

Mapping Functional Group Free Energy Patterns at Protein Occluded Sites: Nuclear Receptors and G-Protein Coupled Receptors

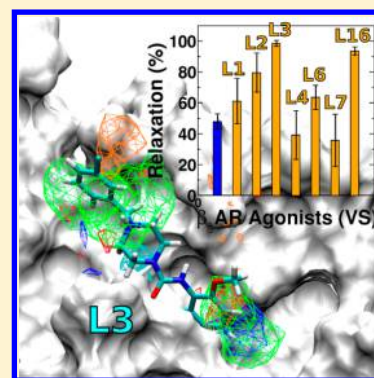
Sirish Kaushik Lakkaraju,[†] Wenbo Yu,[†] E. Prabhu Raman,[†] Alena V. Hershfeld,[‡] Lei Fang,[†] Deepak A. Deshpande,[§] and Alexander D. Mackerell, Jr.*

[†]Department of Pharmaceutical Sciences, School of Pharmacy; [‡]Division of Pulmonary and Critical Care Medicine, School of Medicine, University of Maryland, 20 Penn Street, Baltimore, Maryland 21201, United States

[§]Center for Translational Medicine, Department of Medicine, Thomas Jefferson University, 1020 Locust Street, Philadelphia, Pennsylvania 19107, United States

S Supporting Information

ABSTRACT: Occluded ligand-binding pockets (LBP) such as those found in nuclear receptors (NR) and G-protein coupled receptors (GPCR) represent a significant opportunity and challenge for computer-aided drug design. To determine free energies maps of functional groups of these LBPs, a Grand-Canonical Monte Carlo/Molecular Dynamics (GCMC/MD) strategy is combined with the Site Identification by Ligand Competitive Saturation (SILCS) methodology. SILCS-GCMC/MD is shown to map functional group affinity patterns that recapitulate locations of functional groups across diverse classes of ligands in the LBPs of the androgen (AR) and peroxisome proliferator-activated- γ (PPAR γ) NRs and the metabotropic glutamate (mGluR) and β_2 -adrenergic (β_2 AR) GPCRs. Inclusion of protein flexibility identifies regions of the binding pockets not accessible in crystal conformations and allows for better quantitative estimates of relative ligand binding affinities in all the proteins tested. Differences in functional group requirements of the active and inactive states of the β_2 AR LBP were used in virtual screening to identify high efficacy agonists targeting β_2 AR in Airway Smooth Muscle (ASM) cells. Seven of the 15 selected ligands were found to effect ASM relaxation representing a 46% hit rate. Hence, the method will be of use for the rational design of ligands in the context of chemical biology and the development of therapeutic agents.



INTRODUCTION

Occluded ligand binding pockets (LBP) in proteins with minimal or no accessibility to the surrounding environment represent a significant, yet challenging opportunity for structure-based and computer-aided drug design approaches. LBPs of more than half of all clinical drug targets,¹ including the G-protein coupled receptors (GPCR)² and nuclear receptors (NR),³ are either partially or fully occluded. As the efficacies of ligands of both GPCRs⁴ and NRs⁵ are known to be coupled to small conformational changes in their binding sites, accurate modeling of these sites is critical for future development of therapeutic agents for a wide range of diseases.^{6,7}

The site identification by ligand competitive saturation (SILCS) methodology is a fragment sampling technique that maps free energy affinity patterns of functional groups at protein surfaces, including LBPs.^{8,9} The method accounts for the conformational flexibility of the proteins, chemical space of the ligands, and explicit solvent by running molecular dynamics (MD) of the target protein in an aqueous solution of small solute molecules representative of different chemical functional groups. The affinity patterns of these functional groups are obtained in the form of discretized probability, or, equivalently, free energy maps, called FragMaps. Inclusion of protein

flexibility and explicit solvent representation is particularly important given the known conformational changes within the binding pocket upon ligand binding^{10–12} and competition with and displacement of waters by ligands.¹³ The SILCS method was successful in mapping the functional group requirements of ligands for a range of macromolecules and consequently guided ligand optimization studies.^{14,15} To probe occluded LBPs, SILCS is coupled with an iterative Grand-Canonical Monte Carlo (GCMC) and MD methodology.¹⁶ GCMC drives the sampling of small solutes and explicit solvent in LBPs and MD allows for conformational sampling of the macromolecules in the presence of solutes and water, which is useful in exploring cryptic pockets absent in apo crystal structures that are known to serve as binding sites.¹⁷ In a proof of principle study, FragMaps from the SILCS-GCMC/MD were shown to overlap well with the positions of chemically similar functional groups of known ligands in the occluded LBP of an apolar mutant of the T4-lysozyme.¹⁶

In this work, SILCS-GCMC/MD was used to map the functional group affinity patterns of the occluded pockets of the following therapeutically important NRs and GPCRs for which

Received: December 9, 2014

Published: February 18, 2015

structural data with multiple ligands is available. These include the androgen receptor¹⁸ (AR) and peroxisome proliferator-activated receptor- γ ¹⁹ (PPAR γ) NRs and the metabotropic glutamate receptor²⁰ (mGluR) and β_2 -adrenergic receptor²¹ (β_2 AR) GPCRs. Analysis focused on both the qualitative and quantitative information content of the SILCS FragMaps. The method can predict the relative binding affinities of ligands through a ligand grid free energy (LGFE) scoring scheme (see SI Text, Section S6) in which the inclusion of protein conformational flexibility is found to be important. Also, the method is capable of distinguishing between active and inactive states of the β_2 AR through differences in the affinity patterns across these states, information that is useful in distinguishing the function of ligands. Validation of this capability is the ability of FragMaps differences in identification of new agonists of β_2 AR that have the potential to be developed into therapeutic agents for asthma and other obstructive pulmonary diseases.^{22,23}

RESULTS

Eight representative solutes with different chemical functionalities: benzene, propane, acetaldehyde, methanol, formamide, imidazole, acetate, and methylammonium were chosen to probe the LBPs. Benzene and propane serve as probes for nonpolar functionalities. Methanol, formamide, imidazole, and acetaldehyde are neutral molecules that participate in hydrogen bonding. The positively charged methylammonium and negatively charged acetate molecules serve as probes for charged donor and acceptors, respectively. The normalized probability distributions for selected atoms in these solutes from the SILCS-GCMC/MD simulations were then used to create functional group affinity FragMaps at the respective LBPs on which the analysis described below was performed.

AR. A total of 48 crystal structures of the human wild-type (WT) AR are available, of which 15 structures have distinct ligands. SILCS-GCMC/MD was run with the testosterone-AR crystal structure (PDB 2AM9),¹⁸ after the removal of testosterone. Although the simulations were initiated with a steroid-bound conformation of the AR, FragMaps from the 10 \times 100 ns GCMC/MD simulations recapitulate the locations of the different functional groups of both steroidal and non-steroidal crystallographic ligands (Figure 1). Cycloalkane rings of the steroidal ligands that occupy the largely hydrophobic pocket of the AR have good overlap with the apolar (APOLAR) FragMaps (A1 in Figure 1A). The ketone groups of TES and dihydrotestosterone (DHT) that hydrogen bond with R752 overlap with the negatively charged NEG FragMaps (N1 in Figure 1A). Similarly, hydrogen bond donor (HBDON) and positively charged (POS) FragMaps close to N705 overlap with the 17 β -hydroxyl groups of the steroids.

