Helix-packing motifs in membrane proteins

R. F. S. Walters* and W. F. DeGrado*†‡

Departments of *Biochemistry and Biophysics and †Chemistry, University of Pennsylvania, Philadelphia, PA 19104-6059

Communicated by James A. Wells, University of California, San Francisco, CA, July 14, 2006 (received for review May 11, 2006)

The fold of a helical membrane protein is largely determined by interactions between membrane-imbedded helices. To elucidate recurring helix–helix interaction motifs, we dissected the crystallographic structures of membrane proteins into a library of interacting helical pairs. The pairs were clustered according to their three-dimensional similarity (rmsd ≤1.5 Å), allowing 90% of the library to be assigned to clusters consisting of at least five members. Surprisingly, three quarters of the helical pairs belong to one of five tightly clustered motifs whose structural features can be understood in terms of simple principles of helix–helix packing. Thus, the universe of common transmembrane helix-pairing motifs is relatively simple. The largest cluster, which comprises 29% of the library members, consists of an antiparallel motif with left-handed packing angles, and it is frequently stabilized by packing of small side chains occurring every seven residues in the sequence. Right-handed parallel and antiparallel structures show a similar tendency to segregate small residues to the helix–helix interface but spaced at four-residue intervals. Position-specific sequence propensities were derived for the most populated motifs. These structural and sequential motifs should be quite useful for the design and structural prediction of membrane proteins.

helix–helix packing | protein design | structure prediction

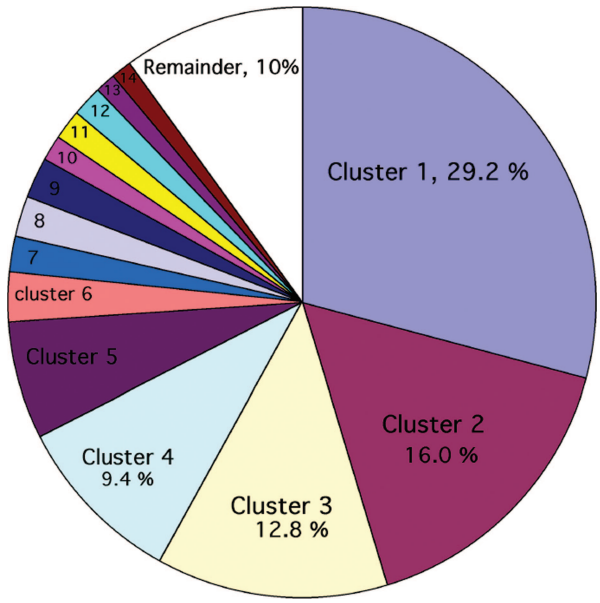

Fig. 1. Pie chart showing the fraction of the total number of pairs that fall within a given cluster.

Helical transmembrane (TM) proteins are a major class of membrane proteins that are critically involved in functionally rich processes, including bioenergetics, signal transduction, ion transmission, and catalysis. The mechanisms by which TM proteins fold into native structures are beginning to be understood from a confluence of structural and biochemical studies (1–3). The determinants of a membrane protein's fold can be understood by dissecting its structure into pairs of interacting helices, which, together with the connecting loops and extramembrane domains, comprise the overall structure. Along these lines, various workers have examined the geometric characteristics of helix–helix interactions in membrane proteins (4) as well as the features in the amino acid sequences that predispose the helices to adopt these geometries. One outstanding success has been the recognition of the GX_3G motif, in which Gly (or other small residues) spaced four residues apart mediate a close approach of TM helices (5–7). This motif was first observed in the TM domain of glycophorin A (GpA) (5) and has subsequently been found in both water-soluble (8) and membrane proteins (9). In the classical GpA GX_3G motif, the helices cross with a right-handed crossing angle of ≈40°. In a different TM motif stabilized by "knobs-into-holes" packing (10), the helices cross with a smaller left-handed crossing angle.

