JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

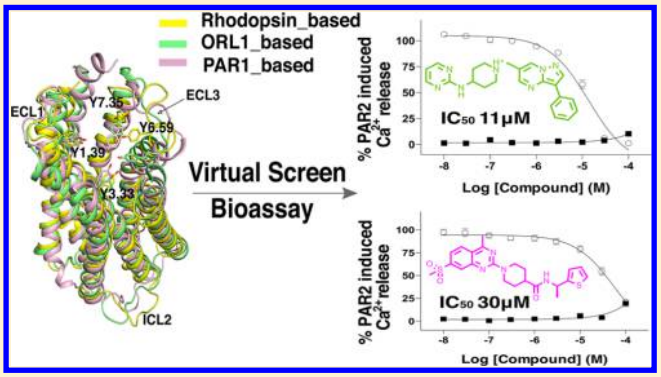
Three Homology Models of PAR2 Derived from Different Templates: Application to Antagonist Discovery

Samuel R. Perry,‡ Weijun Xu,‡ Anna Wirija,‡ Junxian Lim, Mei-Kwan Yau, Martin J. Stoermer, Andrew J. Lucke, and David P. Fairlie*

Division of Chemistry and Structural Biology, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia

Supporting Information

ABSTRACT: Protease activated receptor 2 (PAR2) is an unusual G-protein coupled receptor (GPCR) involved in inflammation and metabolism. It is activated through cleavage of its N-terminus by proteases. The new N-terminus functions as a tethered ligand that folds back and intramolecularly activates PAR2, initiating multiple downstream signaling pathways. The only compounds reported to date to inhibit PAR2 activation are of moderate potency. Three structural models for PAR2 have been constructed based on sequence homology with known crystal structures for bovine rhodopsin, human ORL-1 (also called nociceptin/orphanin FQ receptor), and human PAR1. The three PAR2 model structures were compared and used to predict potential interactions with ligands. Virtual screening for ligands using the Chembridge database, and either ORL-1 or PAR1 derived PAR2 models led to identification of eight new small molecule PAR2 antagonists (IC_{50} 10−100 μM). Notably, the most potent compound **1** (IC_{50} 11 μM) was derived from the less homologous template protein, human ORL-1. The results suggest that virtual screening against multiple homology models of the same GPCR can produce structurally diverse antagonists and that this may be desirable even when some models have less sequence homology with the target protein.



INTRODUCTION

G protein-coupled receptors (GPCRs) have been by far the most successful targets for the development of pharmaceuticals, with 30% of current pharmaceuticals directed at a GPCR on the surface of cells.[1−3] However, finding lead compounds for this target class has had to be achieved through painstakingly iterative structure−activity evaluations of small molecules, usually derived from endogenous activators, hormone fragments, or high-throughput screening of natural products and synthetic libraries.[1,2] An alternative approach to ligand discovery is automated *in silico* docking of virtual chemical libraries into computer-derived structural homology models of GPCRs created by aligning the sequence of a target GPCR with that of a homologous GPCR where there is a known crystal structure.[4−6]

Protease activated receptor 2 (PAR2) is an unusual GPCR where there is no known endogenous ligand. Instead, this GPCR is activated by proteases. Its extracellular N-terminus is cleaved predominately by trypsin-like serine proteases to reveal a "neoepitope" called the tethered ligand (TL), (e.g., SLIGKV, human; SLIGRL, murine), which binds intramolecularly and activates PAR2 at an unknown site.[7,8] PAR2 is expressed in many mammalian tissue types including in the immune, cardiovascular, respiratory, nervous, and musculoskeletal systems. Studies have shown that PAR2 activation is associated with metabolism, inflammation, pain, proliferation, metastasis, and angiogenesis, in many cancers including pancreatic, colon, breast, prostate, and stomach.[9−11] Thus, compounds targeting PAR2 may be suitable for modulating a wide range of diseases.

To date, the development of PAR2 modulators has proceeded predominantly via structural modifications of synthetic peptides that correspond to the protease-cleaved new N-terminal PAR2 sequence leading to a variety of hexapeptide agonists that are called activating peptides (PAR2-APs).[12,13] Although much less potent than trypsin, these PAR2-APs were found to be selective agonists for PAR2 over PAR1.[14] Later, PAR2-APs were further refined by replacing the N-terminal serine with a furoyl group,[15] and the resulting 2f-LIGRLO-NH_2 is a submicromolar PAR2 agonist (EC_{50} ∼ 200 nM, Ca^{2+} HT29 cells).[16] Our group has reported low molecular weight, serum stable, nonpeptidic agonists (e.g., GB110, EC_{50} ∼ 200 nM, Ca^{2+}, HT29 cells)[17] and antagonists (e.g., GB88, IC_{50} 1−10 μM, Ca^{2+})[16] in a variety of human cell lines. GB88 has shown benefits *in vitro* and *in vivo* in rodent models of disease[18−20] and has encouraged subsequent development of PAR2-directed small molecules.

Received: February 17, 2015
Published: May 22, 2015

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181−1191

Many drug discovery programs begin with structure-based virtual screening, which is a process to identify and prioritize molecular binders of a target protein *in silico*, followed by *in vitro* and *in vivo* experimental validation. However, the discovery of new GPCR modulators using this method has usually proven to be difficult as structural characterization for membrane-spanning GPCRs has, until recently, been less successful than for other proteins and enzymes. Creating a homology model of a GPCR, based on a related GPCR for which there is a crystal structure, can be helpful in identifying putative ligands. A decade ago, GPCR modeling was restricted by the availability of a single template (bovine rhodopsin) for structure prediction. However, the number of high-resolution GPCR crystal structures has steadily increased over the last eight years.[21]

Here, we compare three homology structural models for human PAR2, successively constructed over seven years as newer and more homologous high-resolution structures (bovine rhodopsin, ORL-1, PAR1) became available as modeling templates. The model built from different templates was performed by different people and at different times using differing software approaches in our group. The quality of each of the models was assessed using a range of protein structure assessment tools. The models were also assessed by putative binding site analysis for their potential to be useful for docking known PAR2 ligands. The two most reasonable models were then further compared for their efficiency in enriching known PAR2 antagonists (from our lab) using a small set of drug-like decoys from Schrödinger (www.schrödinger.com). These two models were used to screen commercially available Chembridge compounds *in silico*. Hierarchical virtual screening was performed involving database filtering, scoring, and ranking compounds for predicted binding affinity for PAR2, visual inspection, and compound clustering analysis. This paper describes 71 compounds that were purchased and assessed for PAR2 agonist and antagonist activity in a calcium mobilization assay. This work led to the identification of eight nonpeptidic PAR2 antagonists with minimal or no agonist activity, making this the first reported virtual screening study for this target, and it has confirmed the potential of homology models for identifying PAR2 ligands and for studying protein−ligand structures.

METHODS

Focused Sequence Identity and Similarity Comparison. Sequence identity and similarity of the 7-TM region and ligand accessible region between PAR2 and the three templates selected were calculated (Table 1) using GPCRDB tools (http://tools.gpcr.org/). A multiple sequence alignment comparing PAR2 sequence and the three template sequences is shown in Figure S1.

Homology Modeling Based on Bovine Rhodopsin. The human PAR2 protein sequence was obtained from Swiss-Prot Web site (accession number − P55085) and aligned with the sequence of the bovine rhodopsin crystal structure (1U19, TM sequence identity = 21%, resolution = 2.20 Å)[22] using the PAM-250 matrix, which aligns the sequence based on conservation of charged, bulky aliphatic or aromatic residues. Alignment was refined manually by taking the transmembrane (TM) regions and structurally conserved regions into consideration, the study of Bissantz et al. being a reference point for this work.[23] The seven transmembrane regions were identified based on the conservation of residues in each putative transmembrane region. This alignment was used to develop the coordinates of the TM regions using the Modeller program within InsightII Modeling Environment, Release 2000 (Accelrys Inc., San Diego, California, USA, http://www.accelrys.com) using default parameters; a cysteine disulfide bond constraint between C148 on TM3 and C226 on extracellular loop 2 was included. The loop regions were also developed by Modeller using the rhodopsin template. The final model was further refined to remove steric clashes within the molecule by a minor modification of the minimization (the Newton minimization algorithm was not performed because of the large number of atoms in the protein). Additionally, in this minimization protocol the backbone atoms of the TM regions were kept tethered to maintain the helical conformation of the TM regions. This final minimized conformation was used for model validation.

