Journal of Structural Biology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Assessment of the transmembrane domain structures in GPCR Dock 2013 models

Ting Wang[a,b,*], Haiguang Liu[c], Yong Duan[a]

[a] Genome Center, 451 East Health Science Drive, University of California, Davis, CA 95616, United States
[b] Sichuan University of Science and Engineering, 180 Xueyuan Street, Huixing Road, Zigong 643000, Sichuan Province, China
[c] Complex Systems Division, Beijing Computational Science Research Center, 10 W. Dongbeiwang Rd, Haidian District, Beijing 100193, China

ARTICLE INFO

Keywords:
GPCR
Homology model
Transmembrane
Ligand docking
RMSD
Classification

ABSTRACT

The community-wide blind prediction of G-protein coupled receptor (GPCR) structures and ligand docking has been conducted three times and the quality of the models was primarily assessed by the accuracy of ligand binding modes. The seven transmembrane (TM) helices of the receptors were taken as a whole; thus the model quality within the 7TM domains has not been evaluated. Here we evaluate the 7TM domain structures in the models submitted for the last round of prediction – GPCR Dock 2013. Applying the 7 × 7 RMSD matrix analysis described in our prior work, we show that the models vary widely in prediction accuracy of the 7TM structures, exhibiting diverse structural differences from the targets. For the prediction of the 5-hydroxytryptamine receptors, the top 7TM models are rather close to the targets, which however are not ranked top by ligand-docking. On the other hand, notable deviations of the TMs are found in in the previously identified top docking models that closely resemble other receptors. We further reveal reasons of success and failure in ligand docking for the models. This current assessment not only complements the previous assessment, but also provides important insights into the current status of GPCR modeling and ligand docking.

1. Introduction

The GPCR Dock is a community-wide assessment organized by Abagyan and Stevens with the purpose of evaluating the status of the G-protein coupled receptor (GPCR) structural modeling and ligand docking (Kufareva et al., 2014, 2011; Michino et al., 2009). Three rounds of assessment have been conducted since 2008, shortly after the technique breakthroughs in membrane protein crystallography (Cherezov et al., 2007; Jaakola et al., 2008; Rasmussen et al., 2007; Rosenbaum et al., 2007). During each assessment, participants made blind predictions for receptor-ligand complex structures given the information of amino acid sequences of the receptors and ligand chemical structures. The models were then evaluated by comparing with the experimentally solved structures in several aspects, including the seven transmembrane (7TM) domains, the extracellular loops, the ligand-binding pocket definition, the ligand positions and the atomic contacts between receptor residues and ligands. For evaluation of the 7TM domain structures, the seven TMs were taken as a whole and the models were compared by the overall 7TM root-mean-square-deviations (RMSDs) against the target structures. The results showed that the median 7TM RMSDs of the models were around 2 Å for class A receptor targets and around 6 Å for the non-class A receptor target. It appears that it is no challenge to build a GPCR model with reasonable accuracy for the overall structure given the conserved 7TM topology and the availability of template structures. However, similar overall structures may represent rather dissimilar receptors as we have seen that the GPCR structures solved to date typically show overall 7TM RMSDs in the range of 2–3 Å, but the receptors span several phylogenetically distant families, such as from amine to lipid receptors. Looking into the 7TM bundles, one can find that the relative orientations of the seven helices vary significantly with families, functions and activation states of the receptors (Wang et al., 2017). Therefore, it is necessary to evaluate the structural differences within the 7TM bundles of the models.