Other surveys of helix–helix packing in membrane proteins have focused on distributions of the interhelical angles, distances, and the composition of the side chains packed at the interface. The helix–helix interfaces tend to be well packed and richer in small residues than in soluble proteins (11, 12), and polar interactions can stabilize helix–helix association (13, 14). Distributions of interhelical distances show that helices tend to approach somewhat more closely in membrane-soluble proteins than in water-soluble proteins, although the distributions tend to be broad and overlapping (15, 16). Similarly, distributions of interhelical angles are broad and overlapping, and it is not clear to what extent they reflect geometric preferences versus biases associated with random distributions of two vectors (17). However, these statistical surveys considered only two parameters (interhelical distance and crossing angle); six parameters (three Eulerian angles and three distances) are required to define the mutual orientation of two helices (18), and more would be required for nonideal helices. We therefore opted for a different method of analyzing helix–helix interactions, similar to a recent study of Gimpelev *et al*. (16). These workers built a library of helical pairs excised from the three-dimensional crystal structures of membrane proteins and then compared the three-dimensional structure of each water-soluble helical pair to each membrane helical pair, using rmsd as the criterion. The great majority of the membrane-soluble helical pairs had at least one water-soluble counterpart that showed strong three-dimensional similarity.

Here, we focus exclusively on the membrane-soluble database to elucidate the major classes of frequently occurring helix-packing motifs in membrane proteins. The motifs are defined by comparing the three-dimensional structures of each helical pair with all other helical pairs in the library. Motifs are defined as clusters in which each member is within 1.5-Å rmsd of a central reference structure (the centroid). Remarkably, the great majority of the pairs within the database lie within a small number of clusters that have motifs that can be predicted from simple models for packing of α-helices. These results have strong

Conflict of interest statement: No conflicts declared.

Abbreviation: TM, transmembrane.

‡To whom correspondence should be addressed at: 1009B Stellar–Chance Building, University of Pennsylvania, Philadelphia, PA 19104-6059. E-mail: wdegrado@mail.med.upenn.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0605878103

Table 1. Characteristics of the top 14 clusters

Designation*	Rank	No. of members	Crossing angle,† °	Distance,† Å	Interchain, %	End-to-end distance,‡ residues
Frequent left-handed						
Antiparallel			−156.5	8.61		
GAS_{Left}	1	130	(10.1)	(0.89)	14.6	36
Parallel			13.8	9.77		
Left	4	42	(16.6)	(1.18)	31.0	118
Frequent right-handed						
Antiparallel			146.4	8.57		
GAS_{Right}	2	71	(13.6)	(0.99)	15.5	28
Parallel			−37.9	7.93		
GAS_{Right}	3	57	(7.50)	(0.88)	33.3	75
Other						
			178.0	9.14		
	5	29	(20.8)	(1.47)		
			25.5	8.55		
	6	12	(11.2)	(1.05)		
			−161.1	9.30		
	7	10	(10.3)	(1.57)		
			44.8	7.96		
	8	10	(8.8)	(1.13)		
			127.4	9.40		
	9	9	(12.3)	(1.00)		
			−60.2	8.61		
	10	7	(14.8)	(1.04)		
			−129.2	8.97		
	11	7	(12.9)	(1.65)		
			2.4	8.55		
	12	7	(16.2)	(0.78)		
			161.0	8.75		
	13	6	(17.6)	(1.33)		
			93.9	7.73		
	14	5	(4.9)	(7.82)		

*Topological designations are provided for the four most populous clusters to facilitate discussion of the results.
†Standard deviations are provided in parentheses.
‡Median number of residues between the N terminus of the first helix and C terminus of second helix.

implications for membrane protein structure prediction and design.

Results

A Limited Number of Motifs Account for Most Helix-Packing Geometries. A library of 445 helical pairs from 31 proteins were clustered into groups based on their three-dimensional similarity, as described in *Methods*. The pairs were clustered such that each pair within a cluster was within 1.5-Å rmsd from a central structure (the centroid). Average linking clustering was also explored and gave essentially the same results. Pairs of helices in the transmembrane region of membrane proteins tend to interact over at least a 10-residue segment, even when they have relatively wide crossing angles. Therefore, a minimum of 10 residues in each helix was required for the superposition algorithm, which defines a core segment that is common to each pair. Only one core segment was allowed per helical pair examined.