Along with the apolar steroidal pocket, a second apolar cavity between the W741, L873, and T877 in the AR is occupied by some ligands such as the EM-5744 (PDB: 2PNU)²⁴ and the nonsteroidal S-1, an analog of R-bicalutamide²⁵ (PDB: 2AXA). Such a cavity is inaccessible in the PDB 2AM9 (SI Figure S2A). Through the simulation, side chains of W741 and T877 undergo conformational changes (SI Figure S2A,C) that lead to the formation of this cavity without significantly affecting the global conformation of the receptor (backbone RMSD \sim 1.2 Å, SI Figure S2B). Consequently, APOLAR FragMaps were found in this cavity (A2, Figure 1A). No such densities were found in a GCMC-only simulation in which the protein was rigid (Figure 1D), validating that protein flexibility through the

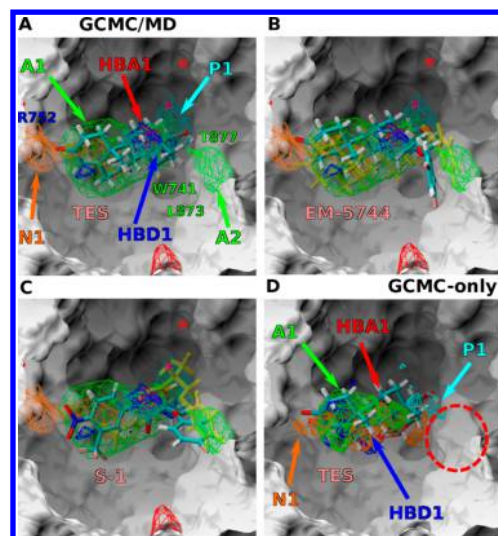


Figure 1. FragMaps overlaid on the LBP of AR (PDB 2AM9) with ligands A) TES, B) EM-5744, and C) S-1 in the crystallographic orientations. D) FragMaps from the GCMC only simulation. Receptor atoms occluding the view of the binding pocket were removed to facilitate visualization. The color for nonpolar (APOLAR), neutral donor (HBDON), neutral acceptor (HBACC), negative acceptor (NEG), and positive donor (POS) FragMaps are green, blue, red, orange, and cyan, respectively. APOLAR, HBACC, and HBDON FragMaps are set to a cutoff of -0.5 kcal/mol, while NEG and POS are set to -1.2 kcal/mol. Distinct FragMap affinities that overlap with the functional groups of the ligands are indicated by arrows colored the same as the FragMaps. D) The absence of protein flexibility in GCMC-only simulations leads to a general decrease in the spatial extent of the FragMaps and omission of the APOLAR FragMap A2 (red dashed circle) that is in the vicinity of the crystallographic conformations of the second phenyl rings of B) EM-5744 and C) S-1. MC sampling of ligands, EM-5744 and S-1 yields conformations (yellow) distinct from the crystal (cyan), in B and C.

inclusion of MD allows for the solutes to sample regions that were not available in the starting conformation. Additional evidence of the importance of inclusion of protein flexibility was the notable increase in the area sampled by APOLAR FragMaps at the A1 site in the GCMC/MD vs the GCMC-only simulation (Figure 1A vs 1D). This increase in sampling is driven by flexibilities of the side chains of the residues that form this pocket (SI Figure S2C). These results point to the qualitative ability of the SILCS-GCMC/MD approach to map the functional group requirements of a fully occluded LBP, including the ability to identify regions accessible to solutes significantly beyond those present in the crystal structure.

PPAR γ . The LBP of PPAR γ is much larger than that of AR. There are a total of 120 crystal structures of human WT PPAR γ , of which 68 structures have distinct ligands. Ten \times 50 ns SILCS-GCMC/MD was initiated with the PDB structure 3U9Q²⁶ after the deletion of decanoic acid. FragMaps not only sample the decanoic acid pocket (marked LBP1 in Figure 2A), but also sample the second pocket flanked between helices H3, H4 and β -sheets B2 and B3 (LBP2). Between the two, LBP1 is more occluded than LBP2 (SI Table S4), and, accordingly, some sampling of the LBP2 occurs with the rigid protein in the GCMC-only simulations, though an increase in the extent of the FragMaps is evident when protein flexibility is included via MD (SI Figure S3A vs B). In addition, FragMaps traced a pathway from the protein surface to the LBP, indicating a possible pathway for ligand binding (SI Figure S4).

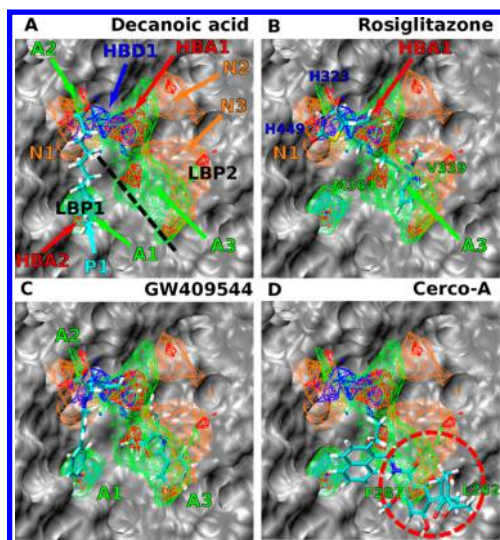


Figure 2. PPAR γ FragMaps overlaid on the LBP of PPAR γ (PDB 3U9Q) with ligands A) decanoic acid, B) Rosiglitazone (PDB: 2PRG,) C) GW409544 (PDB: 1K74), and D) Cerco-A (PDB: 3B1M) in their crystallographic orientations; receptor atoms occluding the view of the binding pocket were removed to facilitate visualization. HBACC and HBDON FragMaps are set to a cutoff of -0.5 kcal/mol, while APOLAR, NEG, and PDON FragMaps are set to a cutoff of -1.2 kcal/mol. No FragMaps were found to overlap with the dibenzofurancaboxamide of Cerco-A (D), marked with a red dashed circle.

Terminal alkyl chains and the carboxylic acid of decanoic acid overlapped well with the favorable APOLAR and the NEG densities (A1, N1, respectively) in the LBP1 (Figure 2A). Different functional groups of Rosiglitazone,²⁷ a known antidiabetic drug that binds to the PPAR γ in the LBP2 also overlap well with the FragMaps (Figure 2): 1) the thiazolidinedione moiety overlaps with the NEG FragMaps in the proximity of H323 and H449 (N1 in Figure 2B), 2) the pyridine overlaps with APOLAR FragMaps (A3) in the proximity of M364 and V339, and 3) the ethoxy linker between the thiazolidinedione and the pyridine overlaps with HBACC FragMaps (HBA1).

Along with these qualitative observations, binding affinities of the ligands were estimated using Ligand Grid Free Energy (LGFE) scoring for 16 ligands whose binding affinity for the human PPAR γ is available and compared against the experimental binding affinity, ΔG_{bind} (SI, Table S5). Both partial and full agonists of PPAR γ are considered and are noted to approximately occupy the pockets shown in Figure 2.