Another major result from the three rounds of assessment was that 7TM domain prediction accuracy did not correlate with ligand docking prediction accuracy (Kufareva et al., 2014, 2011; Michino et al., 2009). Receptor modeling and ligand docking appear to be two distinct steps in the generation of the final receptor-ligand complex models. In addition, factors such as placement of the extracellular loop 2 between TM4 and TM5 (ECL2) heavily affect ligand-docking results. In the assessment, the models were primarily evaluated by ligand binding modes and ranked

* Corresponding author at: Genome Center, 451 East Health Science Drive, University of California, Davis, CA 95616, United States.
E-mail address: twang@ucdavis.edu (T. Wang).

https://doi.org/10.1016/j.jsb.2017.11.007
Received 22 October 2017; Received in revised form 15 November 2017; Accepted 20 November 2017

by how much they reproduced the receptor-ligand contacts as observed in the target structures. It is possible that the top docking models deviate from the targets in the exact arrangements of the seven TMs and the "bad" models may have 7TM structures close to the targets. It is thus interesting to find out how the models distribute in the structural space of the exact arrangements of the 7TM helices and furthermore compare top docking models with top 7TM models.

We have recently developed a novel method, called 7 × 7 RMSD matrix to specifically compare the 7TM bundle structures of GPCRs (Wang et al., 2017). Briefly, a 7 × 7 RMSD matrix records the RMSDs of the backbone atoms between reference and target structures for each TM pair when superposing only one TM in turn. As there are seven TMs, the matrix is composed of seven rows and seven columns, thus 7 × 7 = 49 RMSD values. Specifically, the seven elements in the *i*-th row are the backbone RMSDs calculated for TM1 through TM7 when superposing only the *i*-th TM. If the *i*-th TM transforms significantly relative to the reference structure, large RMSDs may appear in the *i*-th row. This correspondence is very pronounced if the TM undergoes rotational transformation. The 49 parameters in the matrix thus contain information about changes in conformations and relative orientations of the seven TMs: the seven diagonal elements serve as an indicator of helical conformational changes or conservations within the TMs themselves and the off-diagonal elements at each row reveal whether the corresponding TM moves from the references structure relative to the other TMs. Rotational angles of each TM relative to the reference structure can also be computed. The 7 × 7 RMSD matrix has been applied to identify and quantify helix movements in active GPCR structures, compare the X-ray structures of 33 unique GPCR receptors, and derive the structural relationships of different GPCRs by their 7TM arrangements (Wang et al., 2017).

In this work, we apply the 7 × 7 RMSD matrix analysis to evaluate the 7TM domain structures of the receptor models that were submitted for the last round of GPCR Dock, which was performed in 2013. The targets in GPCR Dock 2013 include three receptors, the human 5-hydroxytryptamine (5HT, or serotonin) receptors 1B and 2B (5HT1B and 5HT2B, respectively) and the human smoothened receptor (SMO). Compared to previous assessment rounds, these targets are more challenging in both receptor modeling and ligand docking due to difficulties in modeling of agonist-bound activation states, ligand-interacting ECL2 loops, and a non-class A receptor with distant homology to the structure-known GPCRs. The models have previously been extensively evaluated by using criteria including overall accuracy of the 7TM prediction and ligand-docking (Kufareva et al., 2014), but the structural differences within the 7TM domains have not been evaluated nor the relationships with the experimental structures. Our current assessment complements the previous assessment by focusing on the exact 7TM arrangements of the receptor models. Combined together, the results provide a more detailed and comprehensive picture about the status of GPCR modeling and docking.

2. Methods and Materials

2.1. Source of GPCR Dock 2013 models

In GPCR Dock 2013, the participants submitted 181, 171, 88, and 88 models predicting four receptor-ligand complex structures: 5HT1B and 5HT2B, both with an agonist ergotamine(Wacker et al., 2013; Wang et al., 2013b) and SMO with two distinct antagonists, LY-2940680 (Wang et al., 2013a) and SANT-1 (Wang et al., 2014), respectively. We obtained the PDB files of all models from the GPCR Dock 2013 web site at http://ablab.ucsd.edu/GPCRDock2013.

2.2. Structures of targets and templates

The PDB IDs of the four target structures are 4IAR for the 5HT1B-ergotamine complex, 4IB4 for the 5HT2B-ergotamine complex, 4 JKV for the SMO-LY-2940680 complex (Wang et al., 2013a), and 4N4W for the SMO-SANT-1 complex (Wang et al., 2014), respectively.