Surprisingly, the majority of the helical pairs fell into a small number of well defined clusters: 29% of the pairs fell into the first cluster, and 74% fell within clusters 1–5 (Fig. 1 and Table 1). Although the number of pairs per cluster dropped significantly after the 5th cluster, 80% of the pairs fell within the top 8 clusters and 90% within the first 14. These data indicate that the universe of helix–helix packings is significantly smaller than might have been anticipated. Inspection of the remaining pairs showed that one or both of the helices tended to be irregular, or the crossing angles were unusual.

Simple Geometric Features Account for the Primary Clusters of Helical Pairs. Fig. 2 illustrates the six most frequently populated clusters of helical pairs, each of which shows a high degree of overlap near the common core of the structure used for the superposition, with more variability near the ends. Within the tightly clustered regions, different clusters show varying degrees of homogeneity, as defined visually in Fig. 2 or by examining the variation in the helix crossing angles and distances in Table 1. Clusters with a large number of tightly clustered members presumably represent highly favorable packing arrangements with relatively deep and narrow energy wells. In the following section, we explore how the characteristics of the amino acid sequences and helical geometries define the structures of the four most populated clusters.

Clusters 1 and 4: Antiparallel and Parallel Pairs with Left-Handed Crossing Angles. Cluster 1 is the most populous, representing ≈29% of the helical pairs. In this packing arrangement, the helices lie nearly antiparallel to one another, with a small left-handed crossing angle between the two helices (Table 1). The helices are rather tightly associated (8.6 ± 0.9 Å), and many are involved in a variation of an Ala–coil interaction (19). The Ala–coil is a relatively rare helix-packing motif in water-soluble proteins, in which Ala residues spaced at seven-residue intervals mediate a tight association of antiparallel helices. In the Ala–coil motif, the helices coil around one another, thereby rendering each heptad geometrically identical.

