Predicting protein-membrane interfaces of peripheral membrane proteins using ensemble machine learning

Alexios Chatzigoulas[1,2,*] and Zoe Cournia[1,*]

[1] Biomedical Research Foundation, Academy of Athens, 11527 Athens, Greece

[2] Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, 15784 Athens, Greece

Table S1. The peripheral membrane proteins of the training set, their PDB code, and their experimentally known hydrophobic membrane-penetrating residues after data selection. Residue numbering is the same as in the PDB structures.

Protein	PDB	Membrane-penetrating residues	Reference
Cytosolic phospholipase A2, group IVA	1cjy, chain A	F35, M38, L39, Y96, V97, M98, W464	[1-4]
Synaptotagmin-1, C2A domain	2k45	M34, F95	[5, 6]
Protein kinase C alpha, C2 domain	4dnl	P188, T250, T251	[7, 8]
Synaptotagmin-1, C2B domain	4v11	V305, I368	[9]
Protein kinase C epsilon, C2 domain	1gmi	W23, I89, Y91	[10]
Neutrophil cytosol factor 4 (p40phox)	1h6h	F35, Y94, V95	[11]
Neutrophil cytosol factor 1 (p47phox)	1kq6	I65, W80	[11, 12]
Early endosome antigen 1 FYVE domain	1joc	V1367, T1368, V1369	[13, 14]
Vps27 FYVE domain	1vfy	L185, L186	[15]
C1 domain of protein kinase C delta	3uej, chain A	M239, P241, T242, F243, L250, W252, L254, V255	[16]
Epsin ENTH domain	1h0a	L6, M10, I13, V14	[17]
C2 domain of coagulation factor V	1czs	W26, W27	[18]
C2 domains of coagulation factor Va	1sdd	Y1943, L1944, W2050, W2051	[19]
Coagulation factor VIII	2r7e	W2093, M2199, F2200, L2251, L2252	[20, 21]

Annexin V	1anx, chain A	T74, W187	[22]
Equinatoxin II	1iaz, chain A	W112, Y113	[23-25]
Prostaglandin H2 synthase 1	2ayl	I74, W75, W77, L78, F88, F91, L92, W98, L99, F102	[26]
Pancreatic phospholipase A2, group IB	1hn4, chain A	W3, L19, M20	[27]
Bee venom phospholipase A2	1poc	I1, I2, F24, I78, F82	[28]
Phospholipase A2, group IIA	5g3n, chain A	L2, V3, L19, F23, V30, F63	[28]
Acidic phospholipase A2 1	1poa	Y3, W19, W61, Y63, F64, Y110	[29-31]
Snake phospholipase A2, group II	1vap, chain A	W20, W30, W109	[32]
Snake phospholipase A2, group II, B	4hg9, chain A	Y120, P121, I124, L125	[33]
Phosphatidylinositol-specific phospholipase C	2ptd	I43, W47, W242	[34]
α-toxin (bacterial phospholipase C)	1gyg, chain A	Y331, F334	[35]
Arachidonate 15-lipoxygenase	2p0m, chain A	Y15, F70, L71, W181, L195	[36]
8R-Lipoxygenase	2fnq, chain A	W413, F414, Y448, W449	[37, 38]
Signal peptidase I	3iiq, chain A	W300, W310	[39, 40]
Perfringolysin	1pfo	W466, T490, L491	[41, 42]
Hepatocyte growth factor-regulated tyrosine kinase substrate (Hrs)	1dvp	F173	[15]
Prothrombin, GLA domain	1nl1	F5, L6, V9	[43]
Sorting nexin-3	5f0p	From chain C F103, F110	[44]
Beta-2-glycoprotein 1	1c1z	L313, F315, W316	[45]
Alpha-tocopherol transfer protein	1oiz, chain A	F165, F169, I202, V206, M209	[46]
Annexin 24	1dk5, chain A	W35, W107, Y192	[47]
Protein kinase C gamma type C1B domain	1tbn	Y123, L125	[48]
Protein kinase C gamma type C1A domain	2e73	W57, I59	[48]
RAF-1 proto-oncogene serine/threonine-protein kinase	1faq	L147, L149, F158	[49]
Eosinophil cationic protein	4x08, chain A	W35	[50]
Antimicrobial peptide kalata B1	2mh1	W22, P23, V24, L30, P31, V32	[51]
Kappa-theraphotoxin-Scg1a	1la4	Y4, L5, F6, W30	[52, 53]
Dual adapter for phosphotyrosine and 3-phosphotyrosine and 3-phosphoinositide	1fao	L177, V178	[54]
Pleckstrin homology domain-containing family A member 1	1eaz	V204, M205	[54]
Sticholysin II	1gwy, chain A	W110, Y111, W114, Y136	[55, 56]
Sticholysin I	2ks4	F51, F107, Y109, W111, Y112, W115, M135, Y136, Y137	[57]
Matrix protein VP40	1es6	L295, V298	[58, 59]

Table S2. The peripheral membrane proteins of the test set, their PDB code, and their experimentally known hydrophobic membrane-penetrating residues after data selection. Residue numbering is the same as in the PDB structures.

Protein	PDB	Membrane-penetrating residues	Reference
Retinoid isomerohydrolase	3fsn	F196, F200, I202, F264, L265, W268, L270, W271	[60]
Voltage sensor toxin VSTx1	1s6x	F5, M6, W7, W27, V29, L30	[61-63]
Cytotoxin 2	1ffj	L6, V7, P8, L9, F10, Y22, M24, F25, M26, V27, P30, V32, P33, V34, I39, L47, L48, V49	[64]
Sphingomyelinase C	2ddr, chain A	W284, F285	[65]
Glycolipid transfer protein	3rzn	W142	[66]
Cholesterol-regulated Start protein 4	1jss, chain A	L124	[67]
Ceramide transfer protein, PH domain	2rsg	W33, Y36	[68]
Phosphatidylinositol transfer protein beta isoform	2a1l	W202, W203	[69]

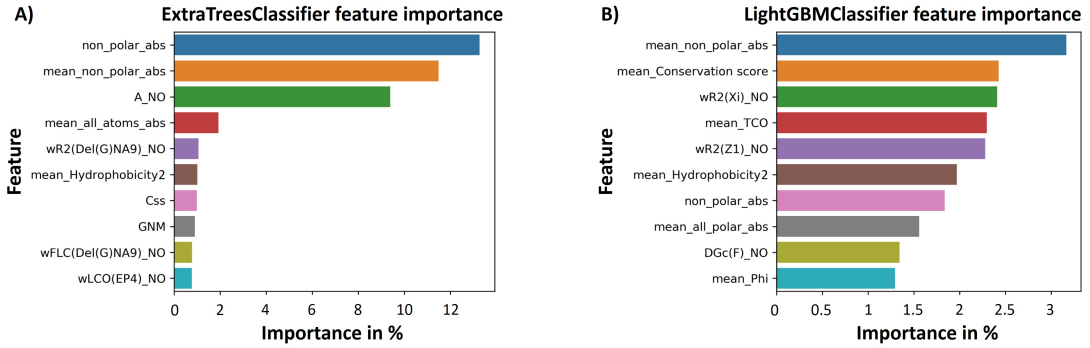

Figure S1. The 10 most important features for A) extremely randomized trees classifier and B) LightGBM classifier.

Table S3. Description of the most important features from the two tree-based approaches. The most important features express hydrophobicity and solvent exposure. The mean values are for each residue and the surrounding residues in a Cα – Cα distance of 7 Å.

Residue Feature Name	Description
non_polar_abs	Non polar SASA
mean_non_polar_abs	Mean non polar SASA
A_NO	SASA
mean_all_atoms_abs	Mean SASA
wR2(Del(G)NA9)_NO	Weighted squared mean radius of the gradient of electronic kinetic energy density perpendicular to the atomic surface
mean_Hydrophobicity2	Mean Wimley-White hydrophobicity scale from water to n-octanol
Css	Coil loop secondary structure
GNM	Squared fluctuations calculated with GNM
wFLC(Del(G)NA9)_NO	Weighted fraction of local contacts of the gradient of electronic kinetic energy density perpendicular to the atomic surface
wLCO(EP4)_NO	Weighted residue local contact order of an electrostatic potential descriptor
mean_Conservation_score	Mean conservation score
wR2(Xi)_NO	Weighted squared mean radius of the torsional angle chi
mean_TCO	Mean cosine of angle between C=O of residue i and C=O of residue i-1
wR2(Z1)_NO	Weighted squared mean radius of a combined measure of hydrophobicity related properties
mean_all_polar_abs	Mean polar SASA
DGc(F)_NO	Index of the contribution to the free energy from the conformational entropy in a folded state
mean_Phi	Mean phi backbone torsion angle

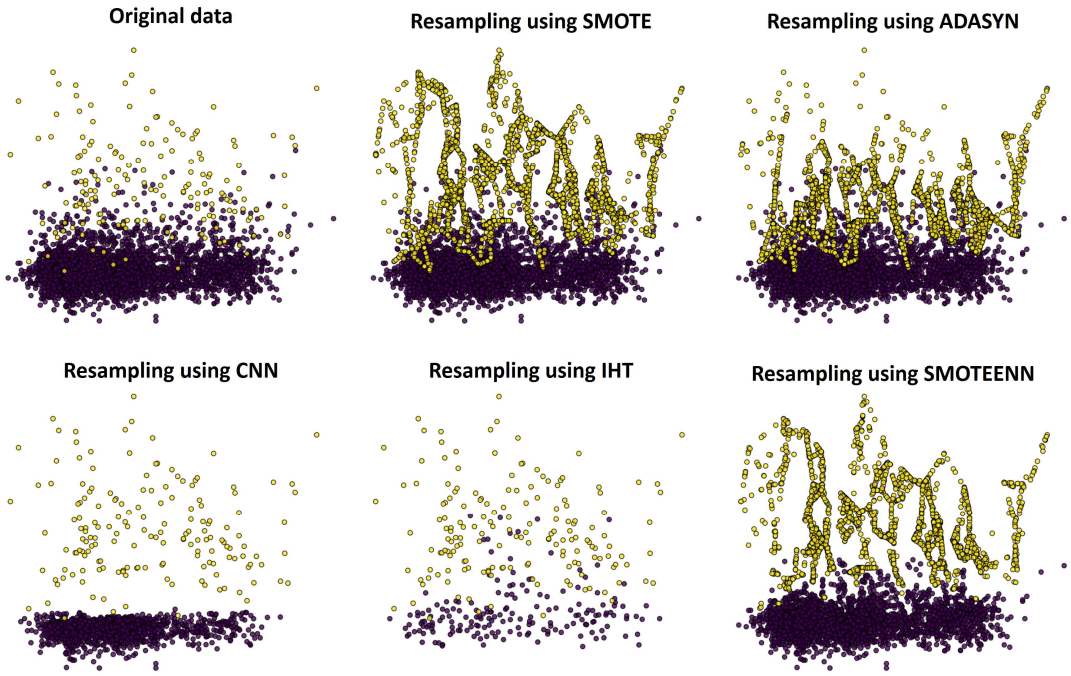

Figure S2. A visual representation of our six training sets reduced to two dimensions. The x-axis is the first principal component from principal component analysis (PCA) and the y-axis the component from linear discriminant analysis (LDA).