According to the methods description provided by the participants (Kufareva et al., 2014), a number of GPCR structures available at the time of prediction were used as templates for model generation. These include rhodopsin (RHO)(PIDB ID: 1U19 (Okada et al., 2004)), adrenoceptors β1AR (PDB ID: 2 VT4 (Warne et al., 2008)) and β2AR (PDB ID: 2RH1 (Cherezov et al., 2007)), muscarinic receptors M2 (PDB ID: 3UON (Haga et al., 2012)) and M3 (PDB ID: 4U15 (Thorsen et al., 2014)), dopamine receptor D3(PDB ID: 3PBL(Chien et al., 2010)), histamine receptor H1(PDB ID: 3RZE (Shimamura et al., 2011)), adenosine receptor A2A (PDB ID: 3EML (Jaakola et al., 2008)), opioid receptors NOP(PDB ID: 4EA3 (Thompson et al., 2012)), δOR(PDB ID: 4N6H (Fenalti et al., 2014)), κOR (PDB ID: 4DJH (Wu et al., 2012)) and μOR (PDB ID: 4DKL (Manglik et al., 2012)), chemokine receptor CXCR4(PDB ID: 3ODU (Wu et al., 2010)), sphingosine 1-phosphate receptor S1P1 (PDB ID: 3V2Y (Hanson et al., 2012)). Some of the structures were solved with low resolution at that time (> 3.0 Å); we replaced them with the recent structures of higher resolution. In addition, several agonist-bound or active structures were also used as templates, including agonist-bound β1AR (PDB ID: 2Y02), agonist-bound A2AR (PDB ID: 3QAK), agonist-G-protein-bound β2AR (PDB ID: 3SN6) and ligand-free rhodopsin (PDB ID: 3CAP).

2.3. TM residue ranges

In our assessment of 5HT1B and 5HT2B models, the TM residues are defined as 1.36–1.60, 2.38–2.65, 3.23–3.55, 4.39–4.62, 5.36–5.63, 6.32–6.58 and 7.33–7.53 in the Ballesteros-Weinstein notion (Ballesteros and Weinstein, 1995). These residues are the maximal TM residues that are solved and present in all target, template and model structures.

In the assessment of SMO models, because of the shorter TM4 in target structures 4 JKV and 4N4W, the TM4 range was changed to 4.45–4.62.

2.4. Ligand binding-pocket residues

We defined the ligand-binding pocket residues to be the protein residues within 4.0 Å distance of any of the ligand atoms. In 5HT1B-ergotamine complex structure (PDB id: 4IAR), the binding pocket residues are W125, L126, D129, I130, C133, T134, C199, V200, V201, S212, A216, W327, F330, F331, S334, L348, F351, D352, T355, and Y359. In 5HT2B-ergotamine complex structure (PDB id: 4IB4), the binding pocket residues are D135, V136, S139, T140, V208, L209, T210, K211, M218, A225, W337, F340, N344, L347, and Q359.

3. 7 × 7 RMSD matrix calculations

Structural superposition and RMSD calculations are performed with the VMD software (Humphrey et al., 1996), including calculations of the overall 7TM RMSDs and the 7 × 7 RMSD matrices. The RMSD values are calculated for the backbone atoms of each TM, i.e. the CA, C, O, N atoms. A 7 × 7 RMSD matrix is obtained by successively superposing one of the seven TMs and recording RMSDs for all seven TMs. For comparison between models and targets, the matrix of each model is computed against the corresponding target structure. For all-against-all structural comparison, each of the structures in the data set serves as a reference structure to compute the 7 × 7 RMSD matrices of all other structures.

3.1. Rotational angle calculations

Rotations of the TMs are defined as around the principal axes of the target structure. It is worth noting that the principal axes of the target structure are not the helical axis of any TM. To compute the rotational