we demonstrated that the two Trp residues (37WXXW49) located within the EpsM wild type TM sequence are essential for dimerization as mutations of both residues to Ala abolishes self-association. Single mutations of Trp to Ala, on the other hand, had only small effect on dimerization (EpsM AW and WA), suggesting that the presence of a single Trp residue is sufficient for stabilizing dimerization. It is possible that in the absence of one bulky Trp residue the adjacent Tyr residue (located at position 36) can contribute to helix-helix interactions. Mutations of the Trp residues to a different aromatic residue, Tyr, had only small effect on the self-assembly of the TM domain, supporting our assumption that the aromatic nature of the residue is
the critical feature that required for self-association. In this study we present the first evidence that aromatic residues are directly involved in the dimerization of a biological TM domain.

The similar dimerization pattern obtained for all aromatic amino acids suggests that this interaction is mediated through \( \pi-\pi \) interactions. However, Trp and Tyr are also polar residues and therefore, could also stabilize association though formation of hydrogen bonds. We observed that the strength of the TM-TM association depends on (i) the residue located at the second position of the interaction interface, together with the aromatic residue, (ii) the backbone sequence in which the aromatic pair is present, and (iii) the proximity of the aromatic residue to the lipid-water interface.

The Effect of the Residue Located at the Second Position of the QXXS Motif on Dimerization—Replacing Ser from the dimerization motif \( ^{22}\text{QXXS}^{25} \) by any aromatic amino acid (Tar Q/W, Tar Q/Y, and Tar Q/F) resulted in a higher propensity of dimer formation compared with the Tar-1 wild type TM domain (Figs. 2 and 3). This result implies that the stacking interactions of the aromatic residue drive stronger homo-oligomerization than the one formed by the serine residue. The activity of the Tar S/W construct (\( \sim70\% \) activity) (Fig. 7), compared with the Tar S/S construct (\( \sim50\% \) activity (32)), provides additional support for this observation. Note that the activity of Tar W/S was similar to that of the Tar S/S construct. However, this result is also influenced by the position of Trp relative to the water-lipid interface, which is discussed below.

Constructs in which Gln, from the QXXS motif, was replaced by an aromatic residue (Tar W/S, Tar Y/S, and Tar F/S) showed weaker dimer formation compared with the Tar-1 WT TM domain (Figs. 2 and 3). These results are consistent with the different contributions of each residue of the QXXS motif to Tar-1 dimer formation, as previously shown (\( \sim70\% \) was stabilized by Gln and only \( \sim30\% \) by the Ser; (32)). The low dimerization activities of Tar-1 I/S and M/S mutants excluded the possibility that the association mediated by aromatic residues is related to the residues’ bulky nature.

Double aromatic residue pairs (Tar W/W, Tar Y/Y, and Tar F/F) as well as single aromatic mutants (Tar A/W and Tar W/A) obtained moderately stable dimerization levels compared with the Tar WT TM domain (Figs. 2, 3, and 7). These results demonstrate the ability of aromatic residues to direct TM self-assembly in the absence of any additional polar residue. The similar assembly pattern of all aromatic residues suggests that this interaction is stabilized through \( \pi-\pi \) stacking interactions, which is a common trait of the aromatic residues.

The Effect of the TM Backbone on the Dimerization Level—To determine whether the WXXW pair can promote TM-TM interactions in a sequence-specific manner, similar to the QXXS motif, we examined its association ability within two non-associated sequences (Leu + Ala backbone and Ala\textsubscript{16}). In contrast to the QXXS motif, which exhibits similar dimeriza-
Aromatic Residues Mediate TM Interactions

| Designation of TM domain | Mutations | Sequence<sup>a,b,c,d,e</sup> |
|--------------------------|-----------|-----------------------------|
| Tar-1 WT<sup>*</sup>     |           |                            |
| Tar-1 S/Q                | Q22S S25Q | 1<sup>24</sup>MVLGVFALLQHGSGIL<sup>28</sup> |
| Tar-1 Q/W<sup>+</sup>     | S25W      | 1<sup>24</sup>MVLGVFALLQHGSGIL<sup>28</sup> |
| Tar-1 W/Q                | Q22W      | 1<sup>24</sup>MVLGVFALLQHGSGIL<sup>28</sup> |
| Tar-1 S/W                | Q22S S25W | 1<sup>24</sup>MVLGVFALLQHGSGIL<sup>28</sup> |
| Tar-1 A/W                | Q22A S25W | 1<sup>24</sup>MVLGVFALLQHGSGIL<sup>28</sup> |
| Tar-1 I/WA               | Q22W S25A | 1<sup>24</sup>MVLGVFALLQHGSGIL<sup>28</sup> |
| Shifted motif            | A9W Q22W  | 1<sup>24</sup>MVLGVFALLQHGSGIL<sup>28</sup> |

<sup>a</sup> Present also in Table 1.
<sup>b</sup> Amino acids are numbered according to their position in the WT protein (Swissprot p07017).
<sup>c</sup> The amino acids in the positions of the original motif are in bold.
<sup>d</sup> Mutations in the Tar-1 TM domain are underlined.
<sup>e</sup> The nomenclature of the TM domains represents the two amino acids at positions 22 and 25 of original polar residues motif, respectively.

