Supporting Material:
Robust Driving Forces for Transmembrane Helix Packing

Ayelet Benjamini
Department of Chemistry,
University of California in Berkeley, Berkeley, CA, USA

Berend Smit
Department of Chemistry,
University of California in Berkeley, Berkeley, CA, USA
Department of Chemical and Biomolecular Engineering,
University of California in Berkeley, Berkeley, CA, USA
Materials Sciences Division,
Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Database analysis

The extraction of experimental helix structure data from the OPM database included a few steps:

- We refine the database to include only proteins with TM helices. This includes *alpha-helical polytopic* and *alpha-helical bitopic* OPM classes as of September 2011 (357 TM proteins, 261 unique). The full list of MPs is available in Table S2.
- We extract the structure of each such protein from the database.
- We extract other supporting information on the protein such as the hydrophobic thickness of the membrane surrounding it (d_{L}) and the list of the protein's transmembrane segments.
- For each TM segment of the protein we analyze the secondary structure of its residues using DSSP (1) to determine whether it has a range of residues with helical geometry.
- For each TM helix we determine the range of residues that span its hydrophobic core. The original TM segment is therefore extended if it includes more hydrophobic-scored helical residues at either end or shortened if it includes hydrophilic-scored helical residues at either end. Details on the hydrophobicity scoring function is available in the main text.

By the aforementioned process we obtained a list of hydrophobic helix cores for each membrane protein. We further analyze each hydrophobic helix core with the following steps:

- We refine the set of TM core helices to include only helices with more than 9 residues, as was done by Walters *et al.* (2).
- We analyze the structure of each helix core to define its major axis using the HELANAL (3) algorithm available through the MDAnalysis (4) package.
- We obtain the length of the hydrophilic part of the helix, d_{H}, by projecting the C_α positions (origins) onto the obtained helix axis vector.

- We define the hydrophobic mismatch of each helix by subtracting the thickness of the membrane ($\Delta d = d_H - d_L$) for each helix.
- We refine the set to analyze helices that are in the bulk range (96%) of hydrophobic mismatch (-10 Å $\leq \Delta d \leq 35$ Å) and have a reasonable length per residue (in range [1.25,1.8] Å/residue; theoretical length per residue is 1.5 Å/residue).
- We eliminate kinked helices by screening helices that are not centered around the bilayer center. Helices whose center-of-mass distance from the bilayer center ($z = 0$) in the bilayer normal direction (z) exceeded 20% of the helix length were defined as part of a kinked helix and were not used in further analysis.
- We define each helix as *unique / non-unique* based on its similarity in sequence and TM segment range to other helices in the same MP. Therefore if a protein was in fact a dimer of two smaller subunits, only helices of the first monomer were defined as *unique* while the other helices were tagged as *non-unique*.

Contacts between same-protein helices were calculated following the analysis in Gimpelev *et al.* (5). Two helices were defined to be a neighboring pair if they contained at least three residues in contact. Residues were defined to be in contact if the distance between any two of their atoms was within 0.6 Å of the sum of their van der Waals radii (6), as calculated by Li & Nussinov (7). Helix pairs were restricted to pairs in which at least one of the helices is *unique*, thereby eliminating over-counting of helix pairs. The number of contacts each helix has with other same-protein helices was also used as a measure of how influenced the helix is by other surrounding helices. Helices with less than one contact per helix residue (average 0.57 contacts / residue) were defined as peripheral helices with minimal helix-helix interactions.

Cross angle distribution comparison

We note that the choice of neighboring helix groups in Fig. 5 of the main text was made by balancing two factors: In order to display a representative cross angle distribution, a large enough number of data points (neighboring pairs) should be accounted for, implying the use of a large range of hydrophobic mismatch. The maximum density of helix pairs is 33 pairs/Δd^2 (see Fig. S3) and so a range of several Å in hydrophobic mismatch is needed to include enough data points in the distribution of pair cross angles. On the other hand, using a too-large hydrophobic mismatch range would include large variability in the tilt angle distribution of the helices (observed change of up to $1.3°/\Delta d$ in average tilt angle) and therefore a variability in the resulting cross angle distribution. We therefore chose windows of an intermediate size of 3Å in mismatch ranges that include a large enough number of data points. We focus on neighboring helix pairs where both helices have a positive mismatch for reasons described in the main text. Figure S3 shows the scatter-plot of neighboring helix pairs and the chosen ranges are shown for comparison.

Reference cross angle distribution

We sampled the reference distribution of cross angle by a Monte Carlo procedure for all neighboring helix pairs in the OPM database. For each pair, we extract the hydrophobic mismatches of both helices, ($\Delta d_1; \Delta d_2$). This imposes the expected tilt angle distribution of each helix, based on the information displayed in Fig. 2(a) of the main text, ($\langle \theta_1 \rangle , \sigma_{\theta_1}; \langle \theta_2 \rangle , \sigma_{\theta_2}$). Note that we do not explicitly use the helices tilt angle but rather the average distribution based on their mismatch, therefore obtaining the reference distribution in the case that hydrophobic mismatch would be the only factor determining the helix tilt. We then produce a set of $N = 1000$ values of tilt angles for each helix in the pair, drawn from the imposed distribution, $\theta_1^i \sim \mathcal{N}(\langle \theta_1 \rangle , \sigma_{\theta_1})$ and $\theta_2^i \sim \mathcal{N}(\langle \theta_2 \rangle , \sigma_{\theta_2})$, for $i = 1 \ldots N$. For each i, a set of $j = 1 \ldots M = 1000$ uniformly distributed projection γ_j^i angles are drawn at random, and M cross angle values, Ω_j^i, are then calculated based on Equation 1 in the main text by plugging in θ_1^i, θ_2^i, γ_j^i. We therefore obtain $MxN = 10^6$

independent cross angle values which we bin into a histogram of narrow $0.1°$ bins. This histogram represents the reference discrete distribution for this specific helix pair.