BIOCHEMISTRY

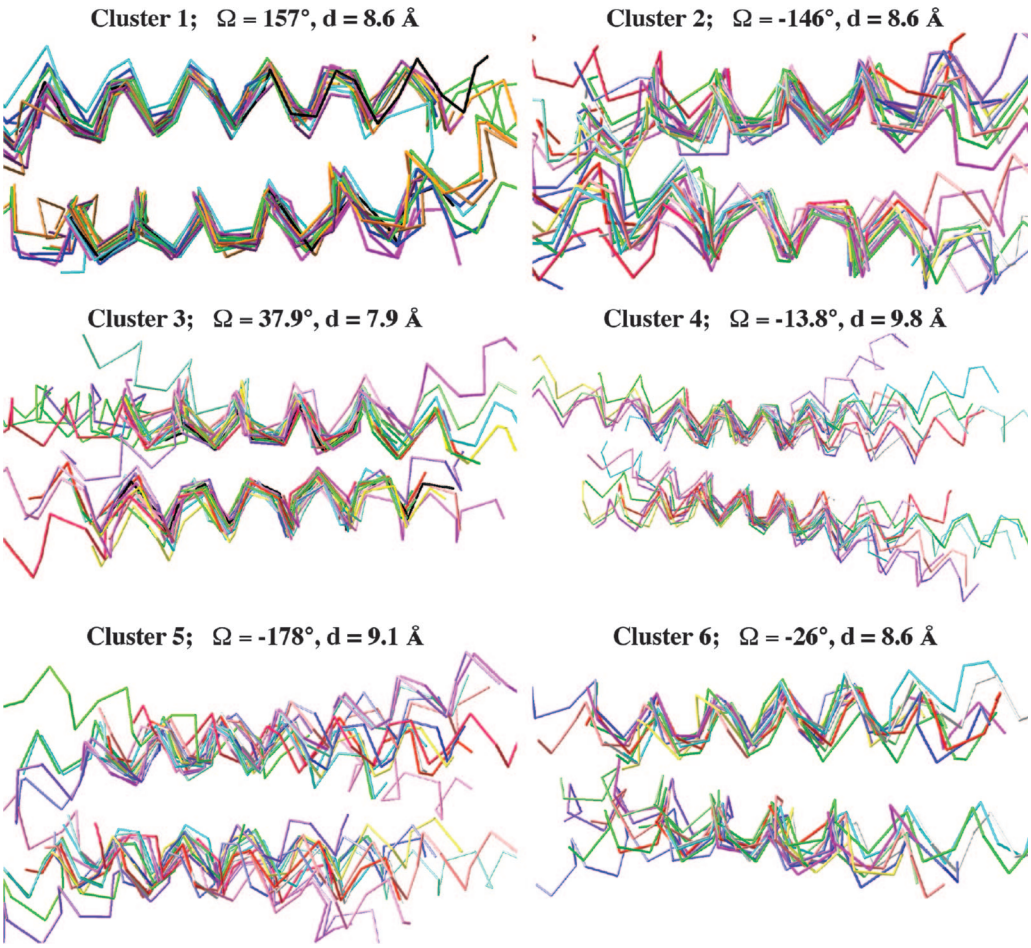

Fig. 2. Overlay of helical pairs in each cluster. Twenty members are shown, including the centroid and the 19 structures that are most similar to the centroid (based on rmsd).

An important feature in the Ala–coil motif is the packing of small residues in a heptad repeat. To test whether a similar motif is present in the cluster 1 motif, we determined sequence-specific amino acid propensities at each position of the aligned helices in cluster 1. A propensity for a given residue type at a specific position of a motif is defined to be the number of times that residue type occurs at that position relative to the number expected if the amino acids were randomly distributed in the motif. Because of the limited number of counts, the propensities of the small residues, Ala, Gly, and Ser, were averaged. Fig. 3*A* illustrates the structure of a typical member of this cluster together with the position-dependent variation of the propensities for Ala, Gly, and Ser in the motif. The graph is annotated according to the typical heptad nomenclature, in which the "a" and "d" positions are designated as those residues that most directly contact the opposing helix. As in the Ala–coil motif, small residues occur frequently at each "a" position. In general, the helices are not coiled but, instead, deviate from a central point of closest approach. Near this point, small residues are generally observed at the "d" or "e" positions too. Because of the high propensity of G, A, and S to occur at the helix–helix interface, we refer to this motif as an antiparallel GAS_{Left} motif, where the subscript refers to the crossing angle between the helices.

Cluster 4 represents a parallel version of the antiparallel packing observed in cluster 1. The packing of side chains in this motif is reminiscent of the packing between helices in parallel coiled coils. Small residues did not appear to be disproportionately represented in this cluster, and there were too few members to draw other sequence–structure correlations,

Clusters 2 and 3: Antiparallel and Parallel GAS_{Right} Motifs. Cluster 2 and 3 together comprise 29% of the helical pairs in the library. Cluster 3 is a parallel pair, which includes many examples of the GX_3G motif, which has been discussed extensively in the literature. We designate this motif the GAS_{Right} motif. In the majority of the examples of cluster 3, the GX_3G (or small-X_3-small) motif occurs on only one of the two helices. This finding is consonant with the view of Bowie, Engelman, and their coworkers (9, 20), who have documented numerous structural and functional variations on this fundamental motif.

An antiparallel version of the parallel GAS_{Right} motif is found in cluster 2. Although this cluster has more members than the glycophorin-like GAS_{Right} motif (71 versus 51 members in the parallel GAS_{Right} cluster), it has not been extensively analyzed.

Fig. 3*B* illustrates the structure of a typical member of the antiparallel GAS_{Right} motif, together with the position-dependent mean propensities for Ala, Gly, and Ser in the motif. As in the left-handed antiparallel GAS_{Left} motif, the right-handed antiparallel GAS_{Right} motif tends to have small residues at the helix–helix interface, but they are spaced at four-residue intervals, rather than the seven-residue spacing observed in cluster 1. The small residues tend to form a flat surface that docks against ridges formed by larger residues that are also spaced at four-residue intervals in the neighboring helix (Fig. 3*B*). Cys residues (also a small residue) sometimes are found at