Table S4. The test set predictions of the 21 classifiers and the chosen voting classifier for the initial dataset using weights. The best score for each metric is highlighted with bold. In the test set the solvent inaccessible residues are excluded, keeping only the interfacial residues (residue depth < 2.5 Å).

Classifier	TN	FN	TP	FP	Precision macro	Recall macro	F_1 score macro	F_2 score macro	PR AUC	g-mean macro	MCC
Decision tree	463	26	18	10	0.79	0.69	0.73	0.71	0.55	0.69	0.48
Extremely randomized trees	462	22	22	11	0.81	0.74	0.77	0.75	0.60	0.74	0.54
Gaussian naive Bayes	437	16	28	36	0.70	0.78	0.73	0.76	0.55	0.78	0.47
Gaussian processes	468	35	9	5	0.79	0.60	0.63	0.61	0.46	0.60	0.33
Gradient boosting	468	27	17	5	**0.86**	0.69	0.74	0.70	0.61	0.69	0.52
k-nearest neighbors	458	27	17	15	0.74	0.68	0.70	0.69	0.48	0.68	0.41
Light gradient boosting machine	463	16	28	10	0.85	0.81	0.83	0.82	0.70	0.81	0.66
Linear discriminant analysis	456	19	25	17	0.78	0.77	0.77	0.77	0.60	0.77	0.54
Linear support vector machine	445	3	41	28	0.79	**0.94**	0.85	**0.89**	**0.77**	**0.94**	**0.72**
Logistic regression	452	6	38	21	0.82	0.91	0.85	**0.89**	0.76	0.91	**0.72**
Multilayer perceptron	435	18	26	38	0.68	0.76	0.71	0.73	0.52	0.76	0.43
Nearest centroid	444	27	17	29	0.66	0.66	0.66	0.66	0.40	0.66	0.32
v-support vector machine	397	29	15	76	0.55	0.59	0.55	0.57	0.28	0.59	0.13
Passive-aggressive	385	22	22	88	0.57	0.66	0.58	0.61	0.37	0.66	0.21
Perceptron	376	22	22	97	0.56	0.65	0.57	0.60	0.36	0.65	0.20
Quadratic discriminant analysis	464	32	12	9	0.75	0.63	0.66	0.64	0.45	0.63	0.36
Radius nearest neighbors	473	44	0	0	0.46	0.50	0.48	0.49	0.54	0.50	0.00
Random forest	462	26	18	11	0.78	0.69	0.73	0.70	0.54	0.69	0.47
Stochastic gradient descent	453	18	26	20	0.76	0.77	0.77	0.77	0.60	0.77	0.54
Support vector machine	463	29	15	10	0.77	0.66	0.70	0.67	0.50	0.66	0.42
Extreme gradient boosting	464	20	24	9	0.84	0.76	0.80	0.78	0.66	0.76	0.60
Voting ensembles	459	10	34	14	0.84	0.87	**0.86**	0.87	0.75	0.87	0.71

Note that because not every sample in our dataset is experimentally tested, many false positives may be actually membrane-penetrating residues, resulting in deceptive score values.

Table S5. Peripheral membrane proteins of the test set, their PDB code, their experimentally known hydrophobic membrane-penetrating residues after data selection, and the true positives and false positives predictions of our ensemble classifier. Residue numbering is the same as in the PDB structures. With † we note the residues that were discarded because their residue depth value was more than 2.5 Å. With * we note the false positive predictions that are on the protein-membrane interface, indicating that are true positives in reality.

Protein	PDB	Membrane-penetrating residues	True positives	False positives
Retinoid isomerohydrolase	3fsn	F196, F200, I202†, F264, L265, W268, L270, W271	Chain A: F200, L265, W268, L270, W271 Chain B: F196, F200, L265, W268, L270, W271	Chain A: F262* Chain B: F262*
Voltage sensor toxin VSTx1	1s6x	F5, M6, W7, W27, V29†, L30	M6, W7	W25, P33
Cytotoxin 2	1ffj	L6, V7, P8, L9, F10, Y22†, M24, F25, M26, V27, P30, V32, P33, V34, I39, L47, L48, V49	L6, V7, L9, F10, F25, M26, V27, P30, V32, V34, L47, L48, V49	V41
Sphingomyelinase C	2ddr, chain A	W284, F285	W284, F285	
Glycolipid transfer protein	3rzn	W142	W142	Y81, I143*, Y153*
Cholesterol-regulated Start protein 4	1jss, chain A	L124	L124	W91, M196*
Ceramide transfer protein, PH domain	2rsg	W33, Y36	W33, Y36	I37*, W40*, F81*
Phosphatidylinositol transfer protein beta isoform	2a1l	W202, W203	W202, W203	M74*

Table S6. The peripheral membrane proteins of the test set, their PDB code, the predictions of our ensemble classifier keeping all residue types, the PPM predictions, and the MODA predictions. Residue numbering is the same as in the PDB structures.

Protein	PDB	Our classifier	PPM	MODA
Retinoid isomerohydrolase	3fsn	Chain A: F200, F262, L265, W268, L270, W271 Chain B: F196, F200, F262, L265, W268, L270, W271	Chain A: - Chain B: P109, F196, N199, F200, F262, L265, S266, W268, S269, L270, W271	Chain A: R33, S54, E55, V99, A107, F108, F196, G197, N199, F200, S201, R234, F235, L261, F262, L265, S266, S267, W268, S269, L270, W271, G272 Chain B: R33, F108, P109, F196, G197, K198, N199, F200, R234, F235, L261, F262, L265, S266, S267, W268, S269, L270, W271, G272
Voltage sensor toxin VSTx1	1s6x	E1, M6, W7, K8, D18, R24, W25, K26, S32, P33	F5, M6, W25, W27	K4, F5, M6, W7, V20, C21, S22, S23, R24, W25, K26, W27, V29, L30, S32, P33, F34
Cytotoxin 2	1ffj	L6, V7, L9, F10, F25, M26, V27, A28, P30, H31, V32, V34, R36, V41, C42, K44, L47, L48, V49	L6, V7, P8, L9, F25, M26, V27, A28, A29, P30, H31, V32, P33, V34, L35, S36, S46, L47, L48, V49	K4, K5, L6, V7, P8, L9, F10, S11, K12, Y22, M24, F25, M26, V27, A28, A29, P30, V32, P33, V34, L35, P43, K44, S45, S46, L47, L48, V49, K50, Y51
Sphingomyelinase C	2ddr, chain A	W284, F285	L128, W284, F285	N23, L24, Y25, P26, N27, C123, L128, Y242, N243, F244, P245, T282, S283, W284, F285, Q286, K287, Y288
Glycolipid transfer protein	3rzn	Y81, W142, I143, K146, Y153	P40, V41, T43, P44, W142, I143, I147	G141, W142, I143, V144, Q145, K146, I147, Q149, A150, Y153, V209
Cholesterol-regulated Start protein 4	1jss, chain A	W91, L124, M196	L124, I126, P198, S200	L124, N125, R222
Ceramide transfer protein, PH domain	2rsg	W33, N35, Y36, I37, H38, G39, W40, F81, E83	Y36, I37, H38, G39, W40, F81	S31, K32, W33, T34, N35, Y36, I37, H38, G39, W40, H79, D80, F81, D82, R85, R98
Phosphatidylinositol transfer protein beta isoform	2a1l	M74, W202, W203	M74, W203	M74, I75, W202, W203, G204

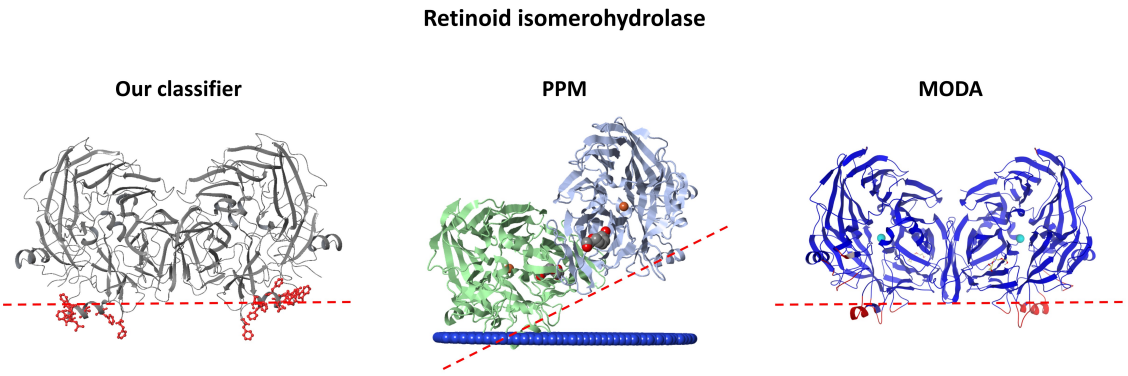

Figure S3. Comparison of the predictions provided from our classifier, PPM, and MODA for the retinoid isomerohydrolase homodimer. Both our classifier and MODA correctly found the protein-membrane regions in both chains, while PPM places only one chain in the membrane. The membrane is depicted with a red dotted line according to Ref. [60].

