ACS Chemical Neuroscience
Cite This: ACS Chem. Neurosci. 2019, 10, 677−689
Research Article
pubs.acs.org/chemneuro

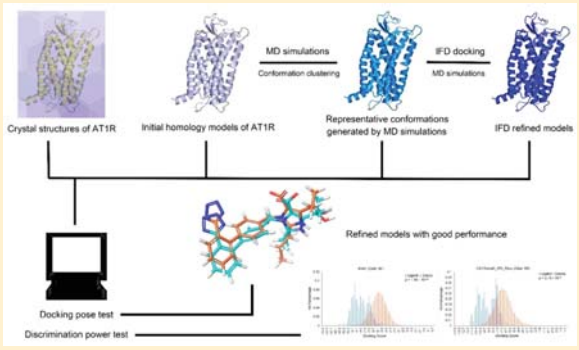
Reliability of Docking-Based Virtual Screening for GPCR Ligands with Homology Modeled Structures: A Case Study of the Angiotensin II Type I Receptor

Haiyi Chen,[†] Weitao Fu,[†] Zhe Wang,[†] Xuwen Wang,[†] Tailong Lei,[†] Feng Zhu,[†] Dan Li,[†] Shan Chang,[§] Lei Xu,[§] and Tingjun Hou*[,†,‡]

[†]College of Pharmaceutical Sciences and [‡]State Key Lab of CAD&CG, Zhejiang University, Hangzhou, Zhejiang 310058, China
[§]Institute of Bioinformatics and Medical Engineering, School of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou 213001, P. R. China

Supporting Information

ABSTRACT: The number of solved G-protein-coupled receptor (GPCR) crystal structures has expanded rapidly, but most GPCR structures remain unsolved. Therefore, computational techniques, such as homology modeling, have been widely used to produce the theoretical structures of various GPCRs for structure-based drug design (SBDD). Due to the low sequence similarity shared by the transmembrane domains of GPCRs, accurate prediction of GPCR structures by homology modeling is quite challenging. In this study, angiotensin II type I receptor (AT1R) was taken as a typical case to assess the reliability of class A GPCR homology models for SBDD. Four homology models of angiotensin II type I receptor (AT1R) at the inactive state were built based on the crystal structures of CXCR4 chemokine receptor, CCR5 chemokine receptor, and δ-opioid receptor, and refined through molecular dynamics (MD) simulations and induced-fit docking, to allow for backbone and side-chain flexibility. Then, the quality of the homology models was assessed relative to the crystal structures in terms of two criteria commonly used in SBDD: prediction accuracy of ligand-binding poses and screening power of docking-based virtual screening. It was found that the crystal structures outperformed the homology models prior to any refinement in both assessments. MD simulations could generally improve the docking results for both the crystal structures and homology models. Moreover, the optimized homology model refined by MD simulations and induced-fit docking even shows a similar performance of the docking assessment to the crystal structures. Our results indicate that it is possible to establish a reliable class A GPCR homology model for SBDD through the refinement by integrating multiple molecular modeling techniques.



KEYWORDS: *Docking-based virtual screening, homology model, class A GPCRs, antagonist, MD simulation, induced-fit docking*

INTRODUCTION

G-protein-coupled receptors (GPCRs), also known as seven-transmembrane domain proteins, mediate most cellular responses to specific signal molecules outside the cell, such as hormones and neurotransmitters.[1] As the largest family of membrane proteins, GPCRs are related with many human diseases, including diabetes, cancers, hypertension, and neurological diseases and are the primary targets of approximately 34% of approved drugs.[2−5] About 800 GPCRs in the human genome are commonly divided into four classes according to their sequence homology: class A (rhodopsin-like), class B (secretin), class C (glutamate), and class F (frizzled/smoothened). The rhodopsin-like family (class A) comprises about 84% of all GPCR members.[6]

Given the therapeutic importance of GPCRs, extensive efforts have been dedicated to screen and identify novel GPCR ligands in the last two decades.[3,7,8] Unfortunately, difficulty associated with growing diffraction-quality crystals of GPCRs brings a big challenge to solving GPCR structures.[1] In 2000, the first crystal structure of GPCR, bovine rhodopsin, was solved through X-ray crystallography,[9] and then, in late 2007, the first high-resolution crystal structure of a human GPCR, β_2-adrenergic receptor, was determined by X-ray crystallography, which initialized the huge progress in structural determination of GPCRs.[10,11] Up to now, the crystal structures of more than 70 unique GPCRs have been solved,[8] but the gap between the available sequences and structures of GPCRs is still quite huge. In the absence of experimental structures, theoretical predictions of GPCR structures by computational techniques, such as homology modeling, are quite demanding.[12] It is well-

Received: September 13, 2018
Accepted: September 28, 2018
Published: September 28, 2018

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

known that the quality of a homology model depends on the sequence identity between the target and template. Once the sequence identity is below 30%, accurate prediction of the target structure is extremely difficult.[13] The overall sequence identity among GPCRs is typically as low as 20%−30%, but the negative impact of relatively low sequence identity on homology modeling can be partially compensated by the highly conserved seven-transmembrane topology and highly conserved residues in the transmembrane helices.[14] With the tremendous expansion of the number of solved GPCR crystal structures in recent years, more diverse templates can be used in homology modeling, which enriches our choice and strategy for building more accurate homology models.[15]

However, it is still a big concern whether a given GPCR homology model can be used as a reliable initial structure for virtual screening (VS) or structure-based drug design (SBDD). In 2011, a docking-based VS campaign was carried out against both the D3 crystal structure and homology model, and the results showed that the homology model performed as well as the crystal structure.[16] However, the transmembrane sequence identity between the target sequence and templates (β1- and β2-adrenergic receptors) is relatively high (42%), which guarantees the reliability of the modeled structure. In 2012, a similar VS campaign was conducted with both the CXCR4 crystal structure and homology model. Compared with D3, the homology modeling for CXCR4 is more challenging, because the sequence identity between the target and four templates (β1- and β2-adrenergic receptors, adenosine A2A receptor, and rhodopsin) is at most 25%. The experimental results show that the docking against the crystal structure identified four novel antagonists while that against the homology model identified only one antagonist with limited novelty and specificity.[17] Apparently, the homology model of CXCR4 is not a good choice for SBDD.

It should be noted that GPCR homology modeling is still an uncertain process, and therefore, a GPCR homology model should be carefully evaluated before it is used in SBDD.[18] The quality of a GPCR homology model is determined by many factors. In addition to reliability of sequence alignment between the template and target and accuracy of variable loop modeling, rigorous structural refinement is quite essential. However, it seems that there is no standard pipeline for the refinement of GPCR homology models. In most studies, the initial structures of GPCR homology models were simply relaxed by molecular mechanics (MM) minimization.[18−20] A number of models were further refined and improved by molecular dynamics (MD) simulations without or with a lipid bilayer environment.[21−24] Moreover, MD simulations can produce a conformational ensemble instead of a single conformation for docking, which may yield more reasonable docking results.[25−28] Furthermore, in several studies, the binding sites of the homology models were optimized by induced-fit docking (IFD) to allow for side-chain flexibility.[18−20] Based on the refined optimized homology models, the capability of a given docking algorithm to recognize near-native-binding poses of known ligands and distinguish known ligands from decoys needs to be evaluated. However, in many studies, the quality of the reported GPCR homology models was not evaluated.[29]

It would be quite interesting to explore the effect of structural refinement on the quality of GPCR homology models and assess the reliability of docking-based VS against the homology models for class A GPCRs. In this study, angiotensin II type 1 receptor (AT1R) was used as the tested case. AT1R is closely related with hypertension and cardiovascular diseases, and AT1R blockade by antagonists can lower blood pressure in hypertensive patients and improves systemic hemodynamics in patients with congestive heart failure.[30−32] In 2015, through serial femtosecond crystallography with an X-ray free-electron laser, the crystal structures of AT1R (PDB entries: 4YAY and 4ZUD) in the inactive state in complex with its selective antagonists ZD7155 and olmesartan were determined,[33] and they were then used in docking-based VS.[34,35] Before the release of the crystal structures of AT1R, a homology model of AT1R was built based on the crystal structure of bovine rhodopsin and used in molecular docking, but the quality of this model was not evaluated.[36] A well-known rule of thumb for homology modeling is that once the target-template sequence identity is below 30%, the quality of the resulting homology model may not be guaranteed. Chemokine and opioid receptors are the two most homologous subtypes to AT1R, and they share ∼30% sequence identity with AT1R. Thus, AT1R is a representative system to investigate how the homology models compare to the crystal structures for ligand discovery. In this study, the homology models of AT1R were built and refined and then evaluated by the following two criteria: prediction accuracy of ligand-binding poses and screening capability of docking-based VS.[37−39] It is expected that according to the calculation results, we can answer the following two questions: First, are the homology models of AT1R reasonably accurate for ligand discovery? Second, can the homology models be effectively improved by the refinement based on different molecular modeling techniques?

RESULTS AND DISCUSSION

Structural Inspection of the Homology Models. The quality of the homology models was verified by Profiles-3D. As shown in Table S1, the scores predicted by Profiles-3D for most homology models are not quite satisfactory. In fact, the Profile-3D scores (61.33 and 59.57) for the two crystal structures (4YAY and 4ZUD) are even slightly lower than the expected low scores (62.67 and 59.99). That is to say, the quality of the crystal structures is not good enough according to the assessment of Profile-3D. It is quite possible that Profile-3D is inappropriate to be used for membrane proteins, because it is optimized based on the data set with mostly globular proteins. Then, the local geometries of each model were assessed by PROCHECK. As shown in Table S2, occupancies of residues (nonglycine and nonproline) in the most favorable regions in the Ramachandran plots for all the homology models are higher than 92%, which are even higher than those for the two crystal structures (85.1% for 4YAY and 88.8% for 4ZUD). Moreover, occupancies of residues in disallowed regions for the homology models are quite low, implying that the homology models are reasonably reliable in terms of their stereochemical properties.