Table 1. Focused Similarity Comparison between PAR2 and Templates

template crystal structure	TM % sequence identity	TM % sequence similarity	ligand accessible site % sequence identity	ligand accessible site % sequence similarity
bovine rhodopsin	21	42	11	39
human ORL-1	29	50	30	50
human PAR1	44	62	48	73

Homology Modeling Based on ORL-1 and PAR1. The human PAR2 protein sequence was obtained from Swiss-Prot (accession number − P55085) and aligned using sALign to the sequence of the crystal structure of the nociceptin/orphanin FQ/ORL-1 receptor (4EA3, TM sequence identity = 29%, resolution = 3.01 Å)[24] and human PAR1 (3VW7, TM sequence identity = 44%, resolution 2.2 Å),[25] respectively. This alignment was then manually inspected and adjusted in Jalview.[26] Model construction was performed using Modeller 9v10;[27] 20 all atom homology models were built for the ORL-1 model, whereas 40 were constructed for the PAR1 based model. The cocrystallized ligand was removed from the ORL-1 crystal structure prior to model construction; however, it was retained in the PAR1 crystal structure to prevent side chains from occupying the putative binding site during model construction. The model with the lowest discrete optimization protein energy (DOPE) score was selected for further refinement. The whole ECL2 in the ORL-1 based model was further refined using the loop refinement option (default setting) in Prime (version 3.1, Schrödinger, LLC, New York, NY, 2012), and the final model that incorporated the optimized ECL2 was minimized in Prime using the truncated-Newton energy minimization (OPLS_2005 force field with restrained helical backbone). Two regions with little homology in ECL2 (V211-I218 and L230-Q233) were further refined in the PAR1 based model using Modeller. Induced-fit docking[28,29] within the Schrödinger Suite v9.3.518 (Schrödinger, LLC, New York, NY, 2012) was performed on the PAR1-derived PAR2 model using the most potent nonpeptidic PAR2 ligand currently available, GB110. The final models were refined using the protein preparation wizard[30] in Schrödinger to optimize hydrogen bond networks and for a restrained energy minimization (OPLS_2005 force field and heavy atom movement <0.5 Å).

Quality Assessment and Validation of Homology Models. The quality of the homology models was assessed

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181−1191

for different physical and chemical parameters. The "structure assessment" module available at Swiss-Model portal (http://swissmodel.expasy.org/) was used for local model quality estimation (QMean score[31]), global model quality estimation (DFIRE energy[32]), and stereochemistry check (Procheck). The ERRAT[33] program was also used for model reliability prediction.

Molecular Docking. Molecular structures of PAR2 agonists, 2f-LIGRLO-NH_2 and GB110, were drawn in ChemDraw Std 13.0, exported as 2D structures, and converted to 3D coordinates using LigPrep in Schrödinger Suite software (version 9.4). LigPrep generated 3D structures of the ligands with that was minimized using the OPLS2005 force field and protonation state corrected to pH 7.4 using Epik. The newly generated 3D structures were saved as mol2 files for docking. GOLD v5.1[34] was used to dock ligands into the putative binding site. In the ORL-1 based model, the active site was centered with a 15 Å radius around residue F155$^{3.32}$, this active site was visually inspected to ensure residues from binding cavity analysis, and hotspot literature surveys were included. Fifty GA runs were performed with flexible residue side chains incorporated for F155, Y156, F243, N304, L307, and Y326. In the PAR1 based model, none of the side chains were required to be flexible to obtain a reasonable pose.

Small Scale Enrichment Study. A small number of PAR2 antagonists (8 in total) developed in our lab were added to the Schrodinger "drug like decoy compound library" and docked into ORL-1 and PAR1 derived PAR2 models, respectively. The ability of the model to discriminate between the actives and decoys was measured by a comparison of their enrichment factors determined at the top 2%, 5%, and 10% of docked poses. This enrichment factor (EF) is an indicator of the potential performance of a model against which a compound library is screened. EF is defined as the ratio of the percentage of active compounds in the subset chosen to the percentage of active compounds in the whole database used for screening.[35] In our study, we define the enrichment factor as

$$\text{Enrichment Factor (EF)} = \frac{\text{Actives}_{x\%\text{sampled}}}{N_{x\%\text{sampled}}} \cdot \frac{N_{\text{total}}}{\text{Actives}_{\text{total}}}$$

where $\text{Actives}_{x\%\text{sampled}}$ is the number of actives found at x% of the database screened, $N_{x\%\text{sampled}}$ is the number of compounds screened at x% of the database, $\text{Actives}_{\text{total}}$ is the number of active compounds in the entire database, and N_{total} is the number of compounds in the entire database.

Virtual Screening. Figure 1 summarizes the schematic workflow for the virtual screening. Briefly, the Chembridge database of purchasable compounds was prepared for virtual screening using the program FILTER[36,37] and OMEGA[38] (OpenEye Scientific Software Inc., Santa Fe, NM) to provide two ligand databases. The first consisted of a library of ~310,000 small lead-like molecules (MW <350), and the second was a library containing ~330,000 drug-like molecules of larger molecular weight (MW 350−600). Each database was separately screened against each model in two rounds, using GOLD v5.1. The initial screen was performed with a single Genetic Algorithm (GA) run at 10% efficiency (standard virtual screen setting). GoldScore was applied to select the best scoring 8000 compounds. The second round consisted of a single GA and 30% efficiency, using GoldScore to select the best scoring 200 compounds. From this list, compounds were visually inspected for predicted interactions with the receptor by taking into account putative key residues during binding pocket analysis and "hotspot" residue searches. Final compound selections were filtered against the pan assay interference compounds substructural filter[39] as implemented in Canvas (version 1.6, Schrödinger, LLC, New York, NY, 2013).

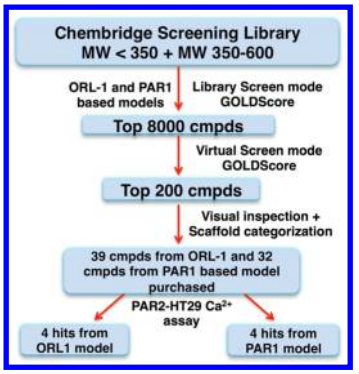


Figure 1. Schematic workflow for virtual screening of putative antagonists against PAR2 homology models.

Purchased Compounds. All compounds were purchased from Chembridge without further purification. According to the vendor, all compounds provided were analyzed by LC-MS and/or NMR spectroscopy to confirm sample identity, and the purity was ≥95%. Stock solutions were prepared by dissolving compounds in DMSO-d_6 and were characterized in house as well by ^{1}H NMR spectroscopy. The exact concentration of each compound was determined by the quantitative NMR integration "PULCON" experiment.[40]

Intracellular Calcium Mobilization Assay. Cell culture was performed according to previously described protocol.[20] Briefly, HT29 colorectal adenocarcinoma cells were cultured in DMEM supplemented with 10% fetal bovine serum, penicillin (100 U/mL), and streptomycin (100 U/mL) at 37 °C with 5% CO_2. Cell dissociation solution was used to dissociate cells from the surface of culture flasks. Cells were grown to 80% confluence. The day before experiments, cells were seeded in 96 well black walled, clear bottom plates at 5×10^4 cells per well. On the day of the experiment, supernatant was removed, and the cells were incubated in dye loading buffer (4 μM Fluo-3, 0.04% Pluronic acid and 1% fetal bovine serum in HBSS) for 1 h at 37 °C. Cells were then washed twice with HBSS. For agonist assays, cells were then left in 50 μL of HBSS and transferred to a FLIPR Tetra (Molecular Device, Sunnyvale CA) for addition of compound. Fluorescence was measured in real time from the bottom of the plate using excitation λ = 480 nm and emission detection at 520 nm. Antagonist assays involved preincubating the cells with the test compounds for 30 min before addition of the agonist (1 μM 2f-LIGRLO-NH_2). Agonist activity was normalized against 100% Ca^{2+} response for 100 μM 2f-LIGRLO-NH_2, while antagonist activity was normalized against 1 μM 2f-LIGRLO-NH_2. The IC_{50} for each antagonist was determined by fitting data to a nonlinear regression curve in GraphPad Prism 6.0, GraphPad Software, San Diego California USA, www.graphpad.com.