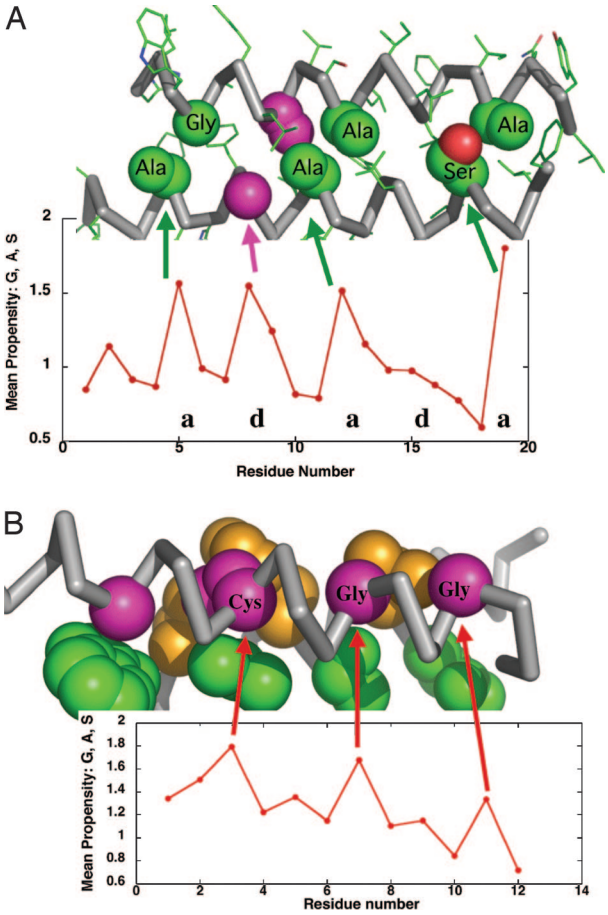

Fig. 3. Structures and sequence-specific propensities for GAS_{Left} (*A*) and GAS_{Right} (*B*) motifs. In each panel, an example of a structure of a helical pair from one of the clusters is shown together with the position-specific mean propensities for Ala, Gly, and Ser below the structure. The arrows connect positions of high propensity with specific locations in the structures. The example of a GAS_{Left} motif is taken from 1jb0 residues B42–64 and B134–155, and the residues at ''a'' positions of the pseudoheptad repeat are labeled. The residues at ''a'' positions are colored (C green and O red), and a ''d'' position is colored purple. The example of a GAS_{Right} motif in *B* is from 1u7g. Note that small residues (purple) occur every four residues on one of the two helices. These small residues fit between two ridges formed by larger residues on the other helix. The side chains that form one of these ridges are shown in green; the side chains occur with four-residue periodicity at positions i, $i + 4$, $i + 8$. . .. The second ridge is formed by residues shown in orange, also with four-residue periodicity but displaced by one residue in sequence (at positions $i + 1$, $i + 5$, $i + 9$. . .).

these sites, and an example is shown in Fig. 3*B*. In general, only one of the two helices has small residues at the interface.

Different Motifs Have Different Numbers of Residues Between the Interacting Helices. We examined the length of the intervening sequence connecting the two helices in different motifs. Relatively short loops tend to connect the antiparallel motifs (21); in the current database, median intervening lengths of 36 and 28 residues were observed for the antiparallel GAS_{Left} and GAS_{Right} versus 118 and 75 residues in the parallel left and parallel GAS_{Right} motifs, respectively. Clearly, one would expect that a parallel pair of helices would have a longer connecting sequence than an antiparallel pair, because, in a parallel pair, the protein chain must loop through the membrane at least once before it can connect to the second helix of the motif. However, the increase in length for parallel versus antiparallel pairs would appear to be longer than the chain length required to span the bilayer once (≈20 residues), indicating that other factors probably come into play.

We also determined whether the two helices in a given motif prefer to occur within the same chain or between two different chains of oligomeric proteins. Approximately 15% of both the antiparallel GAS_{Right} and GAS_{Left} pairs are formed between helices on different chains; the corresponding value for the two most populous parallel motifs is 32% (Table 1). Parallel pairs thus occur ≈2-fold more frequently between different subunits versus within a single subunit. Indeed, parallel motifs occur with approximately the same frequency as antiparallel motifs when the pairing involves two subunits, whereas antiparallel motifs are formed 2-fold more often when the helices occur within a single chain.

