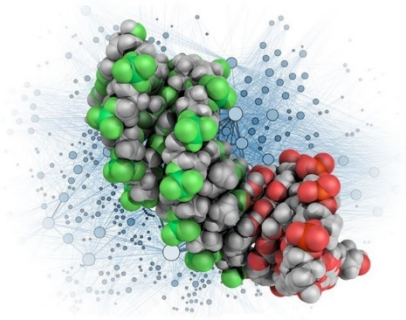


EJIBCE 2017

Encontro de Jovens Investigadores de Biologia Computacional Estrutural
Departamento de Física, Universidade de Coimbra, 22 de Dezembro

MDPI sciforum

MOL2NET, International Conference Series on Multidisciplinary Sciences
http://sciforum.net/conference/mol2net-03

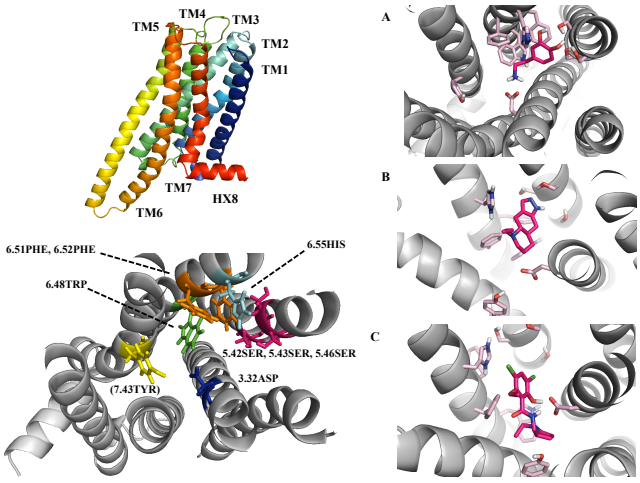
Creating a valid *in silico* Dopamine D2-receptor model for small molecular docking studies

Beatriz Bueschbell (s6bebues@uni-bonn.de)[a], António J. Preto (antonio.gomes@uc.pt)[b], Carlos A.V. Barreto (carlos.barreto@student.uc.pt) [b], Anke C. Schiedel (schiedel@uni-bonn.de) [a], Irina S. Moreira (irina.moreira@cnc.uc.pt)[b]

[a] Pharmaceutical Chemistry I, PharmaCenter Bonn, University of Bonn

[b] - *Structural, Computational and Chemical Biology, CNC - Center for Neuroscience and Cell Biology, University of Coimbra.*

Graphical Abstract



A: Dopamine, B: Quinpirole, C: Raclopride

Abstract

Due to the clinical importance of the Dopamine D2-receptor (D2R) in several brain dysfunctions, the utilization of *in silico* models for drug development is a growing field of investigation. We provided a transparent and reproducible pipeline for creating a valid D2R model for small molecular docking studies. Furthermore, we suggested a binding pocket for the endogenous ligand of D2R, which was attained upon careful consideration of the available experimental data. Molecular docking studies with Dopamine, Quinpirole and Raclopride allowed also a better understanding of the binding pocket characteristics.

Introduction

Dopamine D2-receptor (D2R) is a member of G-Protein Coupled Receptors (GPCR) super-family, consisting in seven transmembranar helices (TM), three extracellular and three intracellular loops (1,2). D2R, along with D3R, are believed to have a more complex function within the dopamine receptor family as they are present in post- and pre-synaptic termina (3). In addition they exert a negative feedback loop in the presynaptic terminal to control firing rate of neurons (4). Along with this functional complexity, it has been shown that alterations in dopaminergic transmission in the brain are correlated to several dysfunctions such as Parkinson's disease and Schizophrenia (5). These distinctive characteristics make D2R an important candidate for drug targeting and in *silico* drug development (3,5). Although some molecular docking studies have been performed with D2R, there are still a lot of problems and open questions that need a new effort in order to be fully understood. First, there is much incongruence in the way new models are created as the use of different templates and modeling softwares makes it hard for a clear comparison between models attained in different research groups. Second, absolute scores of model evaluation programs are often not shown, making model-building not transparent and comparable enough. Another problem rises from the various binding pockets definitions found in literature. Here, we tried to overcome this issue by using a large set of experimental data from Floresca *et al* (6). We aimed to attain a reproducible pipeline to calculate and evaluate correctly GPCR-models in general, and to D2R in particular.