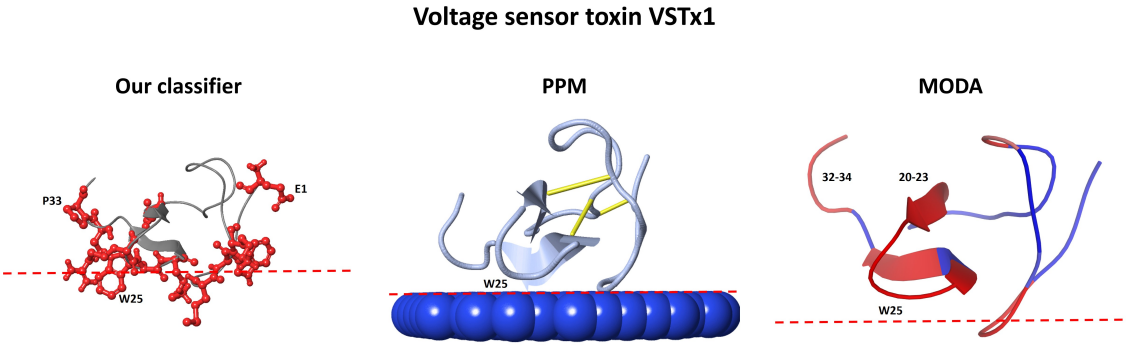

Figure S4. Comparison of the predictions provided from our classifier, PPM, and MODA for the VSTx1 toxin. Correct predictions for every tool, with every tool falsely predicting W25, our classifier falsely predicting the N- and C-terminal, and MODA falsely predicting the beta sheet on the opposite side of the protein-membrane interface and the C-terminal. The membrane is depicted with a red dotted line according to Ref. [61].

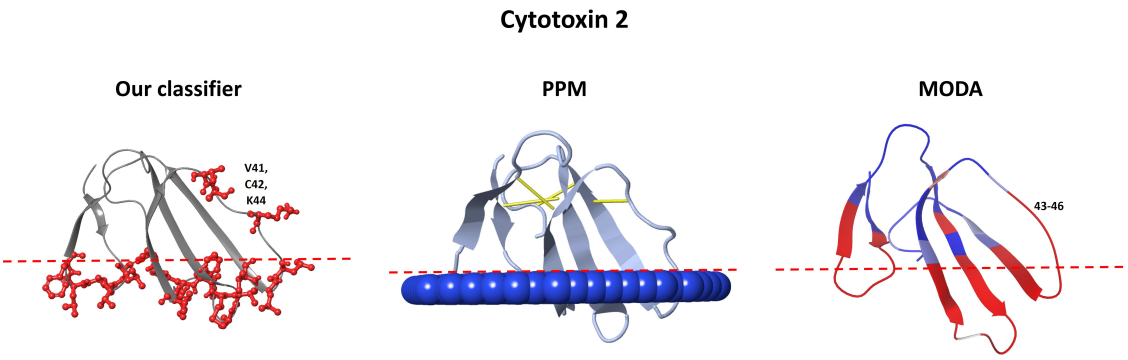

Figure S5. Comparison of the predictions provided from our classifier, PPM, and MODA for the cytotoxin 2. Correct predictions for every tool, with false positives for our classifier and MODA in the 41-46 region. The membrane is depicted with a red dotted line according to Ref. [64].

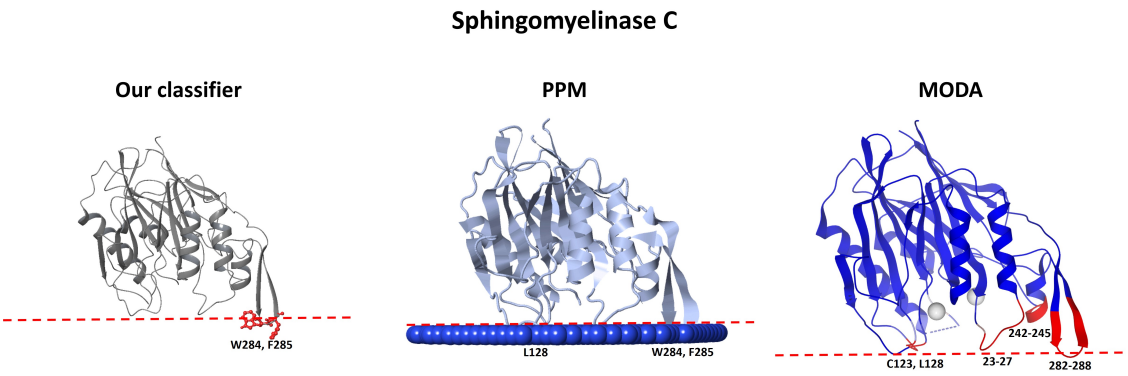

Figure S6. Comparison of the predictions provided from our classifier, PPM, and MODA for the sphingomyelinase C. Correct predictions for every tool, with PPM and MODA predicting the insertion of other loops which are aligned with the experimental residues. The membrane is depicted with a red dotted line according to Ref. [65].

Glycolipid transfer protein

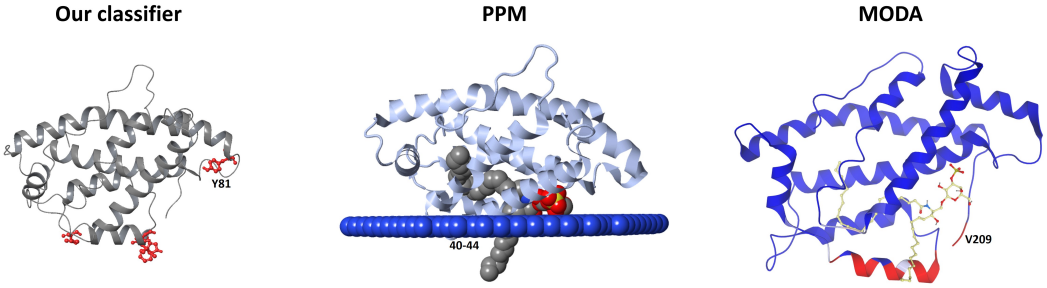

Figure S7. Comparison of the predictions provided from our classifier, PPM, and MODA for the glycolipid transfer protein. Correct predictions for every tool, with our classifier falsely predicting Y81, MODA falsely predicting the C-terminal, and PPM suggesting the insertion of the P40-P44 region.

Cholesterol-regulated Start protein 4

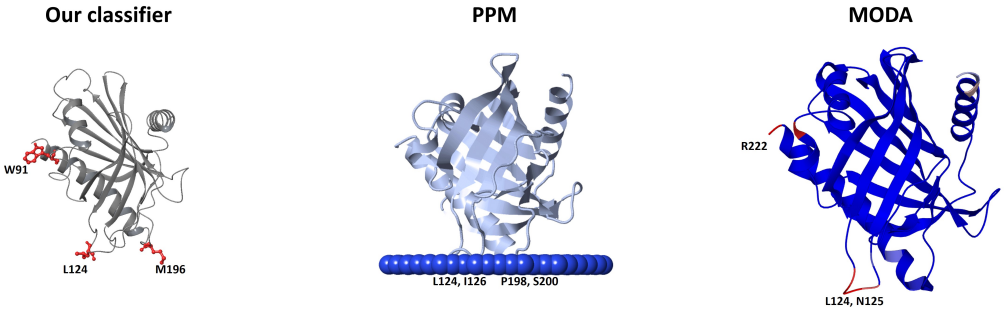

Figure S8. Comparison of the predictions provided from our classifier, PPM, and MODA for the cholesterol-regulated Start protein 4. All tools, predicted the experimentally proven residue L124, with our classifier and PPM additionally predicting the 196-200 region, MODA falsely predicting the C-terminus, and our classifier falsely predicting residue W91.

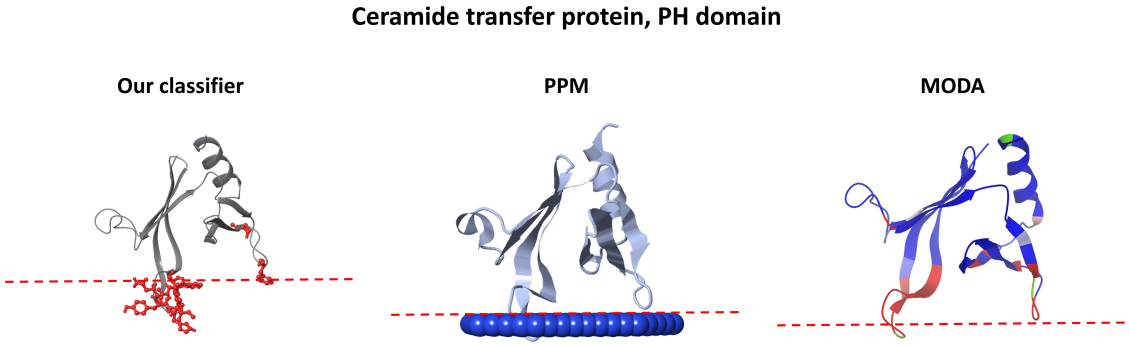

Figure S9. Comparison of the predictions provided from our classifier, PPM, and MODA for the ceramide transfer protein. Correct prediction for every tool. The membrane is depicted with a red dotted line according to Ref. [68].

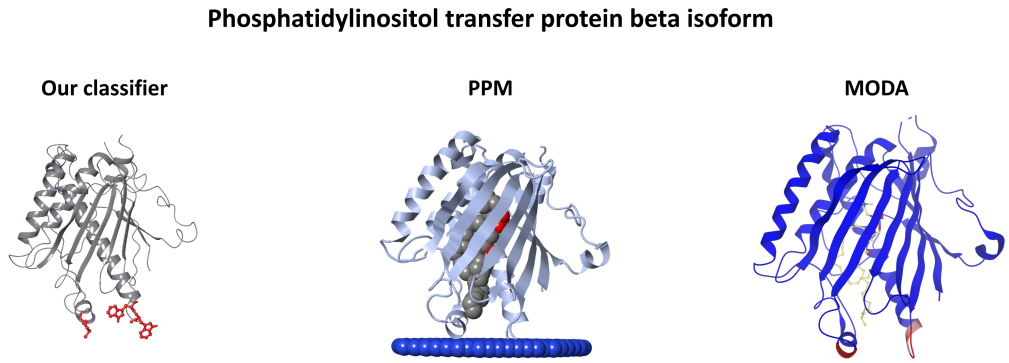

Figure S10. Comparison of the predictions provided from our classifier, PPM, and MODA for the phosphatidylinositol transfer protein beta isoform PH domain. Correct prediction for every tool.

Table S7. The peripheral membrane proteins of the test set, their PDB code, their experimentally known membrane-penetrating regions, the predictions of our ensemble classifier, the PPM predictions, and the MODA predictions. Residue numbering is the same as in the PDB structures.

