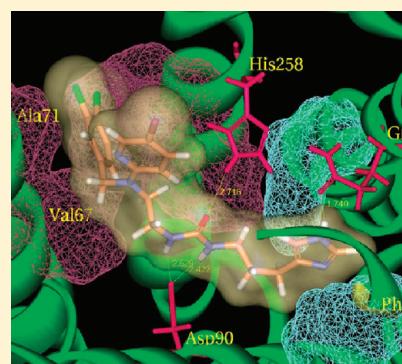


A Structure-Based Approach to Understanding Somatostatin Receptor-4 Agonism (sst4)

Zhaomin Liu,[†] A. Michael Crider,^{†,‡} Daniel Ansbro,[‡] Christina Hayes,[‡] and Maria Kontoyianni*,^{†,‡}

[†]Department of Chemistry and [‡]Department of Pharmaceutical Sciences, Southern Illinois University Edwardsville, Edwardsville, Illinois 62026, United States

ABSTRACT: It has been reported that somatostatin receptor subtypes 4 and 5 would be high-impact templates for homology modeling if their 3D structures became available. We have generated a homology model of the somatostatin receptor subtype 4 (sst4), using the newest active state β_2 adrenoreceptor crystal structure, and subsequently docked a variety of agonists into the model-built receptor to elucidate the binding modes of reported agonists. Using experimental restraints, we were able to explain observed activity profiles. We propose two binding modes that can consistently explain findings for high-affinity agonists and reason why certain structures display low affinities for the receptor.



INTRODUCTION

G protein coupled receptors (GPCRs) are integral membrane proteins targeted by an estimated 27% of the current rule-of-five-compliant experimental and marketed drugs.¹ Besides being the drug targets of many pharmaceuticals, GPCRs also represent an opportunity for future therapeutic intervention based on the druggable genome.^{2,3}

The GPCR superfamily is divided into three main families: A, B, and C, which share no sequence homology with one another, although their structures are similar. The largest is family A (class I or rhodopsin-like) and accounts for almost 85% of the GPCR genes. Over half of the class I GPCRs is predicted to encode olfactory targets, while the remaining are bound to known endogenous compounds or are classified as orphan receptors. Furthermore, the rhodopsin family is subdivided into four groups, α , β , γ , and δ , and 13 sub-branches. Even though there is no overall correlation between phylogeny and the types of ligands that bind to the rhodopsin family,¹ a few general receptor clusters can be observed. Family B (class II) corresponds to the secretin/adhesion receptor family, and family C (class III) is the metabotropic glutamate/pheromone protein family.^{4–6}

GPCRs are involved in signaling cascades that regulate several cellular and physiological functions. The ligands that bind to these receptors range from ions to photons, peptides, proteins, lipids, and nucleotides.⁷ Activation induced by agonists causes changes in the relative orientation of the transmembrane helices (TMs), leading to a wider intracellular surface with an exposure of those intracellular amino acids that are crucial for signal transduction. Inverse agonists and antagonists bind to a number of different sites but prevent the TM helix reorientation. Given that a conformational equilibrium between active and inactive states exists for

receptors, it is believed that the agonist-bound conformation of GPCRs will shift the equilibrium toward the active receptor states with helices undergoing significant movements during activation.^{8,9} Specifically upon agonist binding, the GPCRs undergo conformational changes, leading to regulation of the associated G protein ($G\alpha\beta\gamma$) by stimulating the GDP–GTP exchange. The resulting GTP-bound α subunit dissociates from the β and γ subunits and subsequently modulates several intracellular effectors. Effectors such as adenylyl cyclases, phospholipase C, and phosphoinositide-3 kinases in turn regulate the production of secondary messengers and impact intracellular signaling proteins or target functional proteins.^{10,11}

Structure determination for these systems has been challenging due to their heterologous expression. Recombinant production of GPCRs is still a matter of trial and error, while their folding and stability, when expressed in their native system, are complex and not yet fully understood processes. Furthermore, although milligram amounts of certain GPCRs are attainable, the majority of them are still either produced at very low levels or not at all.¹²

The first crystal structure of a GPCR appeared 11 years ago and corresponded to bovine rhodopsin with the covalently bound endogenous chromophore *cis*-retinal in the inactive state.¹³ Because rhodopsin can be easily obtained in high quantities, several other rhodopsin crystal structures have been published since then. However, rhodopsin is a light-activated GPCR, and therefore such a template is inadequate for any homology modeling of ligand-activated GPCRs.^{14,15} The first non-rhodopsin crystal structure was reported for the β_2 adrenergic receptor bound to a partially inverse agonist,

Received: August 6, 2011

Published: December 12, 2011



carazolol, seven years later.^{16–19} Other non-rhodopsin structures were later resolved: the β_2 adrenergic receptor bound to cholesterol and a partially inverse agonist timolol,²⁰ the mutant version of the β_1 adrenergic receptor with cyanopindolol,²¹ the A_{2A} adenosine receptor with the high-affinity antagonist ZM241385,²² the human dopamine D₃,²³ and CXCR4 chemokine²⁴ receptors. It should be noted that the chemokine receptor belongs to the γ group of rhodopsin-like receptors, whereas the others belong to the α group. The CXCR4 structure has a different topology and binding pocket, which in turn could provide a better template for GPCR homology modeling of the γ group.

While the above structures can be useful as templates for the antagonist-bound conformations of GPCRs, the *trans*-retinal/rhodopsin complex structure is still evasive, as are agonist-bound GPCR conformations, with few exceptions reported this year. Notable examples of informative activated crystal structures include a ligand-free, but with a G protein bound, opsin which is believed to be in an active conformation and has provided a better understanding of the requirements for full activation.²⁵ Also, three agonist-bound structures of the thermostable turkey β_1 agonist-bound adrenergic receptor,²⁶ the human A_{2A} agonist-bound adenosine receptor,²⁷ and the human β_2 agonist bound active state adrenoreceptor²⁸ were deposited in the PDB this year. Notably, β_2 was the first GPCR ever to be crystallized with an agonist and a peptide resembling the G protein, whereas the other two are only agonist-bound. Even more recently, the β_2 adrenergic receptor–Gs protein complex was released.²⁹

Given the preceding discussion, it becomes apparent that additional GPCR crystal structures are needed for drug design and in order to investigate more direct templates for model-building of other closely related GPCR classes. However, with the latest reported agonist-bound structures, studies on agonist-receptor activation and agonist discovery programs have become more feasible.

The somatostatin receptors are rhodopsin-like GPCRs, consisting of five subtypes (sst 1–5). They mediate the inhibitory effects of somatostatin on secretion and proliferation, depending on receptor subtype and tissue localization.³⁰ Most importantly, sst4 has been recognized as an ideal therapeutic target for Alzheimer's disease, due to its high expression in neocortical and hippocampal areas, which are significantly affected by amyloid β accumulation.^{31,32} Since there is no crystal structure, studies on sst4 ligand binding and activation have been scarce. There are, however, reports that aspartic acid in TM3 is essential for ligand binding.^{33,34}

Somatostatin (somatotropin release-inhibiting factor, SRIF) is a hormone peptide, which is normally expressed as a tetradecapeptide (SRIF-14) or an N-terminally extended form (SRIF-28). Both SRIFs contain a disulfide bond between cysteines at positions 3 and 14 which stabilize the structure (Figure 1). SRIF peptidic structure–activity relationship (SAR) studies indicated that the core residues Trp8 and Lys9 are the essential binding sites for all somatostatin receptors.³² Furthermore, alanine scanning studies carried out by Lewis et al. demonstrated that Trp8 and Lys9 were essential for binding to all sst subtypes, whereas Phe6 is specifically important for sst4 activation.^{35,36}

Due to poor oral bioavailability and rapid degradation by peptidases, SRIF is not viable for therapeutic application. Consequently, the development of orally effective, metabolically stable non-peptide compounds has been the focus of many

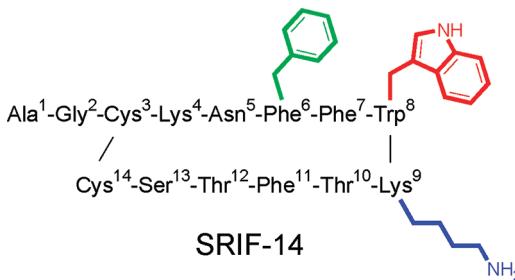


Figure 1. Structure of SRIF-14. The side chains of Phe6, Trp8, and Lys9 are shown.

research groups. On the basis of SAR data, compound series have been designed to mimic Trp8, Lys9, and Phe6. Ankersen et al.³⁷ were the first to report a non-peptide ligand with a 6 nM K_i and over 100-fold selectivity toward sst4. Many other compounds have also been designed to mimic SRIF with varying binding affinities, ranging from 0.7 nM to 10 μ M K_i values. However, it is not clear or obvious why such structurally similar compounds have so diverse binding affinities and how the receptor can also accommodate substantially diverse ligands. Are there distinct binding modes that could possibly explain observed assay data? Is there a possibility that high-affinity compounds adopt more specific modes with an increased number of binding and/or stabilizing interactions, as opposed to low-affinity molecules? To address these questions, we generated a homology model of sst4 and docked a number of reported compounds into the model-built structure. Inspection of the resultant receptor–agonist complexes led us to propose two partially overlapping, but not identical, binding modes for the high- and low-affinity sst4 agonists, which are presented here. It should also be pointed out that the high affinity binding we are proposing is consistent with experimental data for residues thought to be critical for molecular recognition at the sst4 level. Finally, a virtual screening experiment was carried out to validate the effectiveness of the model and its ability to retrieve actives seeded in a compound collection of 996 decoys.