To obtain the overall expected cross angle histogram of all helix pairs we combine the reference histograms by adding the counts of each bin for every helix pair and normalizing the resulting distribution at the end of the procedure. As the reference distribution does not differentiate between negative and positive cross angles we calculate the cross angle value in its positive form, and set its sign randomly such that the percentage of parallel pairs will be identical to its value in the data (58.5%).

We further use the reference histograms to calculate the overall p-value of the experimental results. For each helix pair ($k = 1 \dots 2328$) we calculate the probability of the pair's cross angle, $\Pr^k(\Omega_k)$, according to the pair's corresponding reference histogram. We then sum over the probabilities of all the other bins in that reference histogram with lower probability to obtain the k'th p-value, $p^k = \sum_x \Pr^k(x) \mid \Pr^k(x) \leq \Pr^k(\Omega_k)$. To verify that the discrete binning does not bias the p-value and to get an estimate on the variance of this result, we calculate the same score on $R = 500$ random numbers sampled directly from the histogram $p^{k,l}$ for $l = 1 \dots R$. We then obtain an experimental overall p-value $p = \sum_{k=1}^{2328} p^k = 0.497$ and R random p-values $p^l = \sum_{k=1}^{2328} p^{k,l}$. The average random p-value is calculated by averaging these R values. As expected, it comes out to be 0.500. The standard deviation is also obtain from these R values as 0.006.

CG Model and Simulation Parameters

Our model consists of 3 types of beads: 1. a water like bead, $\textcircled{w}$, used to describe a set of three water molecules; 2. a hydrophilic bead, $\textcircled{h}$, used to model the lipid head group as well as the marginal part of the helix; 3. a hydrophobic tail bead, $\textcircled{t}$, used to model the hydrophobic lipid tail; and 4. a hydrophobic protein bead, $\textcircled{p}$, used to model the hydrophobic core of a TM helix. The lipid model includes a head group consisting of three $\textcircled{h}$ type beads and two tails each containing five $\textcircled{t}$ type beads. See Figure 1 in main text.

Our transmembrane helix model contains three structural features: 1. N_p principal beads, P, that follow the outer radius structure of an α-helix backbone. All structural bonds are associated with those beads. 2. $N_\mathrm{s} = N_\mathrm{p} - 1$ secondary beads, S, located in between every adjacent pair of principal beads and provide excluded volume along the helix exterior. 3. $N_\mathrm{c} = \mathrm{floor}(N_\mathrm{p}/3)$ central beads, C, located along the central axis of the helix and provide excluded volume at the helix interior.

The principal beads of our TM helix model are set along a helix of radius $R_\mathrm{h} = 6$ Å , with angle and height pitch of $\Delta\theta = 100\,^\circ$; $\Delta Z = 1.5$ Å between consecutive principal beads. To generate the positions of N_p principal helix beads we chose an initial angle θ_0 at random. We then place the first bead of the helix at position $\vec{r_0} = (R_\mathrm{h} \times \cos(\theta_0), R_\mathrm{h} \times \sin(\theta_0), 0)$. For each bead thereafter, $\mathsf{P}_{i=1\ldots N_\mathrm{p}-1}$, its position is sequentially determined by $\vec{r_i} = (R_\mathrm{h} \times \cos(\theta_0 + \Delta\theta \cdot i), R_\mathrm{h} \times \sin(\theta_0 + \Delta\theta \cdot i), \Delta Z \cdot i)$.

We next position the secondary beads in between each pair of adjacent principal beads such that for each bead $\mathsf{S}_{j=0\ldots N_\mathrm{s}-1}$; $\vec{r_j} = \left(R_\mathrm{h} \times \cos(\theta_0 + \Delta\theta \cdot (j+\frac{1}{2})), R_\mathrm{h} \times \sin(\theta_0 + \Delta\theta \cdot (j+\frac{1}{2})), \Delta Z \cdot (j+\frac{1}{2})\right)$. Lastly, the central beads are positioned along the major axis of the helix. A central bead's z-component is determined by the average position of the three principal beads it is adjacent to. This will yield for each central bead $\mathsf{C}_{k=0\ldots N_\mathrm{c}-1}$; $\vec{r_k} = (0, 0, \Delta Z \cdot (3k+1))$.

Helix bead types are assigned such that the core of the helix consists of hydrophobic $\textcircled{p}$ type beads. Both edges of the helix, consisting of: three principal beads, two secondary beads and one central bead are assigned $\textcircled{h}$ type. This assignment ensures that both edges of the protein remain in the water phase thus preventing unphysical configurations such as a helix lying within the membrane perpendicular to the membrane normal.

The model helix is placed within a pre-equilibrated membrane such that the z-direction of the helix corresponds to the bilayer normal.

Harmonic bond forces are used to control the inter-bead distance. These are applied between: (*a*) each principal bead P_i and its two adjacent secondary beads $\mathsf{S}_{j=i-1,i}$ (*b*) each two adjacent central beads C_k and C_{k+1} (*c*) each central bead C_k and its three adjacent principal beads $\mathsf{P}_{i=3k,3k+1,3k+2}$, and (*d*) between each principal bead P_i and its consecutive P_{i+4} thus mimicking α-helix hydrogen bonding .

Harmonic angle forces control the stiffness of the helical structure. These are applied between: (*a*) each three consecutive principal beads: $(\mathsf{P}_i, \mathsf{P}_{i+1}, \mathsf{P}_{i+2})$ (*b*) each three principal beads surrounding a hydrogen bond: $(\mathsf{P}_i, \mathsf{P}_{i+4}, \mathsf{P}_{i+5})$, and (*c*) each three consecutive central beads: $(\mathsf{C}_k, \mathsf{C}_{k+1}, \mathsf{C}_{k+2})$.