Protein	PDB	Membrane-penetrating regions	Our classifier	PPM	MODA	Reference
Cholesterol oxidase	1coy	Around residue M81	M332, W333, Y436, Y437	M81, H331, M332, W333, Y372, Y437	R4, H331, M332, W333, G425, Y436, Y437	[70]
Cytochrome P450 3A4	1tqn	F-G loop and A-anchor	F46, F228, I232, V235, L236	H28, L44, P45, F46, L47, L51, P218, F219, S222, V225, F226, F228, L229, I232, L233, L236, I238	H28, P43, L44, P45, F46, L47, L51, P218, S222, V225, F226, P227, F228, L229, P231, I232, L233, V235, L236, N237, I238	[71]
9-cis-epoxycarotenoid dioxygenase 1, chloroplastic	3npe	The two parallel amphipathic helices (85-109, 222-237)	L86, F87, R372, F439	N85, L86, F87, Q88, A90, A91, A94, L95, A97, F98, G101, F102, V106, L107, A232, A235, C236, L371, R372	N85, L86, F87, R89, A90, A91, A93, A94, L95, A97, F98, G101, F102, N105, V106, L107, P110, H221, A225, Y231, A232, A235, C236, I316, K317, L367, L371, R372, G373, G374, I438, F439, R559	[72-74]
Monoglyceride lipase MGLL	3jw8, chain A	Around residues L179, L186	T168, F169, L172, L179, V180, L181	T168, F169, K170, V171, L172, A173, A174, K175, V176, L177, N178, L179, V180, L181, P182, L184	T168, F169, K170, L172, A173, K175, V176, L177, N178, L179, V180, L181, P182, N183, L184, S185, L186, G187	[75, 76]

Dihydroorotate dehydrogenase	3w7r	Region 31-68	F37, H41, L49, W362	F37, E40, H41, L42, P44, T45, L46, Q47, G48, L49, L50, L58, F62	D34, F37, Y38, H41, L42, T45, L46, L49, L50, S54, R57, L58, R61, F62, L65, G66, L67, L68, R70, R245, R246, V247	[77]
Phosphatase PTEN	5bzz, chain A	Regions 263-269, 327-335	L42, K263, M264	V222, Q245, P248, T350, V351	R41, L42, R47, R189, R335, V351	[78, 79]
(S)-mandelate dehydrogenase	6bfg	Region 177-215	Chain A: F6, R53, L54, Y179, L185, L195, H200 Chain B: W87, R172, K174, Y179, K182, H200	Y179, A181, V184, L185, C188, L189, P191, L195	Chain A: Q3, N4, F6, R53, L54, D56, S178, Y179, S180, A181, V184, L185, G187, C188, L189, H190, P191, R192, S194, L195, F197, V198, R199, G201, Q358 Chain B: R172, S178, Y179, S180, A181, V184, L185, G187, C188, L189, P191, R192, S194, L195, V198, R199	[80]
(S)-mandelate dehydrogenase, homotetramer	Homotetramer biological assembly of 6bfg	Region 177-215	Chain A: Y179, L195, H200 Chain B: Y179 Chain C: Y179, L195, H200 Chain D: Y179	Chain A: Y179, A181, V184, L185, C188, L189, P191, L195, V198 Chain B: Y179, A181, V184, L185, C188, L189, P191, L195, V198, R199 Chain C: V184,	Chain A: S178, Y179, S180, A181, V184, L185, G187, C188, L189, H190, P191, R192, S194, L195, F197, V198, R199, G201 Chain B: S178, Y179, S180, A181, V184, L185, G187, C188, L189, P191, R192, S194, L195, V198, R199 Chain C: Y179, S180, A181, V184, L185, G187,	[80]

				L185, C188, L189, P191, L195 Chain D: V184, L185, C188	C188, L189, P191, R192, S194, L195, V198, R199, G201 Chain D: S178, Y179, S180, A181, V184, L185, G187, C188, P191, R192, S194, L195, V198, R199, G201	
L-amino acid deaminase	5hxw, chain A	I345, L347, L351, I352, F355	F326, Y330, L333, L335, F355, M356	V321, V322, K323, S325, F326, T327, Y330, L333, I345, S346, L347, N348, L351, I352, F355, M356	V321, S325, F326, T327, G329, Y330, L333, P334, L336, A337, I345, L347, L351, I352, F355, M356	[81]
Intestinal fatty acid binding protein	3akm, chain A	Around residue K27	V26, K29	V26, K29, L30, H33	N24, I25,V26, K27, K29, L30	[82-84]
Phosphatidylinositol 4,5-bisphosphate 3-kinase	4ovu	Kinase region 720-729, p110a region 863-872, and p85a region iSH2	Chain A: M232, L233, W498, Y508, K867, G868, L870 Chain B: F512, M525	Chain A: M232, L233 Chain B: -	Chain A: M232, L233, N331, R349, S379, N380, W498, S501, R502, E503, A504, G505, F506, Y508, K867, G868, A869 Chain B: F512, K519	[85]
Phosphatidylcholine transfer protein	1ln1	Region 184-193	F107, W186, W190	F107, P108	P106, F107, P108, M109, S110, S147, G148, W186	[86]

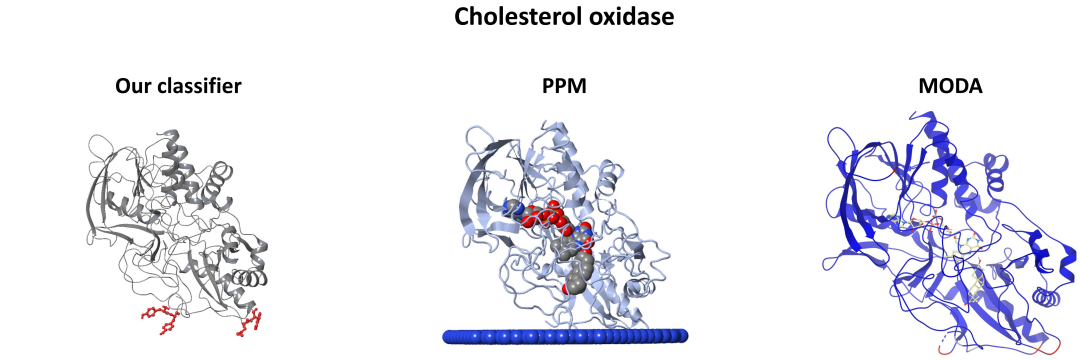

Figure S11. Comparison of the predictions provided from our classifier, PPM, and MODA for the cholesterol oxidase. Correct predictions for every tool.

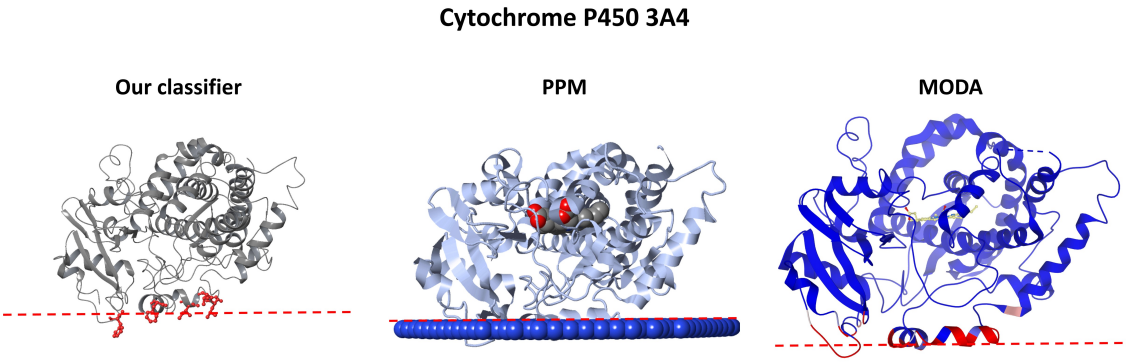

Figure S12. Comparison of the predictions provided from our classifier, PPM, and MODA for the cytochrome P450 3A4. Correct predictions for every tool. The membrane is depicted with a red dotted line according to Ref. [71].

9-cis-epoxycarotenoid dioxygenase 1, chloroplastic

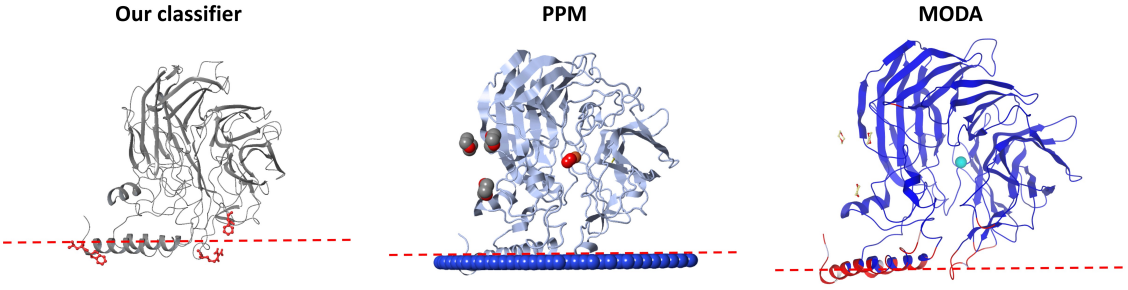

Figure S13. Comparison of the predictions provided from our classifier, PPM, and MODA for the 9-cis-epoxycarotenoid dioxygenase 1, chloroplastic. Correct predictions for every tool, with our classifier predicting the insertion of one of the two parallel amphipathic helices, instead of both. The membrane is depicted with a red dotted line according to Ref. [73].

Monoglyceride lipase MGLL

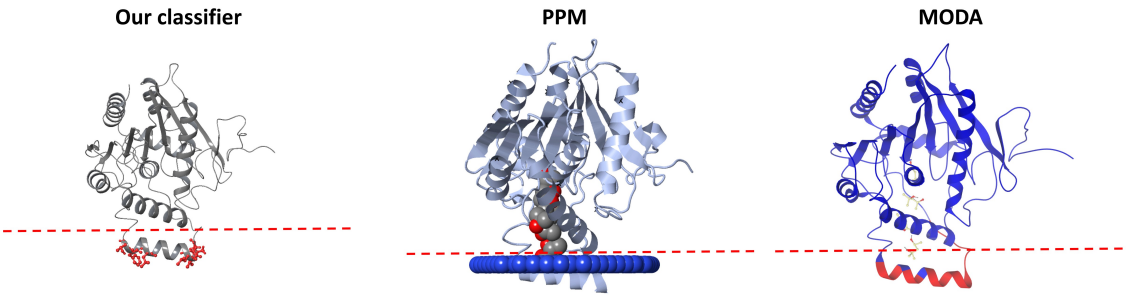

Figure S14. Comparison of the predictions provided from our classifier, PPM, and MODA for the monoglyceride lipase MGLL. Correct predictions for every tool. The membrane is depicted with a red dotted line according to Ref. [75].