■ COMPUTATIONAL METHODS

sst4 Model Building. The amino acid sequence of somatostatin receptor type 4 (sst4) was taken from NCBI; its sequence accession code is NP_001043. All calculations were performed within the Discovery Studio 2.5.5 suite of protocols (Accelrys Inc., San Diego, CA) and Maestro 9.0.211 (Schrödinger, LLC, New York, NY). Computations were run on a quad core Intel 3.0 GHz Xeon X5472 processor.

A protein-BLAST search at NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was employed in order to identify the most suitable templates for model building.^{38–43} The BLAST-P search algorithm resulted in the homologues presented in Table 1. Total scores were used to select the most appropriate template for our purposes, and as such the active state of the β_2 adrenergic receptor (PDB code: 3P0G) was selected for model construction.

Sequence alignment between sst4 and 3P0G sequences was performed using ClustalW³⁸ and was further optimized manually. This alignment was employed for model construction with the comparative modeling software MODELER.^{44,45} Five models were generated per run, with the optimization level set to high. Both the N- and C-terminal residues in the target

Table 1. NCBI BLAST Search Results

accession	description	total score	E value
3OE6_A	chain A, crystal structure of the CXCR4 chemokine receptor in complex with a small molecule antagonist	172	3e-31
3ODU_A	chain A, the 2.5 Å structure of the CXCR4 chemokine receptor in complex with small molecule antagonist	172	3e-31
3OE0_A	chain A, crystal structure of the CXCR4 chemokine receptor in complex with a cyclic peptide antagonist CVX15	171	4e-31
3P0G_A	chain A, structure of a nanobody-stabilized active state of the β_2 adrenoceptor	133	1e-16
2RH1_A	chain A, high resolution crystal structure of human β_2 -adrenergic receptor	133	1e-16
3D4S_A	chain A, cholesterol bound form of human β_2 adrenergic receptor	132	3e-17
3PDS_A	chain A, irreversible agonist- β_2 adrenoceptor complex	122	2e-14
3PBL_A	chain A, structure of the human dopamine d3 receptor in complex with eticlopride	115	3e-10
3KJ6_A	chain A, crystal structure of a methylated β_2 adrenergic receptor–Fab complex	113	2e-25
2R4R_A	chain A, crystal structure of the human β_2 adrenoceptor	113	2e-25
2Y00_A	chain A, turkey β_1 adrenergic receptor with stabilizing mutations and bound partial agonist dobutamine.	108	5e-24
2VT4_A	chain A, turkey β_1 adrenergic receptor with stabilizing mutations	107	1e-23
2KS9_A	chain A, solution conformation of substance p in water complexed with NK1R	104	6e-23
2R4S_A	chain A, crystal structure of the human β_2 adrenoceptor	104	5e-23
1JFP_A	chain A, structure of bovine rhodopsin	95.1	5e-20
3OAX_A	chain A, crystal structure of bovine rhodopsin with β -ionone	95.1	5e-20
2J4Y_A	chain A, crystal structure of a rhodopsin stabilizing mutant expressed in mammalian	92.8	2e-19
2 × 72_A	chain A, crystal structure of the constitutively active E113Q, N2C, D282C rhodopsin mutant with bound galphact peptide	92.4	3e-19
3C9M_A	chain A, structure of a mutant bovine rhodopsin in hexagonal crystal structure	92.4	3e-19
2Z73_A	chain A, crystal structure of squid rhodopsin	88.2	6e-18
2ZIY_A	chain A, crystal structure of squid rhodopsin	88.2	5e-18
3EML_A	chain A, the 2.6 Å crystal structure of a human a2a adenosine receptor bound to ZM241385	82.8	4e-05
2YDO_A	chain A, thermostabilized human a2a receptor with adenosine bound	69.7	2e-12
3E05_A	chain A, crystal structure of precorrin-6y C5,15-methyltransferase	31.2	0.81
2DFL_A	chain A, crystal structure of left-handed rada filament	27.7	9.0

Table 2. DOPE, Verify Protein 3D, and PROCHECK Scores, before and after Minimization

DOPE score	Verify Protein (Profile-3D)			PROCHECK			
	verify score	expected range	core regions	allowed regions	generous regions	disallowed regions	
	before refinement						
Model sst4	-37049.83	55.74	60.19–133.76	96.2%	3.4%	0.0%	0.4%
				after refinement			
Model sst4	-38058.65	128.74	60.19–133.76	96.2%	3.4%	0.0%	0.4%

sequence, which were not aligned with the template, were cut. All models were built with a medium-level loop refinement, and the discrete optimized protein energy (DOPE) potentials⁴⁶ were calculated for subsequent evaluation.

Models were further refined with limited minimization and molecular dynamics (MD) simulations and constrained backbone. The CHARMM forcefield was employed, while the partial charges were calculated using the Momany Rone method.⁴⁷ Specifically, the minimization consisted of 500 steepest descent steps with the RMS gradient set at 0.1, followed by 1000 conjugate gradient steps and an RMS gradient of 0.0001. The MD run included 2000 steps with a gradual increase of the temperature from 50 K to 300 K and a time step of 0.001, followed by an additional 1000 steps of equilibration and a time step of 0.001. Data collection included 1000 steps with the constant temperature/constant volume ensemble (NVT) using a canonical ensemble.⁴⁵ Since sst4 is an integral membrane protein, the effect of the lipid bilayer was taken into consideration by employing the Add Membrane protocol from Discovery Studio. The protocol uses a stepwise search algorithm to optimally position and orient the protein structure relative to an implicit membrane. Two parallel planes define the membrane boundary, while the low-dielectric region between the two planes represents the nonpolar part of the lipid bilayer. The protein structure is treated as a rigid body,

and its optimal orientation corresponds to the minimum solvation energy. The optimization algorithm uses the CHARMM generalized Born model with simple switching.⁴⁸ At the end of the search, the positive-inside rule is utilized to assess the predicted orientation.⁴⁹ Spherical cutoff electrostatic interactions were included, and the nonbonded cutoff was set to 14 Å throughout the run.

Assessment of the constructed models was performed using the DOPE scoring function⁴⁶ and Verify Protein (Profile-3D) protocol⁵⁰ within Discovery Studio. PROCHECK⁵¹ in the automated SWISS-MODEL workspace⁵² was also employed for evaluation of the models. The DOPE score corresponds to the conformational energy that measures the relative stability of one conformation relative to other conformations of the same protein according to a calculated energy. The Profile-3D protocol measures each residue's compatibility by reducing the 3D structure to a one-dimensional representation, called an environment string, in order to check the validity of the generated models. PROCHECK was employed to evaluate the quality of the geometry of the model structures by checking the ϕ and ψ dihedral angles. Results are presented in Table 2.

Ligand Docking. Docking was performed with Glide 5.5^{53–55} within Maestro 9.0 (Schrodinger, LLC, New York, NY). Structures of the docked compounds are shown in Tables 3 and 4. Prior to docking, the receptor was prepared by assigning bond orders, adding hydrogens, and finding overlaps, followed by

Table 3. High Affinity Compounds

Compounds	Structure	Compounds	Structure
Compound 1 $K_i=6 \text{ nM}^{37}$		Compound 11 $K_i=1.2 \text{ nM}^{79}$	
Compound 2 $K_i=14 \text{ nM}^{80}$		Compound 12 $K_i=1.5 \text{ nM}^{79}$	
Compound 3 $K_i=16 \text{ nM}^{80}$		Compound 13 $K_i=3.2 \text{ nM}^{79}$	
Compound 4 $K_i=23 \text{ nM}^{81}$		Compound 14 $K_i=3.6 \text{ nM}^{79}$	
Compound 5 $K_i=53 \text{ nM}^{80}$		Compound 15 $K_i=5.3 \text{ nM}^{79}$	
Compound 6 $K_i=62 \text{ nM}^{82}$		Compound 16 $K_i=6.5 \text{ nM}^{79}$	
Compound 7 $K_i=70 \text{ nM}^{80}$		Compound 17 $K_i=2.9 \text{ nM}^{83}$	
Compound 8 $K_i=104 \text{ nM}^{81}$		Compound 18 $K_i=3.3 \text{ nM}^{84}$	

Table 3. continued

Compounds	Structure	Compounds	Structure
Compound 9 $K_i=113 \text{ nM}^{80}$		Compound 19 $K_i=53 \text{ nM}^{85}$	
Compound 10 $K_i=0.7 \text{ nM}^{86}$		Compound 20 $K_i=100 \text{ nM}^{87}$	

Table 4. Low Affinity Compounds

Compounds	Structure	Compounds	Structure
Compound 21 $K_i=1000^{80}$		Compound 25 $K_i=2080^{88}$	
Compound 22 $K_i=1000^{80}$		Compound 26 $K_i=2200^{89}$	
Compound 23 $K_i=1200^{80}$		Compound 27 $K_i=8600^{90}$	
Compound 24 $K_i=1200^{90}$		Compound 28 $K_i=10000^{91}$	

hydrogen bond optimization and minimization using the Protein Preparation Wizard. Minimization employed the “Impref” utility, which runs a series of constrained impact minimizations with gradually decreasing strength of the heavy-atom restraining potential. Specifically, two minimizations are initially performed with the heavy-atom restraint potential force

constant at 10. In the first minimization, the torsional potential is turned off to improve hydrogen optimization, whereas the second minimization restores the torsional potential. The restraining potential force constant is subsequently reduced to 3, 1, 0.3, and 0.1. If the output structure from a minimization exceeds the specified RMSD threshold, relative to the starting

structure, the program stops and returns the structure from the previous minimization. Thus, the RMSD is checked at the end of each round of minimization. Receptor grid generation was subsequently employed with the van der Waals radius scaling factor set to 1.0 and a partial charge cutoff at 0.25. The binding pocket was defined within 14 Å from Asp90 in TM3 and His258 in TM7. Standard precision was employed for all docking calculations with 15 poses per ligand and postdocking minimization. Poses were selected using experimental data as guides (potential hydrogen bonding with Asp90 in TM3) and the number of feasible bonding interactions.