Materials and Methods

Modelling

The building of the D2R model was performed using MODELLER (7), with D3R complexed with Eticlopride, (Protein DataBank (8) ID 3PBL) as template (9). This crystallographic structure was chosen in accordance with a total sequence similarity of 68%, as calculated with BLAST (10). In addition each TM was then aligned to the TMs of the template, which numeration was obtained at the GPCRdatabase (11). The TMs were checked for sequence similarity and a average value of 77 % was attained for this more relevant and conserved helical bundle. Modelling was performed by specifying the lengths of the TMs and the perimembrane intracellular helix (HX8). Furthermore, disulphide bonds were considered in the pairs of unconserved cysteines at positions 79-154 and 249-251. Further loop refinement was performed when needed, in particular for the extracellular loop 2 (ECL2), since it is a long loop highly determinant for D2R's binding pocket access. Since it is known that the deletion of the intracellular loop (ICL3) middle residues does not affect ligand and G-protein binding (the contact points are normally the N- and C-termini of this loop), residues 214-254 from ICL3 were removed as well as the first 28 residues of the sequence, as there is no template for these regions. A dialanine linker was added to connect TM5 and TM6, which were modelled as helices up to the linker, making the intracellular extension of TM5 and TM6 similar to what is observed in the G-protein bound crystal structure of the β2 receptor (PDBid: 3SN6) (12). 2.50ASP was protonated as it known to interact with K+ binding site, which regulates by allostery the ions that regulate the function of the receptor. Furthermore, a part of the ICL2 (112-114) was set as alpha-helix in MODELLER protocol following the know structure-function experimental data available for this loop from the D3R (9).

Model evaluation

Although visual inspection immediately ruled out some trial models due to biological incoherence on particular and important structural features, selecting the best model was cumbersome. This was particularly true since D2R is a membrane protein, for which the more commonly available metrics are

less reliable. Discrete Optimized Protein Energy (DOPE) (13) and molecules' probability density functions (molpdf) are MODELLER's standard metrics for model assessment, based on the models' free energy and special occupation. These, however, did not allowed us to easily pick the best models, in part since they are mainly directed towards water soluble proteins. For this reason Protein Structure Analysis (ProSA) web service was additionally used for error recognition, in particular the z-score, which indicates overall model quality with respect to an energy distribution derived from random conformations (14). However, as the z-score was still not enough to choose the best model, we extended our analysis by using the online Protein Quality (ProQ) (15) prediction server. This is based on a neural network, the LGScore (16), to predict a p-value for the significance of a structural similarity match and the MaxSub (17) that identifies the largest subset of alpha carbons that superimposes with the template structure. Furthermore ProQ allows for the inclusion of secondary structure information (calculated by PSIPRED (18)) in order to further improve model quality assessment. Together with PSIPRED, ProQ can improve up to 15% its' prediction accuracy. Finally, Ballesteros and Weinstein (19) numbering system for class A GPCRs was applied. The numbering system determines helical numbering (for TM1-7 and HX8) depending on a previously determined most conserved residue in each of the helices, named residue x.50.

Docking Protocol

AutoDockTools, a package of MGLTools was used to perform ligand docking (17). Docking itself was performed using Autodock4.2 (version autodock 4.2.6, released in 2009) (21). D2R hydrogens were added and Kollman united atom charges were assigned. Hydrogens were also added to ligand and Gasteiger-Marsili was used to calculate charges (22). Before docking an energy grid was created using Autogrid (version autogrid 4.2.6, released 2009) with a box-size of 50 Å x 50 Å x 50 Å in dimension with a 0.375 Å-spacing. The grid center was set at 18.235, 17.556, 6.595. For each docking simulation 100 independent Lamarckian genetic algorithm (LGA) runs were performed with the number of energy evaluations set to 10.000.000, the population size set to 200 and the maximum number of generations set to 27.000 (23). Default settings were maintained for the rest of the parameters. Docked conformations within a RMSD of 2 Å were clustered. The most populated and lowest energy cluster was used for conformational binding analysis. According to Floresca *et al.* (6), the following residues are important for ligand binding, and were therefore treated as fully flexible during the docking process, along with all rotatable bonds of ligands: 3.32ASP (ASP86), 5.42SER (SER165), 5.43SER (SER166), 5.46SER (SER169), 6.48TRP (TRP236), 6.51PHE (PHE239), 6.52PHE (PHE240), 6.55HIS (HIS243) and 7.43TYR (TYR266). In this study, we docked Dopamine, the endogenous ligand, Quinpirole, a selective D2R/D3R-agonist and Raclopride, a selective D2R-antagonist. We calculated all distances between the center of mass of the ligand and the alpha-C-atom (Cα) of the flexible residues used in the docking for all the top conformations achieved with AutoDock4.2 in order to attain an initial evaluation of these models (21).