The structural difference between the homology models and crystal structures were visually inspected. As shown in Figure 1, although a number of structural differences exist between the crystal structures and homology models, the majority of the homology models are still quite close to that of the crystal structures. The RMSDs of the TM backbone atoms for the four homology models compared with those of the crystal structure (4YAY) range from 1.80 to2.28 Å. In the transmembrane portion, the upper half of TM5 in the crystal

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

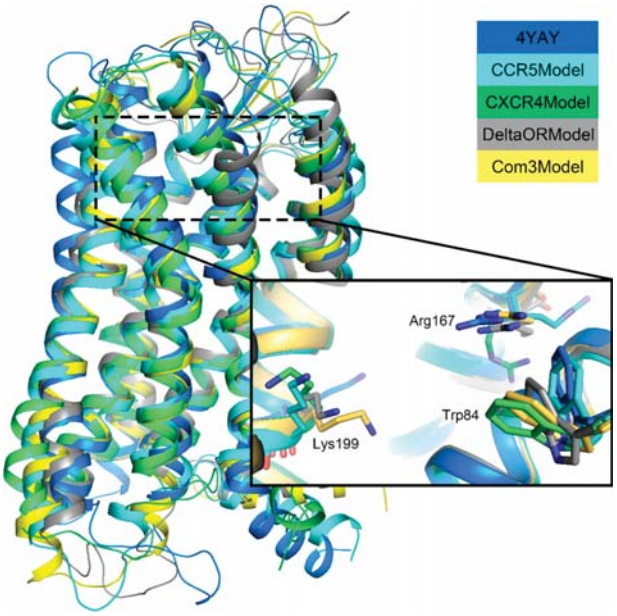


Figure 1. Structural alignment of the homology models and the crystal structure (4YAY). The three important residues for ligand binding were highlighted.

structures adopts a different conformation compared with those in the homology models, for example, the C_α atom of Gly194 in 4YAY shifts outward for 4.25 Å relative to that in CCR5Model. At the intracellular side, helixes VII and ICL3 adopt the most diverse conformations. However, the intracellular portion is not directly involved in ligand binding. To evaluate the similarity of the binding pockets between the crystal structure (4YAY) and homology models (Figure 1), the RMSDs of the heavy atoms of the binding pockets, which were defined as the residues within 5 Å of the ligand in 4YAY, were calculated. Among the four homology models, Com3Model has the highest pocket similarity to 4YAY (RMSD = 1.48 Å), and CXCR4Model also bears relatively high similarity (RMSD = 1.74 Å). DeltaORModel and CCR5Model exhibit lower similarity to the crystal structure, and the conformations of Tyr292, Arg167, and Lys199 are quite different from those in the crystal structure. According to the previous study, Arg167, Trp84, and Lys199 are the three most crucial residues to ligand binding,[33] and a number of nearby residues determine the shape of the pocket. The inward shift of TM5 leads to a significant positional difference of Lys199 between the crystal structure (4YAY) and homology models. An oriented difference exists with respect to Tyr292 in the binding pocket, and the Tyr292 residues in the four models adopt similar orientations that are obviously different from that in the crystal structure. Trp84 in CCR5Model adopts a more rational orientation than those in the other models, but the side chain of Arg167 in CCR5Model takes a reverse orientation compared with that in the crystal structure. The side chains of Trp84 and Arg167 in CXCR4Model adopt different orientations compared with those in the crystal structure, but the RMSD values of these two residues between CXCR4Model and the crystal structure are still acceptable (RMSD = 2.60 Å for Trp84 and RMSD = 3.74 Å for Arg167). Generally, apparent structural differences between the crystal structure (4YAY) and homology models mainly lie at Tyr92, Arg167, Lys199, and Tyr292 (Figure S1). Among the homology models, Com3Model exhibits the highest binding site similarity to the crystal structure, suggesting the homology modeling based on multiple templates may be a better strategy than that based on any single template.

Performance of Molecular Docking Based on the Homology Models. First, the quality of the homology models prior to any refinement was assessed by the reliability for reproducing the experimental-binding modes of the ligands in the crystal structures. As shown in Tables 1 and 2, for 4YAY,

Table 1. AUCs for the ROCs, the *p*-Values Given by the *t*-Test, and the Enrichment Ratio for the Top 20 and Top 100 Molecules Predicted by Molecular Docking

	AUC		*p*-value		enrichment ratio in top 20		enrichment ratio in top 100	
	Glide SP	Vina	Glide SP	Vina	Glide SP	Vina	Glide SP	Vina
YAY	0.940	0.804	1.54×10^{-40}	2.72×10^{-22}	60.0%	0.0%	52.0%	3.0%
4YAY50ns	0.952	0.886	4.38×10^{-44}	3.74×10^{-35}	55.0%	0.0%	57.0%	14.0%
4ZUD	0.867	0.775	2.24×10^{-29}	9.48×10^{-20}	20.0%	0.0%	19.0%	0.0%
4ZUD50ns	0.955	0.890	4.80×10^{-44}	2.83×10^{-34}	80.0%	30.0%	58.0%	27.0%
CCR5Model	0.742	0.720	6.34×10^{-15}	4.36×10^{-14}	0.0%	0.0%	1.0%	0.0%
CXCR4Model	0.725	0.770	4.43×10^{-14}	6.44×10^{-18}	0.0%	0.0%	1.0%	13.0%
DeltaOR	0.493	0.688	0.970	4.66×10^{-10}	0.0%	0.0%	0.0%	0.0%
Com3Model	0.820	0.830	5.18×10^{-25}	3.73×10^{-25}	20.0%	5.0%	10.0%	14.0%
Com3md0	0.750	0.725	3.54×10^{-17}	5.52×10^{-13}	0.0%	5.0%	1.0%	8.0%
Com3md1	0.781	0.711	4.55×10^{-19}	6.00×10^{-12}	5.0%	0.0%	8.0%	2.0%
Com3md2	0.721	0.701	1.96×10^{-12}	1.62×10^{-11}	0.0%	0.0%	1.0%	1.0%
Com3md3	0.787	0.792	3.87×10^{-21}	1.81×10^{-21}	0.0%	5.0%	5.0%	6.0%
Com3md4	0.724	0.710	2.16×10^{-13}	2.14×10^{-12}	5.0%	0.0%	3.0%	0.0%
Com3md5	0.740	0.717	8.18×10^{-15}	2.85×10^{-13}	5.0%	0.0%	5.0%	2.0%
Com3md4_IFD_50ns	0.813	0.696	3.36×10^{-23}	2.48×10^{-10}	0.0%	0.0%	6.0%	6.0%
CXCR4md0	0.800	0.751	7.00×10^{-22}	4.78×10^{-18}	5.0%	0.0%	7.0%	0.0%
CXCR4md1	0.754	0.708	1.54×10^{-16}	1.01×10^{-12}	0.0%	5.0%	7.0%	13.0%
CXCR4md2	0.799	0.738	6.14×10^{-21}	7.74×10^{-15}	5.0%	5.0%	16.0%	7.0%
CXCR4md3	0.870	0.746	1.28×10^{-29}	4.49×10^{-19}	30.0%	0.0%	34.0%	1.0%
CXCR4md3_IFD_50ns	0.890	0.802	2.19×10^{-31}	7.45×10^{-25}	55.0%	0.0%	50.0%	1.0%

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

Table 2. RMSDs of the Poses Predicted by Molecular Docking Relative to the Experimental Poses in the Crystal Structures (Å)

	ZD7155			olmesartan		
	Glide SP	Glide XP	Vina	Glide SP	Glide XP	Vina
4YAY	0.908	0.278	1.056	1.448	1.600	0.845
4YAY50ns	1.007	1.101	0.760	1.700	2.816	1.196
4ZUD	1.134	0.997	9.459	1.422	1.482	8.245
4ZUD50ns	1.754	1.267	3.345	1.553	1.652	1.461
CCR5Model	3.593	4.614	\	8.684	8.807	4.130
CXCR4Model	4.110	4.182	8.423	5.329	5.475	9.021
DeltaORModel	5.768	5.693	3.877	3.846	3.840	3.537
Com3model	6.560	5.115	9.345	7.870	4.833	7.840
Com3md0	4.072	3.163	8.628	9.069	3.425	7.744
Com3md1	3.893	4.247	8.072	8.512	4.152	7.364
Com3md2	6.162	6.170	\	8.853	8.642	7.364
Com3md3	8.681	3.352	9.290	8.226	3.112	8.207
Com3md4	3.358	3.467	3.339	9.316	3.238	3.536
Com3md5	3.053	5.878	8.806	8.709	8.465	8.409
CXCR4md0	7.911	5.750	10.074	8.045	8.598	8.193
CXCR4md1	6.628	6.747	4.603	5.044	5.683	3.716
CXCR4md2	7.760	6.379	6.590	8.558	8.482	8.260
CXCR4md3	5.189	2.079	8.193	2.106	8.558	2.095
Com3md4_IFD	2.157	1.766	7.982	1.760	1.742	3.639
Com3md4_IFD_50ns	2.663	2.095	9.412	2.355	3.136	3.267
CXCR4md3_IFD	6.130	10.032	2.151	1.940	2.011	2.203
CXCR4md3_IFD_50ns	1.965	2.442	1.720	1.991	1.780	0.937

[a]"\" means that the RMSD between the two ligands is beyond 10.5 Å.

two docking algorithms can correctly reproduce the binding poses of ZD7155 and olmesartan in the crystal structures (RMSD < 2 Å). For 4ZUD, Glide can reproduce the binding poses of ZD7155 (RMSD = 1.13 and 1.1 Å) and olmesartan (RMSD = 1.42 and 1.48 Å), but Autodock Vina cannot recognize their correct binding poses successfully (RMSD = 9.46 Å for ZD7155 and 8.24 Å for olmesartan). For the four homology models, Glide can produce better predictions than Autodock Vina, but all of their predictions are unable to accurately predict the binding poses of ZD7155 and olmesartan in the crystal structures.