Virtual Screening. *Ligand Preparation.* The 3D in-house virtual collection was prepared with 996 compounds taken from the Advanced Chemical Directory and the MDDR database, as described earlier.⁵⁶ The 20 high-affinity compounds were added to this database to generate the final library. Special attention was given to the protonation states of the seeded sst4 ligands, with ionizable groups (amines, carboxylic acids, phosphates, amidines) assumed to be ionized at physiological pH. Docking experiments were performed with the default parameters, unless otherwise stated. Following docking, the top 10% of the best-scored compounds were considered.

Docking Protocols. *GOLD 5.1.* Docking was performed using GOLD 5.1^{57,58} (Astex Technology, Cambridge, UK), and both ChemScore⁵⁹ and GoldScore were employed as fitness functions. Standard default settings were used in all calculations. Default cutoff values of 2.5 Å for H-bonds, and 4.0 Å for van der Waal interactions were employed. All single bonds were treated as rotatable. A 10.0 Å radius from the TM3 Asp90 was used to define the active site. Ten poses were saved for each ligand.

Glide 5.7. Glide version 5.7^{53–55} (Maestro 9.2) was used in this study. Glide grids were generated with ligand scaling of 0.8 for the van der Waals radii. Ligands were docked and ranked using standard precision. All docking experiments were performed with default settings, and 10 poses per ligand were saved for further consideration.

LigandFit 2.4. LigandFit⁶⁰ implemented in Discovery Studio Client v2.5.5 (Accelrys, Inc.: San Diego, CA) was used. The binding sites were defined by collecting a cavity within a 10 Å radius from Asp90 in TM3. Subsequently, conformational sampling of the ligands was performed employing Monte Carlo simulations (10 000 trials). Shape similarity with the binding pocket was used as a metric for pose selection, followed by a calculation of the interaction energies for those poses that do not fit with the shape criteria. Both van der Waals and electrostatic energy terms were considered in the calculations. Each docked pose was further fitted into the binding pocket through a number of rigid-body minimizations. A maximum of 10 poses were saved for each molecule to be subsequently scored using LigScore1,⁶¹ LigScore2,⁶¹ PLP1,^{62,63} PLP2,⁶³ and PMF.⁶⁴ LigScores 1 and 2 were calculated using the Dreiding force field.

RESULTS AND DISCUSSION

Despite advances in crystallography of GPCRs in the past 10 years, it was recently reported that the majority of unavailable GPCRs still need better templates than the existing crystal structures in order to generate physically plausible homology models.⁶⁵ Interestingly enough, in the same review, these investigators also concluded that sst4 and sst5 would be high-impact templates for homology modeling, if their 3D structures were resolved. Crystal complexes of the active states of the β_2 ^{19,28} and β_1 ²⁶ adrenergic and A_{2A} adenosine^{24,27} receptors with agonists were published this year; nevertheless, it is doubtful these structures

alone will provide sufficient information for the entire GPCR family, given its breadth. However, in light of the recent additions to the PDB, we felt it would be worthwhile to explore whether these agonist-bound structures would be appropriate templates for sst4 homology modeling. Ultimately, the quality and physical plausibility of the model, coupled with experimental findings pertaining to agonists, are the determining factors in understanding ligand selectivity and potency at the macromolecular level.

Model building is based on the principle that homologous proteins with similar primary sequences are expected to have similar structures. Thus, higher homology between the target and template sequences should lead to a more accurate model of the target protein.⁶⁶ We employed NCBI's BLAST search^{39,40,42,43} in order to identify the most appropriate template(s) for our comparative modeling experiments. Selection criteria included percent homology and levels of activation of respective crystal structures. Results are shown in Table 1. It can be seen that the CXCR4 chemokine receptor is the most similar to sst4. However, this chemokine is complexed with antagonists (peptides and small molecules), which in turn makes it less suitable for modeling the activated state of sst4. We then explored the agonist-bound structures of the β_2 (PDB code: 3P0G) and β_1 (PDB code: 2Y00) adrenoceptors and the A_{2A} adenosine receptor (PDB code: 2YDO) as possible templates. The nanobody-stabilized active state of the β_2 adrenergic receptor seems to be the next choice based on Table 1. Furthermore, being bound with a G protein mimicking peptide, the complex is reflective of a fully activated GPCR. Because it has been reported that multiple templates provide slight or no improvement at all in the TM region of the predicted structural models,⁶⁵ we were not as concerned with our choice of using one template for model generation.

Special attention was given to the alignment because it is the most decisive factor for the quality of the resultant model-built structures. ClustalW was used for sequence alignments,³⁸ followed by manual refinement, as shown in Figure 2. It was pleasing to see that all highly conserved residues and motifs of GPCRs were aligned: N1.50 in TM1, D2.50 in TM2, D/(E)RY in TM3, W4.50 in TM4, P5.50 in TM5, CWxP in TM6, and NPxxY in TM7^{13,25,67–72} (using the Ballesteros and Weinstein numbering system⁷³). Other residues in the transmembrane domains were also well aligned.

Models were generated using the ClustalW alignment of the template and the sequence depicted in Figure 2, followed by refinement with energy minimization and limited constrained MD simulations with an implicit solvent model. To assess the models before and after the simulations, we employed DOPE⁴⁶ and Verify-3D⁵⁰ scores, and PROCHECK⁵¹ (Table 2). DOPE is an atom-based statistical potential, representing the conformational energy that measures the relative stability of a conformation with respect to other conformations of the same protein. Verify-3D scores were used to assess whether a residue is in the desired 3D environment. The “environment” is based on buried area, secondary structure, and fraction of polar contacts. Furthermore, having included the “Add Membrane” routine, the low-dielectric environment of the bilayer was taken into consideration in the Verify-3D calculations; amino acids embedded in the implicit membrane boundary were considered fully buried. A low score is given to a hydrophobic residue on a protein’s surface and a polar residue in the protein’s core. Expected scores are also shown in order to evaluate the structure based on statistical analyses of high-resolution structures from the PDB. A model region with a score higher than the expected values is likely to be correct, but a score lower than the expected values should be carefully examined. Comparison of all constructed

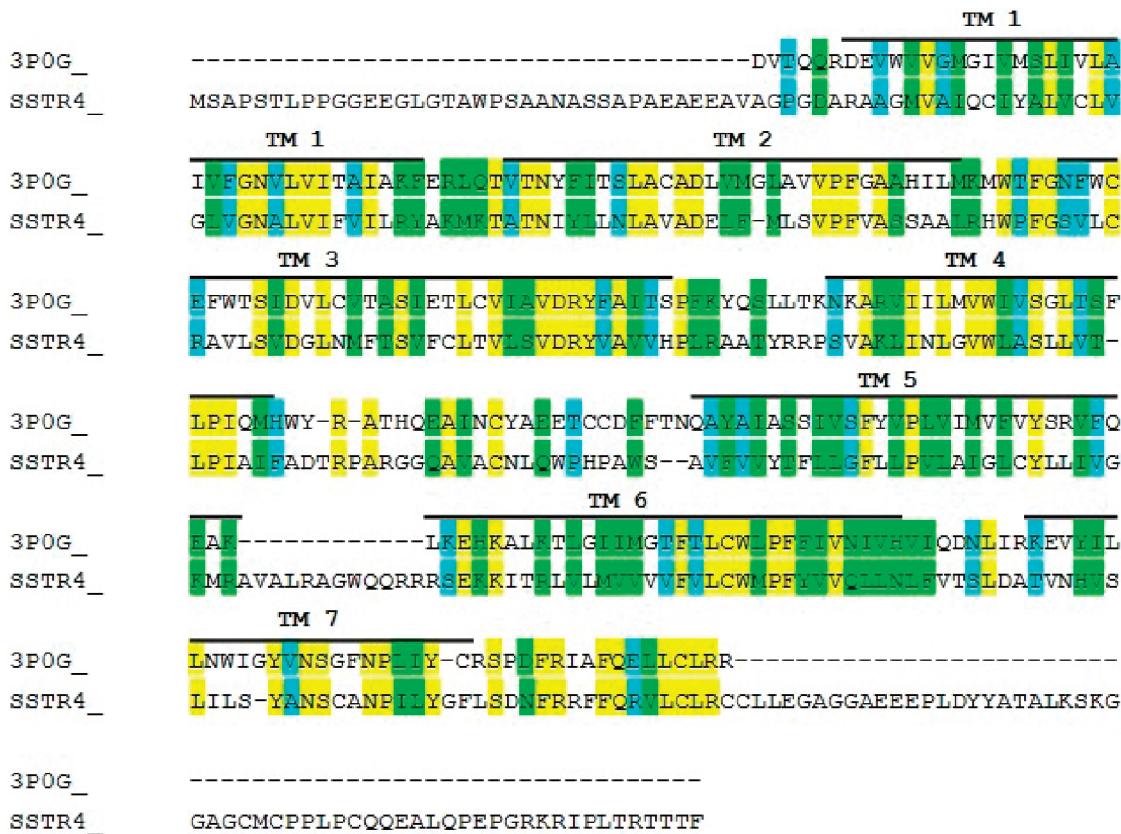


Figure 2. Sequence alignment of sst4 with β_2 -adrenoceptor (PDB code 3P0G). Yellow highlights identical residues. Green indicates similar residues. Blue shows conservative residues. TM domains are labeled.