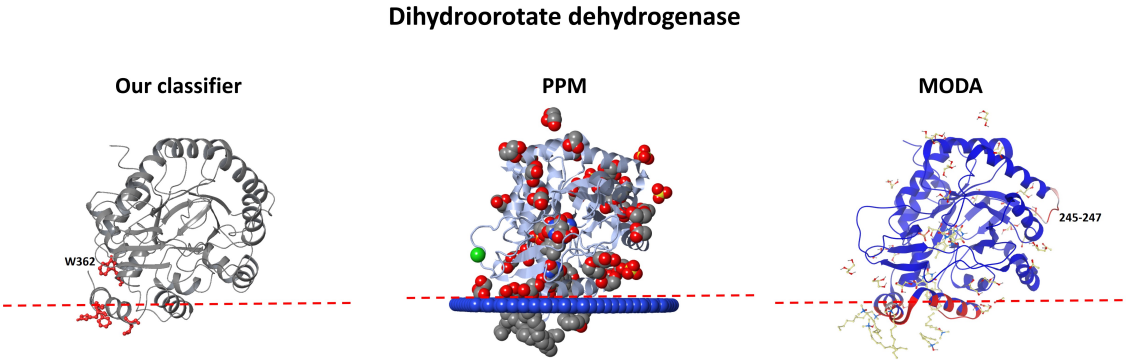

Figure S15. Comparison of the predictions provided from our classifier, PPM, and MODA for the dihydroorotate dehydrogenase. Correct predictions for every tool, with our classifier falsely identifying the residue W362 and MODA the region 245-247. The membrane is depicted with a red dotted line according to Ref. [77].

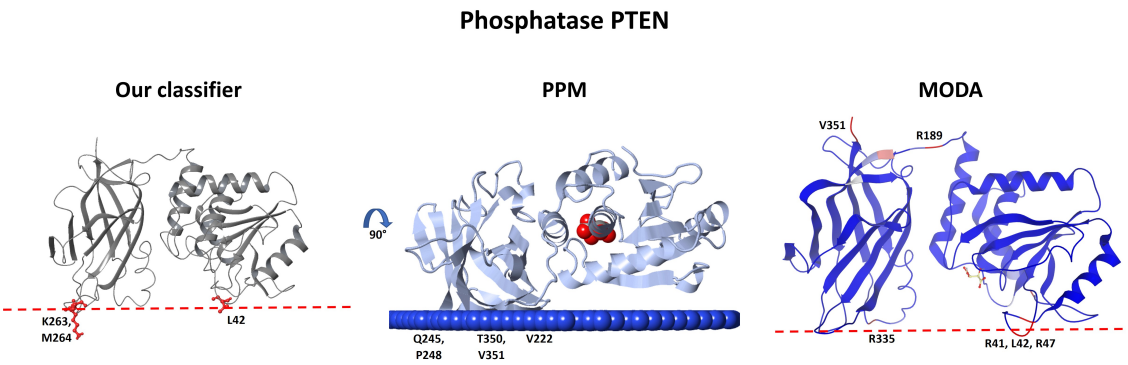

Figure S16. Comparison of the predictions provided from our classifier, PPM, and MODA for the phosphatase PTEN. Correct predictions for our classifier, with MODA correctly identifying the membrane binding region, but falsely identifying the opposite side of the C2 domain. PPM also falsely identified the opposite side of the C2 domain, providing a different orientation. The membrane is depicted with a red dotted line according to Ref. [79].

(S)-mandelate dehydrogenase

Our classifier PPM MODA

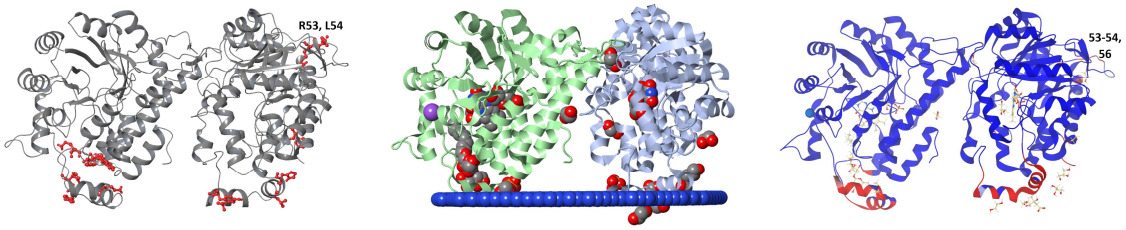

Figure S17. Comparison of the predictions provided from our classifier, PPM, and MODA for the (S)-mandelate dehydrogenase. Correct predictions for every tool, with our classifier and MODA falsely predicting the 53-56 region.

(S)-mandelate dehydrogenase, homotetramer

Our classifier PPM MODA

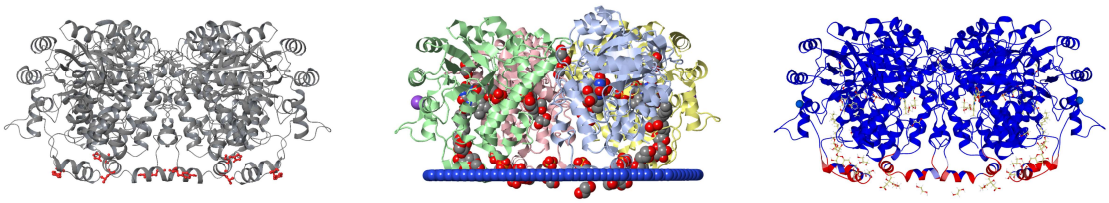

Figure S18. Comparison of the predictions provided from our classifier, PPM, and MODA for the homotetramer (S)-mandelate dehydrogenase (biological assembly). Correct predictions for every tool.

L-amino acid deaminase

Our classifier PPM MODA

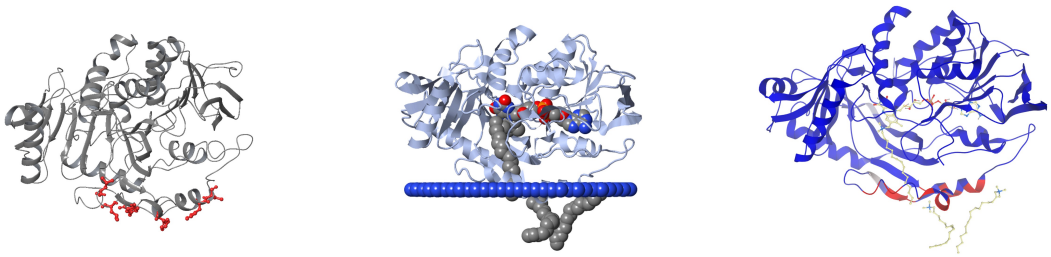

Figure S19. Comparison of the predictions provided from our classifier, PPM, and MODA for the L-amino acid deaminase. Correct predictions for every tool.

Intestinal fatty acid binding protein

Our classifier PPM MODA

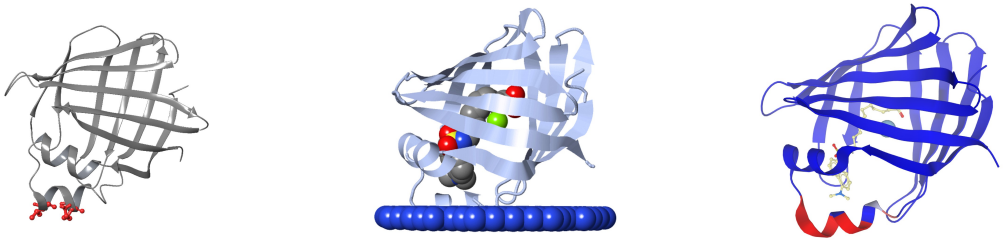

Figure S20. Comparison of the predictions provided from our classifier, PPM, and MODA for the intestinal fatty acid binding protein. Correct predictions for every tool.

Phosphatidylinositol 4,5-bisphosphate 3-kinase

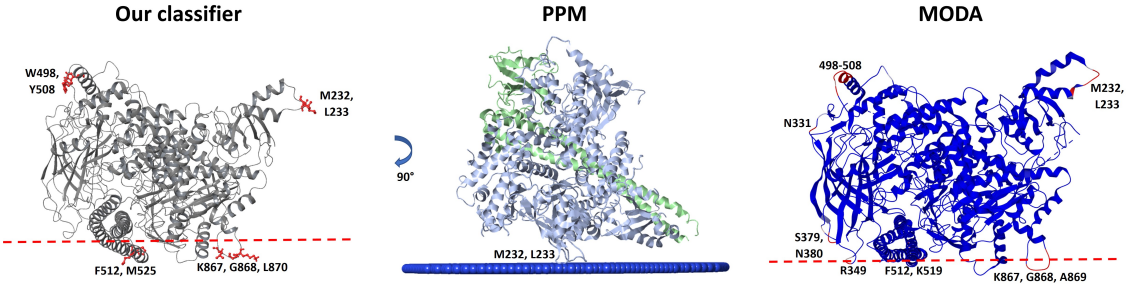

Figure S21. Comparison of the predictions provided from our classifier, PPM, and MODA for the phosphatidylinositol 4,5-bisphosphate 3-kinase. Every tool falsely predicted the region 232-233, with our classifier and MODA successfully identifying the p110a 863-872 region, but also having false positives the 498-508 region. PPM provides a membrane orientation different than the one proposed through mutagenesis experiments [85]. The membrane is depicted with a red dotted line according to Ref. [85].

Phosphatidylcholine transfer protein

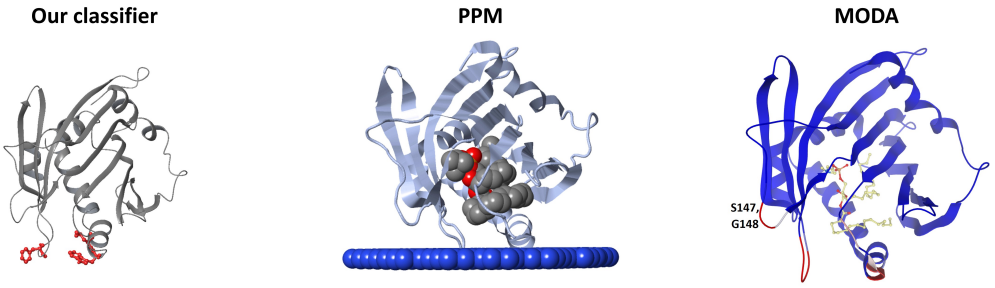

Figure S22. Comparison of the predictions provided from our classifier, PPM, and MODA for the phosphatidylcholine transfer protein. Correct predictions for every tool, with MODA suggesting also the insertion of the 147-148 loop.