K_i s obtained from the different sources were normalized against the K_i of rosiglitazone ($K_i = 120$ nM; bindingDB reported a range of values between 8 and 440 nM).²⁸ Despite the diversity in the ligands and their binding modes as well as confounding contributions from the experimental studies, there is reasonable correlation between the LGFE and ΔG_{bind} values with a predictive index,²⁹ $PI \sim 0.63$ and $R^2 \sim 0.27$ (SI Figure S3C). For instance, GW409544 that binds into both the LBP1 and LBP2 (PDB: 1K74)³⁰ has very good overlap with APOLAR FragMaps A1, A2, and A3 (Figure 2C), leading to a favorable LGFE that correlates with its high binding affinity compared to the partial agonist decanoic acid or the thiazolidinediones such as the rosiglitazone. On the other hand, poor correlations are noted for Cerco-A³¹ (Figure 2D) driven by a lack of APOLAR FragMaps in the hydrophobic cavity between L262 and F287, where the dibenzofurancaboxamide functional group of Cerco-

A binds. This is caused by the loss of this hydrophobic cavity due to 1) the high flexibilities of the side chains of these residues through the simulations (SI Figure S5A) and 2) the conformation of the loop connecting helices H2 and H3 built using MODELLER (SI Text, Section S1). Although some of the side chains are flexible, the overall conformation of the receptor is preserved (SI Figure S5B). LGFE of ligands calculated using FragMaps from the GCMC-only simulations correlated poorly with the experimental ΔG_{bind} (SI Figure S3C vs D). Since SBDD procedures typically screen ligands based on estimation of the binding affinities, this result further validates the importance of incorporating protein flexibility in such studies.

mGluR. A crystal structure of the inactive conformation of the dimeric 7 TM region of the family C mGluR1 in complex with FITM, a negative allosteric modulator (NAM), was recently published.²⁰ To efficiently sample the partially occluded LBP, 10×50 ns SILCS-GCMC/MD was performed on a monomer of mGluR encompassed by palmitoyl oleoylphosphatidyl choline (POPC) membrane bilayer containing ~ 90 lipids with 10% cholesterol (see SI, Section S1 for details of protein preparation), with GCMC of the solutes and water restricted to a 20 \AA cubic region around the LBP. FragMaps correctly recapitulate the different functional groups of the FITM (Figure 3). The largely hydrophobic nature of the

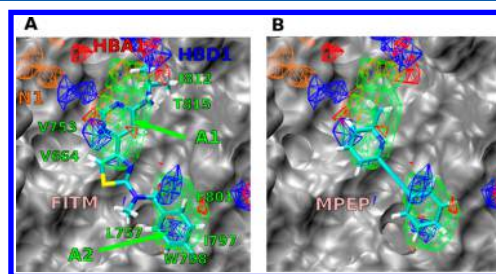


Figure 3. GFE FragMaps overlaid at the partially occluded LBP of mGluR1 with ligands A) FITM and B) MPEP. Protein atoms occluding the view of the pocket were removed to facilitate visualization. All FragMap contours are displayed at -1.2 kcal/mol, and the color scheme is described in Figure 1.

pocket is traced by the high affinities of the APOLAR FragMaps. Two distinct APOLAR densities overlap well with the pyrimidine and the *p*-fluorophenyl moieties of FITM, which are found in the neighborhood of V753, V664, and I812 (A1 in Figure 3A) and F801, I797, W798, and L757 (A2), respectively. HBDON FragMaps in the proximity of T815, noted to be important for binding,³² overlap well with the pyrimidine of FITM. Although some of these side chains are found to be flexible through the GCMC/MD simulation (SI Figure S6A), the overall conformation of the receptor, with the narrow binding pocket, is preserved (SI Figure S6B).

Mutational studies of the mGluR1 and binding activity with other allosteric modulators such as the 2-methyl-6-(phenylethynyl)pyridine (MPEP) (another known NAM) revealed that the binding pocket identified with FITM could be shared with other NAMs. MC sampling of MPEP in the pocket and in the presence of FragMaps (see SI Text, Section S6) yields binding modes similar to FITM (Figure 3). LGFEs for the FITM, MPEP, and other analogs of FITM correlated well with their ΔG_{bind} (SI Figure S7C). Although some sampling of the pocket occurs through a GCMC-only simulation (SI Figure S7A vs B), akin to PPAR γ , the

correlations were poor when LGFE scores were calculated using FragMaps from a GCMC-only simulation of the mGluR (SI Figure S7C vs D).

β_2 AR. Crystal structures of the inactive³³ and the active²¹ conformations of the family A GPCR are available, and two separate 10 × 50 ns SILCS-GCMC/MD were run with each structure. As with mGluR, these simulations were performed in an explicit bilayer. Hereafter B2I refers to the simulations starting with the inactive conformation (PDB: 2RH1) and B2A to simulations starting with the active conformation (PDB: 3P0G).

Good overlaps were obtained between the FragMaps and the crystallographic ligands for both B2I and B2A. APOLAR FragMap affinities close to the hydrophobic region defined by F289, V117, and A200 overlapped well with the benzoxazine and carbazol moieties of the BI-167107 and with carazolol, respectively (A1 in Figure 4A). Both POS and HBDON FragMaps adjacent to D113 overlapped with the amine functional groups in both the ligands (DN2 and D2 in Figure 4A). While the FragMaps recapitulate the location of respective ligands for both B2I and B2A, clear differences were found between the FragMaps for these two conformations. A second

APOLAR density was found close to the second hydrophobic pocket adjacent to W109, I309, and F193 only in B2A (A2, Figure 4A). The lack of this hydrophobic pocket in B2I is due to the relatively high flexibility of the I309 and F193 side chains in B2I as compared to B2A (SI Figure S8A). Despite these side-chain flexibilities, the overall active and inactive conformations of β_2 AR were retained (SI Figure S8B). Agonist interactions with I309 have been identified to be important for both β_2 AR selectivity and ligand activation through mutational studies.³⁴ Extensive interactions between the carbonyl oxygen, amine, and the hydroxyl groups of BI-167107 with S203, S204, and S207 in B2A were correctly recapitulated by HBDON and HBACC FragMaps, while the lone polar nitrogen of the carbazol heterocycle in carazolol was recapitulated by narrower HBDON maps in B2I (HBD1, Figure 4A). These differences may be attributed to the higher flexibilities of the S203 and S204 side chains in B2I (SI Figure S8A). Consistent with these differential FragMaps are mutational studies that have identified S203 and S204 to be important for agonist binding and receptor activation through their catecholamine hydroxyls.³⁵

Notably, the SILCS-GCMC/MD approach is able to quantitatively differentiate between the two states of β_2 AR as well. LGFE scores from MC sampling were obtained for a diverse range of agonists, partial agonists, and antagonists/inverse agonists. Good correlations were obtained between the LGFE scores and binding affinities of agonists and partial agonists (SI Table S7) with B2A ($R^2 \sim 0.46$, PI ~ 0.59 , Figure 4B-(1)), while the same set of ligands yield significantly worse correlations with the B2I FragMaps ($R^2 \sim 0.10$, PI ~ 0.31 , Figure 4B-(2)). Similarly, binding affinities of antagonists/inverse agonists correlated well with the LGFEs scored using B2I FragMaps ($R^2 \sim 0.45$, PI ~ 0.67 , Figure 4B-(4)), while poorer correlation was found with the LGFEs calculated using the B2A FragMaps ($R^2 \sim 0.11$, PI ~ 0.38 , Figure 4B-(3)). Consistent with the quality of the correlations, MC sampling of ligands yielded binding modes similar to the crystallographic BI-167107 and carazolol orientations (SI Figures S9, S10).