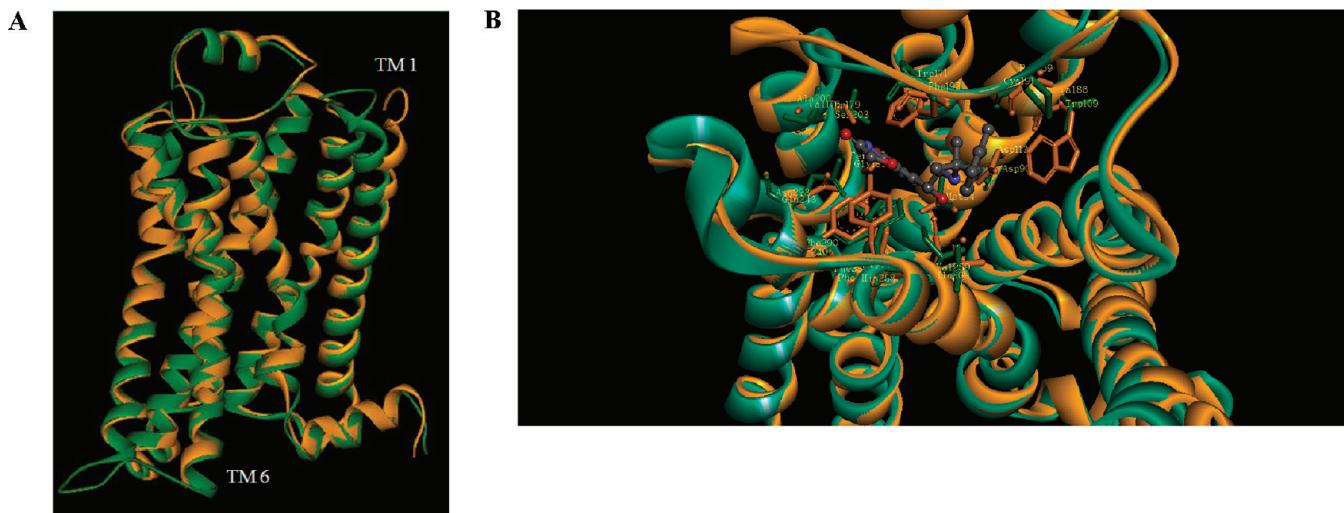


Figure 3. (A) Putative model of sst4 (green) superimposed with template (orange). (B) Overlay of the binding site residues. The only differences are Ser 207, Tyr 308, and Asn 312 of 3P0G versus Gly 183, His 258, and Ile 262 of sst4, respectively.

models led us to select the one with the lowest DOPE and highest Verify-3D scores (data not shown). It should also be mentioned that a substantial improvement of both DOPE and Verify scores was observed for this model following refinement (Verify-3D scores rose from 55.74 to 128.74, as shown in Table 2). It was pleasing to see that the resultant model was within the expected range of the Profiles-3D scoring function, thus suggesting that the residues are placed in a physically acceptable environment after refinement.

PROCHECK was also used in order to check the stereochemical quality of the model-built structures. Ramachandran plots were generated within PROCHECK to highlight regions

with unusual geometry and provide an overall assessment of the quality of the model. It can be seen that almost all residues in the model structures are in favored regions, and only 0.4% falls in the disallowed regions (Table 2). Figure 3A depicts the sst4 model superimposed onto the adrenoreceptor template, while an overlay of the binding site residues, along with the agonist BI-167107 of the crystal 3P0G, are shown in Figure 3B. Regarding differences in respective amino acids, it can be seen that Trp109 of 3P0G is replaced by Val 86 in sst4; because they are both nonpolar residues, besides their different size, no impact is expected on sst4-agonist docked complexes. Similarly, Tyr308, which makes

hydrophobic contacts with the methylene group of BI-167107, is replaced by His 258 in sst4. It would thus be expected that His258 will be involved in binding interactions with agonists, and given its charge, the expected hydrogen-bonding could be crucial in molecular recognition. The remaining two amino acids might be critical, in that both Ser 207 and Asn 312 of the template are polar, while the corresponding residues in sst4 (Gly183 and Ile262, respectively) are nonpolar. Because these are single-residue differences and given that the critical motifs in all GPCRs are aligned well, the aforementioned amino acids might not be major contributors to binding. In summary, biophysical analyses of the sst4 model indicated that the quality of the generated structure was high enough, thus giving us the confidence to proceed with the docking exercises.

In the absence of experimental evidence, we carried out a virtual screening experiment to assess the sst4 model-built structure in terms of its ability to discriminate binders from decoys. Twenty high affinity compounds (Table 3) were seeded into a library of 996 randomly chosen drug-like molecules, and the number of retrieved agonists was evaluated based on yield.¹⁵ The fraction of the agonists recovered relative to the screened database is presented in Figure 4. Only docking/scoring schemes

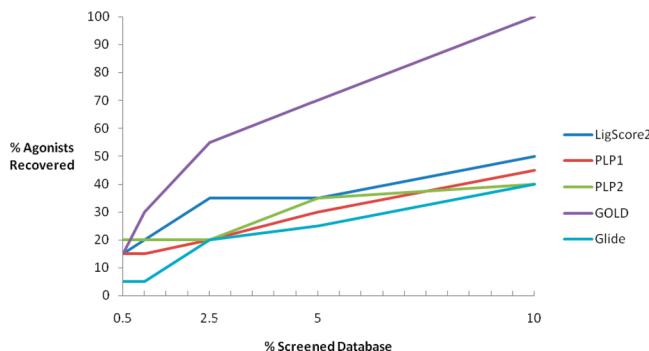


Figure 4. Percentage of high affinity compounds recovered as a function of the screened database. GOLD with its own scoring function retrieves all 20 seeded agonists, followed by LigandFit with LigScore2, which recovers half of the total at 10% of the database.

identifying more than 40% of the high affinity seeded ligands at a 10% fraction of the database are depicted. GOLD with its own scoring function and LigandFit/LigScore2 performed the best, in that they retrieved all 20 and half of the compounds, respectively. LigandFit/PLP1, LigandFit/PLP2, and Glide/GlideScore led to 45%, 40%, and 40% yields, respectively (Figure 4).

The same 20 high affinity compounds (Table 3) were subsequently docked into the model, and all generated poses were assessed visually. Criteria used for sst4–agonist complex selection were as follows: (1) number of bonding interactions such as hydrogen bonds, aromatic, and hydrophobic interactions and (2) published experimental restraints for sst subtypes such as electrostatic interactions with Asp90 in TM3.^{33,34} Emodel scores for the selected poses were subsequently examined. Emodel is used by Glide to rank poses of a given ligand. It combines GlideScore, the nonbonded interaction energy, and the excess internal energy of the generated ligand conformation for flexible docking.^{53,55} Table 5 shows the Emodel rankings of the selected poses, and it was pleasing to see that all were among the top six scored hits. Our next step was to assess patterns and similarities in observed modes of all agonists in accord with the underlying theory of molecular recognition; that is, compounds displaying a high binding affinity should form similar interactions at the

Table 5. Emodel Ranking of High Affinity Compounds

	Emodel ^a	Emodel ^a	
compound 1	1	compound 11	1
compound 2	3	compound 12	3
compound 3	4	compound 13	1
compound 4	1	compound 14	1
compound 5	6	compound 15	5
compound 6	1	compound 16	1
compound 7	1	compound 17	6
compound 8	2	compound 18	2
compound 9	1	compound 19	1
compound 10	6	compound 20	5

^aThe corresponding rank number for each selected pose.

macromolecular level. However, it became apparent that high-affinity agonists did not adopt one binding mode that could possibly explain experimental data. We thus propose herein two binding modes for the high affinity agonists that are seemingly consistent with observed findings as well.

