MEMBRANE PROTEINS STRUCTURES: A review on computational modeling tools

Jose G. Almeida[1#], Antonio J. Preto[1#], Panagiotis I. Koukos[2], Alexandre M.J.J. Bonvin[2], Irina S. Moreira[1,2,*]

# Equal Contribution

[1] CNC - Center for Neuroscience and Cell Biology; Rua Larga, FMUC, Polo I, 1ºandar, Universidade de Coimbra, 3004-517, Coimbra, Portugal.

[2] Bijvoet Center for Biomolecular Research, Faculty of Science - Chemistry, Utrecht University, Utrecht, Padualaan 8, 3584CH, the Netherlands

irina.moreira@cnc.uc.pt

Published at: Biochim Biophys Acta. 2017 Oct;1859(10):2021-2039

DOI: 10.1016/j.bbamem.2017.07.008

SUPPLEMENTARY INFORMATION

Table SI-1. Membrane Proteins Databases

Database	*Description*	*Reference and website*
mpstruc	Membrane-embedded structure database	[1] http://blanco.biomol.uci.edu/mpstruc/
MPNMR	MPs of known structure determined by NMR	[1] http://www.drorlist.com/nmr/MPNMR.html
GPCRdb	G-Protein Coupled Receptor Database	[2] *http://gpcrdb.org/*
MemProtMD	analytical pipeline which embeds MPs in lipid bilayer simulations and stores and publishes the results along with	[3] http://sbcb.bioch.ox.ac.uk/memprotmd/beta/

instructions that allow independent researchers to set up their own simulations

Table SI-2. Biocomputational Tools for protein studies.

Tool	***Purpose***	***Reference and website***
AMBER: Assisted Model Building and Energy Refinement	"(…) package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules."	[4] http://ambermd.org/
BOCTOPUS	"Improved topology prediction of transmembrane β barrel proteins."	[5] https://socbin.org/boctopus/index.php
OCTOPUS	"(…) a new method for predicting transmembrane protein topology is presented and benchmarked using a dataset of 124 sequences with known structures. Using a novel combination of hidden Markov models and artificial neural networks, OCTOPUS predicts the correct topology for 94% of the sequences."	[6] http://octopus.cbr.su.se/
CHARMM: Chemistry at Harvard Molecular Mechanics	Molecular simulation program: "(…) the program provides a large suite of computational tools that include numerous conformational and path sampling methods, free energy estimators, molecular minimization, dynamics and analysis techniques, and model-building capabilities."	[7] https://www.charmm.org/charmm/?CFID=cbe96aec-152d-470a-897d-e69e371b2389&CFTOKEN=0
CPORT	"CPORT is an algorithm for the prediction of protein-protein interface residues. It combines six interface prediction methods into a consensus predictor."	[8]

		http://haddock.science.uu.nl/services/CPORT/
GOMoDo	"This webtool performs automatic homology modeling and ligand docking of GPCR receptors."	[9] http://molsim.sci.univr.it/cgi-bin/cona/begin.php
GPCR-I-TASSER	"A Hybrid Approach to G Protein-Coupled Receptor Structure Modeling and the Application to the Human Genome."	[10] http://zhanglab.ccmb.med.umich.edu/I-TASSER/about.html
GPCR-ModSim	"A comprehensive web based solution for modeling G-protein coupled receptors."	[11] http://open.gpcr-modsim.org/
GPCRM	"(…) novel method for fast and accurate generation of GPCR models using averaging of multiple template structures and profile-profile comparison. "	[12] http://gpcrm.biomodellab.eu/
GROMACS: Groningen Machine for Chemical Simulation	"A parallel message-passing implementation of a molecular dynamics (MD) program that is useful for bio(macro)molecules in aqueous environment (…)"	[13] http://www.gromacs.org/
HomPPI: PS (partner-specific) and NPS (nonpartner specific) HomPPI	" (…) a class of sequence homology-based methods for predicting protein-protein interface residues. "	[14] http://ailab1.ist.psu.edu/PSHOMPPIv1.2/
LOMETS: Local Meta-Threading-Server	" (…) on-line web service for protein structure prediction. It generates 3D models by collecting high-scoring target-to-template alignments from 9 locally-installed threading	[15] http://zhanglab.ccmb.med.umich.edu/LOMETS/