FragMaps Guided Ligand Screening. Differences in FragMaps between the two end states were used to guide virtual screening (VS) studies to identify new agonists for β_2 AR. Following the procedure described in Methods, 15 top scoring chemically diverse ligands (Figure S11) were selected from an *in silico* database containing about 1.8 million compounds. β_2 ARs are expressed on numerous cell types, including airway smooth muscle (ASM). Activation of β_2 AR in ASM causes bronchodilation,³⁶ and inhaled beta-agonists are the main therapy for asthma and other obstructive pulmonary diseases.²² Effect of the selected compounds from VS were studied through relaxation response of tracheal rings from mice lung samples.³⁶ This *ex vivo* method is a relatively high-throughput strategy and a better predictor of *in vivo* macromolecular disposition than *in vitro* studies.³⁷ Isoproterenol was used as positive control. Seven of the 16 ligands effected varying degrees of tracheal relaxation (Figure 4C), representing a 46% hit rate. We note that while these experiments do not confirm the binding of the ligands to the LBP of β_2 AR identified in the crystal structures, they do indicate a β_2 AR mediated relaxation, as ASM relaxation in the trachea from FVB/N mice is mediated via activation of G_s coupled GPCRs and there are only two known G_s -coupled GPCRs predominantly expressed on ASM, including the β_2 AR and E-Prostanoid (EP) receptors (activated by prostaglandin E2 (PGE2)).^{38,39}

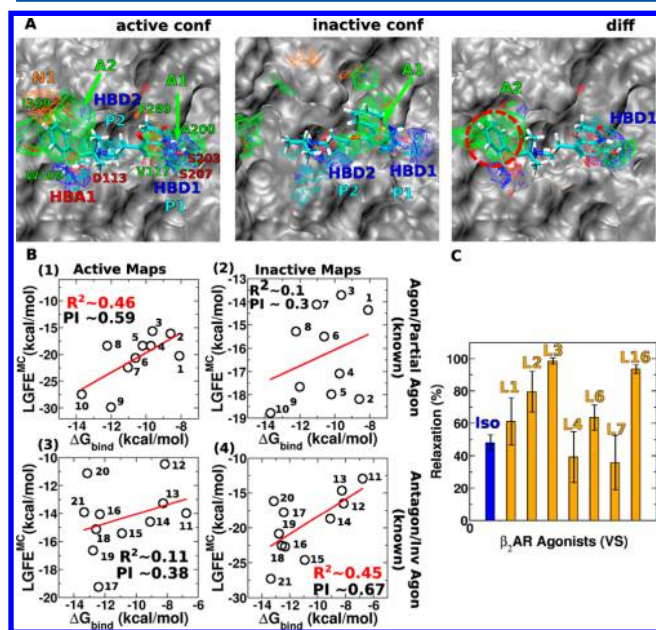


Figure 4. A) GFE FragMaps from the B2A (left) and B2I (center) simulations overlaid on the active (PDB: 3P0G) and inactive (PDB: 2RH1) states of the β_2 AR with ligands BI-167107 and carazolol, respectively; receptors atoms occluding the view of binding pocket were removed. Differential maps (right) highlight differences between the two states; red dashed circle highlights the location of the A2 FragMap that overlaps well with the agonist BI-167107. HBACC and HBDON FragMaps are set to a cutoff of -0.5 kcal/mol, while APOLAR, NEG, and PDON FragMaps are set to a cutoff of -1.2 kcal/mol, and the color scheme is described in Figure 1. B) LGFE scores are obtained from MC conformational ensembles of the known agonists, partial agonists (numbered 1–10), antagonists and inverse agonists (numbered 11–21) (see SI Table S7 for a list) in both the B2A and B2I FragMaps. LGFEs and the experimental ΔG_{bind} values correlate well for agonists and partial agonists with the B2A FragMaps (1), and the antagonists/inverse agonists with the B2I FragMaps (4). C) Relaxation response of tracheal rings with the top 7 of the 15 selected ligands from the virtual screening (VS) studies. Isoproterenol (Iso, blue) is used as control.

Docked conformations of the selected ligands had good overlaps with the distinct affinities in the B2A FragMaps leading to high LGFE scores (SI Figure S12). Since the screening procedure is designed to select ligands that preferentially stabilize the active conformation of β_2 AR by considering the differences between the active and inactive FragMaps, all the shortlisted ligands occupied the hydrophobic cavity defined by residues I309, W109, and F193 (apolar affinity A2 in Figure 4A, dashed red circle) and maintained interactions with Asp 113. Akin to previously known agonists, interaction with Asp113 is through a conserved amine moiety. Ligands that caused tracheal relaxation were seen to also maintain interactions with the S203 and S207 in helix H5 identified to be important for ligand activation and β_2 AR selectivity.^{4,21} While in previously known agonists, these interactions were through the dihydroxybenzene moiety, a range of functional groups are present in the current set of shortlisted ligands that caused tracheal relaxation. These results point to the utility of the SILCS-GCMC/MD methodology in rational ligand design, including identification of agonists. They also lead to future studies where we plan to probe ligand binding to β_2 AR through radioligand binding assays and biochemical methods to assess activation of adenylyl cyclase, generation of cyclic-adenosine monophosphate (cAMP) and activation of protein kinase A (PKA). *Ex vivo* studies will also be extended to validate reduction in tracheal relaxation in the presence of pharmacological antagonists of β -adrenergic receptors. Structural studies are also planned to determine ligand binding modes and receptor conformations in the presence of these ligands.

DISCUSSION

FragMaps from the SILCS-GCMC/MD are shown to successfully map the functional group affinity patterns of the occluded LBPs of the tested systems. This includes recapitulating the location of functional groups of diverse classes of known ligands as well as yielding reasonable correlations with their binding affinities. The importance of the inclusion of protein flexibility in the method is evidenced by the presence of FragMaps in regions of the LBP that were not accessible in the crystallographic conformations. This capability was most evident with the AR, opening up inaccessible sites important for the binding of known ligands, while for PPAR and the GPCRs the inclusion of protein flexibility allowed for fragment sampling of a larger volume of the LBP than that available in the crystal conformations (Figure 1, SI, Figures S3, S7, S13).

Among the different structures available for AR, backbone RMSD of the binding pocket residues with respect to PDB 2AM9¹⁸ varied between 0.29 Å (PDB: 2AX8⁴⁰) and 2.27 Å (PDB: 2PNU²⁴). However, considerable variability was seen in the binding pocket volumes across these different structures. Volume of the completely occluded binding pockets was measured using POVME,⁴¹ after the removal of ligand and water molecules from the crystal structures (SI, Figure S14A, C). Binding pocket cavity is the smallest in PDB 2AM9 with a volume of 127 Å³ and largest in PDB 2PNU²⁴ with a volume of 319 Å³. Despite starting the simulation with the smallest binding pocket structure, distribution of the pocket volumes through the SILCS-GCMC/MD simulation sampled a wide range (SI, Figure S14B) spanning the pocket volumes measured in all the other available crystal structures, driven in part by the opening of the second apolar cavity defined by residues W741, L873, and T877 through the simulation (SI, Figure S2A, C).