Then, the compounds in the validation data set were docked into the crystal structures and homology models to assess the screening power of molecular docking to distinguish the known antagonists from decoys. Compared with the binding pose assessment, the screening power is more important for a practical VS campaign. The screening power for each protein structure was evaluated by the AUC of the ROC curve generated based on the docking scores of the known antagonists and decoys and the *p*-value of the difference between the means of the two distributions of the docking scores for the known antagonists and decoys given by the student's *t*-test. As shown in Figures 2, S2, S3, and S4, as expected, the crystal structure of 4YAY performed much better than the homology models, indicated by the high AUC of 0.940 (Figure 2) and the lowest *p*-value of 1.54×10^{-40} (Figure S2). It is interesting to notice that Com3Model (AUC = 0.830 and $p = 3.73 \times 10^{-25}$) performed significantly better than the other homology models and even better than the crystal structure of 4ZUD when docking with Autodock Vina (AUC = 0.775 and p = to 9.48×10^{-20}). The enrichment ratio for each structure is also shown in Figure 3 and Table 1. The crystal structure of 4YAY possesses a much higher enrichment ratio (~60% among the top 140 molecules) than the other structures. The performance of 4ZUD is slightly better than

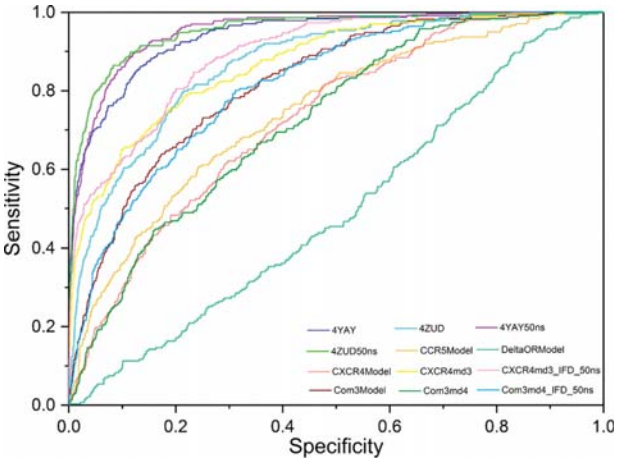


Figure 2. ROC curves based on the Glide SP docking scores of the known antagonists and decoys in the validation data set.

that of Com3Model and significantly better than those of the other homology models. In summary, the crystal structure of 4YAY shows the best performance in the above two assessments. Overall, Com3Model displays similar performance to the crystal structure of 4ZUD, but it still needs further improvement.

MD Simulation Improved Docking Power of Crystal Structures. It is a common strategy to refine protein structures with MD simulations. In this study, the two crystal structures (4YAY and 4ZUD) were submitted to MD simulations in a membrane environment. The RMSDs of the whole structures and those of the binding site residues were monitored to inspect the structural changes during the simulations. As shown in Figure 4, the two structures

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

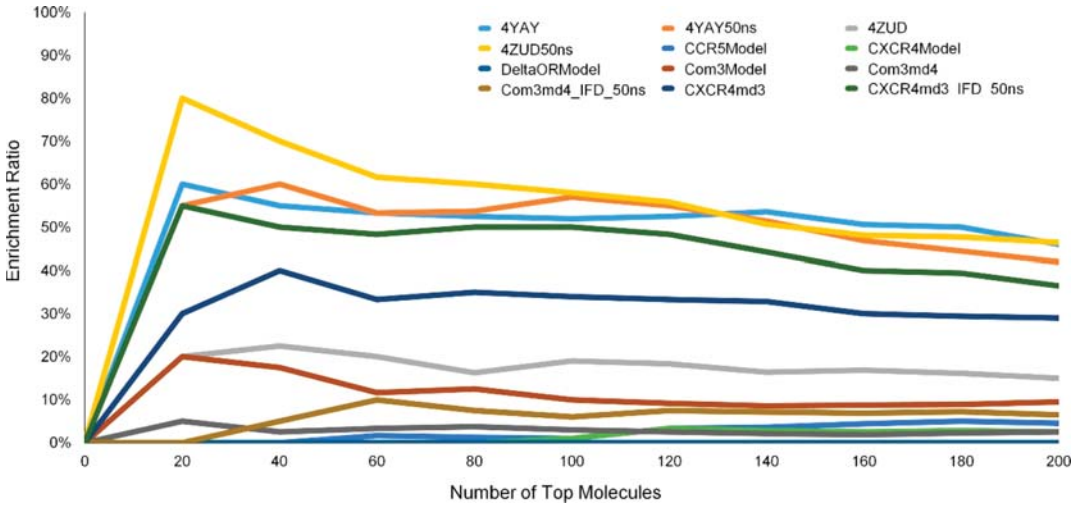


Figure 3. Enrichment ratio for the top 200 molecules predicted by Glide docking based on the SP scoring. Enrichment ratio = number of true ligands among the top N molecules/N × 100%.

underwent obvious conformational rearrangement and achieved stability after about ∼10 ns.

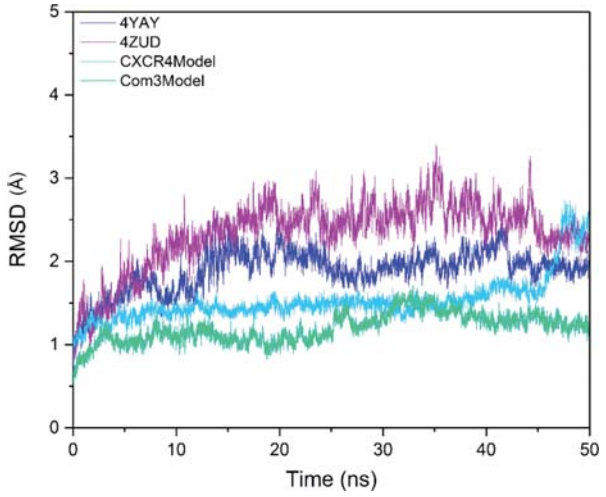


Figure 4. RMSD of the binding pocket residues as the function of simulation time for 4YAY, 4ZUD, Com3md4_IFD, and CXCR4md3_IFD.

After the MD simulations, the last snapshots (4YAY50nsMD and 4ZUD50nsMD) for 4YAY and 4ZUD were extracted from the MD trajectories and prepared by the *Protein Prepare Wizard* in Schrodinger. The screening powers for 4YAY50nsMD and 4ZUD50nsMD were evaluated by using the same strategy mentioned above. The results show that the MD simulations improved the screening powers for both of the two crystal structures (Figures 2, S2, S3, and S4): AUC = 0.952 and $p = 4.38 \times 10^{-44}$ for 4YAY50ns versus AUC = 0.940 and $p = 1.54 \times 10^{-40}$ for 4YAY (Glide SP); AUC = 0.955 and $p = 4.80 \times 10^{-44}$ for 4ZUD50ns versus AUC = 0.867 and $p = 2.24 \times 10^{-29}$ for 4ZUD (Glide SP); AUC = 0.886 and $p = 3.74 \times 10^{-35}$ for 4YAY50ns versus AUC = 0.804 and $p = 2.72 \times 10^{-22}$ for 4YAY (Autodock Vina); and AUC = 0.890 and $p = 2.83 \times 10^{-34}$ for 4ZUD50ns versus AUC = 0.775 and $p = 9.48 \times 10^{-20}$ for 4ZUD (Autodock Vina). The enrichment ratio for 4ZUD50nsMD is much higher than that of 4ZUD and even slightly higher than that of 4YAY in the top 100 molecules (Figure 3). It is surprising to observe that in the top 20 molecules, 80% of the molecules recognized by 4ZUD50nsMD are true antagonists. In addition, it can be observed that, based on the MD refined structure of 4ZUD, the binding poses of ZD7155 and olmesartan predicted by Autodock Vina become much more rational compared with those based on the initial crystal structure of 4ZUD (Tables 1 and 2). The reported AT1R structures were crystallized in a nonmembrane environment,[33] and therefore, the solved structures might be different from the true conformation in a hydrated membrane environment. Theoretically, MD simulations in a virtual membrane environment could transform the experimentally determined conformation to a more realistic one. Therefore, the crystallized structures may not be considered as the gold standard for SBDD, and further refinement may be necessary.

MD Simulation and Conformational Clustering Generated Better Theoretical Models for Docking. According to the evaluation results, the two crystal structures outperformed the homology models. Therefore, the two homology models (Com3Model and CXCR4Model) were refined by MD simulations. The same preparation protocol was employed for Com3Model and CXCR4Model with CHARMM-GUI to build the membrane environment. Because there is no original ligand in both homology models, the MD simulation would be carried out without any ligand. As shown in Figure 4, the RMSDs of Com3Model are quite stable during the whole simulations and even more stable than the two crystal structures. The RMSDs of CXCR4Model are also quite stable between 5−45 ns and increase from ∼1.6 to 2.3 Å at 45 ns. Moreover, according to the conformational-selection theory (population-shift mechanism), a ligand would be recognized by the best conformation among the multiple conformations of a target. Therefore, based on the MD trajectories, multiple representative conformations were extracted for docking assessments.