In this study, we introduced aromatic residues into one or both interacting positions of the QXXS motif. The results clearly indicate that the dimerization activity of the TM domain was influenced by the proximity of the aromatic residue to the membrane-water interface. TM domains that contained aromatic residues closer to the lipid-membrane interface exhibited higher association ability than TM domains with an identical amino acid composition located deeper in the membrane core (Fig. 7). A TM domain that contains an aromatic motif deeper in the membrane core (labeled “shifted motif”) excludes the possibility that the dimerization mediated by aromatic residues is mainly performed by interactions of the aromatic residues with the polar lipid head groups or the water molecules. According to our results, we concluded that the ordered environment of the membrane-water interface supports the helix-packing mediated by the aromatic residues. Therefore, the association at this position is stronger than that within the membrane core. This result correlates with the previous finding that Trp randomized at the different positions of the heptad repeat motif appears at high frequency at the lipid-membrane interface (25).

FIGURE 7. The proximity of the aromatic residue to the lipid-membrane interface affects the dimerization level. Dimerization of ToxR-TM-Malle chimaera proteins containing an aromatic residue in their TM domain at different positions relative to the lipid-membrane interface (the exact sequences are indicated in Table 4).

In summary, we have demonstrated the involvement of all three aromatic amino acids in the association of TM domains. Moreover, we have provided an example of a wild type TM sequence that stabilizes self-association via its native double aromatic pair (WXXW). The interaction mediated by aromatic motifs was found to be greatly influenced by the sequence environment; thus, it may serve as a means for specificity in TM-TM interactions. The strength of association of the WXXW pair was moderate compared with the polar motif QXXS. The Swiss-Prot survey found the WXXW motif to be over-represented within a broad set of bacterial TM domains and identified several TM domains of pathogenic bacterial lipids, which in turn can further weaken the dimer association by interacting with the hydrophobic interface residues.

The Proximity of the Aromatic Residue to the Lipid-Water Interface Affects the Association Propensity—In this study, we introduced aromatic residues into one or both interacting positions of the QXXS motif. The results clearly indicate that the dimerization activity of the TM domain was influenced by the proximity of the aromatic residue to the membrane-water interface. TM domains that contained aromatic residues closer to the lipid-membrane interface exhibited higher association ability than TM domains with an identical amino acid composition located deeper in the membrane core (Fig. 7). A TM domain that contains an aromatic motif deeper in the membrane core (labeled “shifted motif”) excludes the possibility that the dimerization mediated by aromatic residues is mainly performed by interactions of the aromatic residues with the polar lipid head groups or the water molecules. According to our results, we concluded that the ordered environment of the membrane-water interface supports the helix-packing mediated by the aromatic residues. Therefore, the association at this position is stronger than that within the membrane core. This result correlates with the previous finding that Trp randomized at the different positions of the heptad repeat motif appears at high frequency at the lipid-membrane interface (25).

Statistical Analysis of Aromatic Residue Motifs—Statistical analysis of the frequency of occurrence of the double aromatic pair WXXW suggests a general role for this motif in the assembly of bacterial TM domains. A closer look at the TM proteins that contain this aromatic pair revealed two interesting observations. First, the WXXW pattern appears in membrane proteins with different functions, for example, cation efflux system FieF (Swiss-Prot accession number Q83PD6) from *Shigella flexneri*, which mediates resistance to copper and silver; cell volume regulation protein from *Vibrio cholerae* (accession number Q9KNM9), which participates in the control of cell volume under low osmolarity conditions; and lipase chaperone from *Pseudomonas aeruginosa* (accession number Q01725), which is involved in the folding of the extracellular lipase during its passage through the periplasm. These examples show that the WXXW motif is not restricted by function or species. The second observation is that many of the proteins that contain this motif belong to pathogenic bacteria. If the WXXW motif is indeed involved in the assembly of these proteins, this raises the possibility of a new antimicrobial strategy targeting membrane protein assembly.

A Computer Structural Model for the Sequence Specificity of the WXXW Motif—The structural model provides a possible explanation for the sequence-specific association activity of the WXXW pair as opposed to the QXXS motif, which was shown to be insensitive to backbone changes. We observed a significant decrease in the packing of the WXXW motif within the Leu + Ala backbone compared with the WXXW motif within the Tar-1 sequence (Fig. 6). This decrease was manifested in a 2-Å spacing difference between the monomers of the two WXXW-containing TM dimers, shown in the molecular dynamic simulation. In addition, this enlargement of the space between the monomers can facilitate access of the surrounding lipid molecules, which in turn can further weaken the dimer association by interacting with the hydrophobic interface residues.
proteins that contain this motif. These TM proteins are therapeutically attractive targets for inhibition. Overall, our findings advanced our understanding of the oligomerization process within the membrane milieu and may assist us to better understand the basic principles of membrane protein biology. This knowledge, in turn, will help us in the future to rationally modulate the function of membrane proteins for therapeutic purposes.