RESULTS AND DISCUSSION

Construction of PAR2 Homology Models. GPCRs vary enormously in sequence, structure, the sizes of extracellular and intracellular loops, and the lengths of the extracellular N- and intracellular C-termini. However, the conservation of a canonical transmembrane 7-helix bundle in all rhodopsin

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181−1191

Table 2. Validation of Homology Models[a]

homology models	RMSD vs template	Qmean6 score	DFIRE_energy	% residues in favored and allowed regions	ERRAT score
rhodopsin_based	0.37	0.41 (0.49)	−465.5 (−557.8)	100	81.5
ORL-1_based	0.27	0.46 (0.5)	−445.0 (−430.8)	99.2	83.6
PAR1_based	0.09	0.42 (0.47)	−463.09 (−492.4)	99.6	85.7

[a]Values within parentheses are measured for the original template PDB structures.

family members facilitates effective modeling comparisons for the transmembrane domains of GPCRs. Our first PAR2 homology model was created based on the crystal structure of bovine rhodopsin, being the only GPCR crystal structure available at the time. It has a very low TM sequence identity with PAR2, so we waited for crystal structures to be developed before constructing more reliable PAR2 models. More recently crystal structures for other class A GPCRs have been reported and were more homologous templates, so we constructed new PAR2 structural homology models based ORL-1 (PDB ID: 4EA3) and subsequently released PAR1 (PDB ID: 3VW7). These two receptors also engage native peptide ligands, and so they were expected to provide better PAR2 modeling templates than other class A GPCRs.

Assessment of 3 Homology Models for PAR2. The structural homology models derived from each of the three template-derived constructions of PAR2 were submitted to SwissModel portal and ERRAT for evaluation of model quality and stereochemical properties. The assessment of model quality related parameters, such as Qmean6 score, DFIRE energy, Ramachandran statistics, and ERRAT score, is summarized in Table 2. For comparison, the original crystal structures of the three templates were also submitted for SwissProt evaluation. Qmean6 score is a linear combination of six structural descriptors, and a higher Qmean6 (range between 0 and 1) reflects strong reliability of the model.[31] DFIRE is an all-atom statistical potential term used to assess nonbonded atomic interactions in the protein model.[32] All models produced comparable scores for these two components, relative to the templates from which they were constructed. These results suggested that the constructed homology models were both stereochemically and energetically reliable structures.

A Ramachandran plot for each PAR2 model is shown in Figure S2. All residues in the rhodopsin based homology model were located in the favored and allowed $\varphi-\psi$ regions, whereas 99.2% and 99.6% of the residues from the ORL-1 and PAR1 based homology models were located in the same regions. For ORL-1 and PAR1 based models, a closer check of residues located in the nonfavorable regions revealed that these residues belong to the loop regions. For example, A220 and E232^{ECL2} in the ORL-1 based model are residues in extracellular loop 2 (ECL2), or M181 in the PAR1 based model is in intracellular loop 2 (ICL2). It was noticed that ECL2 of PAR2 has 23 residues, whereas the ORL-1 structure has 20. Hence, the inclusion of three extra residues could produce some alteration in the ECL2 region as evidenced by the disallowed backbone angles of these two residues. Overall, the Ramachandran analysis indicated that the backbone stereochemical properties for all three PAR2 models were close to the protein crystal structures. In addition, an ERRAT score was also calculated for each model. ERRAT is a statistical method used to distinguish errors in model building, leading to more randomized distributions of different atom types in protein residues.[33] The higher ERRAT scores of PAR1 (85.7%) and ORL-1 (83.6%) based models, as compared to the rhodopsin based model (81.5%), suggested that they were closer to crystal structures solved at a moderate resolution between 2.5 and 3.0 Å with an average overall ERRAT quality factor of 91%.

Overall Model Comparison. Figures S3−S5 show cartoon representations of all three homology models. After careful PAR2 sequence alignment to its template structure, all models formed intact canonical 7-helical bundles within the transmembrane (TM) region. The TM backbone RMSD between each PAR2 homology model and their template was 0.37 Å (rhodopsin based), 0.27 Å (ORL-1 based) and 0.09 Å (PAR1 based), respectively. The most prominent difference in the three models was the positioning of ECL2 on top of the receptor. This is not surprising as the three templates used for model construction featured different ECL2 residues in both length and identity. Loop prediction and construction represents a major challenge for the modeling of GPCRs, particularly ECL2 due to its large size and high flexibility. Furthermore, ECL2 is commonly involved in ligand binding and stabilization of the exterior of TM bundle.

Studies on both human PAR1 and PAR2 have implicated residues on ECL2 as being important for receptor activation.[41] In the rhodopsin (PDB ID: 1U19) based homology model (Figure S3), ECL2 adopted a random loop that was deeply buried into the surface of the TM bundle, and it fully blocked entry to the ligand-binding TM cleft in PAR2. In the ORL-1 (PDB ID: 4EA3) based model (Figure S4), ECL2 adopted a solvent-exposed β-hairpin, situated on top of TM3, TM4, and TM5, allowing ligand access to the binding cavity. In stark contrast to the location of ECL2 observed in the ORL-1 based homology model, ECL2 in the PAR1 (PDB ID: 3VW7) based model (Figure S5) almost entirely covered the ligand-binding site. Considering the sequence homology in ECL2 residues present between PAR2 and the three templates, PAR1 (65% sequence identity) offers an apparent advantage over ORL-1 (15% sequence identity) and rhodopsin (19% sequence identity). Published crystal structures of class A GPCRs have shown a variety of defined secondary structures present within ECL2.[42] The human PAR1 ECL2 displayed a β-hairpin, and the residues involved in this secondary structure are completely conserved between PAR1 and PAR2, suggesting that this β-hairpin may also be present in the PAR2 structure. However, the secondary structure of ECL2 in the human PAR1 crystal structure could have been induced by the bound antagonist (vorapaxar), permitting interactions between side chains of ECL2 and bound ligand. The actual positioning of this loop in PAR2 is still unknown. Therefore, the ORL-1 based model with a different ECL2 position provides an alternative clue to the potential location of ECL2 in PAR2.

Relative Position of Key Aromatic TM Residues. Mutagenesis studies from our lab have revealed the importance of some TM aromatic residues for ligand binding to PAR2.[43] Among them are Y82$^{1.39}$, Y156$^{3.33}$, Y326$^{7.35}$, and Y311$^{6.59}$, and their relative locations in each homology model were compared. For a clearer view of these residues in the TM region, the top view of the receptor model is shown for all the

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181−1191

models. Furthermore, the superimposed models are shown in Figure 2. Substantial differences in the position and orientation

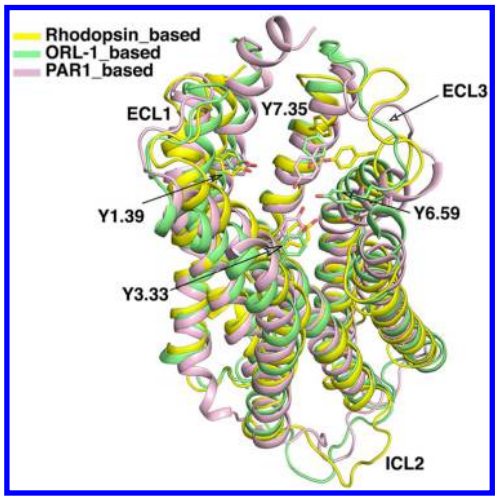


Figure 2. Superimposed homology models of PAR2 derived from bovine rhodopsin, human ORL-1, and human PAR1, with ECL2 removed to clarify comparison of TM residues as highlighted. This figure was generated using PyMOL.[52]

were observed for Y311$^{6.59}$ and Y326$^{7.35}$. The side chain of Y326$^{7.35}$ still projects into the center of the helix bundle; however, it is located one helical turn above in the rhodopsin-based model compared to the other two models. In rhodopsin, P$^{7.35}$ is at the top of TM7. This was in agreement with the finding that ORL-1 and rhodopsin contain several extra residues at the border of ICL3 and TM6, while also containing less amino acids at the border of TM6 and ECL3. Such a difference may lead to dramatic variations in the overall length and folding fidelity of TM6 and TM7, which may explain the differences observed in template dependent models. In contrast, Y82$^{1.39}$ and Y156$^{3.33}$ were modeled in a more similar position with little variation in orientation.

