Supporting Information

Thermodynamics-Based Molecular Modeling of α-Helices in Membranes and Micelles

Andrei L. Lomize, Kevin A. Schnitzer, Spencer C. Todd, and Irina D. Pogozheva

Content

Table S1. Transfer energies of residues in an α-helix ($\Delta G^{\alpha}_{transf}$) and unfolded state ($\Delta G^{coil}_{bind}$) from water to micelles and membranes (kcal/mol).

Residue	$\Delta G^{\alpha}_{transf}$ *	ΔG^{coil}_{bind} **
LEU	-2.57	-0.80
ILE	-2.30	-0.48
VAL	-1.95	-0.40
PHE	-3.06	-1.45
TRP	-2.41	-1.70
MET	-1.35	-0.40
PRO	-1.00	0.00
ALA	-0.63	0.00
CYS	-0.29	0.00
TYR	-1.80	-0.60
GLY	0.75	0.00
THR	-0.29	-0.10
SER	1.29	0.00
HIS	2.65	0.00
LYS	5.69	0.00
GLN	3.59	0.00
GLU	5.45	0.00
ASN	4.11	0.00
ASP	6.77	0.00
ARG	9.14	0.00

* Transfer energy of each residue from water to a micelle ($\Delta G^{\alpha}_{transf}$) was defined based on the transmembrane energy profiles for the corresponding amino acid residues.[1] It was taken as energy of transfer from water to the center of the membrane (all residues except Tyr and Trp) or to the minimum of energy in the interfacial membrane area (for Trp and Tyr).

****** The transfer energy scale of amino acid residues from water to lipid bilayers or micelles (ΔG^{coil}_{bind}) was defined based on the Wimley-White scale for POPC bilayers[2-3] and a similar scale determined for peptide binding to SDS micelles.[4]

Table S2. Performance of the FMAP 2.0 method for predicting α-helices in different data sets

System	Water	SDS or DPC micelles	Various micelles			PC bilayers*	Bitopic proteins**
Approximation	Whole-residue	All-atom	All-atom	All-atom	All-atom	All-atom	Whole-residue
Helix detection method	Boltzmann	Boltzmann	Boltzmann	LEP	Boltzmann	Boltzmann	LEP
Data set	1	2a	2b	2c	2(a-b)	3	11
# of peptides	118	255	152	10	407	34	170
# of observed α-helices	57	274	189	24	463	31	170
# of observed non-helical peptides	65	3	12	0	15	3	NA
Prediction of α-helices							
# (%) of correctly predicted α-helices	54 (95)	274 (100)	188 (99)	23 (96)	462 (100)	31 (100)	170 (100)
# (%) of correctly predicted non-helical peptides	63 (97)	3 (100)	10 (83)	NA	13 (87)	3 (100)	NA
# of falsely-predicted α-helices in non-helical peptides	2	0	2	NA	2	0	NA
# of falsely-predicted α-helices in helical peptides	0	6	9	1	15	3	9***
# of missing α-helices	3	0	1	1	1	0	0
# of merged α-helices	0	1	16	0	17	0	0
# of broken α-helices	0	8	9	4	17	1	0
Prediction of α-helical state per residue (HR and NHR**)**							
# of residues	2239	5255	4014	675	9269	916	20458
# of observed HR	659	3988	2741	503	6729	634	4951
α-helicity (%)	29.4	75.9	68.3	74.5	72.6	69.2	24.2
# of correctly predicted HR (TP)	577	3818	2517	404	6335	502	4546
# of correctly predicted NHR (TN)	1499	862	810	137	1672	284	15018
# of falsely-predicted HR (FP)	81	405	469	35	874	57	489
# of missing HR (FN)	82	170	218	99	388	73	405

* Including bicelles, high-density lipoprotein nanodiscs, and PC bilayers

** Only TM α-helices were analyzed

*** TM α-helices falsely predicted in extramembrane regions of bitopic proteins

**** HR, residue in helical state; NHR, residue in non-helical state. NHR corresponds to coil for sets 1-3 and to extramembrane domains for bitopic proteins (set 11).

Table S3. Adjustable parameters obtained using different data sets

System	Peptides in water	Peptides in micelles*	Bitopic TM proteins
Approximation	**Whole-residue**	**All-atom**	**Whole-residue**
Data set	**1**	**2a**	**11**
ΔH_{bb} (kcal/mol)	-1.24	-1.30	-1.30
ΔS_{bb} (cal/mol K)	4.10	4.10	4.05**
Helix detectability cutoff, P_d (%)	18	20	N/A
$\Delta G_{coil,ref}$ (kcal/mol)	N/A	-0.4 (SDS) -0.6 (DPC)	-0.3
Deformation parameter, C_s (kcal/mol Å^2)	N/A	0.003	0
Reference	[5]	This work	[6]

* Parameters were determined by minimizing deviations of calculated and experimental helix boundaries by grid scan with a gradually decreasing step[5] using set 2a of peptides studied in SDS or DPC micelles with unequivocal assignment of α-helical segments from NMR data. These values of parameters were used in calculations for all peptides in micelles and membranes.