Dihedral angle forces regulate distortions in helix structure and prevent the helix from unfolding. Those dihedral forces are applied between all sets of consecutive principal beads: $(\mathsf{P}_i, \mathsf{P}_{i+1}, \mathsf{P}_{i+2}, \mathsf{P}_{i+3})$. Helix structure features are controlled almost exclusively by interactions of principal beads. secondary and central beads are positioned to serve as exclusive volume and are only weakly bonded to ensure they stay in close proximity to the helical structure. Interaction parameters are available in Table S1.

We simulate the system in the $NP_\perp\gamma T$ ensemble using a hybrid MC-DPD technique identical to what is described by de Meyer *et al.* (8). $N = 2014/2002$ lipids for single/pair helices respectively; reduced temperature $T^* = 0.7$ in which the lipid bilayer is well within the L_α phase; $\gamma = 0$ dyn/cm; Normal pressure is equal to the water bulk pressure, $P^* = 22.3$ in reduced units.

Table S1 - CG Bonding Parameters

Interaction	Interacting Beads / Formula	Force constant	Equilibrium value
U_{bond}	$\frac{1}{2}K_b(r - r_{eq})^2$	$K_b\ (\epsilon_0/{d_0}^2)$	$r_{eq}\ (d_0)$
	$P_i \quad S_{i-1}$	100	0.7936
	$P_i \quad S_i$	100	0.7936
	$P_i \quad P_{i+4}$	100	1.1241
	$C_k \quad C_{k+1}$	100	0.6960
	$C_k \quad P_{3k}$	20	0.9575
	$C_k \quad P_{3k+1}$	20	0.9290
	$C_k \quad P_{3k+2}$	20	0.9575
U_{angle}	$\frac{1}{2}K_a(\varphi - \varphi_{eq})^2$	$K_a\ (\epsilon_0)$	$\varphi_{eq}\ (°)$
	$P_i \quad P_{i+1} \quad P_{i+2}$	600	81.2
	$P_i \quad P_{i+4} \quad P_{i+5}$	100	108.9
	$C_k \quad C_{k+1} \quad C_{k+2}$	20	180.0
$U_{dihedral}$	$\frac{1}{2}K_d[\cos(\chi) - \cos(\chi_{eq})]^2$	$K_d\ (\epsilon_0)$	$\chi_{eq}\ (°)$
	$P_i \quad P_{i+1} \quad P_{i+2} \quad P_{i+3}$	1000	21.7

Table 1: Bonded interaction parameters for the CG helix model. Reduced units correspond to $\epsilon_0 = 1\ k_B T$, $d_0 = 6.46\ \text{Å}$.
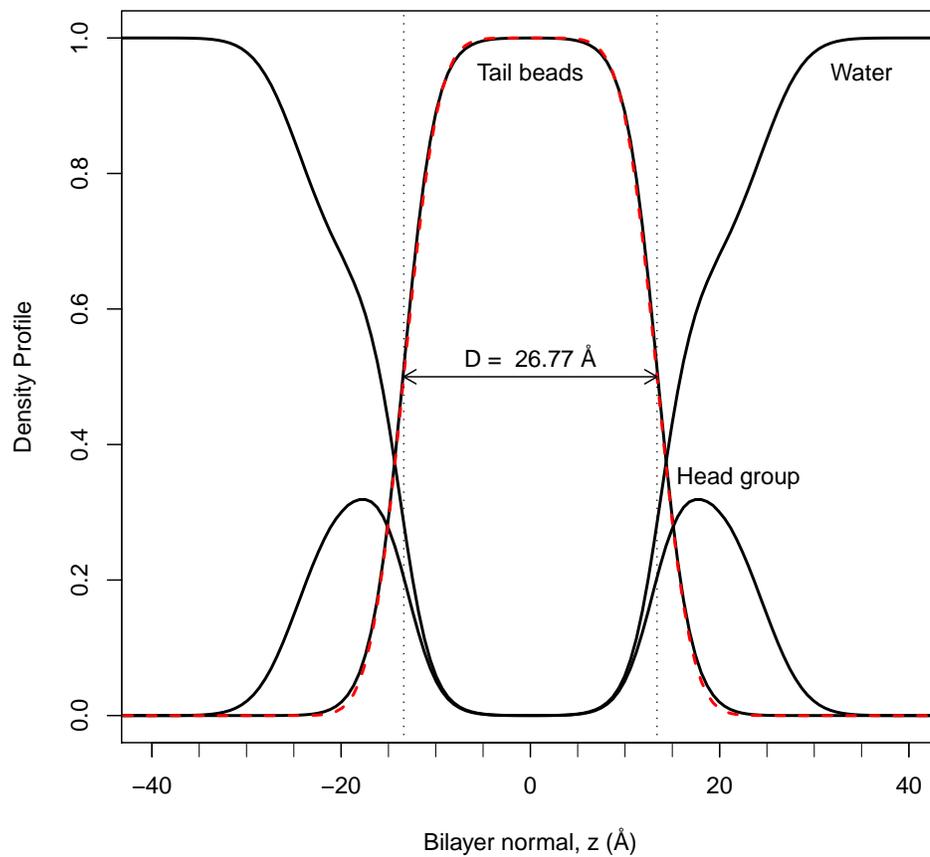
Figure S1