A particular feature of the PAR1 based homology model of PAR2 is the proximity of the side chains of Y156$^{3.33}$ and Y326$^{7.35}$. This close association is stabilized through a side chain hydrogen bond between the two residues and forms the base of a binding pocket in the PAR1 based model. Of note, these two residues also provide a hydrophobic base on which the lipophilic vorapaxar sits in the PAR1 structure. In the PAR2 structures modeled from rhodopsin and ORL-1, these two tyrosine residues were far apart (Tyr side chain hydroxyl group $dist_{OH-OH}$ being 8.6 and 8.1 Å, respectively) from each other.

Putative Ligand Binding Site Comparison. Owing to the different location of ECL2 in the three models, the accessible binding pocket of the models varied accordingly. CASTp[44] calculation showed that the rhodopsin-based model had a restricted binding cavity for ligands. The largest cavity volume for this model was only 275 Å^3, posing a potential challenge for further ligand docking. The binding surface in this model was optimal for docking only when either ECL2 was removed or when applying flexible residues (e.g., D228^{ECL2}) during docking, as most of the docked poses scored negatively, indicating steric clashes between receptor residues and the ligand. This was not ideal for later large-scale virtual screening which was performed with a rigid receptor in order to minimize time and cost. With this caveat, the rhodopsin based homology model will also be reported for our recent study of ligand docking and mutagenesis used in conjunction to characterize PAR2 residues that influence activation by the synthetic agonists 2f-LIGRLO-NH_2 and GB110.[43] The largest pocket in the ORL-1 based model had a much larger volume of 2760 Å^3 (Figure 3). In the PAR1 based model Y156$^{3.33}$ and Y326$^{7.35}$

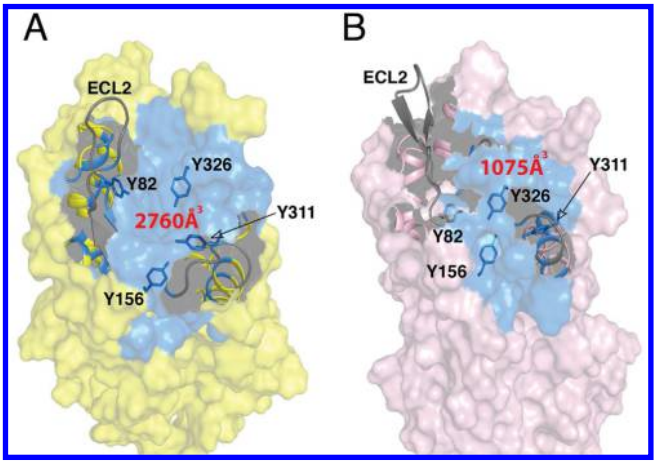


Figure 3. Surface representation of ORL-1 and PAR1 derived homology models of PAR2 with highlighted putative ligand-binding cavity: (A) the ORL-1 based model and (B) the PAR1 based model. ECL2 in each model is colored gray. Residues involved in putative binding cavity calculation from the CASTp program are colored light blue. Surface for residues that are not within binding cavity from ECL2 is not displayed.

act as gatekeepers, dividing the binding pocket into two halves. The left cavity had a relatively larger volume of 1075 Å^3, compared to the smaller 600 Å^3 right cavity. A crystal structure based comparison of common (hot spot) TM residues involved in ligand interaction from several other class A GPCRs (e.g., Turkey β1 adrenergic receptor bound to cyanopindolol (PDB: 2VT4[45]); dopamine D3 receptor bound to eticlopride (PDB: 3PBL[46]); chemokine CXCR4 receptor bound to 1T1t (PDB: 3ODU[47]); chemokine CCR5 receptor bound to maraviroc (PDB: 4MBS[48])), together with information on a ligand accessible site available at GPCRDB, led us to identify and match several analogous residues in PAR2 with those in the three template proteins used here (Table 3). As PAR1 and PAR2 share the highest TM sequence homology (44%), there

Table 3. Comparison of Putative Hotspot Residues between PAR2 and 3 Templates

B.W. no.	PAR2	PAR1	ORL-1	rhodopsin
3.32	F	F	D	A
3.33	Y	Y	Y	T
3.37	Y	Y	F	E
5.39	F	F	A	V
5.42	L	F	I	M
5.43	A	S	F	F
5.47	F	F	F	F
6.44	Y	F	F	F
6.51	S	T	V	Y
6.52	N	N	Q	A
6.55	L	L	V	A
6.59	Y	Y	G	F
7.35	Y	Y	L	M
7.39	L	V	T	A

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181−1191

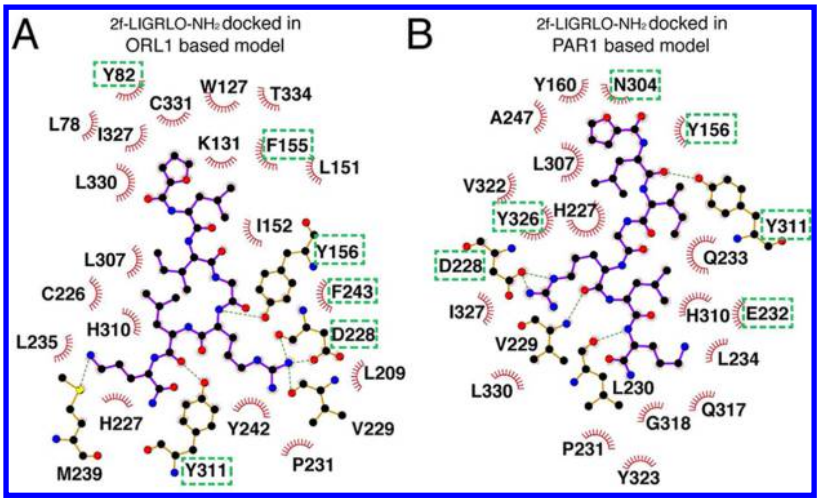


Figure 4. Predicted interactions between PAR2 and 2f-LIGRLO-NH_2: (A) 2f-LIGRLO-NH_2 docked in an ORL-1 derived PAR2 model and (B) 2f-LIGRLO-NH_2 docked in a PAR1 derived PAR2 model. Green dotted boxes represent residues that display at least 10-fold loss in PAR2 activation upon mutation. This figure was generated using Ligplot.[53]

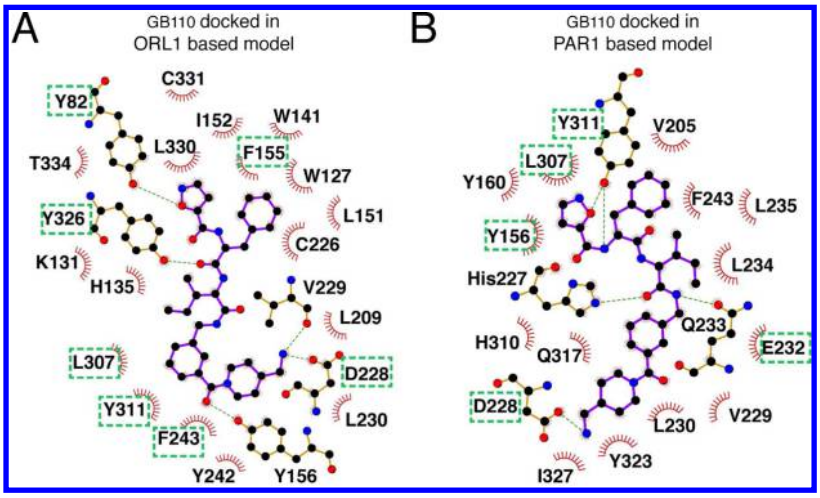


Figure 5. Predicted interactions between PAR2 and agonist GB110: (A) GB110 docked in an ORL-1-based PAR2 model and (B) GB110 docked in a PAR1-based PAR2 model. Green dotted boxes represent residues that display at least 10-fold loss in PAR2 activation upon mutation. This figure was generated using Ligplot.

was also high homology among the class A hotspot residues. The four key aromatic TM residues mentioned above were included in the cavity shown in Figure 3. Considering the positioning of ECL2, the analysis of key aromatic TM residues and the binding site comparison among the three models, it was decided that further more detailed characterization should focus instead on the ORL-1 and PAR1 based homology models of PAR2.