** This value depends on the effective temperature for proteins, which was taken as 293ºK.

N/A, not-applicable

Table S4. Helix end prediction errors for TM α-helices of bitopic proteins from different natural membranes.

Membrane type	Approximation	all-atom	whole-residue
	Number of TM single-helical proteins	Errors (residues/helix)	Errors (residues/helix)
PM, ER/Golgi, Eukarya	18	4.1	3.9
IM, Gram-neg. Bacteria	43	4.4	4.8
Mitochondrial IM	29	7.7	7.7
Thylakoid membrane	71	4.5	4.2
Others*	9	2.0	2.8
All	170	4.8	4.8

* Others: 3 from PM of Archaea, 2 from PM of Gram-positive Bacteria, 2 from mitochondrial outer membrane, 2 designed proteins.

Table S5. Comparison of side-chain conformers calculated with FMAP 2.0 in isolated TM α-helices of 170 bitopic proteins with conformers observed in the corresponding crystal structures of protein complexes.

Residue type	% of correctly predicted χ1 conformers	Number of residues in set 11
ILE	80	436
VAL	88	477
THR	66	253
SER	44	185
LEU	38	732
MET	36	104
PHE	56	365
TYR	66	180
TRP	50	111
HIS	51	47
ASN	55	62
ASP	51	51
GLU	47	85
GLN	48	86
LYS	52	120
ARG	41	133
CYS	38	34
Average	58	3461

Table S6. Membrane deformation parameters that were refined based on calculations of TM α-helices of bitopic proteins from different natural membranes using all-atom "peptide in membrane" model of FMAP 2.0

Membrane type	Number of proteins	D_0 (Å)	f_{mism} (kcal/mol Å^2)	f_{tilt} (kcal/mol Å^2)	$C_{s,\ TM}$ (kcal/mol Å^2)	$C_{s,\ surf}$ (kcal/mol Å^2)
DOPC bilayer	236	28.8	0.020	0.0005	0.001	0.005
PM, Eukarya	7	33.5	0.020	0.0005	0.001	0.005
ER/Golgi, Eukarya	11	30.2	0.020	0.0005	0.001	0.005
IM, Gram-neg. Bacteria	43	30.2	0.015	0.0002	0	0.004
PM, Archaea	3	30.6	0.015	0.0002	0	0.004
PM, Gram-pos. Bacteria	2	31.6	0.015	0.0002	0	0.004
Mitochondrial IM	29	28.6	0.010	0.0002	0	0.004
Thylakoid membrane	71	30.7	0.020	0.0002	0	0.004
Generic membrane	170	30.0	0.020	0.0002	0.001	0.005

PM, plasma membrane; ER, endoplasmic reticulum; IM, inner membrane.

D_0, membrane thickness in equilibrium, values correspond to average hydrophobic thicknesses of polytopic TM proteins in the corresponding biological membrane[7]; f_{mism}, membrane deformation parameter for positive membrane mismatch; f_{tilt}, membrane penalty parameter for the helix tilting; $C_{s,\ TM}$, membrane penalty parameter for TM α-helix insertion; $C_{s,\ surf}$, membrane penalty parameter for insertion of a peptide from the membrane surface.