Binding mode I is based on interactions of compound 2 ($K_i = 14 \text{ nM}$) with sst4, which is selected as a prototype because its respective pose makes the most interactions with the receptor (Figure 5). The violet grid surface depicts a

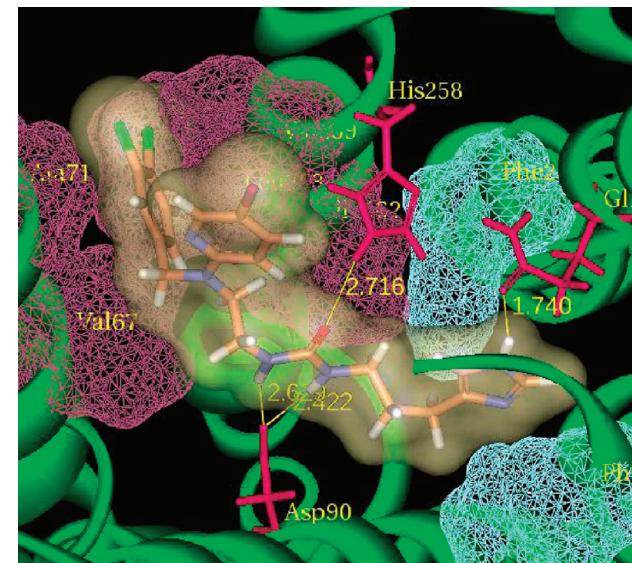


Figure 5. Binding mode I. The violet grid indicates the TM2/TM7 hydrophobic cavity formed by Ala71, Val67 in TM2, Val259, Ile262, and Leu263 in TM7. The cyan grid depicts the aromatic cavity formed by Phe175 and Phe239. Residues involved in hydrogen bonding are shown in red and corresponding distances in yellow dotted lines.

hydrophobic pocket formed by Val67, Ala71 in TM2, Val259, Ile262, and Leu263 in TM7 (TM2/TM7 hydrophobic cavity). The two aromatic rings of the dichlorobenzyl and 5-bromo-2-pyridyl groups are buried in this pocket. The two amine hydrogens of the urea moiety interact the key residue Asp90. The urea carbonyl oxygen points toward the side chain of His258 and forms another hydrogen bond. Finally, the N-1 hydrogen in the imidazole ring forms a hydrogen bond with Gln243. The imidazole ring is involved in aromatic–aromatic (π -stacking) interactions with phenylalanines 175 and 239 on both sides of the compound.

Consistent with binding mode I, though with some slight differences, are compounds **1–9**, all of which carry urea or thiourea

functionalities. Binding modes of these compounds place the two aromatic rings in the hydrophobic pocket, while the thiourea/urea moieties form hydrogen bonds with Asp90 and His258. The imidazole rings are involved in both hydrogen bonding and

aromatic–aromatic interactions. Differences observed among the nine compounds in regard to their binding interactions (Figure 6A) are as follows: (1) The imidazole in compounds **1** and **3** forms a hydrogen bond with His258 in place of the thiourea/urea-His258

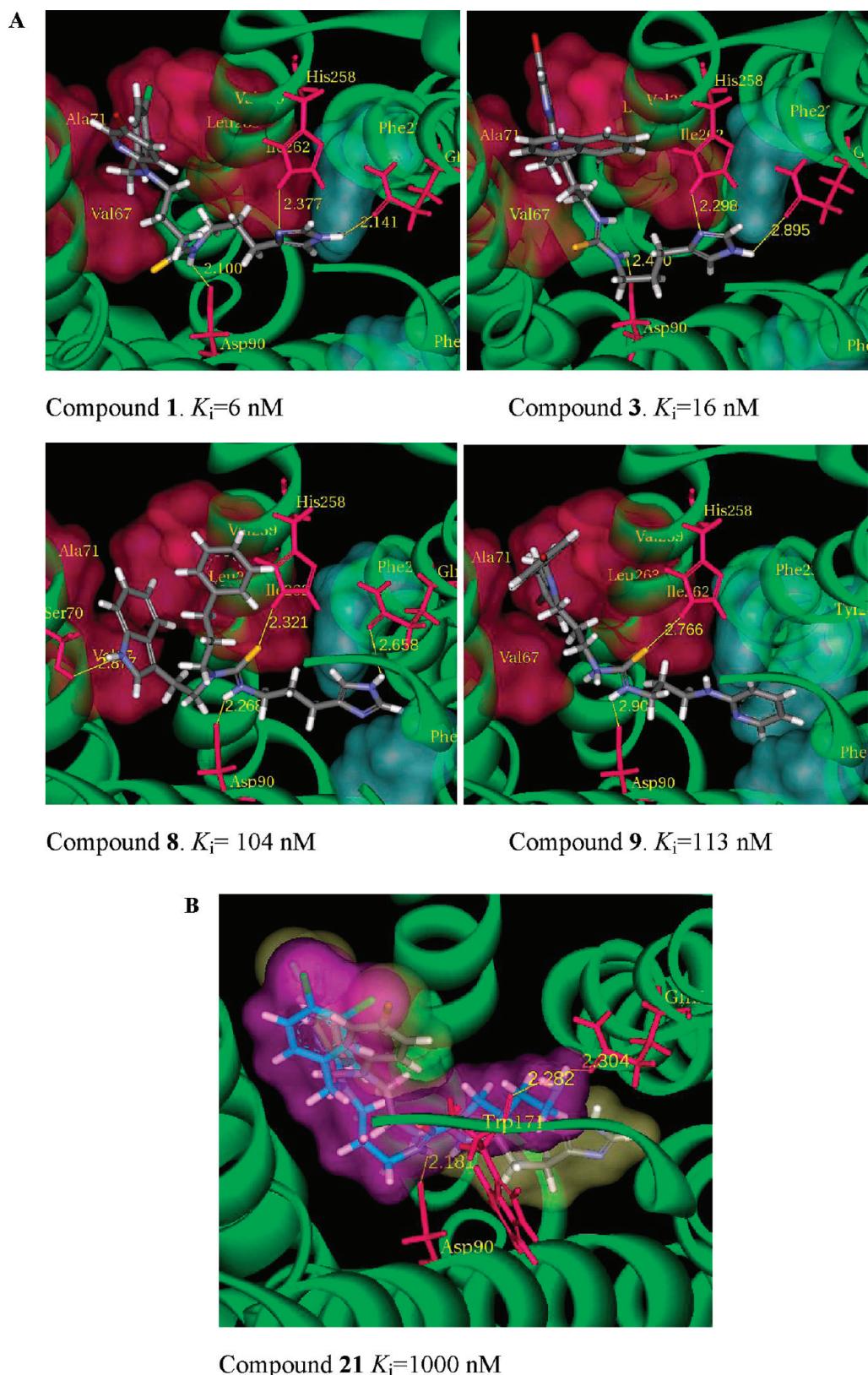


Figure 6. (A) Representative examples of compounds consistent with binding mode I. The red surface represents the TM2/TM7 hydrophobic cavity. Cyan encodes for the aromatic phenylalanines 175 and 239. (B) Putative binding of compound 21 (purple surface) superimposed with binding mode I (yellow surface).

hydrogen bond observed for the other agonists. (2) The pyridine of compound **9** does not form a hydrogen bond with Gln243, due to a lack of hydrogen on the heterocycle; however it maintains the aromatic–aromatic interactions not only with Phe175 and Phe239 but with Tyr240 as well. (3) The indole in compound **8** is positioned inside the hydrophobic cavity and forms a hydrogen bond with Ser70. In summary, the aforementioned agonists form all or most interactions, consistent with binding mode I. We therefore rationalize that sst4–agonist complexes adopting this mode are responsible for the observed high-affinity data.

Mutagenesis experiments have shown that there is a common binding site for small molecules within the TM domain of GPCRs regardless of the nature of the endogenous, physiologically relevant receptor agonist.^{74–76} It was also demonstrated that high-affinity binding involves a salt bridge between a basic functional group of the ligands, such as an amine moiety, with the conserved aspartate in TM3. In accord with these findings, binding to sst4 was thought to involve basic functional groups mimicking the Lys9 of SRIF and interacting with an aspartate in TM3.^{36,77} Consequently, several compounds including basic functionalities were designed in hopes of binding with the TM3 Asp.^{36,77} However, a comparison between the low-affinity compound **21** (Table 4), which carries a primary amine ($K_i = 1000$ nM), and high-affinity compounds **1–8** (Table 3) with the weakly basic imidazoyl moiety disputes the aforementioned hypothesis. In order to explain the above discrepancy, it was proposed that the planar imidazole ring places the NH group in a favorable position to hydrogen-bond with Asp90.^{36,78} However, our docking experiments have provided a sufficient explanation for these contradictory hypotheses. Specifically, binding mode I shows that the basic functional groups of somatostatin ligands, such as imidazole and pyridine, do not necessarily bind with Asp90. Instead, they form hydrogen bonds with Gln243 and fit into the aromatic cavity defined by Phe175 and Phe239. In contrast, compound **21** does not interact the same way (Figure 6B). First of all, no interaction is observed with the aromatic residues, since the primary amine lacks aromaticity. Second, there still is unoccupied space in the binding cavity relative to the volume corresponding to compound **21** (see Figure 6B). Thus, we believe that limited shape complementarity and fewer bonding interactions are responsible for the lower binding affinity of this structure.