Molecular Docking into PAR2. In order to explore the potential applicability of the ORL-1 and PAR1 derived models of PAR2 for *in silico* screening, a peptidic PAR2 agonist 2f-LIGRLO-NH_2 (HT29 cells EC_{50} 210 nM) and a modified nonpeptidic PAR2 agonist GB110 (HT29 cells EC_{50} 280 nM) were docked into each PAR2 model structure. Interactions involving key residues shown to be important for PAR2 activation from our recent mutagenesis study[43] are compared here.

Docking of 2f-LIGRLO-NH_2 into the ORL-1 based model of PAR2 (Figure 4) suggested polar contacts with three residues ($Y156^{3.33}$, $Y311^{6.59}$, $D228^{ECL2}$) known to be important for receptor activation. A potential cation-π interaction between R5 and Y242 was also predicted by docking. $Y311^{6.59}$ and $D228^{ECL2}$ were captured as plausible interacting residues from ligand docking into the PAR1-derived PAR2 model. Although polar contacts between 2f-LIGRLO-NH_2 and $D228^{ECL2}$ and $Y311^{6.59}$ were sampled in both models, this ligand was positioned in opposite orientations in these two models. In the ORL-1 based model, the furoyl group oriented toward a pocket between TM2 and TM7, while in the PAR1 based model the furoyl group was docked between TM5 and TM6.

The proposed binding mode of GB110 in the ORL-1 based model (Figure 5) positions equivalent elements of GB110 in a similar pose as 2f-LIGRLO-NH_2. Polar contacts with $Y156^{3.33}$ and $D228^{ECL2}$ are still present, in addition to polar contacts to $Y82^{1.39}$ and $Y326^{7.35}$. In the human PAR1 based model for PAR2, GB110 is also proposed to adopt a similar pose to 2f-LIGRLO-NH_2. Similar polar contacts with $Y311^{6.59}$ and $D228^{ECL2}$, also $H227^{ECL2}$ and $E232^{ECL2}$, were found. Similarly to the poses of 2f-LIGRLO-NH_2, opposite orientations of GB110 were found in the two models.

Considering the differences observed in the putative binding site, it was expected that one of the models would position the

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181−1191

PAR2 agonists within close proximity to more amino acids known to be important for receptor activation. Docking results predicted residues that have been confirmed by PAR2 mutagenesis studies as being important for PAR2 activity,[43] and this encouraged our use of model-based virtual screening for discovery of novel PAR2 modulators.

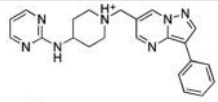
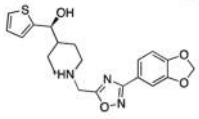
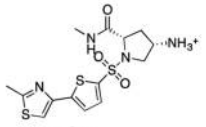
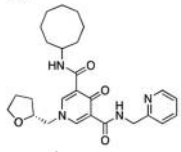
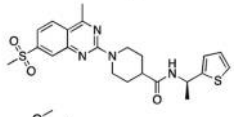
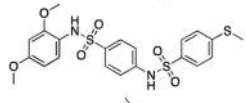
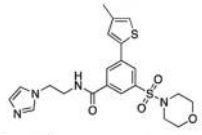
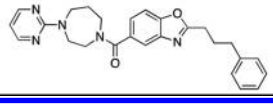
Small Scale Enrichment Study. Prior to virtual screening of a commercial database of potential ligands for PAR2, a preliminary small-scale enrichment calculation was conducted on the ORL-1 and PAR1 based models. This can provide information indicating the potential performance of the models in returning a higher hit rate in virtual screens. The small-scale virtual screening library consisted of 8 PAR2 antagonists and 1000 drug-like Schrödinger decoy compounds with an average molecular weight of 400 Da. Surprisingly, both homology models gave similar enrichment factors of EF2% (0.0), EF5% (2.5), and EF10% (1.25), indicating similar performance of the two models in predicting hits from this training set of ligands. However, the EF study is very dependent on the structural diversity and activities of the compounds used to enrich the data set, the lack of alternative PAR2 antagonists placing a heavy reliance on only in-house synthetic ligands. Consequently, the EF calculation here may not be sufficient to truly reflect the quality of the models. In this context, a prospective virtual assessment of a large chemical library followed by experimental validation of putative ligands in biochemical assays is needed to further validate PAR2 homology models reported herein.

Virtual Library Screening. To utilize the two PAR2 homology models to discover new PAR2 modulators, we performed large-scale virtual library screening of the Chembridge database containing ~700,000 compounds.

The virtual screening workflow was identical for the ORL-1 and PAR1 based models. 71 compounds (Figures S6–S9) were identified as hits from virtual screening using both homology models. These compounds were assayed for PAR2 agonist and antagonist activity in HT29 colon cancer cells, a highly expressing PAR2 cell line. From initial assays, compounds inhibiting PAR2 activation (>50%) at 100 μM were further studied in a concentration-dependent manner for inhibition of PAR2-mediated intracellular Ca^{2+} release induced by 1 μM 2f-LIGRLO-NH_2. Compounds **1**–**4** (IC_{50} 11–44 μM) were identified as PAR2 antagonists from the ORL-1 based model, and four different compounds **5**–**8** (IC_{50} 30–94 μM) were obtained from the PAR1 based model (Table 4). Figure 6A and 6B illustrate the pharmacological profiles of just compounds **1** and **5** (six other antagonists are shown in Figure S10) in both agonist and antagonist assays for PAR2 Ca^{2+} mobilization. Encouragingly, except for compound **5** which exhibited about 25% activation of Ca^{2+} release at the highest concentration tested (100 μM), all the compounds proved to be pure antagonists of 2f-LIGRLO-NH_2 without causing significant PAR2 activation when examined alone with no added agonist (HT29 cells).

These results may reflect a degree of transferability between models of the inactive form of the template crystal structures, supporting functional fidelity of docking ligands to the conformation of the GPCR, which is in good agreement with previous virtual screens conducted with other GPCR homology models.[49] A caveat of GPCR homology modeling lies in the inactive conformational state induced by an antagonist, which is often necessary for conformational stabilization and therefore favors crystallization conditions. This effect has been consistently shown to influence the activity of hits derived from virtual screening GPCR homology models based on these inactive structures. Meanwhile, the higher proportion of active molecules discovered from the library of compounds with MW 350–600, for both models, is likely due to the nature of the binding site of PAR2, which can accommodate peptidic ligands.

Table 4. Structures and Antagonist Potencies of Compounds Derived from Virtual Screening in PAR2 Homology Models

ID	Structure	IC_{50} (μM)[a]
1		11 ± 2
2		16 ± 4
3		41 ± 21
4		44 ± 20
5		30 ± 6
6		40 ± 9
7		93 ± 19
8		94 ± 30

[a]Cell based Ca^{2+} mobilization in HT29 cells. Concentration dependent responses after 30 min preincubation and then addition of agonist, 1 μM 2f-LIGRLO-NH_2. Data are represented as mean ± SEM ($n \geq 3$ independent experiments). Compounds **1**–**4** were derived from the ORL-1 based model of PAR2. Compounds **5**–**8** were derived from the PAR1 based model of PAR2.