These findings are consistent with the known structures and possibly also the kinetic folding mechanisms of membrane proteins. Antiparallel pairs have a high propensity to form between sequentially adjacent helices, in a process that may occur predominantly in the translocon (22), leading to a bias for antiparallel pairs within a protein chain. Upon exiting the translocon, completion of the intrasubunit folding process involves formation of tertiary interactions between preformed antiparallel pairs, formation of new parallel pairs, and, occasionally, rearrangement of antiparallel pairs such that they are no longer formed between sequentially adjacent helices. The formation of interchain interactions, together with insertion of any appropriate cofactors, completes the folding process (9). Because of the lack of biasing interhelical loops, the formation of parallel and antiparallel helical pairs occurs with equal probability when the subunits associate in the final stage of folding.

The Geometry of the α-Helix Dictates the Helix-Packing Motifs. The geometries of the four primary clusters are in excellent agreement with previous theories of helical packing. Crick (23) first noticed that protruding side chains spaced in seven-residue increments form a left-handed spiral that influences the interhelical crossing angle between the two helices. Periodic packing of successive heptad segments on interacting helices provides efficient interdigitation of side chains in a ''knobs-into-holes'' manner and coiling, in the case of long coiled-coils. This theory has been extended to consider the packing of helices in globular proteins that are not necessarily coiled about one another (24, 25) and often show variations on classical knobs into holes packing. The expected crossing angle between two helices is a function of the interhelical distance and can be calculated from Eq. **1** (*Methods*). The predicted value for the antiparallel GAS_{Left} motifs is 158°C, in excellent agreement with the observed value of 157 ± 10°C (Table 1); the corresponding values for the left-handed parallel motif, −25°C and −14°C ± 17°C, respectively, also show reasonable agreement.

The right-handed packing angles are also readily explained by extensions of classical models for packing of helices (24, 25), including GX_3G motifs (26). The packing angle observed for the parallel GAS_{Right} (38 ± 8°C) is in good agreement with the value of 44°C calculated from Eq. **1**, assuming a four-residue repeat and a mean interhelical distance of 8 Å. The corresponding observed and computed values for the antiparallel GAS_{Right} are −146°C ± 14°C and −132°C, respectively.

These data indicate that most helix–helix packings in membrane proteins conform well to the simple principles expected from classical studies of helix packing. At least two features might account for the agreement with theory, which is better than in previous studies of water-soluble globular proteins: We performed the analysis on individual clusters rather than the overall database of helical pairs, allowing one to focus on just one family of structures. Also, the residues at the interface of

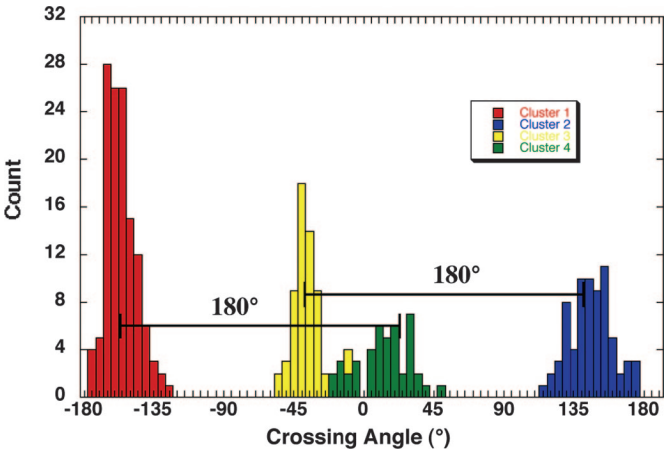

Fig. 4. Helix crossing-angle histograms for each of the four most-populated clusters: cluster 1 (antiparallel GAS_{Left}) is shaded red, cluster 2 (parallel GAS_{Left}) is shaded blue, cluster 3 (parallel GAS_{Right}) is shaded yellow, and cluster 4 is shaded green.

membrane helices tend to have smaller side chains with limited flexibility for accommodating different helical packing geometries. As described previously, the predominance of small residues rather than larger hydrophobic side chains, as in water-soluble proteins, presumably reflects the lack of a hydrophobic effect in membranes and the minimal entropic requirements for packing small side chains with few rotatable bonds.