The prototype for binding mode II is compound **11** (Table 3 and Figure 7). Similarly to binding mode I's cavity, the respective pocket of mode II consists of the TM2/TM7 hydrophobic cavity, the aromatic cavity involving residues Trp171 and Tyr240, in addition to phenylalanines 175 and 239 observed in mode I, and the same hydrogen bonding sites with slight differences: the backbone of Trp171 is also involved in interactions, contrary to Gln243 of binding mode I, which plays only a minor role in mode II. Furthermore, the ligand interactions in mode II differ from those of mode I. The most notable differences are (1) In mode I, a basic group such as imidazole interacts with Gln243 and aromatic residues, whereas in mode II it interacts with Asp90. (2) While in mode I the dichlorobenzyl and 5-bromo-2-pyridyl groups are buried into the hydrophobic pocket, it is the 1-methylnaphthyl that is docked into this pocket in mode II. (3) The urea carbonyl oxygen in mode I interacts with the His258 side chain only, while in mode II the sulfonyl and amide oxygens form hydrogen bonds with the His258 side-chain hydrogen and the

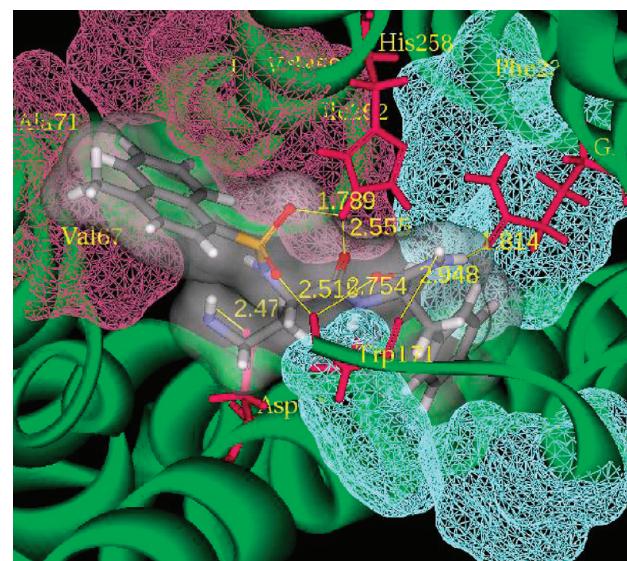


Figure 7. Binding mode II. The violet grid shows the TM2/TM7 hydrophobic cavity and cyan, the aromatic cavity formed by Trp171, Phe175, Phe239, and Tyr240. Hydrogen bonds are also depicted.

backbone hydrogen of Trp171 as well. (4) In mode I, the imidazole is involved in dual π – π and hydrogen-bonding interactions with Gln243, contrary to the observed binding mode II, where the benzene ring inserts into the pocket formed by four aromatic residues (Trp171, Phe175, Phe239, and Tyr240), and Gln243 now forms a hydrogen bond with the amide hydrogen of the prototype structure **11** (see Table 3 and Figure 7).

Consistent with binding mode II are compounds **11–16**, with the exception of compound **14**. These are sulfonamide series with varying basic functionalities and numbers of aromatic rings. Examination and visual inspection of the predicted binding modes revealed the following differences among compounds **11–16**: (1) Either the sulfonamide or the amide nitrogens form hydrogen bonds with His258, whereas the amide or the sulfonamide oxygens interact with the Trp171 backbone. (2) The primary amides form hydrogen bonds with either the backbone Ser172 and Gln243 or with Asp90. Specifically, we observe that in compound **12** the amide hydrogen interacts with Gln243 and Ser172, while in compounds **13**, **15**, and **16**, hydrogen bonding with Asp90 is observed. (3) The basic functionalities, namely, guanidine and piperidine in compounds **12** and **13**, respectively, bind with Asp90 and the Ser63 backbone oxygen (compound **12**) and Ser70 (compound **13**). The predicted complexes for these compounds are depicted in Figure 8A.

Compounds **18** and **19** also fit with binding mode II, even though they are structurally different from compounds **11–16** (Table 3, Figure 8B). Specifically, compound **18** contains an azepinylindole, a primary amine, and benzene rings, which perfectly mimic the SRIF key residues of Trp8, Lys9, and Phe6, respectively. Binding mode analyses of the sst4–compound **18** complex revealed that the primary amine forms a salt bridge with Asp90; the indole makes aromatic–aromatic interactions with Phe239 and Trp171 and also hydrogen bonds with Asp90. One of the two benzene rings makes hydrophobic interactions with Val67, Ala71, Val259, and Ile262 in TM2 and TM7. We also observed hydrogen bonding interactions of the cyclic amide nitrogen with His258 and the

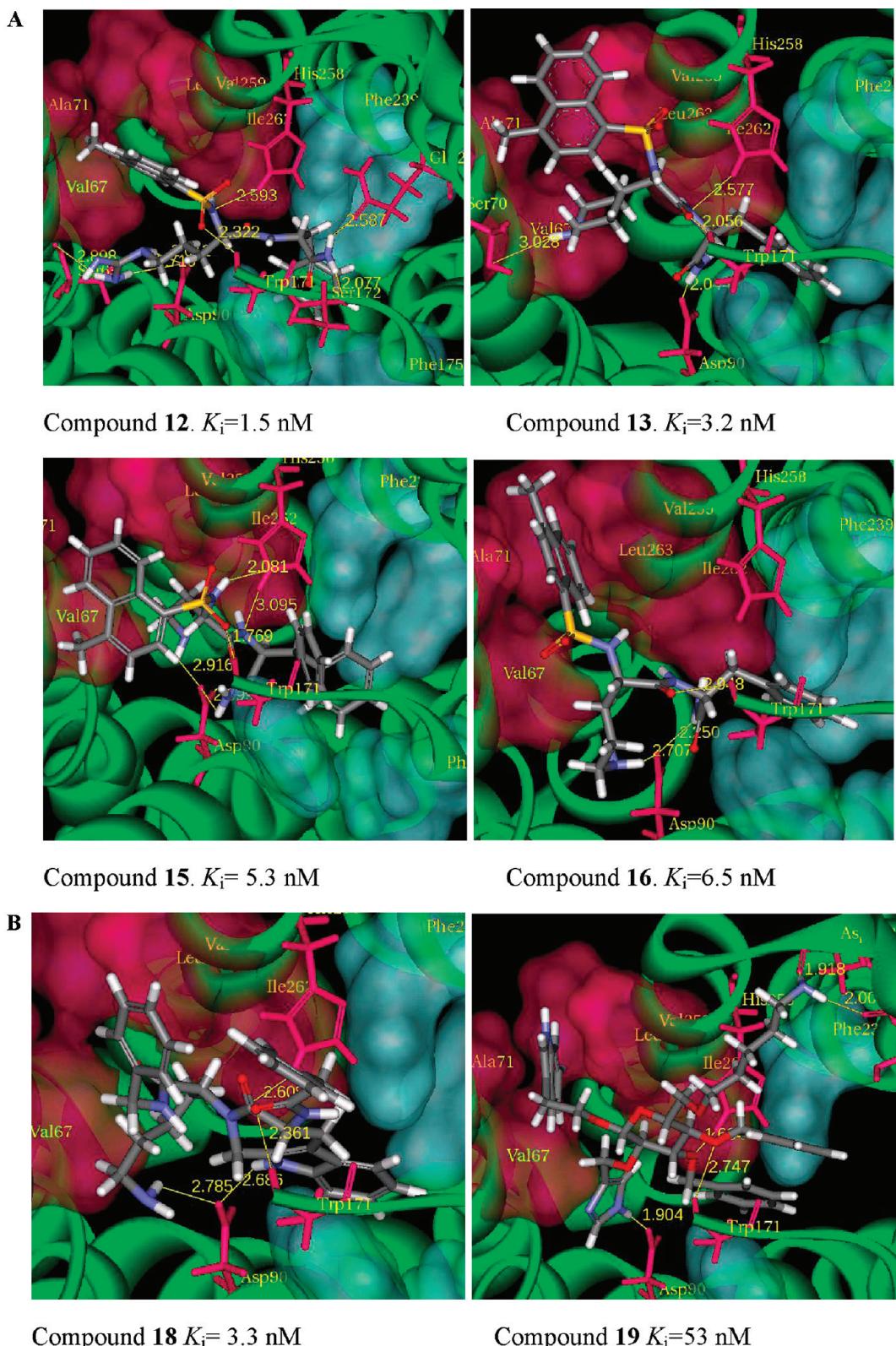


Figure 8. Representative examples of compounds consistent with binding mode II. (A) Sulfonamide-containing compounds and (B) aza-cycloheptane-indole and glucose scaffolds.

amide oxygen of Trp171, which is consistent with mode II. The indole of compound **19** fits well into the hydrophobic pocket, while the other benzene is positioned between Phe239 and Trp171, thus forming aromatic–aromatic interactions. A hydrogen bond interaction was observed between the

imidazole and Asp90, while the two oxygens at C-4 and C-5 of the glucose scaffold bind to His258 and Trp171. It should be pointed out that the primary amine does not bind with Asp90; however, it interacts with the backbone oxygen of the residues Thr250 and Asp253 in the extracellular loop 3. Due to the