Similarly, for PPAR γ , pocket volumes in the different structures bound to the partial and full agonists considered in the LGFE calculations (SI, Table S5, Figure S3C) varied over a wide range: 377 Å³ (PDB: 3U9Q²⁶) to 627 Å³ (PDB: 3T03⁴²). Despite starting the SILCS-GCMC/MD simulation with PDB 3U9Q, distribution of the pocket volume through the simulation spanned the pocket volumes measured across most of these crystal structures (SI, Figure S14D). Conformations with larger pocket cavities, as seen when bound to GQ-16,⁴² Cerco-A⁴³ and a few others are not visited through the length of the simulation (SI, Figure S14D). As discussed before, these could be modulated by the starting conformation of the loop connecting helices H2 and H3 that was inserted into PDB 3U9Q using MODELLER. Additionally, modulation of pocket conformational dynamics due to interactions with retinoic acid receptor at the dimeric interface^{44–46} is not accounted in our simulation with the PPAR γ monomer. These factors could also be leading to poorer LGFE predictions for some of the partial agonists (SI, Figure S3C).

To further validate the effect of conformational variability of the starting structure on the FragMap patterns, a second GCMC/MD SILCS simulation of AR was run using the S-1 bound conformation from PDB 2AXA after the removal of the ligand. In this conformation, the second apolar cavity is accessible at the start of the simulation. FragMaps from this simulation were very similar to those obtained from using PDB 2AM9, as evidenced by reasonable overlap coefficients between the two FragMaps (SI, Figure S15), indicating a limited dependence of the SILCS simulation on the starting conformation of the receptors.

Beyond these qualitative improvements in the FragMaps, the presence of protein flexibility yields good correlations between LGFE scores and ΔG_{bind} across diverse classes of ligands for each receptor (Figure 4C, Figures S3C, S7C). When FragMaps from the rigid protein GCMC-only simulations were used, the correlations became consistently worse (Figures S3D, S7D, S16), validating the importance of incorporating protein flexibility in the SILCS method.

Incorporation of explicit solvent and the consequent competition between the solutes and water for binding sites on the protein in the SILCS-GCMC/MD methods allows for mapping water affinity patterns that is useful in identification of regions where water molecules are thermodynamically unfavorable to displace upon ligand binding.⁴⁷ For instance, in the crystal structure of the AR-S-1 complex (PDB: 2AXA), a water molecule mediates interactions between H874 and the backbone oxygen of W741. Although the AR simulations were performed with the testosterone bound conformation of the LBP and W741 was flexible through these simulations, favorable water affinities occur close to the H874 (SI Figure S17A). Similarly, simulations with the decanoic acid-bound conformation of the LBP of PPAR γ identify favorable water affinities close to S342, which align well with the crystal water that mediates interactions between S342 and GW409544 in PDB 1K74. In the case of AR, no other favorable FragMap regions were found to overlap with the water maps close to the H874, indicating that this water is unfavorable to displace. With the PPAR γ , NEG affinities were found to overlap with the favorable water maps (SI Figure S17B), while neutral HBACC were not present in that region indicating the capability of negatively charged but not neutral acceptors to displace the water. This analysis emphasizes the capability of SILCS-GCMC/MD FragMaps to identify regions where waters can

and cannot be displaced while simultaneously suggesting the chemical composition of functional groups that can displace water, information that will be of great utility in ligand design.

FragMaps were found to distinguish between the LBPs of the active and inactive states of the β_2 AR. This is notable, as conformational changes in the LBP between the two states are subtle, with the largest crystallographically identified structural difference occurring at the intracellular side of helices H5 and H6 that are topologically distant from the LBP. The differences in the FragMaps are related to altered side-chain flexibility that allowed for additional volume sampled in the active B2A state, while the global conformation of the receptor was retained. The differences in the FragMap patterns can distinguish the nature of a particular ligand (agonist vs antagonist/inverse agonist) through differences in correlations of LGFE scores and ΔG_{bind} . FragMaps differences were subsequently used in virtual screening studies to identify new high-efficacy agonists that effected relaxation of airway smooth muscle, further indicating the utility of the method in mapping functional group affinity patterns of different states of the GPCRs.

METHODS

Occludedness of the LBP. The extent by which the binding sites are occluded, referred to as occludedness, is calculated as the ratio of the solvent accessible surface area⁴⁸ (SASA) of the ligands when bound to the protein using the ligand-protein complex X-ray structures to its SASA when alone. As shown in SI Table S4, these ratios vary from 0, for testosterone bound to AR up to 0.13 for both Rosiglitazone bound to PPAR γ and Carazolol bound to β_2 AR. In comparison, this ratio is much higher (>0.25) for solvent exposed binding pockets studied previously using SILCS.^{9,49} Given this level of inaccessibility of LBPs and the importance of these classes of proteins as drug targets, they represent good systems to validate the ability of the SILCS-GCMC/MD approach to map the functional group requirements of the LBPs.

Computational Studies. Empirical force field parameters for the proteins were CHARMM36,⁵⁰ ligand molecules were derived from CGenFF,⁵¹ and water was treated using the TIP3P model.⁵² Protein preparation simulations were performed using CHARMM⁵³ (section S1, Table S1), while the production phase of the GCMC/MD simulations were performed using in-house code for the GCMC and GROMACS⁵⁴ for the MD portions of the calculations. The protein systems are immersed in an aqueous solution (along with the lipid membrane and cholesterol in case of the GPCRs) containing approximately 0.25 M of each of the small solutes: benzene, propane, acetaldehyde, methanol, formamide, imidazole, acetate, and methylammonium.

GCMC/MD. For each of protein system, ten GCMC/MD simulations were run. Each of these 10 simulations constituted 100 cycles of GCMC and MD (200 in the case of AR), with each cycle involving 100,000 steps of GCMC and 0.5 ns of MD, yielding a cumulative 100 million steps of GCMC and 500 ns of MD for every protein system (200 million steps of GCMC and 1 μ s of MD for AR). Random seed numbers used in the GCMC simulations are different across the 10 runs. The GCMC procedure is described in detail in our previous work¹⁶ and SI Text, Section S2. Through the 100,000 steps both the solute molecules (fragments) and waters are exchanged between their gas-phase reservoirs and a subvolume encapsulating the LBP of the protein. The configuration at the end of the GCMC is used as the starting configuration in the MD. Before

the production MD, a 500 step SD minimization and a 100 ps equilibration are run as described during the protein preparation simulations. See SI Text, Section S3 for production step details. The excess chemical potential (μ_{ex}) supplied to drive solute and water exchanges during GCMC are periodically fluctuated over every 3 cycles, such that the average μ_{ex} over the 100 cycles is close to the values in SI Table S2. These values are the magnitude of the μ_{ex} required to maintain 0.25 M of a solute in a bulk aqueous mixture devoid of any protein and are approximately equal to the hydration free energy. We note that by periodically fluctuating μ_{ex} the system is not a formal GC ensemble; however, by maintaining the average μ_{ex} constant over the length of the simulation the extent of deviation is minimal, as previously discussed.¹⁶ For all the proteins, the systems converged as evidenced by the overlap coefficients of the FragMaps (SI Text, Section S8, Figure S1, Table S3).