Discussion

These results differ significantly from earlier surveys of helix-packing geometry, because the helical pairs were first clustered according to three-dimensional similarity to define distinct motifs from which packing geometries were then determined rather than determining packing parameters averaged over the entire database of pairs. As a result, there is very little overlap between the helix crossing angles in the members of different motifs (Fig. 4) in the first four clusters. Remarkably, two-thirds of helix–helix packings are defined by one of only four tightly clustered motifs, representing left- and right-handed parallel and antiparallel helical pairs. Furthermore, some of the less-populated clusters are variants of the primary four motifs. For example, clusters 5 and 6 (Table 1) are, respectively, variants on the parallel left (cluster 4) and antiparallel left motifs (antiparallel GAS_{Left}) clusters that vary slightly in the helical crossing angle, the interhelical distance, and the rotation of the helices about their own axes.

Our results have implications for both the design and the prediction of membrane protein structures. The availability of both idealized packing geometries and preferred sequence motifs will greatly facilitate the design of novel membrane proteins. Furthermore, the availability of position-specific sequence propensities for each of the packing motifs should allow one to predict the preferred packing arrangement for a TM helix, given its amino acid sequence. Although the number of examples of each motif is currently limited, a number of statistically significant trends in the propensity data were observed; for example, the placement of small residues at four- or seven-residue increments shown in Fig. 3 is just one clearly observed trend.

Methods

Initial work was accomplished by using the library of helical pairs described by Gimpelev *et al.* (16). Subsequently, using their method to define helical pairs, we updated their library to include a total of 445 helical pairs from high-resolution membrane proteins in the protein database as of August 2005: 1c3w, 1e12, 1ehk, 1eul, 1fi8, 1h2s, 1j4n, 1jb0, 1k4c, 1kb9, 1kf6, 1kpl, 1kqf, 1l7v, 1l9h, 1m3x, 1m56, 1msl, 1nek, 1ocr, 1pp9, 1pv6, 1pw4, 1q16, 1q90, 1qla, 1rc2, 1rh5, 1u7g, 1xfh, and 1yew (all with <30% sequence homology). The optimal alignment of each helical pair onto another helical pair was accomplished by using the algorithm of Jones (27), which uses a comparison of distance matrices for two pairs of discontinuous fragments to find sets of C-α atoms in the two pairs with similar three-dimensional relationships. Our implementation (see the supporting information, which is published on the PNAS web site) incorporated a strategy to allow rapid all-against-all comparison of the 445 helical pairs considered here. We allow helix 1 of the first pair to match to either helix 1 or helix 2 of the second pair (in which case helix 2 of the first pair would be mapped to helix 2 or 1, respectively, of the second pair). Once the appropriate atoms to be compared in each pair had been identified and their one-to-one relationship established, a standard three-dimensional superposition algorithm was used to find the optimal three-dimensional alignment and the associated rmsd. At least 10 residues in each helix were required for a superposition, and up to 14 residues were considered, if the inclusion of additional residues did not significantly increase the value of the rmsd. Helical angles and interhelical distances were computed from the helical axis vectors identified by using the program HELANAL (28). The angles and distances were computed for the superimposed subsections of the helical pairs, rather than the full helices, which often show irregularities outside of the superimposed regions.

Average linkage clustering (29) and single link clustering were used to cluster the helical pairs based on their rmsd. Both methods gave essentially the same result; the results of single link clustering are presented here. In this approach, pairs were clustered such that each member of a given cluster would be within 1.5-Å rmsd of a central reference pair (the centroid). The use of less stringent cut-offs did not greatly increase the number of structures per cluster, whereas increasing the stringency to 1.0 Å decreased the number of members per cluster, making it difficult to obtain meaningful sequence propensities because of limited counting statistics. The geometric characteristics and sequence propensities of the clusters are provided in the supporting information.

Once clusters had been established, the sequences within the clusters were aligned to the centroid (based on the three-dimensional alignment), and the amino acid propensity for each position in the aligned sequences was computed. Position-dependent propensities were computed as described (30).