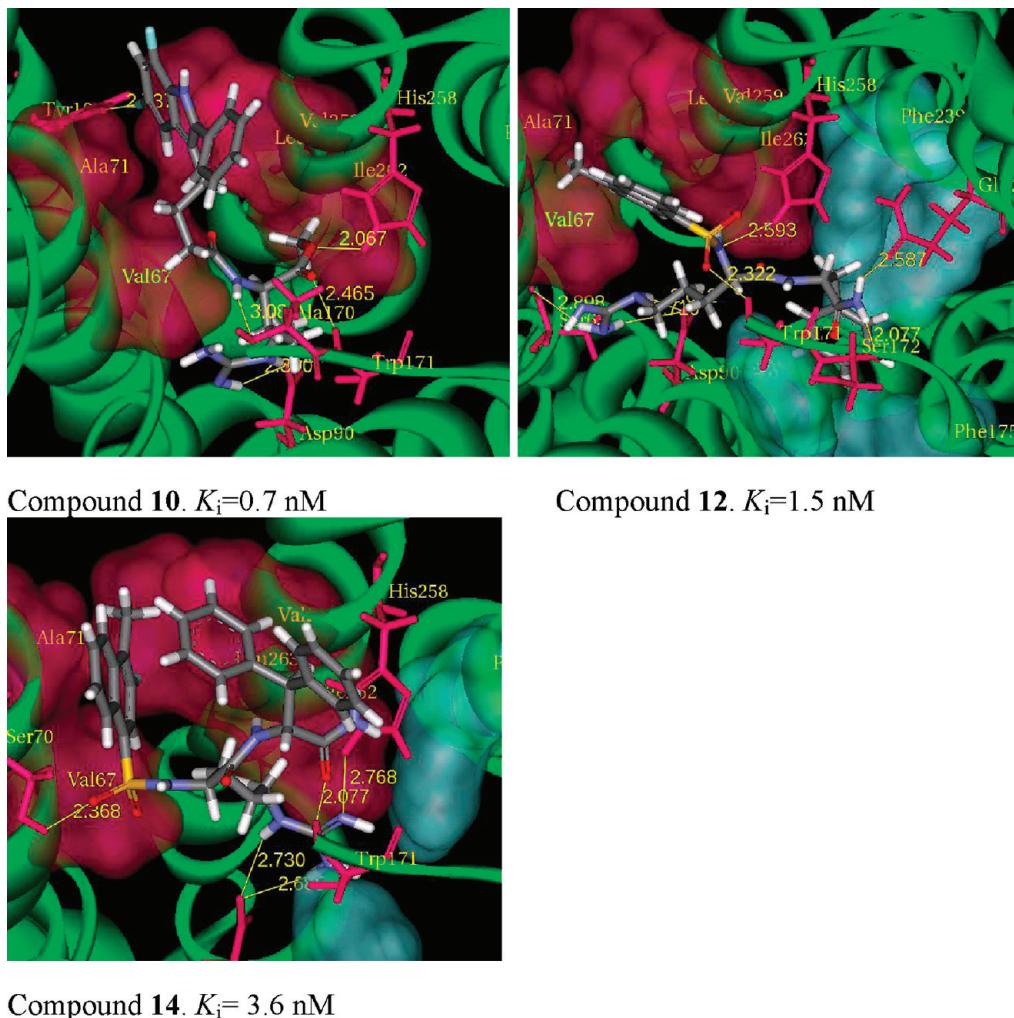


Figure 9. Predicted interactions of guanidium compounds with sst4.

length of the hydrocarbon chain in the amino-pentoxy substituent, the amine is not involved in the typical interactions observed in binding modes I and II. Similarly, compound 17 partially fits with binding mode II (data not shown). We observe that it maintains the hydrogen bonds with Asp90, His258, and the carbonyl oxygen of the Trp171 backbone, as well as the aromatic interactions with Trp171, Phe175, Phe239, and Tyr240. However, no interactions were observed with the TM2/TM7 hydrophobic cavity. Finally, compound 20 fits also partially with binding mode II. One aromatic ring inserts into the hydrophobic cavity, and the primary amine electrostatically interacts with Asp90. No aromatic interactions are observed with Trp171, Phe175, Phe239, and Tyr240; however, two additional hydrogen bonds are formed. One hydrogen bond is formed between the imidazole and the proline backbone, and the other is formed between the indole and the serine.

Among the high affinity compounds, compounds 10, 12, and 14 contain a guanidinium group as a basic functionality (see Table 3). The predicted binding modes of these compounds are presented in Figure 9, and as mentioned in the previous paragraph, compound 14 does not quite fit binding mode II. It can be seen that the guanidine moiety in compound 12 interacts with Asp90. However, the predicted pose of compound 14 places the guanidinium where the aromatic ring in mode II is observed. Thus, this basic functional moiety

Table 6. Emodel Ranking of Low Affinity Compounds

	Emodel ^a	Emodel ^a	
compound 21	1	compound 25	3
compound 22	3	compound 26	2
compound 23	5	compound 27	3
compound 24	1	compound 28	2

^aThe corresponding rank number for each pose with the most interactions with the receptor.

in 14 is within bonding distance from Asp90 and His258. Furthermore, the diphenyl–methyl group, which would be expected to π -stack with the aromatic residues Trp171, Phe175, Phe239, and Tyr240 according to binding mode II, instead forms hydrophobic interactions with Ala71 in TM2 and Leu263 in TM7 in compound 14. The question then arises, why does this compound have high affinity? We speculate that the guanidinium forms three hydrogen bonds (with Asp90 and His258), while the sulfonyl oxygen also interacts with the side chain of Ser 70. Consequently, not only do these electrostatics interactions compensate for the lack of π – π interactions seen in mode II, but they are also stronger. Another interesting receptor–ligand complex is observed with the predicted pose for compound 10. It contains all hydrophobic interactions of binding mode II, and it also satisfies the hydrogen bonding requirements with the Asp90, His258, and Trp171 backbone

Table 7. Proposed Binding Modes for sst4 Agonists and Respective Chemical Structures

Binding mode I			
Compound 1 $K_i=6 \text{ nM}$		Compound 6 $K_i=62 \text{ nM}$	
Compound 2 $K_i=14 \text{ nM}$		Compound 7 $K_i=70 \text{ nM}$	
Compound 3 $K_i=16 \text{ nM}$		Compound 8 $K_i=104 \text{ nM}$	
Compound 4 $K_i=23 \text{ nM}$		Compound 9 $K_i=113 \text{ nM}$	
Compound 5 $K_i=53 \text{ nM}$			
Binding mode II			
Compound 11 $K_i=1.2 \text{ nM}$		Compound 16 $K_i=6.5 \text{ nM}$	
Compound 12 $K_i=1.5 \text{ nM}$		Compound 18 $K_i=3.3 \text{ nM}$	
Compound 13 $K_i=3.2 \text{ nM}$		Compound 19 $K_i=53 \text{ nM}$	
Compound 15 $K_i=5.3 \text{ nM}$			

corresponding to mode II. In addition, there are two extra hydrogen bonds between the fluorine at the C-5 of indole and the side chain of Tyr18 in TM1, as well as between the amide hydrogen and the Ala170 backbone. We believe the tight nature of binding is the reason for the high affinity of compound 10.

However, any structure-based design project is not viable unless observed low-affinity data are also explained. All Glide poses of each compound in Table 4 were visually inspected, and it was pleasing to see that none conforms to either binding mode. Table 6 shows the Emodel rankings of the poses of the

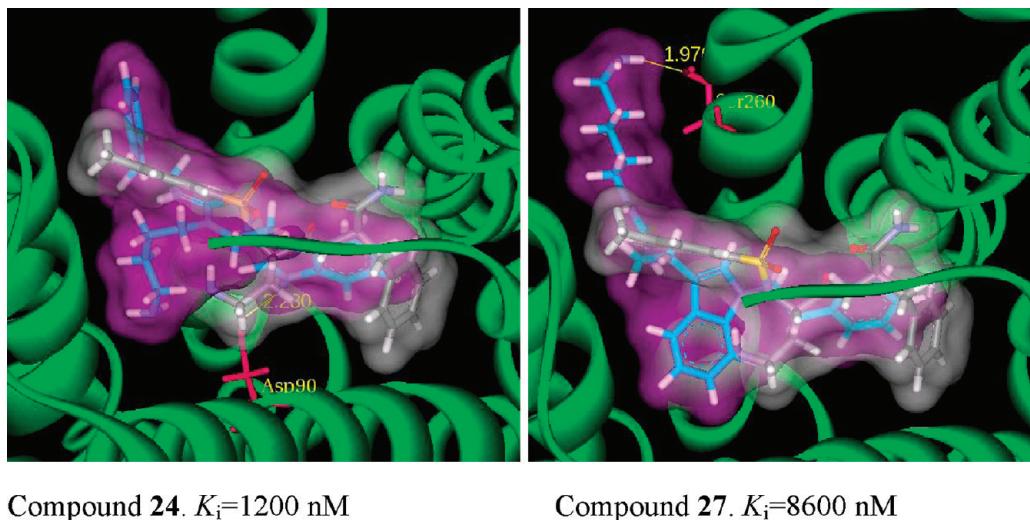


Figure 10. Representative low affinity compounds (purple surface) superimposed with binding mode II (white surface).

low affinity sst4 agonists that showed the most interactions with the receptor. Besides lacking the key interactions observed in the high affinity ligands, the shape of these compounds does not match the volume of the binding pocket. Figure 10 depicts this lack of complementarity. Compound 24 has a low binding affinity with a K_i value of 1200 nM. It can be seen that only one hydrogen bond is formed between Asp90 and the indole nitrogen–hydrogen. Furthermore, few hydrophobic interactions are observed, while the primary amine is not within bonding distance from any residues. Similarly, binding analyses of compound 27 revealed one hydrogen bond formed with residues outside of the binding pocket, and no interaction with Asp90 is seen.

■ CONCLUSION

In this study, we have generated an activated state somatostatin subtype-4 homology model and performed docking of a variety of agonists in order to rationalize affinity data. We have shown that two partially overlapping binding modes can explain the variability observed regarding both the chemistry and the K_i values of reported sst4 agonists (Table 7). Specifically, compounds 1–9, which are urea and thiourea derivates, form complexes characteristic of binding mode I; compounds 11–13 and 15–19 are predicted to form interactions corresponding to binding mode II. One of the fundamental differences between the two modes is that the basic functionality of mode I interacts with Gln243, contrary to hydrogen bonding observed between the primary amine or guanidinium moieties of compounds consistent with mode II and Asp90, a conserved residue on TM3. Our docking results provide an explanation for the low-affinity agonists as well. These agonists bind differently and do not show the extent of interactions characteristic of the high affinity compounds with the receptor, which in turn leads to a less tight complex with the putative sst4 receptor structure.