To understand possible interactions between the most potent antagonists derived from the two parallel virtual screens of PAR2, docked poses of **1** and **5** in the ORL-1 and PAR1 based models respectively were analyzed (Figure 7). Consistent with the observation of the docked pose for 2f-LIGRLO-NH_2 and GB110, the overall binding region and orientation occupied by the two antagonists adopted a similar pattern. In the docked complex of compound **1** and the ORL-1 based model, the phenyl ring was docked between TM2 and TM3, forming interactions with W127, pyrazolol-pyrimidine formed a parallel-displaced pi stacking arrangement with F155, and the piperidine ring was predicted to binding in between $Y156^{3.33}$ and $Y326^{7.35}$. As this group is prone to protonation, we surmise that a potential hydrogen bond might be formed with a receptor residue although this was not observed from docking

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181–1191

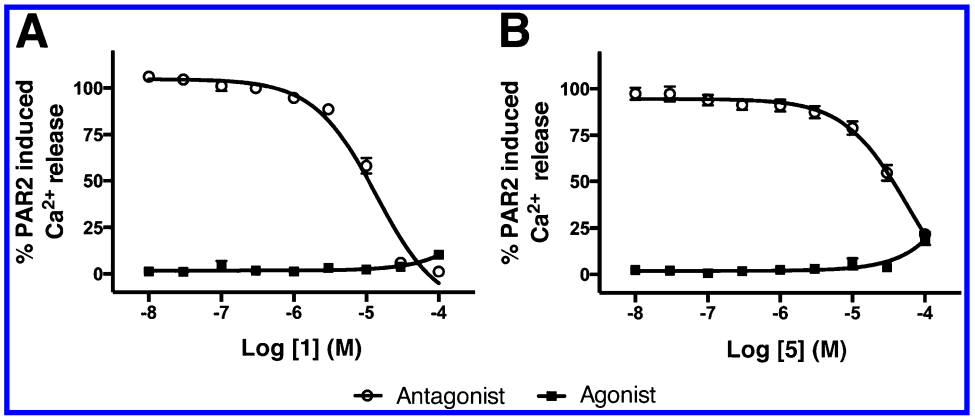


Figure 6. Agonist and antagonist activity of **1** and **5** in HT29 cells. Concentration dependent Ca^{2+} release in HT29 cells induced by **1** (A) and **5** (B) alone at up to 100 μM (agonist assay) or by 1 μM 2f-LIGRLO-NH_2 in the presence of inhibitors **1** and **5** (antagonist assay). Normalized to 100% Ca^{2+} release for 100 μM 2f-LIGRLO-NH_2. Panels reflect data for compounds in Table 4. Data are represented as mean ± SEM ($n \geq 3$ independent experiments).

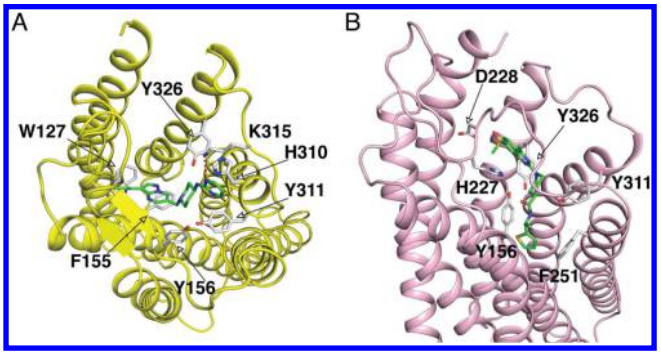


Figure 7. Docked binding mode of compounds **1** and **5** in ORL-1 and PAR1 based homology models of PAR2: (A) compound **1** docked in the ORL-1 based model and (B) compound **5** docked in the PAR1 based model. Hydrogen bond is depicted in red dashed line. Compounds are colored by atom types; carbon: green; oxygen: red; nitrogen: blue; sulfur: yellow.

simulation. The pyrimidine ring intercalated between $H310^{6.58}$ and $Y311^{6.59}$, contributing to aromatic interactions that can further stabilize ligand binding. One of the nitrogen atoms was predicted to form a hydrogen bond interaction with the amine NH on K315, a residue within ECL3. In contrast, compound **5** from the PAR1 based model was docked above the base formed by the two tyrosines, $Y156^{3.33}$ and $Y326^{7.35}$. Intriguingly, this compound also contained a piperidine ring in the central region of the molecule, and it was also positioned in proximity to $Y156^{3.33}$, $Y311^{6.59}$, and $Y326^{7.35}$. The thiophene in compound **5** was predicted to bind between TM3 and TM5, surrounded by $Y156^{3.33}$ and F251. The quinazoline bicyclic ring is situated perpendicular to $Y326^{7.35}$, forming a T-shaped edge-to-face π-aryl interactions with this residue. The terminal methylsulfonyl group is oriented toward ECL2 residues $H227^{ECL2}$ and $D228^{ECL2}$. The putative binding of these two hits was supported by a ligand binding assay (Figure S11). Hits from the ORL-1 based model all showed weak binding affinity, while two hits from the PAR1 based model competed for binding against 2f-LIGRLO(dpta)-NH_2. Compound **5** exhibited the highest competitive binding affinity (70%) toward CHO-hPAR2 at 100 μM, whereas compound **1** showed only 35% competitive binding. Compounds **6** and **8** did not compete at the highest concentration tested. Although this is in agreement with their low potency in the Ca^{2+} assay, it must be said that the competition assay is limited in only identifying binding to the same site as the labeled ligand. Further, we surmise that the binding pocket in the PAR2 homology model derived from ORL-1 and PAR1 may represent different allosteric binding site of PAR2 antagonists identified, which rendered all hits weak antagonists. So far little is known about the orthosteric binding site of the PAR2 tethered ligand, and it is beyond our scope here to further investigate this. However, results presented here do suggest that the dynamic and flexible nature of the PAR2 structure in current models offers alternative and not mutually exclusive templates for discovery of structurally different ligands.

PAR2 antagonists with novel scaffolds and a degree of chemical diversity were discovered from these virtual screening approaches with a surprisingly high hit rate of 10.3% (ORL-1 based model) and 12.5% (PAR1 based model). The majority of PAR2 ligands reported in the literature are peptidic, developed from the canonical tethered ligand.[19] While discovery of potent small molecule antagonists for this receptor has been problematic, we report here the first success in homology modeling-based virtual screening for PAR2. Both models selected for virtual screening returned hits with moderate potency. This approach is clearly useful in providing new structural templates for ligand-based optimization to derive more potent analogues. The most potent compound **1** (IC_{50} 11 μM) has MW = 386 Da, cLogP = 2.55, 2 hydrogen bond donors, 7 hydrogen bond acceptors, tPSA 72 $Å^2$, and 5 rotatable bonds (Table S1), indicating the kinds of properties sought for virtual hits. Encouragingly, the results suggest that the models have a different degree of accuracy in predicting the putative conformation in the binding site of PAR2.

The homology structural model for PAR2 can very likely be further refined to improve the screening rate for PAR2 modulators. Current virtual screens are conducted against a rigid receptor, because flexible virtual screening is more computationally demanding. Interestingly, at the time of preparing this manuscript, Rataj et al. reported the low impact of choosing the most closely related phylogenetic template based on homology model efficiency in virtual screening for four serotonin receptors.[50] We also found from our virtual screening, that the hit rate was not solely dependent on a high overall sequence identity between PAR2 and its template crystal structure or even the sequence homology between residues in the ligand-binding pocket. Indeed the more potent

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181−1191

hits in this study were obtained using the less homologous template (human ORL-1). These results contrast to the GPCR Dock 2010 assessment, which concluded that only closely related proteins are suitable for GPCR homology modeling.[51] That assessment from GPCR Dock focused on predicting the binding modes of ligands in static crystal structures. To some extent, such an assessment is restricted by the central premise that the docked complex in the homology model would be expected to be close to the crystal structure. However, for the virtual screening herein, we were interested in identifying structurally diverse ligands.

Thus, we recommend that future studies investigating *in silico* screening of GPCRs should utilize more than one template in parallel, in order to access wider conformational sampling of the dynamic states of a GPCR and hence potentially sample more structural diversity in ligand hits, rather than relying on a static solid state crystal structure should it be available.