By conformational clustering based on the MD trajectory (Figure 5), six different representative conformations were identified for Com3Model. The population plot of the conformational clusters is shown in Figure 5C, and the

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

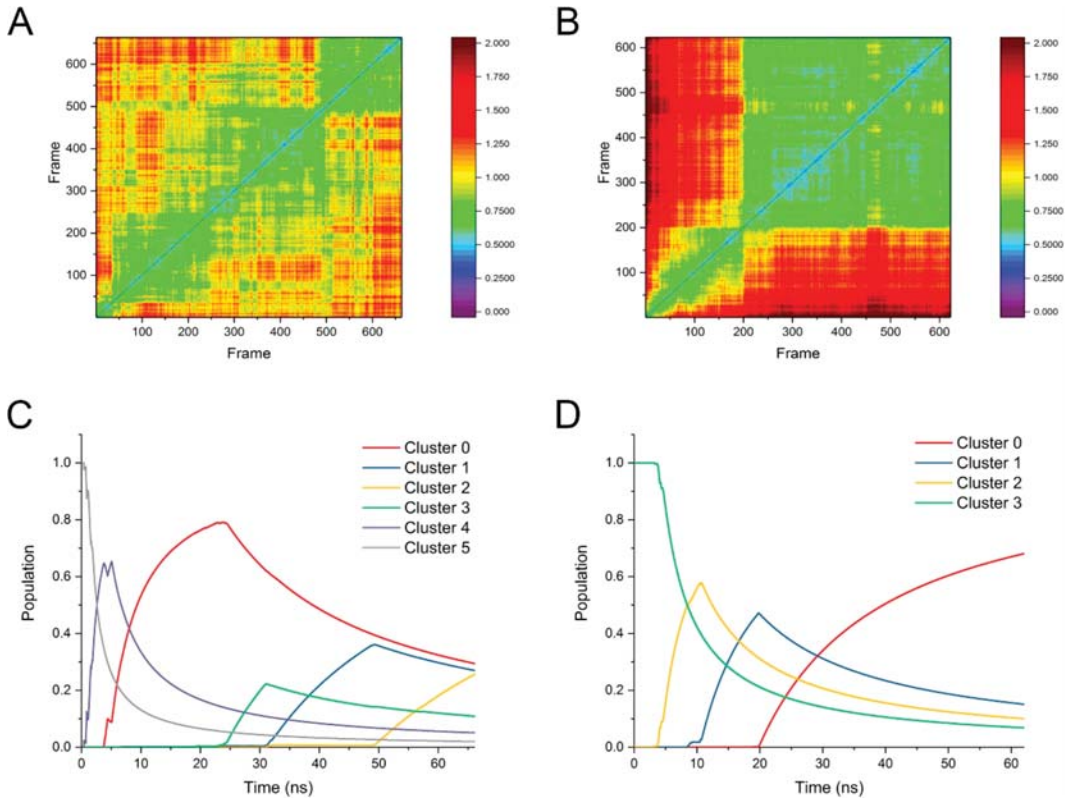


Figure 5. RMSD matrix for the binding pocket residues generated by RMS2D in *cpptraj* for (A) Com3Model and (B) CXCR4Model. The population plot of the conformational clusters for (C) Com3Model and (D) CXCR4Model. The interval between two neighboring snapshots is set to 0.1 ns.

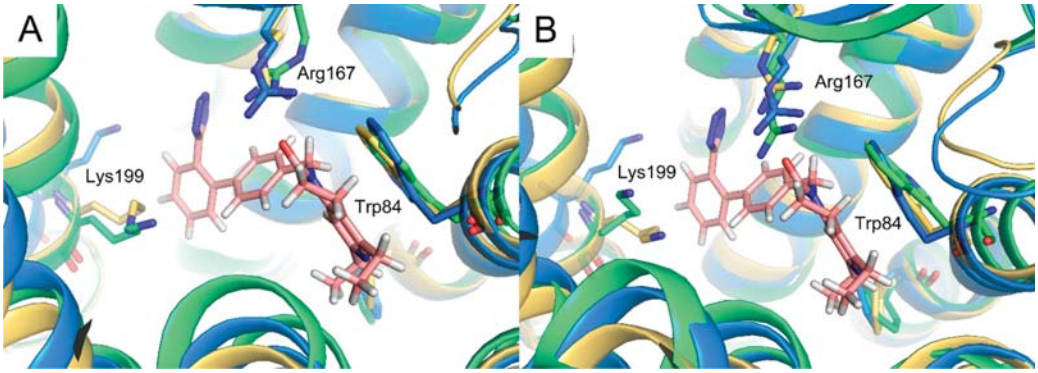


Figure 6. (A) Structural alignment of the crucial residues (Trp84, Arg167, and Lys199) in the binding pockets of 4YAY (marine), Com3md4_IFD (yellow), and CXCR4md3_IFD (green). (B) Structural alignment of the crucial residues (Trp84, Arg167, and Lys199) in the binding pockets of 4YAY (marine), Com3md4_IFD_50ns (yellow), and CXCR4md3_IFD_50ns (green).

distributions of the different conformational populations are approximately consistent with the RMSD matrix for the binding pocket (Figure 5A). The six representative structures prepared by the *Protein Prepare Wizard* in Schrodinger were used for molecular docking. The docking results are illustrated in Tables 1 and 2 and Figures 2, S2, S3, and S4. The accuracy of reproducing the experimental-binding poses of the original ligands is not improved significantly, but the screening power becomes much better. According to the structural alignments of the binding site residues between 4YAY and the six representative structures, Com3md4 is more similar to 4YAY than the others, indicated by a relatively lower RMSD of the binding site residues (1.57 Å). Among the six conformations, Com3md4 can roughly recognize the near-native-binding poses for ZD7155 (RMSD = 3.36 Å for Glide SP, 3.47 Å for Glide XP, and 3.34 Å for Autodock Vina) and olmesartan (RMSD = 9.32 Å for Glide SP, 3.24 Å for Glide XP, and 3.54 Å for Autodock Vina), as shown in Table 2. As mentioned above, Arg167, Trp84, and Lys199 are the most important residues for ligand binding.[33] The Arg167 and Trp84 residues in Com3md4 and 4YAY overlap well (RMSD = 0.53 Å for Arg167 and RMSD = 0.88 Å for Trp84), but the Lys199 residues overlap poorly (RMSD = 4.33 Å). As shown in Figures 6 and S7, the MD simulations do not improve the

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

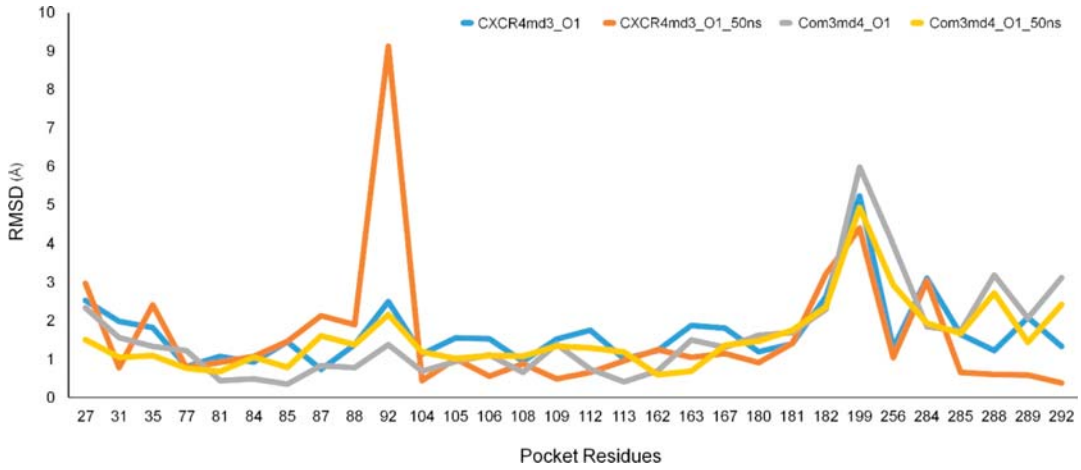


Figure 7. RMSDs of the binding site residues in CXCR4md3_IFD, CXCR4md3_IFD_50ns, Com3md4_IFD, and Com3md4_IFD_50ns, relative to those in 4YAY.

inward shift of TM5. Summing up the above analyses, Com3md4 is structurally more similar to 4YAY than the other conformations, but it still does not show good screening power (AUC = 0.724 and $p = 2.16 \times 10^{-13}$ for Glide SP, AUC = 0.710 and $p = 2.14 \times 10^{-12}$ for Autodock Vina).

As CXCR4Model outperformed the other two homology models constructed based on a single template in the docking assessment, it was also submitted to MD simulations. Then, based on the MD trajectory, the conformational cluster population plot was generated (Figure 5) and compared with the RMSD matrix to verify the validity of conformational clustering. Four representative conformations were identified, and the prepared structures were used for the docking assessment. The docking results are shown in Tables 1 and 2 and Figures 2, S2, S3, and S4. It is interesting to observe that CXCR4md3 exhibits acceptable capability to discriminate the known antagonists from decoys (AUC = 0.870 and $p = 1.28 \times 10^{-29}$ for Glide SP, AUC = 0.746 and $p = 4.49 \times 10^{-19}$ for Autodock Vina). According to the enrichment ratios in Figure 3, CXCR4md3 performed slightly worse than the crystal structure. Comparison of the binding site residues shown in Figure S illustrates that most of the binding site residues of CXCR4md3 have better structural alignment with those of the crystal structure and those of the other three conformations extracted from the MD trajectory. In the docking pose assessment, the binding pose of ZD7155 predicted by Glide XP (RMSD = 2.08 Å) and those of olmesartan predicted by Glide SP (RMSD = 2.11 Å) and Autodock Via (RMSD = 2.10 Å) in CXCR4md3 are quite reasonable. The ligand-binding pocket of CXCR4md3 extracted from the MD trajectory is more rational compared with that of the original homology model (CXCR4Model), and therefore, the docking calculations based on CXCR4md3 illustrate better screening power and higher enrichment than those based on CXCR4Model, suggesting that the combination of MD simulations and conformational clustering is an effective strategy to refine a homology model and identify the best conformation for docking-based VS.