Acknowledgment—We thank Dr. D. Langosch at Heidelberg University for providing the ToxR-glycophorin A and Ala16 plasmids as well as the FHK12 and PD28 E. coli strains.

REFERENCES

1. von Heijne, G. (1999) Q. Rev. Biophys. 32, 285–307
2. Gouldson, P. R., Higgs, C., Smith, R. E., Dean, M. K., Gkoutos, G. V., and Reynolds, C. A. (2000) Neuropsychopharmacology 23, Suppl. 2, S60–S77
3. Li, R., Bennett, J. S., and Degrado, W. F. (2004) Biochem. Soc. Trans. 32, 412–415
4. Lemmon, M. A., Flanagan, J. M., Hunt, J. F., Adair, B. D., Bormann, B. J., Dempsey, C. E., and Engelman, D. M. (1992) J. Biol. Chem. 267, 7683–7689
5. Arkin, I. T. (2002) Biochim. Biophys. Acta 1565, 347–363
6. Lemmon, M. A., Tresat, H., Adams, P. D., Brunger, A. T., and Engelman, D. M. (1999) J. Biol. Chem. 274, 9265–9270
7. Humphrey, W., Dalke, A., and Schulten, K. (1996) J. Mol. Graph. 14, 33–38
8. Lemmon, M. A., Tresat, H., Adams, P. D., Brunger, A. T., and Engelman, D. M. (1999) Nat. Struct. Biol. 1, 157–163
9. Russ, W. P., and Engelman, D. M. (2000) J. Mol. Biol. 296, 911–919
10. Zhou, F. X., Merianos, H. J., Brunger, A. T., and Engelman, D. M. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 2250–2255
11. Choma, C., Gratkowski, H., Lear, J. D., and DeGrado, W. F. (2000) Nat. Struct. Biol. 7, 161–166
12. Dawson, J. P., Weinger, J. S., and Engelman, D. M. (2002) J. Mol. Biol. 316, 799–805
13. Gratkowski, H., Lear, J. D., and DeGrado, W. F. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 880–885
14. Sal-Man, N., Gerber, D., and Shai, Y. (2005) J. Biol. Chem. 280, 27449–27457
15. Belbeoch, S., Falasca, C., Leroy, J., Ayon, A., Massoulié, J., and Bon, S. (2004) Eur. J. Biochem. 271, 1476–1487
16. Ramachandran, R., Tweten, R. K., and Johnson, A. E. (2004) Nat. Struct. Mol. Biol. 11, 697–705
17. Azriel, R., and Gazit, E. (2001) J. Biol. Chem. 276, 34156–34161
18. Soto, C., Sigurdsson, E. M., Morelli, L., Kumar, R. A., Castano, E. M., and Frangione, B. (1998) Nat. Med. 4, 822–826
19. Pallitto, M. M., Ghanta, J., Heinzelman, P., Kiessling, L. L., and Murphy, R. M. (1999) Biochemistry 38, 3570–3578
20. Findeis, M. A., Musso, G. M., Arico-Muendel, C. C., Benjamin, H. W., Hundal, A. M., Lee, J. J., Chin, J., Kelley, M., Wakefield, J., Hayward, N. J., and Moloney, S. M. (1999) Biochemistry 38, 6791–6800
21. Bernstein, E. R., and Sun, S. (1996) J. Phys. Chem. 100, 13348–13366
22. McGaughey, G. B., Gagne, M., and Rape, A. K. (1998) J. Biol. Chem. 273, 15458–15463
23. Gazit, E. (2002) FASEB J. 16, 77–83
24. Campbell, K. S., Backstrom, B. T., Tiefenthaler, G., and Palmer, E. (1994) Semin. Immunol. 6, 393–410
25. Ritter, A., Skupien, P., Unterreitmeier, S., and Langosch, D. (2005) J. Mol. Biol. 354, 894–902
26. Abendroth, J., Rice, A. E., McLuskey, K., Bagdasarian, M., and Hol, W. G. (2004) J. Mol. Biol. 338, 585–596
27. Langosch, D., Brosig, B., Kolmar, H., and Frits, H. J. (1996) J. Mol. Biol. 263, 525–530
28. Kolmar, H., Frisch, C., Kleemann, G., Gotze, K., Stevens, F. J., and Frits, H. J. (1994) Biol. Chem. Hoppe-Seyler 375, 61–70
29. Brosig, B., and Langosch, D. (1998) Protein Sci. 7, 1052–1056
30. Senes, A., Gerstein, M., and Engelman, D. M. (2000) J. Mol. Biol. 296, 921–936
31. Tsang, S., and Saier, M. H., Jr. (1996) J. Comput. Biol. 3, 185–190
32. Sal-Man, N., Gerber, D., and Shai, Y. (2004) Biochemistry 43, 2309–2313
33. Landolt-Marticorena, C., Williams, K. A., Deber, C. M., and Reithmeier, R. A. (1993) J. Mol. Biol. 229, 602–608