Results and Discussion

Due to the high degree of homology between the D2R and D3R (total similarity 68%), it was suitable to use the recently solved crystal structure of the D3R as template (9,24). The ClustalOmega alignment showed that the chosen TMs for the D2R model were conserved compared to D3R (similarity for TM1: 60%, for TM2: 92.86%, for TM3: 87.88%, for TM4: 68.18%, for TM5: 64.71%, for TM6: 70.59%, for TM7: 84.00% and for HX8: 92.31%). For the D2R model a DOPE-score of -39622.63 and z-score of -2.55 (MODELLER and ProSA-web) were attained. In addition, the model was evaluated with ProQ with and without a PSIPRED secondary prediction LGscore: 3.43/3.78 and MaxSub: 0.12

/0.57, where accurate scores were achieved. Although other D2R models were built, definite scores are not provided here. The calculated z-scores outcome is comparable to the z-score of the template structure (PDBid: 3pbl) (9), -2.37, concluding that the created D2R-model showed no faulty regions and was overall in line with the D3R crystal structure. The utilization of ProQ provides general scores categories for a batter comparison. The following classifications were defined for LGscore: "correct" scores >1.50, "good" scores >3 and "very good" scores >5.00. For MaxSub scores >0.1 are defined as "correct", >0.50 is "good" and scores >0.80 are considered "very good". According to that gradation the created D2R is "good" regarding the LGscore and "correct/good" regarding the MaxSub. Interestingly, the MaxSub improved when applying PSIPRED prediction. All in all, the created D2R model seems valid among various protein evaluation programs.

One of the challenges in building *in silico* models is the identification of the binding pocket. In the case of D2R there was enough experimental data to get a close idea about the 3D localization of the pocket. (6) For binding Dopamine, the following residues were considered as interacting residues: 3.32ASP, 5.42SER, 5.43SER, 5.46SER, 6.48TRP, 6.51PHE, 6.52PHE and 6.55HIS according to Floresca *et al*. (6). Additionally, these residues were considered also in others studies involving molecular docking, but in a different constellation, concluding that they are important for general ligand binding to the D2R (26,28,29). To compare the performance between the endogenous ligand, the selective D2R/D3R-agonist Quinpirole and the selective D2R-antagonist Raclopride, an additional conserved residue, 7.43TYR, was considered as part of the flexible residues in the docking procedure, since some ligands, especially antagonists, are believed to access a Secondary Binding Pocket (SBP) (30,31).

With 100 runs performed, AutoDock4.2 clusters similar docking conformations into groups with the same binding energies. The more conformations are found in one cluster, the higher the possibility of this specific docked position to be the closest to reality. Furthermore, the clusters are ranked by lowest binding energy. For Dopamine, the lowest binding energy achieved was -10.51 kcal/mol (22 conformations in this cluster). Moreover, this was the cluster with the most conformations sorted, concluding that the docking performance of Dopamine was reliable. For Quinpirole the lowest binding energy was -8.11 kcal/mol with 17 conformations counted in this cluster. Again, this was the cluster with the higher number of possible conformations. The selective antagonist Raclopride achieved 17 conformations in the cluster with the lowest binding energy of -8.72 kcal/mol. These conformations for all three ligands are shown in the graphical abstract. Regarding literature for similar results, an average docking score of -10.63 kcal/mol was achieved for Dopamine from Durdagi *et al.* (26) using the software Glide (32) and of -6.6 kcal/mol for Quinpirole in a study of Platania *et al* using AutoDock4.2 (27). For further evaluation, the distance between the center of mass of the ligand and the alpha-carbon (Cα) of each residue was measured. First of all, a strong interaction with 3.32ASP was found for Dopamine as well as for the other two ligands, since the average distance between the Cα-atom of 3.32ASP and the center of mass of the ligand was around 8.47 Å for Dopamine, 7.00 Å for Quinpirole and 6.57 Å for Raclopride. According to literature, the positive aspartic acid performs a salt-bridge with protonable amines of the ligands (26,27). We also observed a strong interaction between the by Floresca *et al.* (6) defined serine microdomain (5.42SER, 5.43SER and 5.46SER) and the catecholamine hydroxyl-groups of Dopamine verified by distances of 8.60 Å, 7.57 Å and 8.50 Å between these serines and the ligand. Compared to that, Quinpirole and Raclopride do not seem to interact directly with that microdomain as greater distances were attained: 9.40 Å, 9.14 Å and 9.97 Å and 10.03 Å, 9.37 Å, 10.60 Å, respectively. Other *in silico* docking studies with Dopamine hypothesized that the serine microdomain maintains a H-bonding network when the receptor is activated by its endogenous ligand (6). According to that, Raclopride should not interact with this microdomain at all, which was confirmed by the elevated distances obtained here.