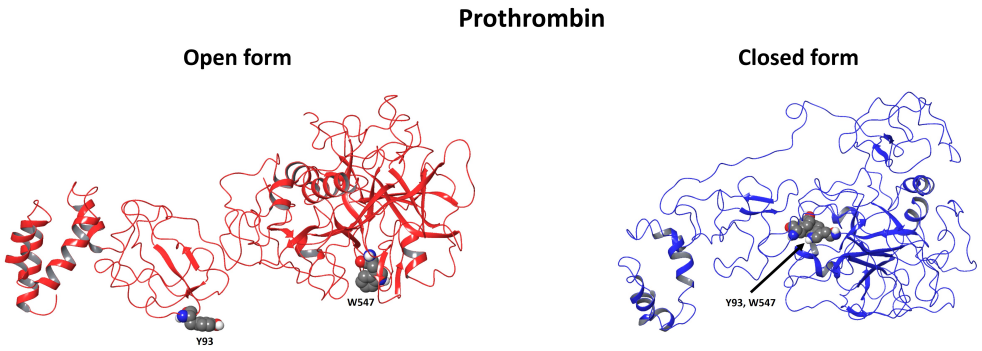

Figure S23. The open (PDB: 5EDM [87]) and closed (PDB: 6BJR [88]) forms of the prothrombin protein. The residues Y93 and W547 (W533 of 5EDM) are highlighted in CPK representation. In the closed form, Y93 inserts into the active site of the protease domain where it interacts with W547, while in the open form is distant from W547, suggested from our classifier and MODA to insert in the membrane.

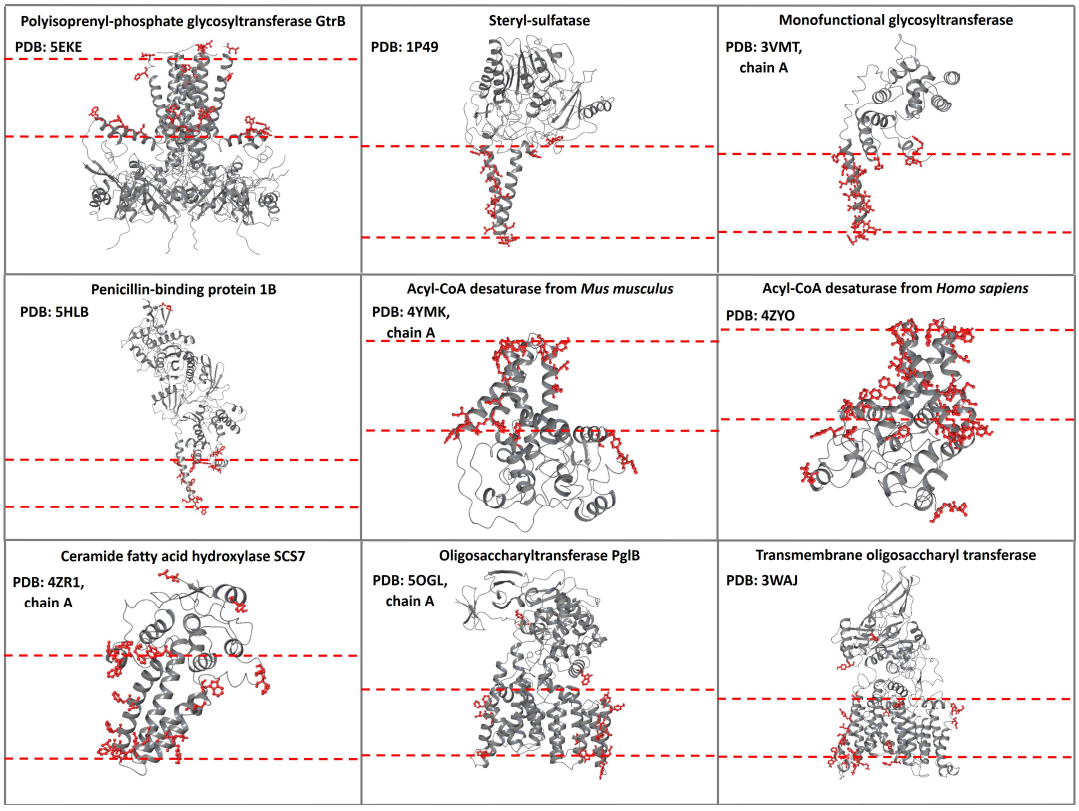

Figure S24. Our classifier's predictions in nine transmembrane enzymes with soluble domain performing extracellular catalysis. The membrane-penetrating residues predicted from our classifier are depicted with red, and the membrane is depicted with red dotted lines. The proteins are placed in the membrane according to available structural information [89]. Overall, the predictions are in agreement with the experiments, except for some use cases where a few predictions are false positives.

References

1. Bittova L., Sumandea M., Cho W. A structure-function study of the c2 domain of cytosolic phospholipase a2. Identification of essential calcium ligands and hydrophobic membrane binding residues. J Biol Chem. 1999; 274(14):9665-9672, 10.1074/jbc.274.14.9665
2. Frazier A.A., Wisner M.A., Malmberg N.J., Victor K.G., Fanucci G.E., Nalefski E.A., Falke J.J., Cafiso D.S. Membrane orientation and position of the c2 domain from cpla2 by site-directed spin labeling. Biochemistry. 2002; 41(20):6282-6292, 10.1021/bi0160821
3. Malmberg N.J., Van Buskirk D.R., Falke J.J. Membrane-docking loops of the cpla2 c2 domain: Detailed structural analysis of the protein-membrane interface via site-directed spin-labeling. Biochemistry. 2003; 42(45):13227-13240, 10.1021/bi035119+
4. Evans J.H., Leslie C.C. The cytosolic phospholipase a2 catalytic domain modulates association and residence time at golgi membranes. J Biol Chem. 2004; 279(7):6005-6016, 10.1074/jbc.M311246200
5. Gerber S.H., Rizo J., Sudhof T.C. Role of electrostatic and hydrophobic interactions in ca(2+)-dependent phospholipid binding by the c(2)a-domain from synaptotagmin i. Diabetes. 2002; 51 Suppl 1:S12-18, 10.2337/diabetes.51.2007.s12
6. Frazier A.A., Roller C.R., Havelka J.J., Hinderliter A., Cafiso D.S. Membrane-bound orientation and position of the synaptotagmin i c2a domain by site-directed spin labeling. Biochemistry. 2003; 42(1):96-105, 10.1021/bi0268145
7. Chen C.H., Malkova S., Pingali S.V., Long F., Garde S., Cho W., Schlossman M.L. Configuration of pkcalpha-c2 domain bound to mixed sopc/sops lipid monolayers. Biophys J. 2009; 97(10):2794-2802, 10.1016/j.bpj.2009.08.037
8. Igumenova T.I. Dynamics and membrane interactions of protein kinase c. Biochemistry. 2015; 54(32):4953-4968, 10.1021/acs.biochem.5b00565
9. Rufener E., Frazier A.A., Wieser C.M., Hinderliter A., Cafiso D.S. Membrane-bound orientation and position of the synaptotagmin c2b domain determined by site-directed spin labeling. Biochemistry. 2005; 44(1):18-28, 10.1021/bi048370d
10. Corbalan-Garcia S., Sanchez-Carrillo S., Garcia-Garcia J., Gomez-Fernandez J.C. Characterization of the membrane binding mode of the c2 domain of pkc epsilon. Biochemistry. 2003; 42(40):11661-11668, 10.1021/bi034850d
11. Stahelin R.V., Burian A., Bruzik K.S., Murray D., Cho W. Membrane binding mechanisms of the px domains of nadph oxidase p40phox and p47phox. J Biol Chem. 2003; 278(16):14469-14479, 10.1074/jbc.M212579200
12. Karathanassis D., Stahelin R.V., Bravo J., Perisic O., Pacold C.M., Cho W., Williams R.L. Binding of the px domain of p47(phox) to phosphatidylinositol 3,4-bisphosphate and phosphatidic acid is masked by an intramolecular interaction. EMBO J. 2002; 21(19):5057-5068, 10.1093/emboj/cdf519
13. Kutateladze T., Overduin M. Structural mechanism of endosome docking by the fyve domain. Science. 2001; 291(5509):1793-1796, 10.1126/science.291.5509.1793
14. Kutateladze T.G., Capelluto D.G., Ferguson C.G., Cheever M.L., Kutateladze A.G., Prestwich G.D., Overduin M. Multivalent mechanism of membrane insertion by the fyve domain. J Biol Chem. 2004; 279(4):3050-3057, 10.1074/jbc.M309007200
15. Stahelin R.V., Long F., Diraviyam K., Bruzik K.S., Murray D., Cho W. Phosphatidylinositol 3-phosphate induces the membrane penetration of the fyve domains of vps27p and hrs. J Biol Chem. 2002; 277(29):26379-26388, 10.1074/jbc.M201106200