Figure 1: Density profile of a pure bilayer simulation. Solid lines show the volume fraction of the various components in the simulation: lipid head beads (head group), lipid tail beads and water. Red dashed line corresponds to the symmetric error function: $P_{HC}(z) = \frac{1}{2}\left[erf(z, -z_{HC}, \sigma_{HC}) - erf(z, z_{HC}, \sigma_{HC}) \right]$ as defined by Eq. 3 in Kucerka *et al.* (9). Fit values correspond to $z_{HC} = 13.385$ Å and $\sigma_{HC} = 2.85$ Å yielding a hydrophobic thickness of $d_{\mathrm{L}} = 26.77$ Å.
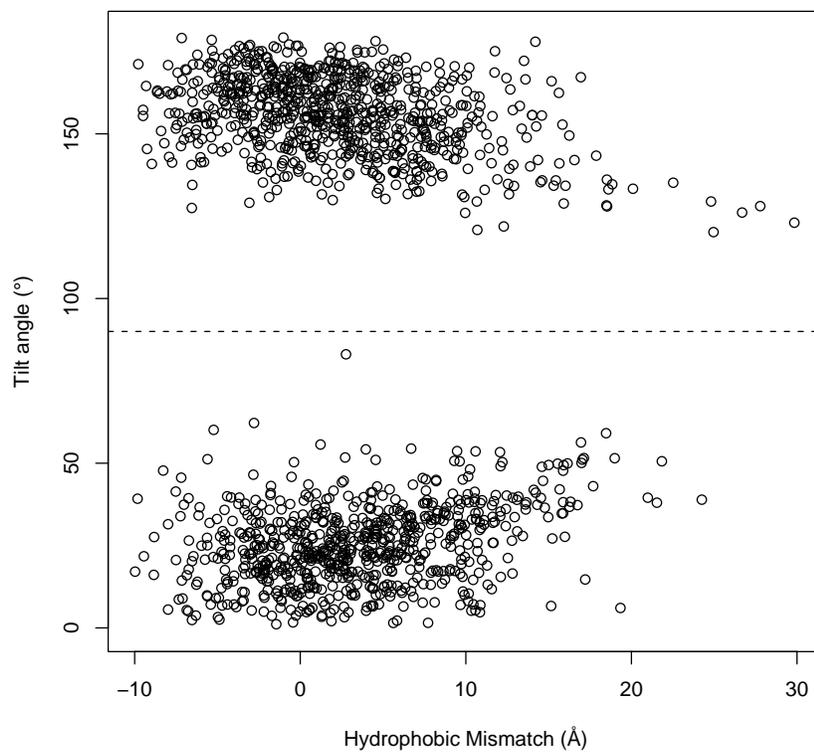
Figure S2

Figure 2: Unadjusted tilt angles versus hydrophobic mismatch scatter plot shows a mirror effect around $\theta = 90°$ (dashed horizontal line).
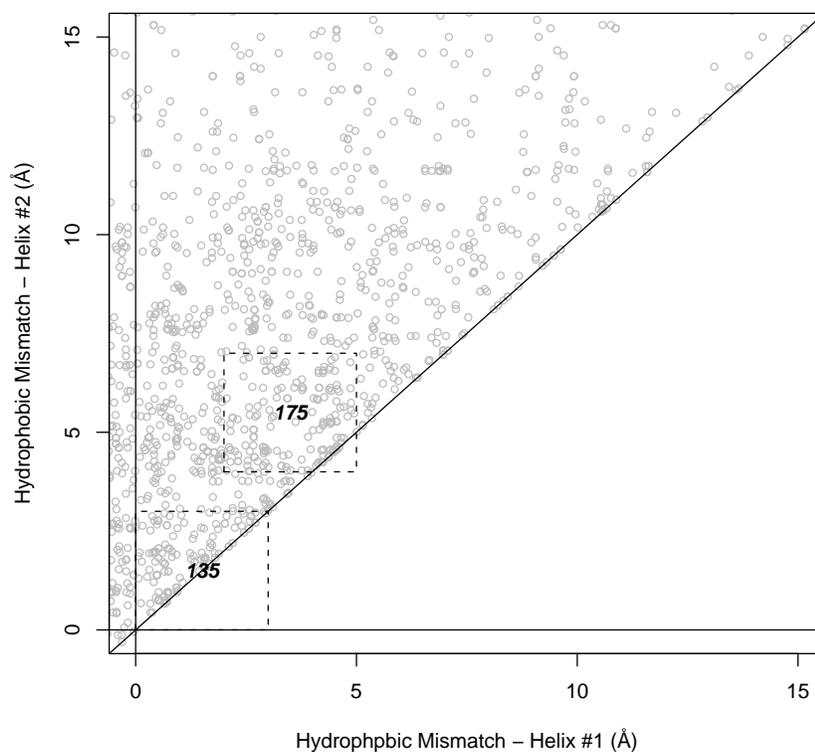
Figure 3: Scatter plot of neighboring helix pairs with positive hydrophobic mismatch as a function of helix hydrophobic mismatch. X-axis represents the hydrophobic mismatch of the helix with the lower mismatch value in the pair. Y-axis represents the hydrophobic mismatch of the helix with the higher mismatch value in the pair. Diagonal line represents pairs with equal mismatch values for both helices. Dashed rectangles correspond to mismatch ranges used in Fig. 5 in the main text. The number on each rectangle represents the number of pairs in that rectangle.
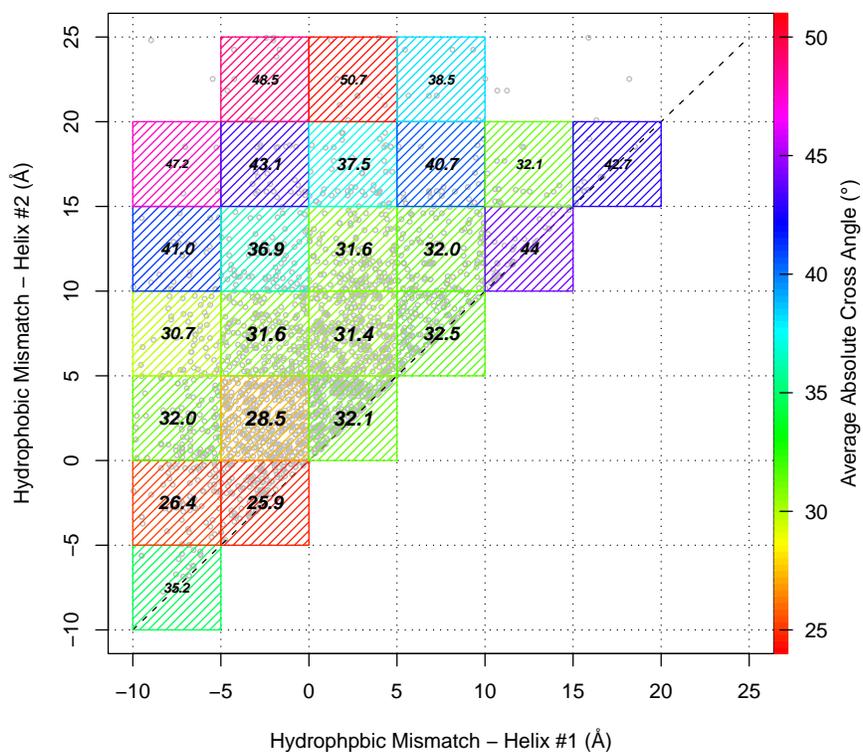
Figure 4: Scatter plot of neighboring helix pairs as function of helix hydrophobic mismatch. Available data points represented by grey circles. Colors and text correspond to average absolute cross angle for pairs in each mismatch range. Text size scales with the number of data points in each range, representing the accuracy in average value calculation. Larger text therefore corresponds to smaller error in average calculation. The plot shows the dependence of cross angle on the hydrophobic mismatch of both helices. We observe smaller cross angle values for helices with small hydrophobic mismatch, and larger cross angle values for helices with large hydrophobic mismatch. We additionally observe larger cross angle values for helices whose mismatch ranges differ (increase in average cross angle towards upper left corner of the plot). Both phenomena are expected from our reference model.