Refinement of Ligand-Binding Pocket by Induced-Fit Docking. The RMSDs of the binding poses predicted by IFD relative to the experimental poses of ZD7155 and olmesartan are listed in Table S3. For the crystal structures (4YAY and 4ZUD) and the other two refined structures (4YAY50nsMD and 4ZUD50nsMD), the RMSDs of the binding poses are around 1−2 Å, suggesting that the experimental ligand-binding poses are well-predicted by IFD. However, based on the initial homology models, molecular docking could not recognize near-native conformations (RMSDs = 3.252−13.021 Å). As for the multiple conformations retrieved from the MD simulations and conformational clustering, the RMSDs of most of the predicted poses are higher than 2 Å, and only several poses exhibit relatively low RMSDs. For Com3md4 (the receptor optimized by IFD was named as Com3md4_IFD), the first binding pose of olmesartan predicted by IFD is approximately the desired conformation (RMSD = 2.101 Å).

Two complexes predicted by IFD, including olmesartan−Com3md4_IFD and olmesartan−CXCR4md3_IFD, were inserted into the membrane environment and then submitted to 50 ns MD simulations. The RMSDs of the binding site residues during the simulations are shown in Figure 4. The refined structures (Com3md4_IFD_50ns and CXCR4md3_IFD_50ns) extracted from the MD trajectories were prepared and used in the docking calculations. As shown in Tables 1 and 2 and Figures 2, S2, S3, and S4, the refined receptor CXCR4md3_IFD_50ns can not only reproduce the correct binding poses for ZD7155 and olmesartan but also shows extraordinary screening powers (AUC = 0 0.890 and $p = 2.19 \times 10^{-31}$ for Glide SP, AUC = 802 and $p = 7.45 \times 10^{-25}$ for Autodock Vina). For Com3md4_IFD_50ns, the binding poses predicted by Glide improve slightly compared with those for Com3md4. Overall, CXCR4md3_IFD_50ns outperformed Com3md4_IFD_50ns greatly on the basis of the docking assessments.

To explore the binding site similarity between the crystal structure and the four models predicted by IFD, the RMSDs of the binding site residues were calculated and are shown in Figure 7. Obviously, CXCR4md3_IFD_50ns aligns well with 4YAY for most binding site residues except Tyr92 (RMSD = 9.13 Å). Tyr92 lies at the edge of the pocket, and it may be not quite important to ligand binding. CXCR4md3_IFD_50ns overlaps well with 4YAY at Ala104, Ser105, Ala106, Val108, Ser109, Leu112, Cys180, Hid256, Pro285, Ile288, Cys289, and Tyr292 (RMSD < 1 Å), while CXCR4md3_IFD and Com3md4_IFD_50ns cannot overlap well with 4YAY at several other residues close to the ligand (RMSD > 2 Å for His256, Ile288, and Tyr292 and RMSD > 1.5 Å for Met284

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

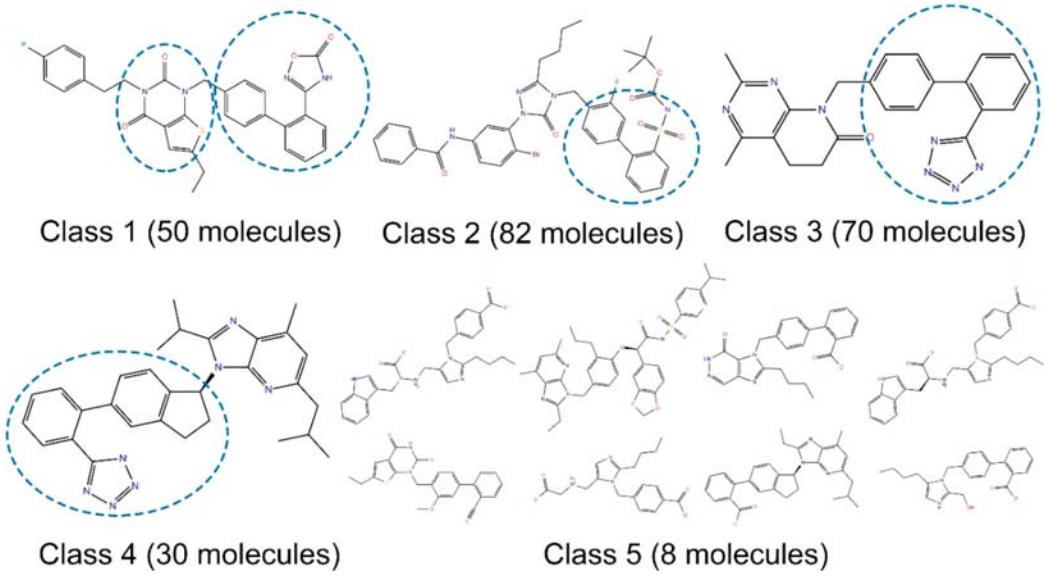


Figure 8. Classification of 240 known ligands based on the structural features of the scaffolds highlighted by blue dashed ovals. The representative ligands for classes 1, 2, 3, and 4 are listed, and the molecules that do not contain any scaffold for classes 1, 2, 3, and 4 are categorized into class 5.

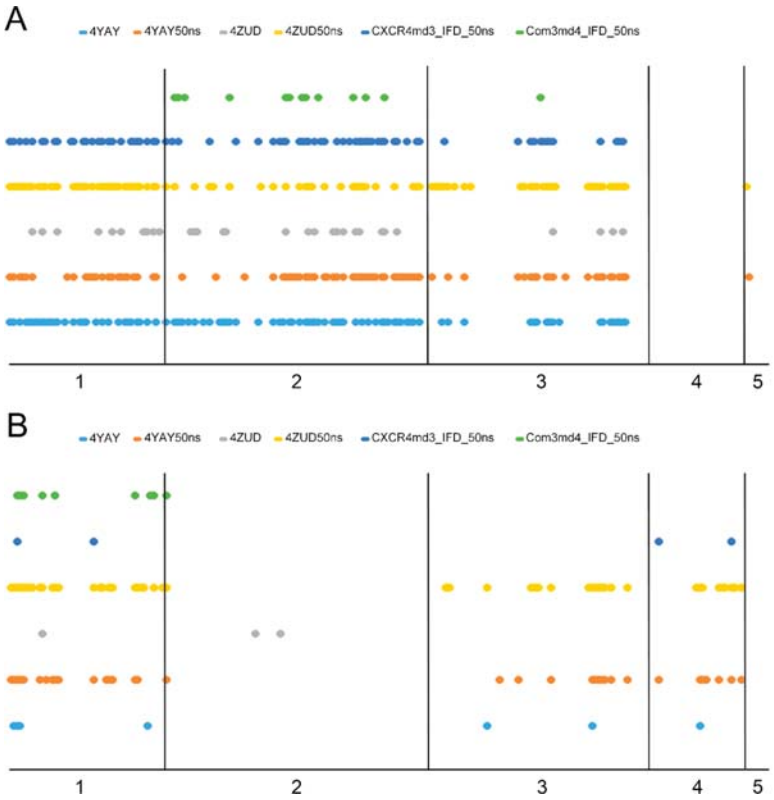


Figure 9. Distributions of the known antagonists in the top 200 hits predicted by (A) Glide SP docking and (B) Vina docking (B).

and Pro285). Therefore, it is not surprising to understand why CXCR4md3_IFD_50ns outperformed the other models. Even compared with 4YAY, 4YAY50ns, and 4ZUD50ns, CXCR4md3_IFD_50ns shows a higher enrichment ratio (Figure 3). In conclusion, the refined homology model, CXCR4md3_IFD_50ns, has high binding site similarity toward 4YAY and similar performance to the optimized crystal structure according to the docking assessments, suggesting that it is possible to obtain reliable GPCR models for SBDD through the combination of homology modeling, MD simulations, conformational clustering, and induced-fit docking.

Structural Preference of Top Hits Identified by Different Structures. To investigate the structural preference of the top hits selected by different structures, the structural features of the known antagonists in the top 200 hits identified

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

by the molecular docking based on the crystal structure, Com3md4_IFD_50ns and CXCR4md3_IFD_50ns, were analyzed.

First, the 240 ligands in the validation data set were roughly classified into five classes based on the structural features (Figure 8). Then, the known ligands found in the top 200 hits for each structure were categorized (Figure 9). Analyses of the results show that the Glide docking based on Com3md4_IFD_50ns tends to find the class 2 ligands, while those based on the crystal structure and CXCR4md3_IFD_50ns tend to find more diverse ligands (classes 1, 2, and 3), suggesting that the binding site of CXCR4md3_IFD_50ns can accommodate ligands with more structural diversity. To investigate whether structural preference is related to docking algorithms, the known antagonists in the top 200 hits selected by Glide were compared with those selected by Autodock Vina. Ignoring enrichment ratio, the known ligands recognized by different protein structures do not show a large difference in structural preference. However, it is apparent that Glide prefers to identify classes 1, 2, and 3 antagonists, and Autodock Vina prefers to identify classes 1, 3, and 4 antagonists. The reason for this difference of the class preference probably lies in the difference between the two docking algorithms. Generally, in our study, the performance of Autodock Vina is worse than that of Glide, but Autodock Vina exhibits obvious preference for class 4 ligands, which Glide does not find in the top hits. Therefore, to identify more diverse ligands of AT1R, an integrated VS by combining Glide and Autodock Vina may be a feasible strategy.