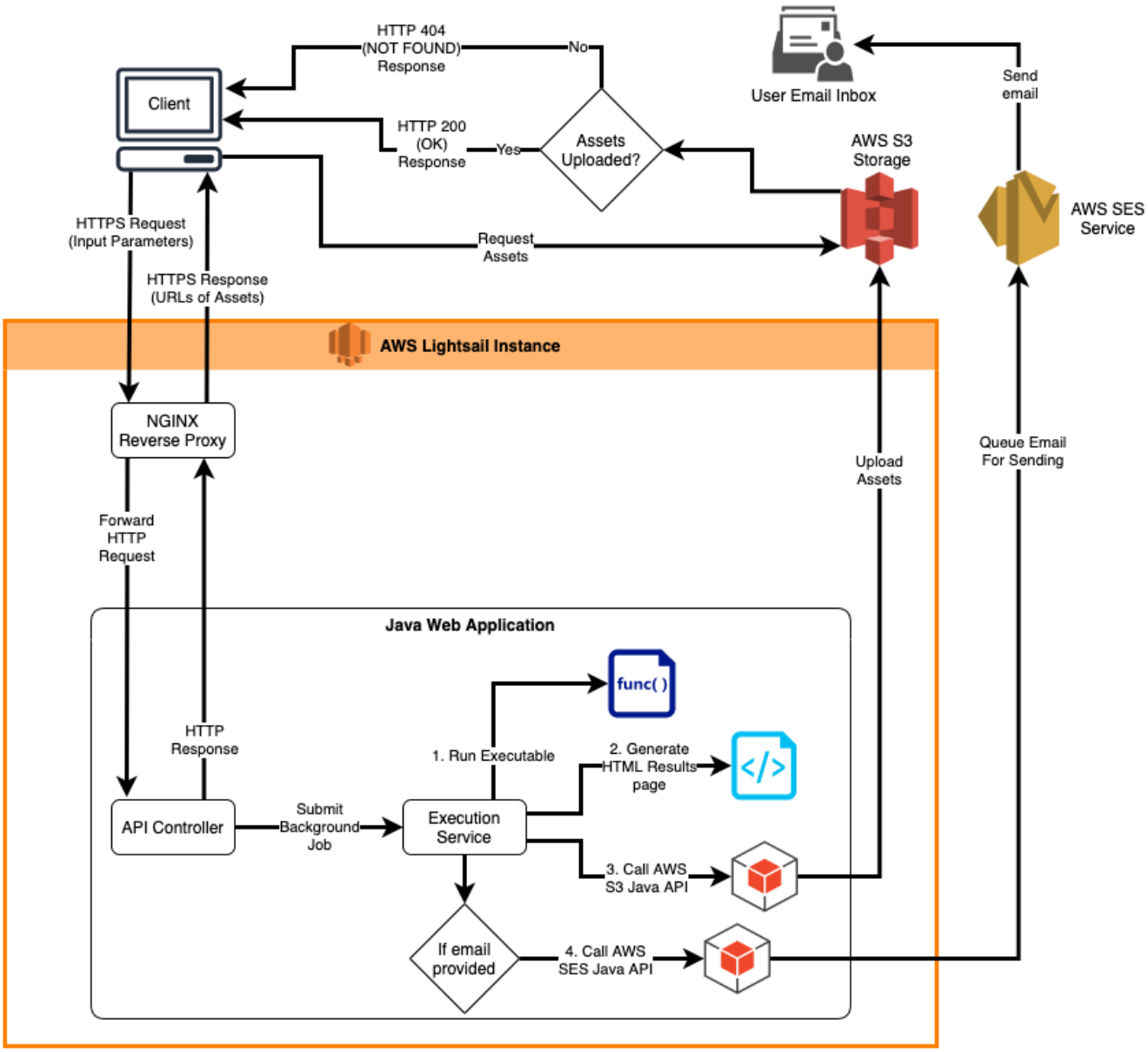

Figure S1. The FMAP 2.0 server architecture and data flow.

The application is written in Java and uses Spring Boot, an open-source, Java-based web application framework. The running application is hosted on Amazon Web Services (AWS) using a single AWS Lightsail instance, which runs the Ubuntu distribution of the Linux. The AWS Java SDK is used to access AWS APIs from Java code. Computation results, including PDB files and HTML results pages, are stored in AWS Simple Storage Service (S3) buckets. Results are publicly accessible through the URL of the results files. The application sends emails using AWS Simple Email Service.

REFERENCES

1. Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I., Anisotropic solvent model of the lipid bilayer. 2. Energetics of insertion of small molecules, peptides, and proteins in membranes. *J Chem Inf Model* **2011,** *51*, 930-946.
2. Wimley, W. C.; White, S. H., Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol* **1996,** *3*, 842-848.
3. Hristova, K.; White, S. H., An Experiment-Based Algorithm for Predicting the Partitioning of Unfolded Peptides into Phosphatidylcholine Bilayer Interfaces. *Biochemistry* **2005,** *44*, 12614-12619.
4. Imamura, T.; Konishi, K., Interaction of Tryptophan Dipeptides with Sodium Dodecyl Sulfate Micelles. *J Colloid Interface Sci* **1998,** *198*, 300-307.
5. Lomize, A. L.; Mosberg, H. I., Thermodynamic model of secondary structure for alpha-helical peptides and proteins. *Biopolymers* **1997,** *42*, 239-269.
6. Lomize, A. L.; Lomize, M. A.; Krolicki, S. R.; Pogozheva, I. D., Membranome: a database for proteome-wide analysis of single-pass membrane proteins. *Nucleic Acids Res* **2017,** *45*, D250-D255.
7 Pogozheva, I. D.; Tristram-Nagle, S.; Mosberg, H. I.; Lomize, A. L., Structural adaptations of proteins to different biological membranes. *Biochim Biophys Acta* **2013,** *1828*, 2592-2608.