Figure S5
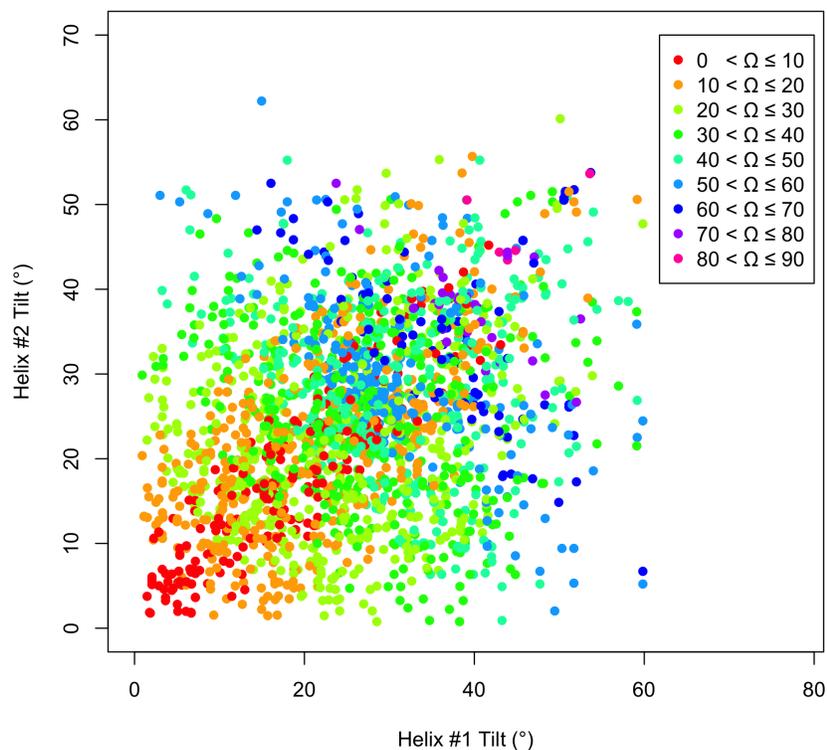
Figure 5: Scatter plot of tilt angles for both helices across all neighboring helix pairs. Colors represent the cross angle of the pair, according to the legend. The plot shows the dependence of cross angle on the tilt angle of both pairs, as expected by Eq. 1 in the main text. Small cross angles ($\Omega \leq 10$, red circles) are available only to pairs with a small difference in tilt angles (along the diagonal). While moving away from the diagonal, the cross angles are limited to growingly larger values. The dependence of cross angle value on the sum of tilt angles is also apparent. Each value of cross angles is limited in the bottom left direction of the plot, suggesting a minimal limit on the sum of tilt angles.