GCMC-Only Simulations. To probe the role of protein flexibility in modulating the FragMaps in the LBPs, a second set of GCMC-only simulations was run for all the protein systems. Like the GCMC/MD simulation, 10 independent runs are set up. Through the 100 cycles, no MD was run at the end of 100,000 steps of GCMC; the last configuration at the end of the 100,000 steps of GCMC was used as the input configuration for the next cycle. The rest of the protocol is maintained the same as with the GCMC/MD, including oscillating μ_{ex} around the values shown in SI Table S2 over every 3 cycles.

Virtual Screening (VS). An *in silico* database of about 1.8 million compounds which contains all accessible tautomers and protonation state of each compound from the CHEMBRIDGE (<http://chembridge.com>) and MAYBRIDGE (<http://maybridge.com>) databases was used for screening. 1) Distinct B2A FragMap affinities (A1, A2, P1, P2, HBD1, and HBD2 in Figure 4B) were converted into pharmacophore features (SILCS-Pharm) using a method that is an extension to our recently published work.⁵⁵ PHARMER⁵⁶ was used to annotate each of the ligands from the database with the SILCS-PHARM features and then match these with the features from the FragMaps based on the RMSD of an alignment of the matched features. A cutoff of RMSD < 1.6 Å yielded 11,119 ligands. These ligands were docked into both the active (PDB: 3P0G) and inactive (PDB: 2RH1) structures of β_2 AR using Autodock Vina,⁵⁷ and differences in scores of the top ranked conformations were calculated as $\Delta E_{\text{dock}} = |E_{\text{dock}}^{\text{act}}| - |E_{\text{dock}}^{\text{inact}}|$, where $E_{\text{dock}}^{\text{act}}$ and $E_{\text{dock}}^{\text{inact}}$ are scores from top ranked poses in the active and inactive structures, respectively. A second selection criterion of $\Delta E_{\text{dock}} > 0$, yielded 906 ligands. LGFEs of these 906 ligands were calculated as Boltzmann averages over 1000 steps MC sampling in fields of both active (LGFE^{act}) and inactive (LGFE^{inact}) FragMaps. Differences in these LGFE scores were calculated as $\Delta \text{LGFE} = |\text{LGFE}^{\text{act}}| - |\text{LGFE}^{\text{inact}}|$, and 109 ligands were selected with $\Delta \text{LGFE} > 0$. Chemical clustering⁵⁸ of these ligands was performed in MOE using the MACCS fingerprints and setting the compound similarity coefficient (overlap index) to 0.7 and the Tanimoto cutoff to 0.2. Representative structures from the resulting clusters were handpicked based on their chemical diversity for functional assessment studies. SI, Figure S11 shows the chemical structures of the shortlisted ligands along with those of some of known β_2 AR agonists (epinephrine, isoproterenol, and BI-1617107).

Ex Vivo Intact Airway Physiology. All mouse studies were approved by the Animal Care and Use Committee of UMB.

Five mm sections of trachea from FVB/N mice were excised and studied in an isometric myograph system (ADInstrument) as previously described.³⁶ A passive tension of 0.5 g was applied for each ring for a baseline. Rings were contracted with 1 μ M acetylcholine, followed by the addition of 100 μ M of isoproterenol (iso) and the selected ligands from VS studies. Percentage of relaxation in the presence of a ligand was measured as change in tension from acetylcholine-stimulated peak contraction. Rings were washed for reuse. Relaxation was calculated as an average over 9 runs with isoproterenol and 4 runs with each of the selected ligands.

■ ASSOCIATED CONTENT

■ Supporting Information

Protein preparation simulations details; GCMC and MD simulations details; bulk-phase GCMC/MD for calculation of μ_{ex} of SILCS fragments; FragMaps preparation; LGFE scoring methodology; overlap coefficient to estimate convergence of GCMC/MD; List of known ligands and experimental binding affinities for PPAR γ , mGluR, and β_2 AR; Structural analysis for AR, PPAR γ , mGluR, and β_2 AR; FragMaps and LGFE vs ΔG_{bind} between GCMC-only and GCMC/MD simulations for PPAR γ , mGluR, and β_2 AR; passage of ligand entry in PPAR γ ; list of selected ligands from VS of β_2 AR; binding modes of some known and shortlisted ligands from VS studies for the β_2 AR; pocket volume distributions from the AR and PPAR γ SILCS simulations; AR FragMaps obtained from separate SILCS-GCMC/MD simulations initiated with different PDB structures; FragMaps tracing water oxygens in AR and PPAR γ . This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: alex@outerbanks.umaryland.edu.

Author Contributions

S.K.L. developed and implemented the GCMC/MD method, ran simulations with the GPCRs, and wrote the manuscript with help from all the authors. W.Y. and E.P.R. ran simulations with the NRs. A.V.H. performed *ex vivo* studies. L.F. ran PHARMER simulations in the virtual screening studies. D.A.D. developed the protocol and analyzed data from *ex vivo* studies. S.K.L, W.Y, E.P.R, and A.D.M analyzed computational studies data. A.D.M. conceived and initiated the research.

Notes

The authors declare the following competing financial interest(s): A.D.M. is cofounder and Chief Scientific Officer of SilcsBio LLC. A provisional patent has been filed for the β_2 AR ligands identified through this study.

■ ACKNOWLEDGMENTS

This work was supported by NIH grants GM051501 and CA107331 (ADM), HL104119, AG041265 (DAD), and Maryland Industrial Partnerships Award 5212. The authors acknowledge computer time and resources from the Computer Aided Drug Design (CADD) Center at the University of Maryland, Baltimore.