CONCLUSIONS

In this study, the homology models of AT1R were built and refined by MD simulations, conformational clustering, and ligand-fit docking. The initial homology models and the refined models were then evaluated by the prediction accuracy of ligand-binding poses and the capability to distinguish known antagonists from decoys. The structures of the homology models are generally consistent with the crystal structures, but without any refinement, the crystal structure significantly outperformed the homology models in docking power and screening power assessments. MD simulations could improve the docking results for the crystal structures and homology models. Through the optimization by combining MD simulations, conformational clustering, and induced-fit docking, the refined homology model even shows a similar performance to the optimized crystal structure according to the docking assessment, suggesting that it is possible to establish a reliable class A GPCR homology model for VS through refinement by integrating multiple molecular modeling techniques. However, it should be noted that the crystal structure of AT1R, as well as all available class A GPCR crystal structures with high sequence identity to AT1R (CCR5, CCR9, CXCR4, δ-receptor, κ-receptor, and NOP receptor),[40] adopt inactive conformations, and therefore, this study can only provide guidelines to construct class A GPCR models at the inactive states.

In our study, the homology modeling based on multiple templates is not proven to be superior. Although Com3Model exhibits the best docking power among the homology models prior to any refinement, the refined Com3Model models show worse predictions than several refined models. On the other hand, the sequence identity between template and target is more likely to be a more crucial factor that determines the reliability of a homology model. For example, in this study, CXCR4 shares the highest transmembrane sequence identity with AT1R among the three templates, and the refined model, CXCR4md3_IFD_50ns, possesses the best docking power among all the models.

METHODS

Homology Modeling. According to the previous study, chemokine receptors and opioid receptors are the closest homologies to AT1R.[33] The homology models of AT1R at the inactive state were built using the ligand-bound crystal structures of three GPCRs at the inactive states, including CXCR4 chemokine receptor (PDB entry: 3ODU), CCR5 chemokine receptor (PDB entry: 4MBS) and δ-opioid receptor (PDB entry: 4N6H) as the templates.[5,41,42] The profiles of the target and templates were obtained through the BLAST sequence alignment to the nonredundant sequence database, and the top 500 sequences were used as the profile.[43] The profiles of the target and templates were aligned using the MUSCLE program (version 3.8).[44] The AT1R-template sequence identities for the transmembrane domains are 37.7% for CXCR4 chemokine receptor, 33.7% for CCR5 chemokine receptor, and 32.7% for δ-opioid receptor, respectively. Then, for each template, five homology models were built using the *modeler* program[45] implemented in the *Create Homology Models* module in Discovery Studio 3.1,[46] and the best scored model was selected. The obtained models were named as CCR5Model, CXCR4Model, and DeltaORModel, respectively. Furthermore, the multiple template strategy was also employed, in which the three templates were all regarded as the reference sequences to generate five homology models, and the best scored model was selected and named as Com3Model. The qualities of all the selected models were assessed by Profile-3D[47,48] and PROCHECK[49,50] in Discovery Studio 3.1.

System Preparation for MD Simulations. The two crystal structures (4YAY and 4ZUD) and four homology models were used as the starting structures for MD simulations. The missing loop regions in the crystal structures were modeled and refined by *Prime*[29,51] in Schrodinger 2017.[52] Each structure was embedded into the periodic box of pre-equilibrated 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC) molecules, which are the main component for the membrane in vascular smooth muscle cells, where AT1R is widely distributed.[53] Each AT1R structure, along with DPPC molecules, TIP3P waters, and KCl at a concentration of 0.15 M were assembled together using CHARMM-GUI.[54−58] Subsequently, each prepared system was translated by the *charmmlipid2amber.py* protocol in AMBER16.[59] For the two original ligands (ZD7155 and olmesartan) in the crystal structures, their partial atom charges were calculated with the restrained electrostatic potential (RESP)[60,61] protocol by fitting the electrostatic potentials computed by the HF/6-31G level in Gaussian09.[62] The force field parameters for the ligands were created by using the *Antechamber* module in AMBER16.[63] The ff14SB force field,[64] General Amber Force field (GAFF),[65] and lipid14 force field[66] were used for the proteins, ligands, and membrane, respectively. Finally, the topology and coordinate files of each system were prepared for the simulations.

MD Simulations and Conformation Clustering. The minimization, heating, and equilibration stages were executed with the *pmemd* program in AMBER16.[59] First, a restraint force of 5.0 kcal·mol^{-1}·Å^{-2} was exerted on the protein, ligand, and membrane atoms, and the solvent and ion molecules were optimized by 2500 cycles of steepest descent and 2500 cycles of conjugate gradient minimizations. Second, the protein backbone and ligand atoms were restrained by an elastic force of 4.0 kcal·mol^{-1}·Å^{-2}, and the side chain and DPPC molecules were optimized by 2500 cycles of steepest descent and 2500 cycles of conjugate gradient minimizations. Finally, all the atoms were optimized by 2500 cycles of steepest descent and 2500 cycles of conjugate gradient minimizations without any restraint. Each optimized system was heated from 0 to 300 K over a period of 0.1 ns with the protein, ligand, and membrane atoms restrained by an elastic force of 5.0 kcal·mol^{-1}·Å^{-2}. Then, each system was equilibrated

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

over 0.45 ns in the NPT ensemble (T = 300 K and P = 1 bar). Finally, 50−70 ns MD simulations were conducted in the NPT ensemble. The temperature was controlled by the Langevin temperature equilibration scheme with a collision frequency of 2.0 ps^{-1}. The particle mesh Ewald (PME) algorithm was used to handle the long-range electrostatic interactions under the periodic boundary condition, and a cutoff of 8 Å was used for the real-space interactions.[67] The SHAKE algorithm was used to constrain all covalent bonds involving hydrogen atoms, and the time step was set to 2 fs.[68]

The *cpptraj* module in Amber16 was used to analyze the MD trajectories.[59] The root-mean-square deviation (RMSD) of backbone atoms was calculated to monitor the changes of the whole structure during the MD simulations. The *RMS2D* program was also used to generate the RMSD map for the binding site region. Subsequently, the clustering function based on the Hieragglo algorithm implemented in *cpptraj* was used to cluster the binding site region based on the snapshots extracted from the MD trajectories. The most representative conformation for each cluster was exported and named as CXCR4md0-3 and Com3md0-5. The cluster population plot for each MD trajectory was generated based on the clustering results. The atom coordinates of all the representative conformations were aligned to the crystal structure (4YAY), and the binding site of each receptor was defined as the residues within 5 Å from the original ligand ZD7155.

Construction of Validation Data Set. A validation data set was prepared and applied to assess the discrimination capability to distinguish known antagonists from decoys for the crystal structures and homology models. The known antagonists of AT1R were retrieved from the BindingDB database,[69] and those with weak biological activity (IC_{50} or K_i > 1 μM) were eliminated. The inactive compounds (decoys) were generated using DUD-E.[70] The final data set contains 240 known antagonists with 3D coordinates and 14 450 decoys without 3D coordinates. All active and decoys were prepared using the *Ligprep* module in Schrodinger by adding 3D coordinates for decoys, generating possible ionic states at pH = 7.0 ± 2.0 and enumerating all possible stereoisomers for each molecule. Finally, the data set contains 331 actives and 26 916 decoys.

Molecular Docking. Two popular docking algorithms, including Glide[71,72] and Autodock Vina,[73] were used to assess the prediction accuracy of ligand-binding poses and discrimination capability of docking-based VS for the crystal structures and homology models. First, to evaluate the docking reliability for reproducing the experimental-binding modes of ligands, the original antagonists, ZD7155 and olmesartan, were extracted from the crystal structures (4YAY and 4ZUD) and then redocked into the crystal structures and homology models. The RMSD values of the heavy atoms between the docked-binding pose and the experimental-binding pose was used as a key criterion to evaluate the prediction accuracy of ligand-binding poses. If the RMSD is less than 2.0 Å, the docking pose was considered as near-native. Then, to evaluate the screening power to distinguish the known antagonists from decoys, all the molecules in the validation data set were docked into each AT1R structure and then ranked by the docking scores. The area under curve (AUC) of a receiver operating characteristic (ROC) curve was used to measure the overall performance of docking enrichment, and it has a value ranging from 0 for a complete failure to 1 for a perfect enrichment. Moreover, the discrimination capability was evaluated by the p-value of the difference between the means of the two distributions of the docking scores for the known inhibitors and decoys in the data set given by the student's t-test with a 95% confidence interval.

Glide Docking. For the Glide calculations, each protein structure was prepared by the *Protein Prepare Wizard*[74] module in Schrodinger 2017 by adding hydrogens and disulfide bridges, removing crystallographic waters and ions, fixing bond orders, assigning partial charges with the OPLS force field, and minimizing the structure until the RMSD reached a maximum value of 0.3 Å. The binding box with the size of 10 × 10 × 10 Å centered on the cocrystallized ligand or ZD7155 copied from the crystal structure of 4YAY was generated by using the *Receptor Grid Generation* component of *Glide*. For the docking calculations of ZD7155 and olmesartan, the standard precision (SP) and extra precision (XP) scoring functions of *Glide* were used. For the docking calculations of the molecules in the validation data set, only the SP scoring function was used, because the XP scoring function costs much more computational resource but does not improve docking performance significantly.[37]

Autodock Vina Docking. Proteins and ligands were preprocessed by AutoDockTools,[73] including format conversion, addition of hydrogens, assignment of Gasteiger charges, and cleanup of unwanted elements. The grid box was centered on the coordinates of ZD7155 in 4YAY, and the size of the search space was set to 22 × 22 × 22 Å. The docking scores were calculated by the default scoring function, and the best docking score for each molecule was saved.