16. Stewart M.D., Cole T.R., Igumenova T.I. Interfacial partitioning of a loop hinge residue contributes to diacylglycerol affinity of conserved region 1 domains. J Biol Chem. 2014; 289(40):27653-27664, 10.1074/jbc.M114.585570
17. Yoon Y., Tong J., Lee P.J., Albanese A., Bhardwaj N., Kallberg M., Digman M.A., Lu H., Gratton E., Shin Y.K., Cho W. Molecular basis of the potent membrane-remodeling activity of the epsin 1 n-terminal homology domain. J Biol Chem. 2010; 285(1):531-540, 10.1074/jbc.M109.068015
18. Kim S.W., Quinn-Allen M.A., Camp J.T., Macedo-Ribeiro S., Fuentes-Prior P., Bode W., Kane W.H. Identification of functionally important amino acid residues within the c2-domain of human factor v using alanine-scanning mutagenesis. Biochemistry. 2000; 39(8):1951-1958, 10.1021/bi992256r
19. Peng W., Quinn-Allen M.A., Kane W.H. Mutation of hydrophobic residues in the factor va c1 and c2 domains blocks membrane-dependent prothrombin activation. J Thromb Haemost. 2005; 3(2):351-354, 10.1111/j.1538-7836.2004.01083.x
20. Gilbert G.E., Kaufman R.J., Arena A.A., Miao H., Pipe S.W. Four hydrophobic amino acids of the factor viii c2 domain are constituents of both the membrane-binding and von willebrand factor-binding motifs. J Biol Chem. 2002; 277(8):6374-6381, 10.1074/jbc.M104732200
21. Meems H., Meijer A.B., Cullinan D.B., Mertens K., Gilbert G.E. Factor viii c1 domain residues lys 2092 and phe 2093 contribute to membrane binding and cofactor activity. Blood. 2009; 114(18):3938-3946, 10.1182/blood-2009-01-197707
22. Campos B., Mo Y.D., Mealy T.R., Li C.W., Swairjo M.A., Balch C., Head J.F., Retzinger G., Dedman J.R., Seaton B.A. Mutational and crystallographic analyses of interfacial residues in annexin v suggest direct interactions with phospholipid membrane components. Biochemistry. 1998; 37(22):8004-8010, 10.1021/bi973142n
23. Hong Q., Gutierrez-Aguirre I., Barlic A., Malovrh P., Kristan K., Podlesek Z., Macek P., Turk D., Gonzalez-Manas J.M., Lakey J.H., Anderluh G. Two-step membrane binding by equinatoxin ii, a pore-forming toxin from the sea anemone, involves an exposed aromatic cluster and a flexible helix. J Biol Chem. 2002; 277(44):41916-41924, 10.1074/jbc.M204625200
24. Anderluh G., Razpotnik A., Podlesek Z., Macek P., Separovic F., Norton R.S. Interaction of the eukaryotic pore-forming cytolysin equinatoxin ii with model membranes: 19f nmr studies. J Mol Biol. 2005; 347(1):27-39, 10.1016/j.jmb.2004.12.058
25. Bakrac B., Gutierrez-Aguirre I., Podlesek Z., Sonnen A.F., Gilbert R.J., Macek P., Lakey J.H., Anderluh G. Molecular determinants of sphingomyelin specificity of a eukaryotic pore-forming toxin. J Biol Chem. 2008; 283(27):18665-18677, 10.1074/jbc.M708747200
26. Spencer A.G., Thuresson E., Otto J.C., Song I., Smith T., DeWitt D.L., Garavito R.M., Smith W.L. The membrane binding domains of prostaglandin endoperoxide h synthases 1 and 2. Peptide mapping and mutational analysis. J Biol Chem. 1999; 274(46):32936-32942, 10.1074/jbc.274.46.32936
27. Tatulian S.A., Qin S., Pande A.H., He X. Positioning membrane proteins by novel protein engineering and biophysical approaches. J Mol Biol. 2005; 351(5):939-947, 10.1016/j.jmb.2005.06.080
28. Pande A.H., Qin S., Nemec K.N., He X., Tatulian S.A. Isoform-specific membrane insertion of secretory phospholipase a2 and functional implications. Biochemistry. 2006; 45(41):12436-12447, 10.1021/bi060898q
29. Sumandea M., Das S., Sumandea C., Cho W. Roles of aromatic residues in high interfacial activity of naja naja atra phospholipase a2. Biochemistry. 1999; 38(49):16290-16297, 10.1021/bi9921384

30. Stahelin R.V., Cho W. Differential roles of ionic, aliphatic, and aromatic residues in membrane-protein interactions: A surface plasmon resonance study on phospholipases a2. Biochemistry. 2001; 40(15):4672-4678, 10.1021/bi0020325
31. Burke J.E., Karbarz M.J., Deems R.A., Li S., Woods V.L., Jr., Dennis E.A. Interaction of group ia phospholipase a2 with metal ions and phospholipid vesicles probed with deuterium exchange mass spectrometry. Biochemistry. 2008; 47(24):6451-6459, 10.1021/bi8000962
32. Lathrop B., Gadd M., Biltonen R.L., Rule G.S. Changes in ca2+ affinity upon activation of agkistrodon piscivorus piscivorus phospholipase a2. Biochemistry. 2001; 40(11):3264-3272, 10.1021/bi001901n
33. Wijewickrama G.T., Albanese A., Kim Y.J., Oh Y.S., Murray P.S., Takayanagi R., Tobe T., Masuda S., Murakami M., Kudo I., Ucker D.S., Murray D., Cho W. Unique membrane interaction mode of group iif phospholipase a2. J Biol Chem. 2006; 281(43):32741-32754, 10.1074/jbc.M606311200
34. Feng J., Bradley W.D., Roberts M.F. Optimizing the interfacial binding and activity of a bacterial phosphatidylinositol-specific phospholipase c. J Biol Chem. 2003; 278(27):24651-24657, 10.1074/jbc.M301207200
35. Jepson M., Bullifent H.L., Crane D., Flores-Diaz M., Alape-Giron A., Jayasekeera P., Lingard B., Moss D., Titball R.W. Tyrosine 331 and phenylalanine 334 in clostridium perfringens alpha-toxin are essential for cytotoxic activity. FEBS Lett. 2001; 495(3):172-177, 10.1016/s0014-5793(01)02385-7
36. Walther M., Wiesner R., Kuhn H. Investigations into calcium-dependent membrane association of 15-lipoxygenase-1. Mechanistic roles of surface-exposed hydrophobic amino acids and calcium. J Biol Chem. 2004; 279(5):3717-3725, 10.1074/jbc.M309564200
37. Oldham M.L., Brash A.R., Newcomer M.E. Insights from the x-ray crystal structure of coral 8r-lipoxygenase: Calcium activation via a c2-like domain and a structural basis of product chirality. J Biol Chem. 2005; 280(47):39545-39552, 10.1074/jbc.M506675200
38. Gilbert N.C., Neau D.B., Newcomer M.E. Expression of an 8r-lipoxygenase from the coral plexaura homomalla. Methods Enzymol. 2018; 605:33-49, 10.1016/bs.mie.2018.02.010
39. Kim Y.T., Muramatsu T., Takahashi K. Identification of trp300 as an important residue for escherichia coli leader peptidase activity. Eur J Biochem. 1995; 234(1):358-362, 10.1111/j.1432-1033.1995.358_c.x
40. Bhanu M.K., Kendall D.A. Fluorescence spectroscopy of soluble e. Coli spase i delta2-75 reveals conformational changes in response to ligand binding. Proteins. 2014; 82(4):596-606, 10.1002/prot.24429
41. Sekino-Suzuki N., Nakamura M., Mitsui K.I., Ohno-Iwashita Y. Contribution of individual tryptophan residues to the structure and activity of theta-toxin (perfringolysin o), a cholesterol-binding cytolysin. Eur J Biochem. 1996; 241(3):941-947, 10.1111/j.1432-1033.1996.00941.x
42. Farrand A.J., LaChapelle S., Hotze E.M., Johnson A.E., Tweten R.K. Only two amino acids are essential for cytolytic toxin recognition of cholesterol at the membrane surface. Proc Natl Acad Sci U S A. 2010; 107(9):4341-4346, 10.1073/pnas.0911581107
43. Huang M., Rigby A.C., Morelli X., Grant M.A., Huang G., Furie B., Seaton B., Furie B.C. Structural basis of membrane binding by gla domains of vitamin k-dependent proteins. Nat Struct Biol. 2003; 10(9):751-756, 10.1038/nsb971
44. Lenoir M., Ustunel C., Rajesh S., Kaur J., Moreau D., Gruenberg J., Overduin M. Phosphorylation of conserved phosphoinositide binding pocket regulates sorting nexin membrane targeting. Nat Commun. 2018; 9(1):993, 10.1038/s41467-018-03370-1

45. Mehdi H., Naqvi A., Kamboh M.I. A hydrophobic sequence at position 313-316 (leu-ala-phe-trp) in the fifth domain of apolipoprotein h (beta2-glycoprotein i) is crucial for cardiolipin binding. Eur J Biochem. 2000; 267(6):1770-1776, 10.1046/j.1432-1327.2000.01174.x
46. Zhang W.X., Thakur V., Lomize A., Pogozheva I., Panagabko C., Cecchini M., Baptist M., Morley S., Manor D., Atkinson J. The contribution of surface residues to membrane binding and ligand transfer by the alpha-tocopherol transfer protein (alpha-ttp). J Mol Biol. 2011; 405(4):972-988, 10.1016/j.jmb.2010.11.028
47. Hofmann A., Proust J., Dorowski A., Schantz R., Huber R. Annexin 24 from capsicum annuum. X-ray structure and biochemical characterization. J Biol Chem. 2000; 275(11):8072-8082, 10.1074/jbc.275.11.8072
48. Ananthanarayanan B., Stahelin R.V., Digman M.A., Cho W. Activation mechanisms of conventional protein kinase c isoforms are determined by the ligand affinity and conformational flexibility of their c1 domains. J Biol Chem. 2003; 278(47):46886-46894, 10.1074/jbc.M307853200
49. Travers T., Lopez C.A., Van Q.N., Neale C., Tonelli M., Stephen A.G., Gnanakaran S. Molecular recognition of ras/raf complex at the membrane: Role of raf cysteine-rich domain. Sci Rep. 2018; 8(1):8461, 10.1038/s41598-018-26832-4
50. Garcia-Mayoral M.F., Moussaoui M., de la Torre B.G., Andreu D., Boix E., Nogues M.V., Rico M., Laurents D.V., Bruix M. Nmr structural determinants of eosinophil cationic protein binding to membrane and heparin mimetics. Biophys J. 2010; 98(11):2702-2711, 10.1016/j.bpj.2010.02.039
51. Shenkarev Z.O., Nadezhdin K.D., Sobol V.A., Sobol A.G., Skjeldal L., Arseniev A.S. Conformation and mode of membrane interaction in cyclotides. Spatial structure of kalata b1 bound to a dodecylphosphocholine micelle. FEBS J. 2006; 273(12):2658-2672, 10.1111/j.1742-4658.2006.05282.x
52. Wang J.M., Roh S.H., Kim S., Lee C.W., Kim J.I., Swartz K.J. Molecular surface of tarantula toxins interacting with voltage sensors in k(v) channels. J Gen Physiol. 2004; 123(4):455-467, 10.1085/jgp.200309005
53. Phillips L.R., Milescu M., Li-Smerin Y., Mindell J.A., Kim J.I., Swartz K.J. Voltage-sensor activation with a tarantula toxin as cargo. Nature. 2005; 436(7052):857-860, 10.1038/nature03873
54. Manna D., Albanese A., Park W.S., Cho W. Mechanistic basis of differential cellular responses of phosphatidylinositol 3,4-bisphosphate- and phosphatidylinositol 3,4,5-trisphosphate-binding pleckstrin homology domains. J Biol Chem. 2007; 282(44):32093-32105, 10.1074/jbc.M703517200
55. Garcia-Linares S., Maula T., Rivera-de-Torre E., Gavilanes J.G., Slotte J.P., Martinez-Del-Pozo A. Role of the tryptophan residues in the specific interaction of the sea anemone stichodactyla helianthus's actinoporin sticholysin ii with biological membranes. Biochemistry. 2016; 55(46):6406-6420, 10.1021/acs.biochem.6b00935
56. Garcia-Linares S., Alm I., Maula T., Gavilanes J.G., Slotte J.P., Martinez-Del-Pozo A. The effect of cholesterol on the long-range network of interactions established among sea anemone sticholysin ii residues at the water-membrane interface. Mar Drugs. 2015; 13(4):1647-1665, 10.3390/md13041647
57. Lopez-Castilla A., Pazos F., Schreier S., Pires J.R. Solution nmr analysis of the interaction between the actinoporin sticholysin i and dhpc micelles--correlation with backbone dynamics. Proteins. 2014; 82(6):1022-1034, 10.1002/prot.24475
58. Adu-Gyamfi E., Soni S.P., Xue Y., Digman M.A., Gratton E., Stahelin R.V. The ebola virus matrix protein penetrates into the plasma membrane: A key step in viral protein 40 (vp40)