The crossing angles computed for this library of helical pairs were compared with the expected values obtained from the packing of idealized α-helices. The idealized interhelical crossing angles (Ω) for the right-handed and left-handed crossing angles can be calculated from the geometry of the α-helix by using Eq. **1** (23, 26, 31):

$$\Omega = -2\ \tan^{-1}[(\omega_1 - \omega_\alpha)(d/2h)], \qquad [1]$$

in which ω_1 is the difference in the helical angular frequency of the integral repeat ($4\pi/55$ radians for a seven-residue repeat and $2\pi/4$ radians for a four-residue repeat); ω_α is the α-helical repeat, $\approx 2\pi/3.64$ radians per residue for an idealized membrane helix (32); d is the interhelical distance; and h is the rise per residue of the α-helix ($\approx$1.5 Å per residue).

We thank Donald Engel for helpful discussions. This work was supported, in part, by National Institutes of Health Grants GM60610 and GM56423 and by National Science Foundation Grant DMR05-20020.

1. Bowie JU (2005) *Nature* 438:581–589.
2. Von Heijne G (2003) *Adv Protein Chem* 63:1–18.
3. DeGrado WF, Gratkowski H, Lear JD (2003) *Protein Sci* 12:647–665.
4. Chamberlain AK, Faham S, Yohannan S, Bowie JU (2003) *Adv Protein Chem* 63:19–46.
5. Lemmon MA, Flanagan JM, Treutlein HR, Zhang J, Engelman DM (1992) *Biochemistry* 31:12719–12725.
6. Senes A, Gerstein M, Engelman DM (2000) *J Mol Biol* 296:921–936.
7. Senes A, Engel DE, DeGrado WF (2004) *Curr Opin Struct Biol* 14:465–479.
8. Kleiger G, Grothe R, Mallick P, Eisenberg D (2002) *Biochemistry* 41:5990–5997.
9. Engelman DM, Chen Y, Chin CN, Curran AR, Dixon AM, Dupuy AD, Lee AS, Lehnert U, Matthews EE, Reshetnyak YK, *et al.* (2003) *FEBS Lett* 555:122–125.
10. Langosch D, Heringa J (1998) *Proteins* 31:150–159.
11. Adamian L, Liang J (2001) *J Mol Biol* 311:891–907.
12. Javadpour MM, Eilers M, Groesbeek M, Smith SO (1999) *Biophys J* 77:1609–1618.
13. Gratkowski H, Lear JD, DeGrado WF (2001) *Proc Natl Acad Sci USA* 98:880–885.
14. Zhou FX, Merianos HJ, Brunger AT, Engelman DM (2001) *Proc Natl Acad Sci USA* 98:2250–2255.
15. Bowie JU (1997) *J Mol Biol* 272:780–789.
16. Gimpelev M, Forrest LR, Murray D, Honig B (2004) *Biophys J* 87:4075–4086.
17. Bowie JU (1997) *Nat Struct Biol* 4:915–917.
18. Engel DE, DeGrado WF (2005) *Proteins* 61:325–337.
19. Gernert KM, Surles MC, Labean TH, Richardson JS, Richardson DC (1995) *Protein Sci* 4:2252–2260.
20. Kim S, Jeon TJ, Oberai A, Yang D, Schmidt JJ, Bowie JU (2005) *Proc Natl Acad Sci USA* 102:14278–14283.
21. Spencer RH, Rees DC (2002) *Annu Rev Biophys Biomol Struct* 31:207–233.
22. White SH, von Heijne G (2004) *Curr Opin Struct Biol* 14:397–404.
23. Crick FHC (1953) *Acta Crystallogr* 6:689–697.
24. Chothia C (1984) *Annu Rev Biochem* 53:537–572.
25. Walther D, Eisenhaber F, Argos P (1996) *J Mol Biol* 255:536–553.
26. Dieckmann GR, DeGrado WF (1997) *Curr Opin Struct Biol* 7:486–494.
27. Jones TA (2004) *Acta Crystallogr D* 60:2115–2125.
28. Bansal M, Kumar S, Velavan R (2000) *J Biomol Struct Dyn* 17:811–819.
29. King B (1967) *J Am Stat Assoc* 69:86–101.
30. Engel DE, DeGrado WF (2004) *J Mol Biol* 337:1195–1205.
31. North B, Summa CM, Ghirlanda G, DeGrado WF (2001) *J Mol Biol* 311:1081–1090.
32. Mohapatra PK, Khamari A, Raval MK (2004) *J Mol Model (Online)* 10:393–398.