Conclusions

Since a high-resolution crystal structure of D2R is not yet available, it is a common approach in rational drug design to use *in silico*-generated three-dimensional (3D). To attain an accurate 3D model to be used for molecular docking, it is fundamental to correctly define the disulphide-bridges and lengths of the TMs when utilizing the MODELLER-software. Furthermore, loop refinement is a helpful tool to improve the model afterwards. While incorrectly modelled loop positions at the TMs domains can be easily detected visually (e.g. if the loop is positioned into the TMs region), others errors or more problematic regions could be harder to detect. For that, the assessment of the created models with different protein validation programs like ProSA and ProQ is an important step to attain a better initial model for a particular system. Selecting reasonable flexible residues in the docking protocol is also a crucial step to achieve lower binding energies within the clustered conformations. Molecular docking results obtained here were similar to other studies in literature and makes us confident to have a reproducible pipeline to attain a 3D D2R model suitable for docking small ligands with different activation roles. In brief, this approach is a promising base for small molecular docking studies on the D2R.

References

1. Beaulieu J-M, Gainetdinov RR. The Physiology, Signaling, and Pharmacology of Dopamine Receptors. Pharmacol Rev. 2011;63(1):182–217.
2. Moreira IS. Structural features of the G-protein/GPCR interactions. Biochim Biophys Acta - Gen Subj [Internet]. 2014;1840(1):16–33. Available from: http://dx.doi.org/10.1016/j.bbagen.2013.08.027
3. Beaulieu JM, Espinoza S, Gainetdinov RR. Dopamine receptors - IUPHAR review 13. Br J Pharmacol. 2015;172(1):1–23.
4. Johnson KA, Lovinger DM. Presynaptic G Protein-Coupled Receptors: Gatekeepers of Addiction? Front Cell Neurosci. 2016 Nov;10:264.
5. Jaber M, Robinson SW, Missale C, Caron MG. Dopamine receptors and brain function. Neuropharmacology [Internet]. 1997;35(11):1503–19. Available from: http://www.sciencedirect.com/science/article/pii/S0028390896001001
6. Floresca CZ, Schetz JA. Dopamine Receptor Microdomains Involved in Molecular Recognition and the Regulation of Drug Affinity and Function [Internet]. Vol. 24, Journal of Receptors and Signal Transduction. 2004. 207-239 p. Available from: http://www.tandfonline.com/doi/full/10.1081/RRS-200032088
7. Webb B, Sali A, Francisco S. Comparative Protein Structure Modeling Using MODELLER. Curr Protoc Bioinforma. 2017;54(ii):1–55.
8. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. Nucleic Acids Res. 2000;28(1):235–42.
9. Chien EYT, Liu W, Zhao Q, Katritch V, Won Han G, Hanson MA, et al. Structure of the Human Dopamine D3 Receptor in Complex with a D2/D3 Selective Antagonist. Science (80-) [Internet]. 2010;330(6007):1091–5. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1197410
10. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol [Internet]. 1990;215(3):403–10. Available from: http://www.sciencedirect.com/science/article/pii/S0022283605803602
11. Isberg V, Mordalski S, Munk C, Rataj K, Harpsøe K, Hauser AS, et al. GPCRdb: An information system for G protein-coupled receptors. Nucleic Acids Res. 2016;44(D1):D356–64.
12. Rasmussen SGF, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, et al. Crystal structure of the β2 adrenergic receptor-Gs protein complex. Nature [Internet]. 2011;477(7366):549–55. Available from: http://dx.doi.org/10.1038/nature10361
13. Shen M-Y, Sali A. Statistical potential for assessment and prediction of protein structures. Protein Sci. 2006;15:2507–24.