Binding Site Refinement by Induced-Fit Docking. The induced-fit docking (IFD) algorithm was employed to refine the side chains of the ligand-binding site residues for all the models.[75−77] The original ligands, ZD7155 and olmesartan, were docked into the protein structures by using the *induced-fit docking* module in Schrodinger. First, each ligand was docked into the active site of the receptor by the rigid receptor docking in *Glide*,[78] and the top 20 poses of each ligands were retained. Then, the side chains of the residues within 5 Å of each ligand were refined by *Prime* in Schrödinger 9.0. Finally, each pose was redocked into the relaxed proteins and evaluated by the *Glide* Extra Precision (XP) scoring. All the predicted complexes were recorded. The obtained best two conformations were named as CXCR4md3_IFD and Com3md4_IFD, and the corresponding structures obtained from the further MD refinement were named as CXCR4md3_IFD_50ns and Com3md4_IFD_50ns.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acschemneuro.8b00489.

Table S1. Profile-3D scores for the theoretical models; Table S2. PROCHECK data for the theoretical models; Table S3. The RMSDs of the top 10 binding poses predicted by IFD relative to the experimental poses of ZD7155 and olmesartan (Å); Figure S1. RMSDs of the binding site residues of four homology models relative to the crystal structure (4YAY); Figure S2. The ROC curves based on the Autodock Vina docking scores of the known antagonists and decoys in the validation data set; Figure S3. Distributions of the Glide SP docking scores of the known antagonists and decoys; Figure S4. Distributions of the Autodock Vina docking scores of the known antagonists and decoys (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: tingjunhou@zju.edu.cn; Phone: +86-517-8820-8412.

ORCID

Tailong Lei: 0000-0003-2067-1787
Feng Zhu: 0000-0001-8069-0053
Tingjun Hou: 0000-0001-7227-2580

Author Contributions

T.H. designed the research. H.C., W.F., and Z.W. conducted the simulations. H.C., W.F., X.W., T.L., F.Z., D.L., S.C., and L.X. analyzed the data. H.C. drafted the manuscript. T.H. edited the manuscript. All authors read and approved the final manuscript.

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

Funding

This study was supported by the National Key R&D Program of China (2016YFA0501701, 2016YFB0201700) and the National Science Foundation of China (21575128, 81773632).

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Rosenbaum, D. M., Rasmussen, S. G. F., and Kobilka, B. K. (2009) The structure and function of G-protein-coupled receptors. *Nature 459*, 356−363.

(2) Ahren, B. (2009) Islet G protein-coupled receptors as potential targets for treatment of type 2 diabetes. *Nat. Rev. Drug Discovery 8*, 369−385.

(3) Conn, P. J., Christopoulos, A., and Lindsley, C. W. (2009) Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. *Nat. Rev. Drug Discovery 8*, 41−54.

(4) Penela, P., Murga, C., Ribas, C., Tutor, A. S., Peregrin, S., and Mayor, F., Jr (2006) Mechanisms of regulation of G protein-coupled receptor kinases (GRKs) and cardiovascular disease. *Cardiovasc. Res. 69*, 46−56.

(5) Wu, B. L., Chien, E. Y. T., Mol, C. D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P., Bi, F. C., Hamel, D. J., Kuhn, P., Handel, T. M., Cherezov, V., and Stevens, R. C. (2010) Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science 330*, 1066−1071.

(6) Gloriam, D. E., Fredriksson, R., and Schioth, H. B. (2007) The G protein-coupled receptor subset of the rat genome. *BMC Genomics 8*, 338.

(7) Lappano, R., and Maggiolini, M. (2011) G protein-coupled receptors: novel targets for drug discovery in cancer. *Nat. Rev. Drug Discovery 10*, 47−60.

(8) Munk, C., Isberg, V., Mordalski, S., Harpsøe, K., Rataj, K., Hauser, A. S., Kolb, P., Bojarski, A. J., Vriend, G., and Gloriam, D. E. (2016) GPCRdb: the G protein-coupled receptor database - an introduction. *Br. J. Pharmacol. 173*, 2195−2207.

(9) Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science 289*, 739−745.

(10) Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H. J., Yao, X. J., Weis, W. I., Stevens, R. C., and Kobilka, B. K. (2007) GPCR engineering yields high-resolution structural insights into beta(2)-adrenergic receptor function. *Science 318*, 1266−1273.

(11) Huber, T., Menon, S., and Sakmar, T. P. (2008) Structural basis for ligand binding and specificity in adrenergic receptors: Implications for GPCR-targeted drug discovery. *Biochemistry 47*, 11013−11023.

(12) Loo, J. S. E., Emtage, A. L., Ng, K. W., Yong, A. S. J., and Doughty, S. W. (2018) Assessing GPCR homology models constructed from templates of various transmembrane sequence identities: Binding mode prediction and docking enrichment. *J. Mol. Graphics Modell. 80*, 38−47.

(13) Baker, D., and Sali, A. (2001) Protein structure prediction and structural genomics. *Science 294*, 93−96.

(14) Forrest, L. R., Tang, C. L., and Honig, B. (2006) On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins. *Biophys. J. 91*, 508−517.

(15) Mobarec, J. C., Sanchez, R., and Filizola, M. (2009) Modern Homology Modeling of G-Protein Coupled Receptors: Which Structural Template to Use? *J. Med. Chem. 52*, 5207−5216.

(16) Carlsson, J., Coleman, R. G., Setola, V., Irwin, J. J., Fan, H., Schlessinger, A., Sali, A., Roth, B. L., and Shoichet, B. K. (2011) Ligand discovery from a dopamine D-3 receptor homology model and crystal structure. *Nat. Chem. Biol. 7*, 769−778.

(17) Mysinger, M. M., Weiss, D. R., Ziarek, J. J., Gravel, S., Doak, A. K., Karpiak, J., Heveker, N., Shoichet, B. K., and Volkman, B. F. (2012) Structure-based ligand discovery for the protein-protein interface of chemokine receptor CXCR4. *Proc. Natl. Acad. Sci. U. S. A. 109*, 5517−5522.

(18) McRobb, F. M., Capuano, B., Crosby, I. T., Chalmers, D. K., and Yuriev, E. (2010) Homology Modeling and Docking Evaluation of Aminergic G Protein-Coupled Receptors. *J. Chem. Inf. Model. 50*, 626−637.

(19) Beuming, T., and Sherman, W. (2012) Current Assessment of Docking into GPCR Crystal Structures and Homology Models: Successes, Challenges, and Guidelines. *J. Chem. Inf. Model. 52*, 3263−3277.

(20) Loo, J. S. E., Emtage, A. L., Ng, K. W., Yong, A. S. J., and Doughty, S. W. (2018) Assessing GPCR homology models constructed from templates of various transmembrane sequence identities: Binding mode prediction and docking enrichment. *J. Mol. Graphics Modell. 80*, 38−47.

(21) Zhang, G., Wang, K., Li, X. D., Zhang, D. L., and Xu, F. (2016) Discovery of novel antagonists of human neurotensin receptor 1 on the basis of ligand and protein structure. *Biomed. Pharmacother. 84*, 147−157.

(22) Bresso, E., Togawa, R., Hammond-Kosack, K., Urban, M., Maigret, B., and Martins, N. F. (2016) GPCRs from fusarium graminearum detection, modeling and virtual screening - the search for new routes to control head blight disease. *BMC Bioinf. 17*, 463.

(23) Singh, K. D., and Muthusamy, K. (2013) Molecular modeling, quantum polarized ligand docking and structure-based 3D-QSAR analysis of the imidazole series as dual AT(1) and ETA receptor antagonists. *Acta Pharmacol. Sin. 34*, 1592−1606.

(24) Yap, B. K., Buckle, M. J. C., and Doughty, S. W. (2012) Homology modeling of the human 5-HT1A, 5-HT2A, D1, and D2 receptors: model refinement with molecular dynamics simulations and docking evaluation. *J. Mol. Model. 18*, 3639−3655.

(25) Antunes, D. A., Devaurs, D., and Kavraki, L. E. (2015) Understanding the challenges of protein flexibility in drug design. *Expert Opin. Drug Discovery 10*, 1301−1313.

(26) Kong, X., Pan, P., Li, D., Tian, S., Li, Y., and Hou, T. (2015) Importance of protein flexibility in ranking inhibitor affinities: modeling the binding mechanisms of piperidine carboxamides as Type I1/2 ALK inhibitors. *Phys. Chem. Chem. Phys. 17*, 6098−6113.

(27) Kong, X., Sun, H., Pan, P., Zhu, F., Chang, S., Xu, L., Li, Y., and Hou, T. (2018) Importance of protein flexibility in molecular recognition: a case study on Type-I1/2 inhibitors of ALK. *Phys. Chem. Chem. Phys. 20*, 4851−4863.

(28) Tian, S., Sun, H., Pan, P., Li, D., Zhen, X., Li, Y., and Hou, T. (2014) Assessing an Ensemble Docking-Based Virtual Screening Strategy for Kinase Targets by Considering Protein Flexibility. *J. Chem. Inf. Model. 54*, 2664−2679.