■ REFERENCES

- (1) Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. How Many Drug Targets Are There? *Nat. Rev. Drug Discovery* **2006**, *5*, 993–996.
- (2) Venkatakrishnan, A.; Deupi, X.; Lebon, G.; Tate, C. G.; Schertler, G. F.; Babu, M. M. Molecular Signatures of G-Protein-Coupled Receptors. *Nature* **2013**, *494*, 185–194.
- (3) Markov, G. V.; Laudet, V. Origin and Evolution of the Ligand-Binding Ability of Nuclear Receptors. *Mol. Cell. Endocrinol.* **2011**, *334*, 21–30.
- (4) 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 β -2 Adrenergic G Protein Coupled Receptor. *ACS Chem. Biol.* **2013**, *8*, 1018–1026.
- (5) Park, S.-J.; Kufareva, I.; Abagyan, R. Improved Docking, Screening and Selectivity Prediction for Small Molecule Nuclear Receptor Modulators Using Conformational Ensembles. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 459–471.
- (6) Gronemeyer, H.; Gustafsson, J.-Å.; Laudet, V. Principles for Modulation of the Nuclear Receptor Superfamily. *Nat. Rev. Drug Discovery* **2004**, *3*, 950–964.
- (7) Kobilka, B. The Structural Basis of G-Protein-Coupled Receptor Signaling (Nobel Lecture). *Angew. Chem., Int. Ed.* **2013**, *52*, 6380–6388.
- (8) Guvench, O.; MacKerell, A. D. Computational Fragment-Based Binding Site Identification by Ligand Competitive Saturation. *PLOS Comput. Biol.* **2009**, *5*, e1000435.
- (9) Raman, E. P.; Yu, W.; Lakkaraju, S. K.; MacKerell, A. D., Jr. Inclusion of Multiple Fragment Types in the Site Identification by Ligand Competitive Saturation (SILCS) Approach. *J. Chem. Inf. Model.* **2013**, *53*, 3384–3398.
- (10) Carlson, H. A. Protein Flexibility is an Important Component of Structure-Based Drug Discovery. *Curr. Pharm. Des.* **2002**, *8*, 1571–1578.
- (11) Mobley, D. L.; Chodera, J. D.; Dill, K. A. Confine-and-Release Method: Obtaining Correct Binding Free Energies in the Presence of Protein Conformational Change. *J. Chem. Theory Comput.* **2007**, *3*, 1231–1235.
- (12) Nygaard, R.; Zou, Y.; Dror, R. O.; Mildorf, T. J.; Arlow, D. H.; Manglik, A.; Pan, A. C.; Liu, C. W.; Fung, J. J.; Bokoch, M. P. The Dynamic Process of β_2 -Adrenergic Receptor Activation. *Cell* **2013**, *152*, 532–542.
- (13) Nguyen, C. N.; Young, T. K.; Gilson, M. K. Grid Inhomogeneous Solvation Theory: Hydration Structure and Thermodynamics of the Miniature Receptor Cucurbit [7] Uril. *J. Chem. Phys.* **2012**, *137*, 044101.
- (14) Yu, W.-M.; Guvench, O.; MacKerell, A. D., Jr.; Qu, C.-K. Identification of Small Molecular Weight Inhibitors of Src Homology 2 Domain-Containing Tyrosine Phosphatase 2 (SHP-2) Via In Silico Database Screening Combined with Experimental Assay. *J. Med. Chem.* **2008**, *51*, 7396–7404.
- (15) Cao, X.; Yap, J. L.; Newell-Rogers, M. K.; Peddaboina, C.; Jiang, W.; Papaconstantinou, H. T.; Jupiter, D.; Rai, A.; Jung, K.-Y.; Tubin, R. P. The Novel BH3 α -helix Mimetic JY-1-106 Induces Apoptosis in a Subset of Cancer Cells (Lung Cancer, Colon Cancer and Mesothelioma) by Disrupting Bcl-xL and Mcl-1 Protein–Protein Interactions with Bak. *Mol. Cancer* **2013**, *12*, 42.
- (16) Lakkaraju, S. K.; Raman, E. P.; Yu, W.; MacKerell, A. D. Sampling of Organic Solutes in Aqueous and Heterogeneous Environments using Oscillating Excess Chemical Potentials in Grand Canonical-Like Monte Carlo-Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 2281–2290.
- (17) Eyrisch, S.; Helms, V. Transient Pockets on Protein Surfaces Involved in Protein-Protein Interaction. *J. Med. Chem.* **2007**, *50*, 3457–3464.
- (18) Pereira de Jesus-Tran, K.; Côté, P. L.; Cantin, L.; Blanchet, J.; Labrie, F.; Breton, R. Comparison of Crystal Structures of Human Androgen Receptor Ligand-Binding Domain Complexed with Various Agonists Reveals Molecular Determinants Responsible For Binding Affinity. *Protein Sci.* **2006**, *15*, 987–999.
- (19) Nolte, R. T.; Wisely, G. B.; Westin, S.; Cobb, J. E.; Lambert, M. H.; Kurokawa, R.; Rosenfeld, M. G.; Willson, T. M.; Glass, C. K.

Milburn, M. V. Ligand Binding and Co-Activator Assembly of the Peroxisome Proliferator-Activated Receptor- γ . *Nature* **1998**, *395*, 137–143.

(20) Wu, H.; Wang, C.; Gregory, K. J.; Han, G. W.; Cho, H. P.; Xia, Y.; Niswender, C. M.; Katritch, V.; Meiler, J.; Cherezov, V. Structure of a Class C GPCR Metabotropic Glutamate Receptor 1 Bound to an Allosteric Modulator. *Science* **2014**, *344*, 58–64.

(21) Rasmussen, S. G.; Choi, H.-J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; DeVree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S. Structure of a Nanobody-Stabilized Active State of the β_2 Adrenoceptor. *Nature* **2011**, *469*, 175–180.

(22) Liggett, S. B.; Meyers, D. A. *The genetics of asthma*; CRC Press: 1996.

(23) Green, S. A.; Turki, J.; Bejarano, P.; Hall, I. P.; Liggett, S. B. Influence of β_2 -Adrenergic Receptor Genotypes on Signal Transduction in Human Airway Smooth Muscle Cells. *Am. J. Respir. Cell Mol. Biol.* **1995**, *13*, 25–33.

(24) Cantin, L.; Faucher, F.; Couture, J.-F.; de Jésus-Tran, K. P.; Legrand, P.; Ciobanu, L. C.; Fréchette, Y.; Labrecque, R.; Singh, S. M.; Labrie, F. Structural Characterization of the Human Androgen Receptor Ligand-Binding Domain Complexed with EM5744, A Rationally Designed Steroidal Ligand Bearing a Bulky Chain Directed Toward Helix 12. *J. Biol. Chem.* **2007**, *282*, 30910–30919.

(25) Masiello, D.; Cheng, S.; Buble, G. J.; Lu, M. L.; Balk, S. P. Bicalutamide Functions As An Androgen Receptor Antagonist by Assembly of a Transcriptionally Inactive Receptor. *J. Biol. Chem.* **2002**, *277*, 26321–26326.

(26) Malapaka, R. R.; Khoo, S.; Zhang, J.; Choi, J. H.; Zhou, X. E.; Xu, Y.; Gong, Y.; Li, J.; Yong, E.-L.; Chalmers, M. J. Identification and Mechanism of Ten-Carbon Fatty Acid As Modulating Ligand of Peroxisome Proliferator-Activated Receptors. *J. Biol. Chem.* **2012**, *287*, 183–195.

(27) Mayerson, A. B.; Hundal, R. S.; Dufour, S.; Lebon, V.; Befroy, D.; Cline, G. W.; Enoksson, S.; Inzucchi, S. E.; Shulman, G. I.; Petersen, K. F. The Effects of Rosiglitazone On Insulin Sensitivity, Lipolysis, and Hepatic and Skeletal Muscle Triglyceride Content in Patients with Type 2 Diabetes. *Diabetes* **2002**, *51*, 797–802.

(28) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: A Web-Accessible Database of Experimentally Determined Protein–Ligand Binding Affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201.

(29) Pearlman, D. A.; Charifson, P. S. Are Free Energy Calculations Useful In Practice? A Comparison with Rapid Scoring Functions for the p38 MAP Kinase Protein System. *J. Med. Chem.* **2001**, *44*, 3417–3423.

(30) Xu, H. E.; Lambert, M. H.; Montana, V. G.; Plunket, K. D.; Moore, L. B.; Collins, J. L.; Oplinger, J. A.; Klier, S. A.; Gampe, R. T.; McKee, D. D. Structural Determinants of Ligand Binding Selectivity Between the Peroxisome Proliferator-Activated Receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 13919–13924.