14. Wiederstein M, Sippl MJ. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35(SUPPL.2):407–10.
15. Wallner B, Elofsson A. Identification of correct regions in protein models using structural, alignment, and consensus information. Protein Sci. 2006;15:900–13.
16. Cristobal S, Zemla a, Fischer D, Rychlewski L, Elofsson a. A study of quality measures for protein threading models. BMC Bioinformatics. 2001;2:5.
17. Siew N, Elofsson a, Rychlewski L, Fischer D. MaxSub: an automated measure for the assessment of protein structure prediction quality. Bioinformatics. 2000;16(9):776–85.
18. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics [Internet]. 2000;16(4):404–5. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/16.4.404
19. Ballesteros JA, Weinstein H. Integrated methods for the construction of three dimensional models and computational probing of structure-function relations in G-protein coupled receptors. Methods Neurosci [Internet]. 1995;25:366–428. Available from: http://www.sciencedirect.com/science/article/pii/S1043947105800497
20. Morris GM, Ruth H, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem. 2009 Dec;30(16):2785–91.
21. Stefano Forli, Ruth Huey, Michael E Pique, Michel F Sanner DSG& AJO. Computational protein–ligand docking and virtual drug screening with the AutoDock suite. Nat Protoc [Internet]. 2016;11(5). Available from: https://www.nature.com/articles/nprot.2016.051
22. Gasteiger J, Marsili M. Iterative partial equalization of orbital electronegativity-a rapid access to atomic charges. Tetrahedron. 1980 Jan;36(22):3219–28.
23. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem. 1998 Nov;19(14):1639–62.
24. Zanatta G, Barroso-neto IL, Bambini-junior V, Dutra MF, Bezerra EM, Costa RF, et al. Quantum Biochemistry Description of the Human Dopamine D3 Receptor in Complex with the Selective Antagonist Eticlopride. 2012;5(7):155–62.
25. Salmas RE, Yurtsever M, Stein M, Durdagi S. Modeling and protein engineering studies of active and inactive states of human dopamine D2 receptor (D2R) and investigation of drug/receptor interactions. Mol Divers. 2015;19(2):321–32.
26. Durdagi S, Salmas RE, Stein M, Yurtsever M, Seeman P. Binding Interactions of Dopamine and Apomorphine in D2High and D2Low States of Human Dopamine D2 Receptor Using Computational and Experimental Techniques. ACS Chem Neurosci. 2016;7(2):185–95.
27. Platania CBM, Salomone S, Leggio GM, Drago F, Bucolo C. Homology Modeling of Dopamine D 2 and D 3 Receptors : Molecular Dynamics Refinement and Docking Evaluation. 2012;7(9).
28. Tschammer N, Dörfler M, Hübner H, Gmeiner P. Engineering a GPCR-ligand pair that simulates the activation of D 2L by dopamine. ACS Chem Neurosci. 2010;1(1):25–35.
29. Männel B, Jaiteh M, Zeifman A, Randakova A, Möller D, Hübner H, et al. Structure-guided screening for functionally selective D2 dopamine receptor ligands from a virtual chemical library. ACS Chem Biol [Internet]. 2017;acschembio.7b00493. Available from: http://pubs.acs.org/doi/abs/10.1021/acschembio.7b00493%0Ahttp://www.ncbi.nlm.nih.gov/pubmed/28846380
30. Duan X, Zhang M, Zhang X, Wang F, Lei M. Molecular modeling and docking study on dopamine D2-like and serotonin 5-HT2A receptors. J Mol Graph Model [Internet]. 2015;57:143–55. Available from: http://dx.doi.org/10.1016/j.jmgm.2015.01.014
31. Leggio GM, Bucolo C, Platania CBM, Salomone S, Drago F. Current drug treatments targeting dopamine D3 receptor. Pharmacol Ther [Internet]. 2016;165:164–77. Available from: http://dx.doi.org/10.1016/j.pharmthera.2016.06.007
32. Halgren TA, Murphy RB, Friesner RA, Hege SB, Klicic JJ, Mainz DT, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. J Med Chem. 2004;47(7):1739–49.