oligomerization and viral egress. J Biol Chem. 2013; 288(8):5779-5789, 10.1074/jbc.M112.443960
59. Soni S.P., Adu-Gyamfi E., Yong S.S., Jee C.S., Stahelin R.V. The ebola virus matrix protein deeply penetrates the plasma membrane: An important step in viral egress. Biophys J. 2013; 104(9):1940-1949, 10.1016/j.bpj.2013.03.021
60. Kiser P.D., Farquhar E.R., Shi W., Sui X., Chance M.R., Palczewski K. Structure of rpe65 isomerase in a lipidic matrix reveals roles for phospholipids and iron in catalysis. Proc Natl Acad Sci U S A. 2012; 109(41):E2747-2756, 10.1073/pnas.1212025109
61. Jung H.J., Lee J.Y., Kim S.H., Eu Y.J., Shin S.Y., Milescu M., Swartz K.J., Kim J.I. Solution structure and lipid membrane partitioning of vstx1, an inhibitor of the kvap potassium channel. Biochemistry. 2005; 44(16):6015-6023, 10.1021/bi0477034
62. Ozawa S., Kimura T., Nozaki T., Harada H., Shimada I., Osawa M. Structural basis for the inhibition of voltage-dependent k+ channel by gating modifier toxin. Sci Rep. 2015; 5:14226, 10.1038/srep14226
63. Lau C.H.Y., King G.F., Mobli M. Molecular basis of the interaction between gating modifier spider toxins and the voltage sensor of voltage-gated ion channels. Sci Rep. 2016; 6:34333, 10.1038/srep34333
64. Dubovskii P.V., Dementieva D.V., Bocharov E.V., Utkin Y.N., Arseniev A.S. Membrane binding motif of the p-type cardiotoxin. J Mol Biol. 2001; 305(1):137-149, 10.1006/jmbi.2000.4283
65. Ago H., Oda M., Takahashi M., Tsuge H., Ochi S., Katunuma N., Miyano M., Sakurai J. Structural basis of the sphingomyelin phosphodiesterase activity in neutral sphingomyelinase from bacillus cereus. J Biol Chem. 2006; 281(23):16157-16167, 10.1074/jbc.M601089200
66. Ohvo-Rekila H., Mattjus P. Monitoring glycolipid transfer protein activity and membrane interaction with the surface plasmon resonance technique. Biochim Biophys Acta. 2011; 1808(1):47-54, 10.1016/j.bbamem.2010.08.018
67. Iaea D.B., Dikiy I., Kiburu I., Eliezer D., Maxfield F.R. Stard4 membrane interactions and sterol binding. Biochemistry. 2015; 54(30):4623-4636, 10.1021/acs.biochem.5b00618
68. Sugiki T., Takeuchi K., Yamaji T., Takano T., Tokunaga Y., Kumagai K., Hanada K., Takahashi H., Shimada I. Structural basis for the golgi association by the pleckstrin homology domain of the ceramide trafficking protein (cert). J Biol Chem. 2012; 287(40):33706-33718, 10.1074/jbc.M112.367730
69. Phillips S.E., Ile K.E., Boukhelifa M., Huijbregts R.P., Bankaitis V.A. Specific and nonspecific membrane-binding determinants cooperate in targeting phosphatidylinositol transfer protein beta-isoform to the mammalian trans-golgi network. Mol Biol Cell. 2006; 17(6):2498-2512, 10.1091/mbc.e06-01-0089
70. Chen X., Wolfgang D.E., Sampson N.S. Use of the parallax-quench method to determine the position of the active-site loop of cholesterol oxidase in lipid bilayers. Biochemistry. 2000; 39(44):13383-13389, 10.1021/bi001407j
71. Baylon J.L., Lenov I.L., Sligar S.G., Tajkhorshid E. Characterizing the membrane-bound state of cytochrome p450 3a4: Structure, depth of insertion, and orientation. J Am Chem Soc. 2013; 135(23):8542-8551, 10.1021/ja4003525
72. Tan B.C., Cline K., McCarty D.R. Localization and targeting of the vp14 epoxy-carotenoid dioxygenase to chloroplast membranes. Plant J. 2001; 27(5):373-382, 10.1046/j.1365-313x.2001.01102.x
73. Messing S.A., Gabelli S.B., Echeverria I., Vogel J.T., Guan J.C., Tan B.C., Klee H.J., McCarty D.R., Amzel L.M. Structural insights into maize viviparous14, a key enzyme in the biosynthesis of the phytohormone abscisic acid. Plant Cell. 2010; 22(9):2970-2980, 10.1105/tpc.110.074815

74. Sui X., Kiser P.D., Lintig J., Palczewski K. Structural basis of carotenoid cleavage: From bacteria to mammals. Arch Biochem Biophys. 2013; 539(2):203-213, 10.1016/j.abb.2013.06.012
75. Nasr M.L., Shi X., Bowman A.L., Johnson M., Zvonok N., Janero D.R., Vemuri V.K., Wales T.E., Engen J.R., Makriyannis A. Membrane phospholipid bilayer as a determinant of monoacylglycerol lipase kinetic profile and conformational repertoire. Protein Sci. 2013; 22(6):774-787, 10.1002/pro.2257
76. Karageorgos I., Silin V.I., Zvonok N., Marino J., Janero D.R., Makriyannis A. The role of human monoacylglycerol lipase (hmagl) binding pocket in breakup of unsaturated phospholipid membranes. Anal Biochem. 2017; 536:90-95, 10.1016/j.ab.2017.08.009
77. Costeira-Paulo J., Gault J., Popova G., Ladds M., van Leeuwen I.M.M., Sarr M., Olsson A., Lane D.P., Lain S., Marklund E.G., Landreh M. Lipids shape the electron acceptor-binding site of the peripheral membrane protein dihydroorotate dehydrogenase. Cell chemical biology. 2018; 25(3):309-317 e304, 10.1016/j.chembiol.2017.12.012
78. Lee J.O., Yang H., Georgescu M.M., Di Cristofano A., Maehama T., Shi Y., Dixon J.E., Pandolfi P., Pavletich N.P. Crystal structure of the pten tumor suppressor: Implications for its phosphoinositide phosphatase activity and membrane association. Cell. 1999; 99(3):323-334, 10.1016/s0092-8674(00)81663-3
79. Nguyen H.N., Afkari Y., Senoo H., Sesaki H., Devreotes P.N., Iijima M. Mechanism of human pten localization revealed by heterologous expression in dictyostelium. Oncogene. 2014; 33(50):5688-5696, 10.1038/onc.2013.507
80. Xu Y., Mitra B. A highly active, soluble mutant of the membrane-associated (s)-mandelate dehydrogenase from pseudomonas putida. Biochemistry. 1999; 38(38):12367-12376, 10.1021/bi990996u
81. Ju Y., Liu Z., Zhang Z., Duan L., Liu Q., Gu Q., Zhang C., Xu J., Zhou H. Membrane binding of the insertion sequence of proteus vulgaris l-amino acid deaminase stabilizes protein structure and increases catalytic activity. Sci Rep. 2017; 7(1):13719, 10.1038/s41598-017-14238-7
82. Corsico B., Cistola D.P., Frieden C., Storch J. The helical domain of intestinal fatty acid binding protein is critical for collisional transfer of fatty acids to phospholipid membranes. Proc Natl Acad Sci U S A. 1998; 95(21):12174-12178, 10.1073/pnas.95.21.12174
83. Corsico B., Liou H.L., Storch J. The alpha-helical domain of liver fatty acid binding protein is responsible for the diffusion-mediated transfer of fatty acids to phospholipid membranes. Biochemistry. 2004; 43(12):3600-3607, 10.1021/bi0357356
84. Falomir-Lockhart L.J., Laborde L., Kahn P.C., Storch J., Corsico B. Protein-membrane interaction and fatty acid transfer from intestinal fatty acid-binding protein to membranes. Support for a multistep process. J Biol Chem. 2006; 281(20):13979-13989, 10.1074/jbc.M511943200
85. Gabelli S.B., Huang C.H., Mandelker D., Schmidt-Kittler O., Vogelstein B., Amzel L.M. Structural effects of oncogenic pi3kalpha mutations. Curr Top Microbiol Immunol. 2010; 347:43-53, 10.1007/82_2010_53
86. Feng L., Chan W.W., Roderick S.L., Cohen D.E. High-level expression and mutagenesis of recombinant human phosphatidylcholine transfer protein using a synthetic gene: Evidence for a c-terminal membrane binding domain. Biochemistry. 2000; 39(50):15399-15409, 10.1021/bi001076a
87. Pozzi N., Chen Z., Di Cera E. How the linker connecting the two kringles influences activation and conformational plasticity of prothrombin. J Biol Chem. 2016; 291(12):6071-6082, 10.1074/jbc.M115.700401

88. Chinnaraj M., Chen Z., Pelc L.A., Grese Z., Bystranowska D., Di Cera E., Pozzi N. Structure of prothrombin in the closed form reveals new details on the mechanism of activation. Sci Rep. 2018; 8(1):2945, 10.1038/s41598-018-21304-1
89. Dufrisne M.B., Petrou V.I., Clarke O.B., Mancia F. Structural basis for catalysis at the membrane-water interface. Biochim Biophys Acta Mol Cell Biol Lipids. 2017; 1862(11):1368-1385, 10.1016/j.bbalip.2016.11.011