(29) Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J. F., Honig, B., Shaw, D. E., and Friesner, R. A. (2004) A hierarchical approach to all-atom protein loop prediction. *Proteins: Struct., Funct., Genet. 55*, 351−367.

(30) Savoia, C., Burger, D., Nishigaki, N., Montezano, A., and Touyz, R. M. (2011) Angiotensin II and the vascular phenotype in hypertension. *Expert Rev. Mol. Med. 13*, e11.

(31) Hunyady, L., and Catt, K. J. (2006) Pleiotropic AT(1) receptor signaling pathways mediating physiological and pathogenic actions of angiotensin II. *Mol. Endocrinol. 20*, 953−970.

(32) Sugawara, A., Takeuchi, K., Uruno, A., Ikeda, Y., Arima, S., Kudo, M., Sato, K., Taniyama, Y., and Ito, S. (2001) Transcriptional suppression of type 1 angiotensin II receptor gene expression by peroxisome proliferator-activated receptor-gamma in vascular smooth muscle cells. *Endocrinology 142*, 3125−3134.

(33) Zhang, H. T., Unal, H., Gati, C., Han, G. W., Liu, W., Zatsepin, N. A., James, D., Wang, D. J., Nelson, G., Weierstall, U., Sawaya, M. R., Xu, Q. P., Messerschmidt, M., Williams, G. J., Boutet, S., Yefanov, O. M., White, T. A., Wang, C., Ishchenko, A., Tirupula, K. C., Desnoyer, R., Coe, J., Conrad, C. E., Fromme, P., Stevens, R. C., Katritch, V., Karnik, S. S., and Cherezov, V. (2015) Structure of the

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

Angiotensin Receptor Revealed by Serial Femtosecond Crystallography. *Cell 161*, 833−844.
(34) Kritsi, E., Matsoukas, M. T., Potamitis, C., Karageorgos, V., Detsi, A., Magafa, V., Liapakis, G., Mavromoustakos, T., and Zoumpoulakis, P. (2016) Exploring new scaffolds for angiotensin II receptor antagonism. *Bioorg. Med. Chem. 24*, 4444−4451.
(35) Zhang, J., Liu, X., Wang, S. Q., Liu, G. Y., Xu, W. R., Cheng, X. C., and Wang, R. L. (2017) Identification of dual ligands targeting angiotensin II type 1 receptor and peroxisome proliferator-activated receptor-gamma by core hopping of telmisartan. *J. Biomol. Struct. Dyn. 35*, 2665−2680.
(36) Wu, Q., Kang, H., Wang, H. H., Gao, J., Zhu, R. X., and Kang, T. G. (2012) AT1R-based Virtual Screening Model for Bioactive Components from Traditional Chinese Medicines and Its Mechanism Study. *Huaxue Xuebao 70*, 796−802.
(37) Wang, Z., Sun, H., Yao, X., Li, D., Xu, L., Li, Y., Tian, S., and Hou, T. (2016) Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power. *Phys. Chem. Chem. Phys. 18*, 12964−12975.
(38) Wang, Z., Kang, Y., Li, D., Sun, H., Dong, X., Yao, X., Xu, L., Chang, S., Li, Y., and Hou, T. (2018) Benchmark Study Based on 2P2IDB to Gain Insights into the Discovery of Small-Molecule PPI Inhibitors. *J. Phys. Chem. B 122*, 2544−2555.
(39) Wang, R. X., Lu, Y. P., and Wang, S. M. (2003) Comparative evaluation of 11 scoring functions for molecular docking. *J. Med. Chem. 46*, 2287−2303.
(40) Isberg, V., Mordalski, S., Munk, C., Rataj, K., Harpsøe, K., Hauser, A. S., Vroling, B., Bojarski, A. J., Vriend, G., and Gloriam, D. E. (2017) GPCRdb: an information system for G protein-coupled receptors (vol 44, pg D356, 2016). *Nucleic Acids Res. 45*, 2936.
(41) Wu, H. X., Wacker, D., Mileni, M., Katritch, V., Han, G. W., Vardy, E., Liu, W., Thompson, A. A., Huang, X. P., Carroll, F. I., Mascarella, S. W., Westkaemper, R. B., Mosier, P. D., Roth, B. L., Cherezov, V., and Stevens, R. C. (2012) Structure of the human kappa-opioid receptor in complex with JDTic. *Nature 485*, 327−U369.
(42) Tan, Q. X., Zhu, Y., Li, J., Chen, Z. X., Han, G. W., Kufareva, I., Li, T. T., Ma, L. M., Fenalti, G., Li, J., Zhang, W. R., Xie, X., Yang, H. Y., Jiang, H. L., Cherezov, V., Liu, H., Stevens, R. C., Zhao, Q., and Wu, B. L. (2013) Structure of the CCR5 Chemokine Receptor-HIV Entry Inhibitor Maraviroc Complex. *Science 341*, 1387−1390.
(43) Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids. Res. 25*, 3389−3402.
(44) Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids. Res. 32*, 1792−1797.
(45) Webb, B., and Sali, A. (2014) Comparative Protein Structure Modeling Using MODELLER. *Curr. Prot. Bioinformatics 47*, 5.6.1−5.6.32.
(46) (2010) *Discovery Studio, version 3.1*, Accelrys Software Inc., San Diego, CA; http://www.accelrys.com.
(47) Bowie, J. U., Luthy, R., and Eisenberg, D. (1991) A Method to Identify Protein Sequences That Fold into a Known 3-Dimensional Structure. *Science 253*, 164−170.
(48) Luthy, R., Bowie, J. U., and Eisenberg, D. (1992) Assessment of Protein Models with 3-Dimensional Profiles. *Nature 356*, 83−85.
(49) Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr. 26*, 283−291.
(50) Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR 8*, 477−486.
(51) Jacobson, M. P., Friesner, R. A., Xiang, Z., and Honig, B. (2002) On the Role of Crystal Packing Forces in Determining Protein Sidechain Conformations. *J. Mol. Biol. 320*, 597−608.
(52) (2017) *Schrödinger, version 2017*, Schrödinger, LLC, New York, NY; https://www.schrodinger.com.
(53) Maganti, L., Open Source Drug Discovery Consortium, and Ghoshal, N. (2014) Probing the structure of Mycobacterium tuberculosis MbtA: model validation using molecular dynamics simulations and docking studies. *J. Biomol. Struct. Dyn. 32*, 273−288.
(54) Jo, S., Kim, T., Iyer, V. G., and Im, W. (2008) Software news and updates - CHARNIM-GUI: A web-based grraphical user interface for CHARMM. *J. Comput. Chem. 29*, 1859−1865.
(55) 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., and Karplus, M. (2009) CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem. 30*, 1545−1614.
(56) Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., Jeong, J. C., Qi, Y. F., Jo, S., Pande, V. S., Case, D. A., Brooks, C. L., MacKerell, A. D., Klauda, J. B., and Im, W. (2016) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput. 12*, 405−413.
(57) Wu, E. L., Cheng, X., Jo, S., Rui, H., Song, K. C., Davila-Contreras, E. M., Qi, Y. F., Lee, J. M., Monje-Galvan, V., Venable, R. M., Klauda, J. B., and Im, W. (2014) CHARMM-GUI Membrane Builder Toward Realistic Biological Membrane Simulations. *J. Comput. Chem. 35*, 1997−2004.
(58) Jo, S., Lim, J. B., Klauda, J. B., and Im, W. (2009) CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J. 97*, 50−58.
(59) Salomon-Ferrer, R., Case, D. A., and Walker, R. C. (2013) An overview of the Amber biomolecular simulation package. *Wiley Interdiscip. Rev. Comput. Mol. Sci. 3*, 198−210.
(60) Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model. *J. Phys. Chem. 97*, 10269−10280.
(61) Fox, T., and Kollman, P. A. (1998) Application of the RESP methodology in the parametrization of organic solvents. *J. Phys. Chem. B 102*, 8070−8079.
(62) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, O., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. *Gaussian 09*; Gaussian, Inc., Wallingford, CT, 2009.
(63) Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell. 25*, 247−260.
(64) Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput. 11*, 3696−3713.
(65) Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem. 25*, 1157−1174.

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689

(66) Dickson, C. J., Madej, B. D., Skjevik, A. A., Betz, R. M., Teigen, K., Gould, I. R., and Walker, R. C. (2014) Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput. 10*, 865−879.
(67) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: an NLog(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.
(68) Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J. Comput. Phys. 23*, 327−341.
(69) Gilson, M. K., Liu, T., Baitaluk, M., Nicola, G., Hwang, L., and Chong, J. (2016) BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Res. 44*, D1045−D1053.
(70) Mysinger, M. M., Carchia, M., Irwin, J. J., and Shoichet, B. K. (2012) Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem. 55*, 6582−6594.
(71) Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., and Shenkin, P. S. (2004) Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem. 47*, 1739−1749.
(72) Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., and Mainz, D. T. (2006) Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem. 49*, 6177−6196.
(73) Trott, O., and Olson, A. J. (2010) AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem. 31*, 455−461.
(74) Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des. 27*, 221−234.
(75) Farid, R., Day, T., Friesner, R. A., and Pearlstein, R. A. (2006) New insights about HERG blockade obtained from protein modeling, potential energy mapping, and docking studies. *Bioorg. Med. Chem. 14*, 3160−3173.
(76) Sherman, W., Beard, H. S., and Farid, R. (2006) Use of an induced fit receptor structure in virtual screening. *Chem. Biol. Drug Des. 67*, 83−84.
(77) Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A., and Farid, R. (2006) Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem. 49*, 534−553.
(78) Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., and Shenkin, P. S. (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem. 47*, 1739−1749.

DOI: 10.1021/acschemneuro.8b00489
ACS Chem. Neurosci. 2019, 10, 677−689