CONCLUSIONS

Three structurally reliable and energetically reasonable homology models of the structure of PAR2 are reported here. Detailed comparisons of TM residues, ECL2, and the binding cavities led us to prioritize the ORL-1 and PAR1 based models for further molecular docking and virtual screening of putative ligands. Eight novel, structurally diverse, PAR2 antagonists (with negligible or no agonist activity) were identified, with IC_{50} values ranging from 10 to 100 μM against 1 μM 2f-LIGRLO-NH_2 in HT29 colon cancer cells. Despite a lower TM sequence identity and similarity to PAR2, the ORL-1 based model returned an almost equal hit rate to that obtained from the more closely related and more sequence-homologous template, PAR1. Thus, the common practice of using the template structure with the highest sequence identity for GPCR modeling may not necessarily be ideal, and templates of lower homology might produce equipotent hits in virtual screening campaigns. It may be valuable to use a set of homology models derived from templates with different sequence similarity and binding pocket variations in the TM region, and thus different conformations to that stabilized in the solid state as elucidated by a crystal structure, in order to capture structurally diverse sets of potential ligands through virtual screening discovery programs. Subsequent hit-to-lead campaigns can benefit greatly from structurally diverse hits as alternative starting points for medicinal chemistry optimization programs.

ASSOCIATED CONTENT

Supporting Information

Multiple sequence alignment of PAR2 with 3 template GPCRs is shown in Figure S1. Ramachandran plots of 3 homology models for PAR2 is illustrated in Figure S2. Cartoon representation of 3 homology models for PAR2 are shown in Figures S3–S5. Chemical structures of purchased and assayed virtual hits are shown in Figures S6–S9. Pharmacological profiles for concentration dependent Ca^{2+} release in HT29 cells by eight virtual screening hits (compounds **1**–**8**) are shown in Figure S10. Bar chart showing competitive binding of hits is shown in Figure S11. Physicochemical properties of all hits are summarized in Table S1. Details of the database filtering procedures used to generate the two virtual databases and of the competitive binding assay used to measure PAR2 affinity of compounds are given under Methods in the Supporting Information. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.5b00087.

AUTHOR INFORMATION

Corresponding Author

*Fax: +61733462990. Email: d.fairlie@imb.uq.edu.au.

Author Contributions

‡Joint first authors. Authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We sincerely thank Dr. Jacky Suen for assistance and training early in this research. We acknowledge receipt of an academic license to the OpenEye Suite of software at no cost. We also acknowledge the Australian Research Council for grant DP130100629 and for a Centre of Excellence in Advanced Molecular Imaging CE140100011, the National Health and Medical Research Council for grants 1084083 and 1047759, a Senior Principal Research Fellowship to D.F. (1027369), and the Qld Government for a CIF grant.

REFERENCES

(1) Tyndall, J. D.; Pfeiffer, B.; Abbenante, G.; Fairlie, D. P. Over One Hundred Peptide-Activated G Protein-Coupled Receptors Recognize Ligands with Turn Structure. *Chem. Rev.* **2005**, *105*, 793–826.

(2) Blakeney, J. S.; Fairlie, D. P. Nonpeptide Ligands That Target Peptide-Activated GPCRs in Inflammation. *Curr. Med. Chem.* **2005**, *12*, 3027–3042.

(3) Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. How Many Drug Targets Are There? *Nat. Rev. Drug Discovery* **2006**, *5*, 993–996.

(4) Carlsson, J.; Coleman, R. G.; Setola, V.; Irwin, J. J.; Fan, H.; Schlessinger, A.; Sali, A.; Roth, B. L.; Shoichet, B. K. Ligand Discovery from a Dopamine D3 Receptor Homology Model and Crystal Structure. *Nat. Chem. Biol.* **2011**, *7*, 769–778.

(5) Mysinger, M. M.; Weiss, D. R.; Ziarek, J. J.; Gravel, S.; Doak, A. K.; Karpiak, J.; Heveker, N.; Shoichet, B. K.; Volkman, B. F. Structure-Based Ligand Discovery for the Protein-Protein Interface of Chemokine Receptor CXCR4. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 5517–5522.

(6) Langmead, C. J.; Andrews, S. P.; Congreve, M.; Errey, J. C.; Hurrell, E.; Marshall, F. H.; Mason, J. S.; Richardson, C. M.; Robertson, N.; Zhukov, A.; Weir, M. Identification of Novel Adenosine A(2A) Receptor Antagonists by Virtual Screening. *J. Med. Chem.* **2012**, *55*, 1904–1909.

(7) Vu, T. K.; Hung, D. T.; Wheaton, V. I.; Coughlin, S. R. Molecular Cloning of a Functional Thrombin Receptor Reveals a Novel Proteolytic Mechanism of Receptor Activation. *Cell* **1991**, *64*, 1057–1068.

(8) Nystedt, S.; Emilsson, K.; Larsson, A. K.; Strombeck, B.; Sundelin, J. Molecular Cloning and Functional Expression of the Gene Encoding the Human Proteinase-Activated Receptor 2. *Eur. J. Biochem.* **1995**, *232*, 84–89.

(9) Nishibori, M.; Mori, S.; Takahashi, H. K. Physiology and Pathophysiology of Proteinase-Activated Receptors (PARs): PAR-2-Mediated Proliferation of Colon Cancer Cell. *J. Pharmacol. Sci.* **2005**, *97*, 25–30.

(10) Matej, R.; Mandakova, P.; Netikova, I.; Pouckova, P.; Olejar, T. Proteinase-Activated Receptor-2 Expression in Breast Cancer and the Role of Trypsin on Growth and Metabolism of Breast Cancer Cell Line MDA MB-231. *Physiol. Res.* **2007**, *56*, 475–484.

(11) Iwaki, K.; Shibata, K.; Ohta, M.; Endo, Y.; Uchida, H.; Tominaga, M.; Okunaga, R.; Kai, S.; Kitano, S. A Small Interfering RNA Targeting Proteinase-Activated Receptor-2 Is Effective in Suppression of Tumor Growth in a Panc1 Xenograft Model. *Int. J. Cancer* **2008**, *122*, 658–663.