Table S2 - Experimental Structures

PDB	Experimental Method	Temperature	# Unique Helices	# Neigh. Pairs
1A11	SOLUTION NMR	313	1	0
1A91	SOLUTION NMR	300	2	1
1AFO	SOLUTION NMR	313	1	1
1AR1	X-RAY DIFFRACTION	N/A	12	18
1DXR	X-RAY DIFFRACTION	N/A	9	8
1E12	X-RAY DIFFRACTION	N/A	6	13
1EHK	X-RAY DIFFRACTION	293	13	21
1EYS	X-RAY DIFFRACTION	277	11	15
1FFT	X-RAY DIFFRACTION	277	13	12
1FJK	SOLUTION NMR	300	1	0
1H2S	X-RAY DIFFRACTION	N/A	9	18
1H6I	ELECTRON CRYSTALLOGRAPHY	N/A	6	15
1HGZ	FIBER DIFFRACTION	300	2	2
1IFP	FIBER DIFFRACTION	298	1	1
1IJD	X-RAY DIFFRACTION	291	1	2
1J4N	X-RAY DIFFRACTION	277	6	15
1JB0	X-RAY DIFFRACTION	277	26	25
1KF6	X-RAY DIFFRACTION	295	6	8
1KPL	X-RAY DIFFRACTION	293	8	11
1KQF	X-RAY DIFFRACTION	277	4	6
1L9B	X-RAY DIFFRACTION	292	11	14
1LDF	X-RAY DIFFRACTION	298	6	19
1LGH	X-RAY DIFFRACTION	N/A	1	2
1M0L	X-RAY DIFFRACTION	295	6	11
1M56	X-RAY DIFFRACTION	277	21	35
1MHS	ELECTRON CRYSTALLOGRAPHY	277	7	6
1N7L	SOLUTION NMR	323	1	0
1NEK	X-RAY DIFFRACTION	293	6	7
1NKZ	X-RAY DIFFRACTION	289	1	3
1OKC	X-RAY DIFFRACTION	N/A	6	7
1OQW	X-RAY DIFFRACTION	293	1	0
1OTS	X-RAY DIFFRACTION	298	8	10
1P49	X-RAY DIFFRACTION	298	2	1
1PP9	X-RAY DIFFRACTION	277	11	18
1PW4	X-RAY DIFFRACTION	293	11	18
1PY6	X-RAY DIFFRACTION	310	6	7
1Q90	X-RAY DIFFRACTION	277	13	25
1QL1	FIBER DIFFRACTION	283	2	0
1R3J	X-RAY DIFFRACTION	293	2	3
1RC2	X-RAY DIFFRACTION	300	6	15
1RKL	SOLUTION NMR	298	1	0
1RWT	X-RAY DIFFRACTION	291	3	1

PDB	Experimental Method	Temperature	# Unique Helices	# Neigh. Pairs
1S5H	X-RAY DIFFRACTION	293	2	3
1SU4	X-RAY DIFFRACTION	N/A	7	8
1T5S	X-RAY DIFFRACTION	292	8	7
1U7G	X-RAY DIFFRACTION	298	9	28
1UAZ	X-RAY DIFFRACTION	280	6	7
1V55	X-RAY DIFFRACTION	N/A	22	47
1VGO	X-RAY DIFFRACTION	283	6	7
1WPG	X-RAY DIFFRACTION	N/A	9	12
1WU0	SOLUTION NMR	298	2	1
1XIO	X-RAY DIFFRACTION	293	7	12
1XL6	X-RAY DIFFRACTION	293	2	5
1XRD	SOLUTION NMR	298	1	0
1YCE	X-RAY DIFFRACTION	290	2	7
1YEW	X-RAY DIFFRACTION	298	7	5
1YMG	X-RAY DIFFRACTION	298	6	17
1YQ3	X-RAY DIFFRACTION	278	6	8
1ZCD	X-RAY DIFFRACTION	279	7	9
1ZLL	SOLUTION NMR	303	1	2
1ZOY	X-RAY DIFFRACTION	290	5	6
1ZRT	X-RAY DIFFRACTION	277	10	13
2A0L	X-RAY DIFFRACTION	293	2	3
2A65	X-RAY DIFFRACTION	291	9	18
2A9H	SOLUTION NMR	315	2	3
2AGV	X-RAY DIFFRACTION	283	5	4
2B2F	X-RAY DIFFRACTION	298	9	28
2B5F	X-RAY DIFFRACTION	N/A	5	9
2B6O	ELECTRON CRYSTALLOGRAPHY	300	6	17
2BBJ	X-RAY DIFFRACTION	297	1	0
2BG9	ELECTRON MICROSCOPY	277.20	16	28
2BHW	X-RAY DIFFRACTION	N/A	1	0
2BL2	X-RAY DIFFRACTION	N/A	4	11
2BS2	X-RAY DIFFRACTION	N/A	5	9
2BS3	X-RAY DIFFRACTION	N/A	4	6
2E74	X-RAY DIFFRACTION	N/A	10	20
2F2B	X-RAY DIFFRACTION	295	6	19
2FYN	X-RAY DIFFRACTION	288.20	10	15
2GFP	X-RAY DIFFRACTION	N/A	10	15
2H8A	ELECTRON CRYSTALLOGRAPHY	277	3	4
2HIL	ELECTRON MICROSCOPY	368	1	0
2HYD	X-RAY DIFFRACTION	N/A	4	4
2IFO	FIBER DIFFRACTION	298	2	0
2IQL	THEORETICAL MODEL	298	6	9
2IQO	THEORETICAL MODEL	298	6	9
2IQR	THEORETICAL MODEL	298	5	6