	programs (FFAS-3D, HHsearch, MUSTER, pGenTHREADER, PPAS, PRC, PROSPECT2, SP3, and SPARKS-X)."	
MEMBRAIN	"A method to derive transmembrane inter-helix contacts from amino acid sequences by combining correlated mutations and multiple machine learning classifiers."	[16] https://omictools.com/membrain-tool
MEMOIR	"Template-based prediction for membrane proteins (…) produces homology models using alignment and coordinate generation software that has been designed specifically for transmembrane proteins."	[17] http://opig.stats.ox.ac.uk/webapps/memoir/php/index.php
MEMSAT3	"(…) a new method for the prediction of the secondary structure and topology of integral membrane proteins based on the recognition of topological models. "	[18] http://bioinf.cs.ucl.ac.uk/software_downloads/memsat/
PAIRPred: Partner Aware Interacting Residue Predictor	"(…) partner specific protein-protein interaction site predictor that can make accurate predictions of whether a pair of residues from two different proteins interact or not."	[19] http://combi.cs.colostate.edu/supplements/pairpred/
PPIPP: prediction of protein-protein interaction pairs	"This web server takes fasta sequences of two proteins and predicts pairs of interacting residues between them. The input is a pair of amino acid sequences in the fasta format."	[20] http://mizuguchilab.org/netasa/ppipp/
PRIMSIPLR	"Prediction of inner-membrane situated pore-lining residues for alpha-helical transmembrane proteins."	[21]
ProMate	"(…) a structure based prediction program to identify the location of protein-protein binding sites."	[22]
PSIPRED	"A two-stage neural network has been used to predict protein secondary structure based on the position specific scoring matrices generated by PSI-BLAST."	[23] http://bioinf.cs.ucl.ac.uk/psipred/

PSIVER	"(…) server for the prediction of protein-protein interaction sites in protein sequences. This server is free and open to all users and there is no login requirement."	[24] http://mizuguchilab.org/PSIVER/
SPPIDER	"The SPPIDER protein interface recognition server can be used to: (1) predict residues to be at the putative protein interface(s) by considering single protein chain with resolved 3D structure; (2) analyse protein-protein complex with given 3D structural information and identify residues that are being in interchain contact."	[25] http://sppider.cchmc.org/
TMkink	"a method to predict transmembrane helix kinks"	[26] http://tmkinkpredictor.mbi.ucla.edu/
VGIchan	"prediction and classification of voltage-gated ion channels"	[27] http://www.imtech.res.in/raghava/vgichan/

REFERENCES

1. Capponi, S.; Freites, J. A.; Tobias, D. J.; White, S. H., Interleaflet mixing and coupling in liquid-disordered phospholipid bilayers. *Biochimica et biophysica acta* **2016,** *1858* (2), 354-62.
2. (a) Munk, C.; Isberg, V., GPCRdb: the G protein-coupled receptor database - an introduction. **2016,** *173* (14), 2195-207; (b) Isberg, V.; Mordalski, S.; Munk, C.; Rataj, K.; Harpsøe, K.; Hauser, A. S.; Vroling, B.; Bojarski, A. J.; Vriend, G.; Gloriam, D. E., GPCRdb: an information system for G protein-coupled receptors. *Nucleic acids research* **2016,** *44* (D1), D356-D364.
3. Stansfeld, P. J.; Goose, J. E.; Caffrey, M.; Carpenter, E. P.; Parker, J. L.; Newstead, S.; Sansom, M. S., MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. *Structure (London, England : 1993)* **2015,** *23* (7), 1350-61.
4. Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P., AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Computer Physics Communications* **1995,** *91* (1-3), 1-41.
5. Hayat, S.; Elofsson, A., BOCTOPUS: improved topology prediction of transmembrane beta barrel proteins. *Bioinformatics (Oxford, England)* **2012,** *28* (4), 516-22.
6. Viklund, H.; Elofsson, A., OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics (Oxford, England)* **2008,** *24* (15), 1662-1668.

7. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: the biomolecular simulation program. *Journal of computational chemistry* **2009,** *30* (10), 1545-614.
8. de Vries, S. J.; Bonvin, A. M. J. J., CPORT: A Consensus Interface Predictor and Its Performance in Prediction-Driven Docking with HADDOCK. *PLoS ONE* **2011,** *6* (3), e17695-e17695.
9. Sandal, M.; Duy, T. P.; Cona, M.; Zung, H.; Carloni, P.; Musiani, F.; Giorgetti, A., GOMoDo: A GPCRs online modeling and docking webserver. *PloS one* **2013,** *8* (9), e74092.
10. Zhang, J.; Yang, J.; Jang, R.; Zhang, Y., GPCR-I-TASSER: A Hybrid Approach to G Protein-Coupled Receptor Structure Modeling and the Application to the Human Genome. *Structure (London, England : 1993) 23* (8), 1538-1549.
11. Esguerra, M.; Siretskiy, A.; Bello, X.; Sallander, J.; Gutierrez-de-Teran, H., GPCR-ModSim: A comprehensive web based solution for modeling G-protein coupled receptors. **2016,** *44* (W1), W455-62.
12. Latek, D.; Pasznik, P.; Carlomagno, T.; Filipek, S., Towards improved quality of GPCR models by usage of multiple templates and profile-profile comparison. *PloS one* **2013,** *8* (2), e56742.
13. Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R., GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **1995,** *91* (1-3), 43-56.
14. Xue, L. C.; Dobbs, D.; Honavar, V., HomPPI: a class of sequence homology based protein-protein interface prediction methods. *BMC Bioinformatics* **2011,** *12* (1), 1-24.
15. Wu, S.; Zhang, Y., LOMETS: a local meta-threading-server for protein structure prediction. *Nucleic Acids Res* **2007,** *35* (10), 3375-82.
16. Shen, H.; Chou, J. J., MemBrain: Improving the Accuracy of Predicting Transmembrane Helices. *PLoS ONE* **2008,** *3* (6), e2399-e2399.
17. Ebejer, J. P.; Hill, J. R.; Kelm, S.; Shi, J.; Deane, C. M., Memoir: template-based structure prediction for membrane proteins. *Nucleic Acids Res* **2013,** *41* (Web Server issue), W379-83.
18. Jones, D. T., Improving the accuracy of transmembrane protein topology prediction using evolutionary information. *Bioinformatics (Oxford, England)* **2007,** *23* (5), 538-544.
19. ul Amir Afsar Minhas, F.; Geiss, B. J.; Ben-Hur, A., PAIRpred: Partner-specific prediction of interacting residues from sequence and structure. *Proteins* **2014,** *82* (7), 1142-1155.
20. Ahmad, S.; Mizuguchi, K., Partner-Aware Prediction of Interacting Residues in Protein-Protein Complexes from Sequence Data. *PLoS ONE* **2011,** *6* (12), e29104-e29104.
21. Nguyen, D.; Helms, V.; Lee, P.-H., PRIMSIPLR: prediction of inner-membrane situated pore-lining residues for alpha-helical transmembrane proteins. *Proteins* **2014,** *82* (7), 1503-1511.
22. Neuvirth, H.; Raz, R.; Schreiber, G., ProMate: a structure based prediction program to identify the location of protein-protein binding sites. *J Mol Biol* **2004,** *338*.
23. McGuffin, L. J.; Bryson, K.; Jones, D. T., The PSIPRED protein structure prediction server. *Bioinformatics (Oxford, England)* **2000,** *16* (4), 404-405.
24. Murakami, Y.; Mizuguchi, K., Applying the Naive Bayes classifier with kernel density estimation to the prediction of protein-protein interaction sites. *Bioinformatics (Oxford, England)* **2010,** *26*.
25. Porollo, A.; Meller, J., Prediction-based fingerprints of protein-protein interactions. *Proteins* **2007,** *66*.
26. Meruelo, A. D.; Samish, I.; Bowie, J. U., TMKink: a method to predict transmembrane helix kinks. *Protein science : a publication of the Protein Society* **2011,** *20* (7), 1256-1264.
27. Saha, S.; Zack, J.; Singh, B.; Raghava, G. P., VGIchan: prediction and classification of voltage-gated ion channels. *Genomics, proteomics & bioinformatics* **2006,** *4* (4), 253-8.