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181–1191

(12) Kawabata, A.; Kanke, T.; Yonezawa, D.; Ishiki, T.; Saka, M.; Kabeya, M.; Sekiguchi, F.; Kubo, S.; Kuroda, R.; Iwaki, M.; Katsura, K.; Plevin, R. Potent and Metabolically Stable Agonists for Protease-Activated Receptor-2: Evaluation of Activity in Multiple Assay Systems in Vitro and in Vivo. *J. Pharmacol. Exp. Ther.* **2004**, *309*, 1098−1107.
(13) Barry, G. D.; Suen, J. Y.; Low, H. B.; Pfeiffer, B.; Flanagan, B.; Halili, M.; Le, G. T.; Fairlie, D. P. A Refined Agonist Pharmacophore for Protease Activated Receptor 2. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 5552−5557.
(14) Maryanoff, B. E.; Santulli, R. J.; McComsey, D. F.; Hoekstra, W. J.; Hoey, K.; Smith, C. E.; Addo, M.; Darrow, A. L.; Andrade-Gordon, P. Protease-Activated Receptor-2 (PAR-2): Structure-Function Study of Receptor Activation by Diverse Peptides Related to Tethered-Ligand Epitopes. *Arch. Biochem. Biophys.* **2001**, *386*, 195−204.
(15) McGuire, J. J.; Saifeddine, M.; Triggle, C. R.; Sun, K.; Hollenberg, M. D. 2-Furoyl-LIGRLO-amide: A Potent and Selective Proteinase-Activated Receptor 2 Agonist. *J. Pharmacol. Exp. Ther.* **2004**, *309*, 1124−1131.
(16) Suen, J. Y.; Barry, G. D.; Lohman, R. J.; Halili, M. A.; Cotterell, A. J.; Le, G. T.; Fairlie, D. P. Modulating Human Proteinase Activated Receptor 2 with a Novel Antagonist (GB88) and Agonist (GB110). *Br. J. Pharmacol.* **2012**, *165*, 1413−1423.
(17) Barry, G. D.; Suen, J. Y.; Le, G. T.; Cotterell, A.; Reid, R. C.; Fairlie, D. P. Novel Agonists and Antagonists for Human Protease Activated Receptor 2. *J. Med. Chem.* **2010**, *53*, 7428−7440.
(18) Lohman, R. J.; Cotterell, A. J.; Suen, J.; Liu, L.; Do, A. T.; Vesey, D. A.; Fairlie, D. P. Antagonism of Protease-Activated Receptor 2 Protects against Experimental Colitis. *J. Pharmacol. Exp. Ther.* **2012**, *340*, 256−265.
(19) Lohman, R. J.; Cotterell, A. J.; Barry, G. D.; Liu, L.; Suen, J. Y.; Vesey, D. A.; Fairlie, D. P. An Antagonist of Human Protease Activated Receptor-2 Attenuates PAR2 Signaling, Macrophage Activation, Mast Cell Degranulation, and Collagen-Induced Arthritis in Rats. *FASEB J.* **2012**, *26*, 2877−2887.
(20) Suen, J. Y.; Cotterell, A.; Lohman, R. J.; Lim, J.; Han, A.; Yau, M. K.; Liu, L.; Cooper, M. A.; Vesey, D. A.; Fairlie, D. P. Pathway Selective Antagonism of Proteinase Activated Receptor 2. *Br. J. Pharmacol.* **2014**, *171*, 4112−4124.
(21) Tautermann, C. S. GPCR Structures in Drug Design, Emerging Opportunities with New Structures. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 4073−4079.
(22) Okada, T.; Sugihara, M.; Bondar, A. N.; Elstner, M.; Entel, P.; Buss, V. The Retinal Conformation and Its Environment in Rhodopsin in Light of a New 2.2 A Crystal Structure. *J. Mol. Biol.* **2004**, *342*, 571−583.
(23) Bissantz, C.; Bernard, P.; Hibert, M.; Rognan, D. Protein-Based Virtual Screening of Chemical Databases. II. Are Homology Models of G-Protein Coupled Receptors Suitable Targets? *Proteins* **2003**, *50*, 5−25.
(24) Thompson, A. A.; Liu, W.; Chun, E.; Katritch, V.; Wu, H.; Vardy, E.; Huang, X. P.; Trapella, C.; Guerrini, R.; Calo, G.; Roth, B. L.; Cherezov, V.; Stevens, R. C. Structure of the Nociceptin/Orphanin FQ Receptor in Complex with a Peptide Mimetic. *Nature* **2012**, *485*, 395−399.
(25) Zhang, C.; Srinivasan, Y.; Arlow, D. H.; Fung, J. J.; Palmer, D.; Zheng, Y.; Green, H. F.; Pandey, A.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Coughlin, S. R.; Kobilka, B. K. High-Resolution Crystal Structure of Human Protease-Activated Receptor 1. *Nature* **2012**, *492*, 387−392.
(26) Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J. Jalview Version 2–a Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics* **2009**, *25*, 1189−1191.
(27) Sali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.
(28) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *J. Med. Chem.* **2006**, *49*, 534−553.
(29) Sherman, W.; Beard, H. S.; Farid, R. Use of an Induced Fit Receptor Structure in Virtual Screening. *Chem. Biol. Drug. Des.* **2006**, *67*, 83−84.
(30) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221−234.
(31) Benkert, P.; Kunzli, M.; Schwede, T. Qmean Server for Protein Model Quality Estimation. *Nucleic Acids Res.* **2009**, *37*, W510−514.
(32) Zhang, C.; Liu, S.; Zhou, Y. Accurate and Efficient Loop Selections by the Dfire-Based All-Atom Statistical Potential. *Protein Sci.* **2004**, *13*, 391−399.
(33) Colovos, C.; Yeates, T. O. Verification of Protein Structures: Patterns of Nonbonded Atomic Interactions. *Protein Sci.* **1993**, *2*, 1511−1519.
(34) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* **1997**, *267*, 727−748.
(35) Chen, H.; Lyne, P. D.; Giordanetto, F.; Lovell, T.; Li, J. On Evaluating Molecular-Docking Methods for Pose Prediction and Enrichment Factors. *J. Chem. Inf. Model.* **2006**, *46*, 401−415.
(36) Wang, R.; Fu, Y.; Lai, L. A New Atom-Additive Method for Calculating Partition Coefficients. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 615−621.
(37) *FILTER, Version 2.0.2*; OpenEye Scientific Software, Inc.: Santa Fe, NM, USA, 2010. http://www.eyesopen.com (accessed May 1, 2015). Open Eye Scientific Software kindly provided the software at no cost.
(38) *OMEGA, Version 2.5.1.4*; Openeye Scientific Software, Inc.: Santa Fe, NM, USA, 2010. http://www.eyesopen.com (accessed May 1, 2015). Open Eye Scientific Software kindly provided the software at no cost.
(39) Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53*, 2719−2740.
(40) Dreier, L.; Wider, G. Concentration Measurements by Pulcon Using X-Filtered or 2D NMR Spectra. *Magn. Reson. Chem.* **2006**, *44* (Spec No), S206−212.
(41) Lerner, D. J.; Chen, M.; Tram, T.; Coughlin, S. R. Agonist Recognition by Proteinase-Activated Receptor 2 and Thrombin Receptor. Importance of Extracellular Loop Interactions for Receptor Function. *J. Biol. Chem.* **1996**, *271*, 13943−13947.
(42) Wheatley, M.; Wootten, D.; Conner, M. T.; Simms, J.; Kendrick, R.; Logan, R. T.; Poyner, D. R.; Barwell, J. Lifting the Lid on GPCRs: The Role of Extracellular Loops. *Br. J. Pharmacol.* **2012**, *165*, 1688−1703.
(43) Suen, J. Y.; Adams, M. N.; Lim, J.; Madala, P. K.; Xu, W.; Cotterell, A.; He, Y.; Hooper, J. D.; Fairlie, D. P. Native Versus Synthetic Regulators of Proteinase Activated Receptor 2; manuscript submitted for publication.
(44) Binkowski, T. A.; Naghibzadeh, S.; Liang, J. CASTp: Computed Atlas of Surface Topography of Proteins. *Nucleic Acids Res.* **2003**, *31*, 3352−3355.
(45) Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G.; Tate, C. G.; Schertler, G. F. Structure of a Beta1-Adrenergic G-Protein-Coupled Receptor. *Nature* **2008**, *454*, 486−491.
(46) Chien, E. Y.; Liu, W.; Zhao, Q.; Katritch, V.; Han, G. W.; Hanson, M. A.; Shi, L.; Newman, A. H.; Javitch, J. A.; Cherezov, V.; Stevens, R. C. Structure of the Human Dopamine D3 Receptor in Complex with a D2/D3 Selective Antagonist. *Science* **2010**, *330*, 1091−1095.
(47) Wu, B.; Chien, E. Y.; 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.; Stevens, R. C. Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science* **2010**, *330*, 1066−1071.

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181−1191

(48) Tan, Q.; Zhu, Y.; Li, J.; Chen, Z.; Han, G. W.; Kufareva, I.; Li, T.; Ma, L.; Fenalti, G.; Li, J.; Zhang, W.; Xie, X.; Yang, H.; Jiang, H.; Cherezov, V.; Liu, H.; Stevens, R. C.; Zhao, Q.; Wu, B. Structure of the CCR5 Chemokine Receptor-HIV Entry Inhibitor Maraviroc Complex. *Science* **2013**, *341*, 1387−1390.
(49) Weiss, D. R.; Ahn, S.; Sassano, M. F.; Kleist, A.; Zhu, X.; Strachan, R.; Roth, B. L.; Lefkowitz, R. J.; Shoichet, B. K. Conformation Guides Molecular Efficacy in Docking Screens of Activated Beta-2 Adrenergic G Protein Coupled Receptor. *ACS Chem. Biol.* **2013**, *8*, 1018−1026.
(50) Rataj, K.; Witek, J.; Mordalski, S.; Kosciolek, T.; Bojarski, A. J. Impact of Template Choice on Homology Model Efficiency in Virtual Screening. *J. Chem. Inf. Model.* **2014**, *54*, 1661−1668.
(51) Kufareva, I.; Rueda, M.; Katritch, V.; Stevens, R. C.; Abagyan, R.; participants, G. D. Status of GPCR Modeling and Docking as Reflected by Community-Wide GPCR Dock 2010 Assessment. *Structure* **2011**, *19*, 1108−1126.
(52) *The Pymol Molecular Graphics System, Version 1.5.0.4*; Schödinger, LLC.
(53) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. Ligplot: A Program to Generate Schematic Diagrams of Protein-Ligand Interactions. *Protein Eng.* **1995**, *8*, 127−134.

DOI: 10.1021/acs.jcim.5b00087
J. Chem. Inf. Model. 2015, 55, 1181−1191