PDB	Experimental Method	Temperature	# Unique Helices	# Neigh. Pairs
2IQV	THEORETICAL MODEL	298	6	9
2JLN	X-RAY DIFFRACTION	N/A	7	6
2JWA	SOLUTION NMR	313	2	1
2K1K	SOLUTION NMR	313	1	1
2K74	SOLUTION NMR	313	2	1
2K9P	SOLUTION NMR	320	2	1
2K9Y	SOLUTION NMR	313	1	1
2KA2	SOLUTION NMR	313	1	1
2KB7	SOLID-STATE NMR - SOLUTION NMR	N/A	1	0
2KDC	SOLUTION NMR	318	1	0
2KIX	SOLUTION NMR	305	1	2
2KNC	SOLUTION NMR	298	2	1
2KQT	SOLID-STATE NMR	243	1	2
2KSD	SOLUTION NMR	318	2	1
2KSE	SOLUTION NMR	313	2	1
2KSF	SOLUTION NMR	318	3	2
2KSR	SOLUTION NMR	313	4	3
2KV5	SOLUTION NMR	303	1	0
2KWX	SOLUTION NMR	303	1	2
2KYH	SOLUTION NMR	318	3	2
2L35	SOLUTION NMR	303	3	3
2L9U	SOLUTION NMR	313	1	1
2LAT	SOLUTION NMR	298	1	0
2LCK	SOLUTION NMR	306	4	2
2NQ2	X-RAY DIFFRACTION	277	8	12
2NR9	X-RAY DIFFRACTION	298	4	4
2NRF	X-RAY DIFFRACTION	295	4	4
2NS1	X-RAY DIFFRACTION	293	9	27
2NWL	X-RAY DIFFRACTION	277	5	3
2O01	X-RAY DIFFRACTION	278	6	0
2OAR	X-RAY DIFFRACTION	N/A	2	5
2OAU	X-RAY DIFFRACTION	277	1	0
2Q7R	X-RAY DIFFRACTION	298	2	2
2QI9	X-RAY DIFFRACTION	293	14	18
2QKS	X-RAY DIFFRACTION	293	2	5
2QTS	X-RAY DIFFRACTION	277	2	4
2RDD	X-RAY DIFFRACTION	293	10	13
2RH1	X-RAY DIFFRACTION	293	7	10
2RLF	SOLUTION NMR	303.10	1	2
2UUH	X-RAY DIFFRACTION	N/A	3	8
2V50	X-RAY DIFFRACTION	N/A	22	41
2V8N	X-RAY DIFFRACTION	N/A	5	2
2VL0	X-RAY DIFFRACTION	N/A	2	0
2VPZ	X-RAY DIFFRACTION	N/A	14	25

PDB	Experimental Method	Temperature	# Unique Helices	# Neigh. Pairs
2VT4	X-RAY DIFFRACTION	N/A	7	11
2W2E	X-RAY DIFFRACTION	N/A	5	18
2W5J	X-RAY DIFFRACTION	N/A	1	2
2WCD	X-RAY DIFFRACTION	N/A	1	2
2WIT	X-RAY DIFFRACTION	N/A	12	26
2WLL	X-RAY DIFFRACTION	N/A	2	5
2WSW	X-RAY DIFFRACTION	N/A	7	8
2WSX	X-RAY DIFFRACTION	N/A	10	19
2WWB	ELECTRON MICROSCOPY	357	7	4
2XKM	SOLID-STATE NMR; X-RAY DIFFRACTION	N/A	1	0
2XOK	X-RAY DIFFRACTION	N/A	2	7
2XQ2	X-RAY DIFFRACTION	N/A	11	13
2XQU	X-RAY DIFFRACTION	N/A	2	7
2XUT	X-RAY DIFFRACTION	N/A	8	9
2XZB	ELECTRON CRYSTALLOGRAPHY	277.50	7	4
2YDV	X-RAY DIFFRACTION	N/A	7	12
2YL4	X-RAY DIFFRACTION	N/A	5	7
2YVX	X-RAY DIFFRACTION	293	3	3
2YXR	X-RAY DIFFRACTION	277	8	7
2ZBD	X-RAY DIFFRACTION	283	8	12
2ZJS	X-RAY DIFFRACTION	293	6	5
2ZT9	X-RAY DIFFRACTION	N/A	11	22
2ZW3	X-RAY DIFFRACTION	277	4	9
2ZXE	X-RAY DIFFRACTION	298	6	3
2ZZ9	ELECTRON CRYSTALLOGRAPHY	293	6	15
3A0B	X-RAY DIFFRACTION	293	24	21
3A7K	X-RAY DIFFRACTION	293	6	11
3AM6	X-RAY DIFFRACTION	293	6	7
3AQP	X-RAY DIFFRACTION	293	12	23
3AR8	X-RAY DIFFRACTION	283	5	4
3AR9	X-RAY DIFFRACTION	283	7	8
3ARC	X-RAY DIFFRACTION	285	27	32
3B4R	X-RAY DIFFRACTION	295	4	2
3B60	X-RAY DIFFRACTION	277	5	7
3B8C	X-RAY DIFFRACTION	N/A	2	0
3B8E	X-RAY DIFFRACTION	292	5	1
3B9B	X-RAY DIFFRACTION	292	8	7
3B9Y	X-RAY DIFFRACTION	298	9	22
3BEH	X-RAY DIFFRACTION	298	5	9
3C02	X-RAY DIFFRACTION	291	5	14
3CAP	X-RAY DIFFRACTION	277	7	13
3CHX	X-RAY DIFFRACTION	293	7	3
3CX5	X-RAY DIFFRACTION	277	10	14
3D31	X-RAY DIFFRACTION	293.15	10	16