(31) Thor, M.; Beierlein, K.; Dykes, G.; Gustavsson, A.-L.; Heidrich, J.; Jendeborg, L.; Lindqvist, B.; Pegurier, C.; Roussel, P.; Slater, M. Synthesis and Pharmacological Evaluation of a New Class of Peroxisome Proliferator-Activated Receptor Modulators. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3565–3567.

(32) Suzuki, G.; Kimura, T.; Satow, A.; Kaneko, N.; Fukuda, J.; Hikichi, H.; Sakai, N.; Maehara, S.; Kawagoe-Takaki, H.; Hata, M. Pharmacological Characterization of a New, Orally Active and Potent Allosteric Metabotropic Glutamate Receptor 1 Antagonist, 4-[1-(2-fluoropyridin-3-yl)-5-methyl-1H-1,2,3-triazol-4-yl]-N-isopropyl-N-methyl-3,6-dihydropyridine-1(2H)-carboxamide (FTIDC). *J. Pharmacol. Exp. Ther.* **2007**, *321*, 1144–1153.

(33) Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K. High-Resolution Crystal Structure of an Engineered Human β_2 -Adrenergic G Protein–Coupled Receptor. *Science* **2007**, *318*, 1258–1265.

(34) Kikkawa, H.; Isogaya, M.; Nagao, T.; Kurose, H. The Role of the Seventh Transmembrane Region in High Affinity Binding of a β_2 -Selective Agonist TA-2005. *Mol. Pharmacol.* **1998**, *53*, 128–134.

(35) Strader, C. D.; Sigal, I. S.; Register, R. B.; Candelore, M. R.; Rands, E.; Dixon, R. Identification of Residues Required For Ligand Binding To the Beta-Adrenergic Receptor. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 4384–4388.

(36) Deshpande, D. A.; Wang, W. C.; McIlmoyle, E. L.; Robinett, K. S.; Schillinger, R. M.; An, S. S.; Sham, J. S.; Liggett, S. B. Bitter Taste Receptors on Airway Smooth Muscle Bronchodilate by Localized Calcium Signaling and Reverse Obstruction. *Nat. Med.* **2010**, *16*, 1299–1304.

(37) Deshpande, D. A.; Yan, H.; Kong, K.-C.; Tiegs, B. C.; Morgan, S. J.; Pera, T.; Panettieri, R. A.; Eckhart, A. D.; Penn, R. B. Exploiting Functional Domains of GRK2/3 To Alter the Competitive Balance of Pro- and Anticontractile Signaling in Airway Smooth Muscle. *FASEB J.* **2014**, *28*, 956–965.

(38) Burgess, J. K.; Ge, Q.; Boustany, S.; Black, J. L.; Johnson, P. R. Increased Sensitivity of Asthmatic Airway Smooth Muscle Cells to Prostaglandin E2 Might be Mediated by Increased Numbers of E-Prostanoid Receptors. *J. Allergy Clin. Immunol.* **2004**, *113*, 876–881.

(39) Billington, C. K.; Penn, R. B. Signaling and Regulation of G Protein-Coupled Receptors in Airway Smooth Muscle. *Respir. Res.* **2003**, *4*.

(40) Bohl, C. E.; Miller, D. D.; Chen, J.; Bell, C. E.; Dalton, J. T. Structural Basis For Accommodation of Nonsteroidal Ligands in the Androgen Receptor. *J. Biol. Chem.* **2005**, *280*, 37747–37754.

(41) Durrant, J. D.; Votapka, L.; Sørensen, J.; Amaro, R. E. POVME 2.0: An Enhanced Tool for Determining Pocket Shape and Volume Characteristics. *J. Chem. Theory Comput.* **2014**, *10*, 5047–5056.

(42) Amato, A. A.; Rajagopalan, S.; Lin, J. Z.; Carvalho, B. M.; Figueira, A. C.; Lu, J.; Ayers, S. D.; Mottin, M.; Silveira, R. L.; Souza, P. C. GQ-16, A Novel Peroxisome Proliferator-Activated Receptor γ (PPAR γ) Ligand, Promotes Insulin Sensitization Without Weight Gain. *J. Biol. Chem.* **2012**, *287*, 28169–28179.

(43) Wakabayashi, K.; Hayashi, S.; Matsui, Y.; Matsumoto, T.; Furukawa, A.; Kuroha, M.; Tanaka, N.; Inaba, T.; Kanda, S.; Tanaka, J. Pharmacology and in Vitro Profiling of a Novel Peroxisome Proliferator-Activated Receptor. GAMMA. Ligand, Cerco-A. *Biol. Pharm. Bull.* **2011**, *34*, 1094–1104.

(44) Hughes, T. S.; Chalmers, M. J.; Novick, S.; Kuruvilla, D. S.; Chang, M. R.; Kamenecka, T. M.; Rance, M.; Johnson, B. A.; Burris, T. P.; Griffin, P. R. Ligand and Receptor Dynamics Contribute to the Mechanism of Graded PPAR γ Agonism. *Structure* **2012**, *20*, 139–150.

(45) Pirat, C.; Farce, A.; Lebègue, N.; Renault, N.; Furman, C.; Millet, R.; Yous, S.; Specq, S.; Berthelot, P.; Desreumaux, P. Targeting Peroxisome Proliferator-Activated Receptors (PPARs): Development of Modulators. *J. Med. Chem.* **2012**, *55*, 4027–4061.

(46) Harmon, G. S.; Lam, M. T.; Glass, C. K. PPARs and Lipid Ligands in Inflammation and Metabolism. *Chem. Rev.* **2011**, *111*, 6321–6340.

(47) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817–2831.

(48) Richmond, T. J. Solvent Accessible Surface Area and Excluded Volume in Proteins: Analytical Equations for Overlapping Spheres and Implications for the Hydrophobic Effect. *J. Mol. Biol.* **1984**, *178*, 63–89.

(49) Raman, E. P.; Yu, W.; Guvench, O.; MacKerell, A. D., Jr. Reproducing Crystal Binding Modes of Ligand Functional Groups Using Site-Identification by Ligand Competitive Saturation (SILCS) Simulations. *J. Chem. Inf. Model.* **2011**, *51*, 877–896.

(50) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell, A. D., Jr. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(51) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.

CHARMM General Force Field: A Force Field For Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

(52) Durell, S. R.; Brooks, B. R.; Ben-Naim, A. Solvent-Induced Forces Between Two Hydrophilic Groups. *J. Phys. Chem.* **1994**, *98*, 2198–2202.

(53) Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.

(54) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(55) Yu, W.; Lakkaraju, S.; Raman, E. P.; MacKerell, A., Jr. Site-Identification by Ligand Competitive Saturation (SILCS) Assisted Pharmacophore Modeling. *J. Comput.-Aided Mol. Des.* **2014**, 1–17.

(56) Koes, D. R.; Camacho, C. J. Pharmer: Efficient and Exact Pharmacophore Search. *J. Chem. Inf. Model.* **2011**, *51*, 1307–1314.

(57) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.

(58) Godden, J. W.; Stahura, F. L.; Bajorath, J. Anatomy of Fingerprint Search Calculations on Structurally Diverse Sets of Active Compounds. *J. Chem. Inf. Model.* **2005**, *45*, 1812–1819.