PDB	Experimental Method	Temperature	# Unique Helices	# Neigh. Pairs
3D9S	X-RAY DIFFRACTION	281	5	11
3DDL	X-RAY DIFFRACTION	295	7	10
3DIN	X-RAY DIFFRACTION	298	10	12
3DWW	ELECTRON CRYSTALLOGRAPHY	373	3	4
3E86	X-RAY DIFFRACTION	293	2	3
3EAM	X-RAY DIFFRACTION	298	3	6
3EML	X-RAY DIFFRACTION	293	6	8
3F5W	X-RAY DIFFRACTION	298	2	1
3F7V	X-RAY DIFFRACTION	298	2	3
3F7Y	X-RAY DIFFRACTION	298	2	1
3FB5	X-RAY DIFFRACTION	298	2	3
3G5U	X-RAY DIFFRACTION	278	12	18
3GD8	X-RAY DIFFRACTION	298	6	17
3GIA	X-RAY DIFFRACTION	293	8	11
3H1J	X-RAY DIFFRACTION	277	9	9
3H9V	X-RAY DIFFRACTION	277	1	2
3HD6	X-RAY DIFFRACTION	277	6	12
3HD7	X-RAY DIFFRACTION	293	2	1
3HFX	X-RAY DIFFRACTION	289	7	8
3HQK	X-RAY DIFFRACTION	293	9	19
3HZQ	X-RAY DIFFRACTION	277	1	2
3J01	ELECTRON MICROSCOPY	N/A	6	2
3JYC	X-RAY DIFFRACTION	277.10	2	5
3K03	X-RAY DIFFRACTION	293	2	3
3K3F	X-RAY DIFFRACTION	293	6	11
3KBC	X-RAY DIFFRACTION	277	4	1
3KCU	X-RAY DIFFRACTION	291	6	23
3KG2	X-RAY DIFFRACTION	277	3	5
3KLY	X-RAY DIFFRACTION	291	13	48
3KP9	X-RAY DIFFRACTION	298	5	6
3KZI	X-RAY DIFFRACTION	291	28	35
3L1L	X-RAY DIFFRACTION	291.15	15	18
3LBW	X-RAY DIFFRACTION	N/A	1	2
3LDC	X-RAY DIFFRACTION	293	2	1
3LLQ	X-RAY DIFFRACTION	295	6	16
3LUT	X-RAY DIFFRACTION	N/A	5	6
3M73	X-RAY DIFFRACTION	298	9	20
3MK7	X-RAY DIFFRACTION	294	10	13
3MKT	X-RAY DIFFRACTION	295	10	11
3MP7	X-RAY DIFFRACTION	273	9	11
3MRA	SOLUTION NMR	303	1	0
3N23	X-RAY DIFFRACTION	292	6	3
3ND0	X-RAY DIFFRACTION	293	7	9
3NE2	X-RAY DIFFRACTION	N/A	6	23

PDB	Experimental Method	Temperature	# Unique Helices	# Neigh. Pairs
3NE5	X-RAY DIFFRACTION	298	6	6
3O0R	X-RAY DIFFRACTION	277	12	23
3O7Q	X-RAY DIFFRACTION	291	12	24
3OAX	X-RAY DIFFRACTION	283	7	11
3OB6	X-RAY DIFFRACTION	298	8	12
3ODU	X-RAY DIFFRACTION	293	5	6
3ORG	X-RAY DIFFRACTION	298	6	7
3OUF	X-RAY DIFFRACTION	293	2	3
3P03	X-RAY DIFFRACTION	291	19	32
3P0G	X-RAY DIFFRACTION	293	7	12
3P5N	X-RAY DIFFRACTION	298	4	4
3PBL	X-RAY DIFFRACTION	293	6	9
3PCQ	X-RAY DIFFRACTION	277	24	23
3PJZ	X-RAY DIFFRACTION	298	5	2
3PL9	X-RAY DIFFRACTION	291	3	1
3PQR	X-RAY DIFFRACTION	277	6	8
3PUY	X-RAY DIFFRACTION	293	11	18
3PUZ	X-RAY DIFFRACTION	293	10	16
3PXO	X-RAY DIFFRACTION	277	7	13
3Q7K	X-RAY DIFFRACTION	293	6	23
3QE7	X-RAY DIFFRACTION	291	9	11
3QNQ	X-RAY DIFFRACTION	293	7	14
3RCE	X-RAY DIFFRACTION	298	6	4
3RFU	X-RAY DIFFRACTION	293	5	3
3RHW	X-RAY DIFFRACTION	277	3	4
3RKO	X-RAY DIFFRACTION	296	41	67
3RLB	X-RAY DIFFRACTION	279	4	3
3RVY	X-RAY DIFFRACTION	298	8	16
3RZE	X-RAY DIFFRACTION	293	7	11
3S0X	X-RAY DIFFRACTION	298	4	2

Table 2: PDB entries of all experimental structures used in the OPM database analysis. Experimental Method and Temperature was obtained for each PDB entry from the RCSB protein data bank (10). In X-ray diffraction experiments the temperature refers to the crystallization temperature. The two right-most columns provide the number of unique helices and the number of neighboring helix pairs extracted for each protein structure by our method.

Supporting References

1. Kabsch, W., and C. Sander, 1983. Dictionary of protein secondary structure - pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2577–2637.

2. Walters, R. F. S., and W. F. DeGrado, 2006. Helix-packing motifs in membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America* 103:13658–13663.

3. Bansal, M., S. Kumar, and R. Velavan, 2000. HELANAL: A program to characterize helix geometry in proteins. *Journal of Biomolecular Structure & Dynamics* 17:811–819.

4. Michaud-Agrawal, N., E. J. Denning, T. B. Woolf, and O. Beckstein, 2011. Software News and Updates MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *Journal of Computational Chemistry* 32:2319–2327.

5. Gimpelev, M., L. R. Forrest, D. Murray, and B. Honig, 2004. Helical packing patterns in membrane and soluble proteins. *Biophysical Journal* 87:4075–4086.

6. Chothia, C., M. Levitt, and D. Richardson, 1981. Helix to helix packing in proteins. *Journal of Molecular Biology* 145:215–250.

7. Li, A. J., and R. Nussinov, 1998. A set of van der Waals and Coulombic radii of protein atoms for molecular and solvent-accessible surface calculation, packing evaluation, and docking. *Proteins-Structure Function and Genetics* 32:111–127.

8. de Meyer, F. J. M., J. M. Rodgers, T. F. Willems, and B. Smit, 2010. Molecular Simulation of the Effect of Cholesterol on Lipid-Mediated Protein-Protein Interactions. *Biophysical Journal* 99:3629–3638.

9. Kucerka, N., J. F. Nagle, J. N. Sachs, S. E. Feller, J. Pencer, A. Jackson, and J. Katsaras, 2008. Lipid bilayer structure determined by the simultaneous analysis of neutron and x-ray scattering data. *Biophysical Journal* 95:2356–2367.

10. Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank. http://www.rcsb.org.