■ AUTHOR INFORMATION

Corresponding Author

*Telephone: 618-650-5166. Fax: 618-650-5145. E-mail: mkontoy@siue.edu.

■ REFERENCES

- (1) Lagerstrom, M. C.; Schioth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discovery* **2008**, *7*, 339–357.
- (2) Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discovery* **2002**, *1*, 727–730.
- (3) Russ, A. P.; Lampel, S. The druggable genome: an update. *Drug Discovery Today* **2005**, *10*, 1607–1610.
- (4) Kolakowski, L. F. Jr. GCRDb: a G-protein-coupled receptor database. *Receptors Channels* **1994**, *2*, 1–7.
- (5) Attwood, T. K.; Findlay, J. B. Fingerprinting G-protein-coupled receptors. *Protein Eng.* **1994**, *7*, 195–203.
- (6) Kroese, W. K.; Sheffler, D. J.; Roth, B. L. G-protein-coupled receptors at a glance. *J. Cell Sci.* **2003**, *116*, 4867–4869.
- (7) Schioth, H. B.; Fredriksson, R. The GRAFS classification system of G-protein coupled receptors in comparative perspective. *Gen. Comp. Endocrinol.* **2005**, *142*, 94–101.
- (8) Swaminath, G.; Xiang, Y.; Lee, T. W.; Steenhuis, J.; Parnot, C.; Kobilka, B. K. Sequential binding of agonists to the beta2 adrenoceptor. Kinetic evidence for intermediate conformational states. *J. Biol. Chem.* **2004**, *279*, 686–691.
- (9) Kobilka, B. K.; Deupi, X. Conformational complexity of G-protein-coupled receptors. *Trends Pharmacol. Sci.* **2007**, *28*, 397–406.
- (10) Eglen, R. M. Emerging concepts in GPCR function—the influence of cell phenotype on GPCR pharmacology. *Proc. West Pharmacol. Soc.* **2005**, *48*, 31–34.
- (11) Eglen, R. M. Functional G protein-coupled receptor assays for primary and secondary screening. *Comb. Chem. High Throughput Screen.* **2005**, *8*, 311–318.
- (12) McCusker, E. C.; Bane, S. E.; O’Malley, M. A.; Robinson, A. S. Heterologous GPCR expression: a bottleneck to obtaining crystal structures. *Biotechnol. Prog.* **2007**, *23*, 540–547.
- (13) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289*, 739–745.
- (14) Archer, E.; Maigret, B.; Escrieut, C.; Pradayrol, L.; Fourmy, D. Rhodopsin crystal: new template yielding realistic models of G-protein-coupled receptors? *Trends Pharmacol. Sci.* **2003**, *24*, 36–40.
- (15) 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.
- (16) 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.; Stevens, R. C.; Takeda, S.; Kadokawa, S.; Haga, T.;

- Takaesu, H.; Mitaku, S.; Fredriksson, R.; Lagerstrom, M. C.; Lundin, L. G.; Schiøth, H. B.; Pierce, K. L.; Premont, R. T.; Lefkowitz, R. J.; Shenoy, S. K.; Rosenbaum, D. M. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **2007**, *318*, 1258–1265.
- (17) Rasmussen, S. G. F.; Choi, H.-J.; Rosenbaum, D. M.; Kobilka, T. S.; Sun, F. S.; Edwards, P. C.; Burghammer, M.; Ratnala, V. R. P.; Sanishvili, R.; Fischetti, R. F.; Schertler, G. F. X.; Weis, W. I.; Kobilka, B. K. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* **2007**, *450*, 383–387.
- (18) Rosenbaum, D. M.; Cherezov, V.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Yao, X.-J.; Weis, W. I.; Stevens, R. C.; Kobilka, B. K.; Pierce, K. L.; Premont, R. T.; Lefkowitz, R. J.; Kobilka, B. K.; Deupi, X.; Cherezov, V.; Rasmussen, S. G. F. GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* **2007**, *318*, 1266–1273.
- (19) Rosenbaum, D. M.; Zhang, C.; Lyons, J. A.; Holl, R.; Aragao, D.; Arlow, D. H.; Rasmussen, S. G. F.; Choi, H.-J.; DeVree, B. T.; Sunahara, R. K.; Chae, P. S.; Gellman, S. H.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Martin, C.; Gmeiner, P.; Kobilka, B. K. Structure and function of an irreversible agonist-beta(2) adrenoceptor complex. *Nature* **2011**, *469*, 236–240.
- (20) Hanson, M. A.; Cherezov, V.; Griffith, M. T.; Roth, C. B.; Jaakola, V.-P.; Chien, E. Y. T.; Velasquez, J.; Kuhn, P.; Stevens, R. C. A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor. *Structure* **2008**, *16*, 897–905.
- (21) Serrano-Vega, M. J.; Magnani, F.; Shibata, Y.; Tate, C. G. Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 877–882.
- (22) Jaakola, V. P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; Ilzerman, A. P.; Stevens, R. C. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* **2008**, *322*, 1211–1217.
- (23) Chien, E. Y.; Liu, W.; Zhao, Q.; Katritch, V.; Won Han, G.; 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.
- (24) Xu, F.; Wu, H.; Katritch, V.; Han, G. W.; Jacobson, K. A.; Gao, Z.-G.; Cherezov, V.; Stevens, R. C. Structure of an Agonist-Bound Human A2A Adenosine Receptor. *Science* **2011**, *332*, 322–327.
- (25) Scheerer, P.; Park, J. H.; Hildebrand, P. W.; Kim, Y. J.; Krauss, N.; Choe, H.-W.; Hofmann, K. P.; Ernst, O. P. Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **2008**, *455*, 497–502.
- (26) Warne, T.; Moukhametzianov, R.; Baker, J. G.; Nehme, R.; Edwards, P. C.; Leslie, A. G. W.; Schertler, G. F. X.; Tate, C. G. The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor. *Nature* **2011**, *469*, 241–244.
- (27) Lebon, G.; Warne, T.; Edwards, P. C.; Bennett, C.; Langmead, C. J.; Leslie, A. G. W.; Tate, C. G. Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature* **2011**, *474*, 521–525.
- (28) Rasmussen, S. G.; Choi, H. J.; Fung, J. J.; Pardon, E.; Caarosa, P.; Chae, P. S.; DeVree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; Schnapp, A.; Konetzki, A.; Sunahara, R. K.; Gellman, S. H.; Pautsch, A.; Steyaert, J.; Weis, W. I.; Kobilka, B. K. Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **2011**, *469*, 175–180.
- (29) Rasmussen, S. G.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K.-Y.; Kobilka, T.-S.; Thian, F.-S.; Chae, P.-S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **2011**, *477*, 549–555.
- (30) Simon, A. Comparison of the binding modes of TT-232 in somatostatin receptors type 1 and 4. *THEOCHEM* **2007**, *816*, 73–76.
- (31) Bruno, J. F.; Xu, Y.; Song, J.; Berelowitz, M. Molecular cloning and functional expression of a brain-specific somatostatin receptor. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 11151–11155.
- (32) Moller, L. N.; Stidsen, C. E.; Hartmann, B.; Holst, J. J. Somatostatin receptors. *Biochim. Biophys. Acta* **2003**, *1616*, 1–84.
- (33) Nehring, R. B.; Meyerhof, W.; Richter, D. Aspartic acid residue 124 in the third transmembrane domain of the somatostatin receptor subtype 3 is essential for somatostatin-14 binding. *DNA Cell Biol* **1995**, *14*, 939–944.
- (34) Patel, Y. C. Somatostatin and its receptor family. *Front. Neuroendocrinol.* **1999**, *20*, 157–198.
- (35) Lewis, I.; Bauer, W.; Albert, R.; Chandramouli, N.; Pless, J.; Weckbecker, G.; Bruns, C. A novel somatostatin mimic with broad somatotropin release inhibitory factor receptor binding and superior therapeutic potential. *J. Med. Chem.* **2003**, *46*, 2334–2344.
- (36) Crider, A. M.; Witt, K. A. Somatostatin sst4 ligands: chemistry and pharmacology. *Mini Rev. Med. Chem.* **2007**, *7*, 213–220.
- (37) Ankersen, M.; Crider, M.; Liu, S.; Ho, B.; Andersen, H. S.; Stidsen, C. Discovery of a Novel Non-Peptide Somatostatin Agonist with SST4 Selectivity. *J. Am. Chem. Soc.* **1998**, *120*, 1368–1373.
- (38) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
- (39) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410.
- (40) Altschul, S. F.; Lipman, D. J. Protein database searches for multiple alignments. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 5509–5513.
- (41) Gish, W.; States, D. J. Identification of protein coding regions by database similarity search. *Nat. Genet.* **1993**, *3*, 266–272.
- (42) Madden, T. L.; Tatusov, R. L.; Zhang, J. Applications of network BLAST server. *Methods Enzymol.* **1996**, *266*, 131–141.
- (43) Zhang, J.; Madden, T. L. PowerBLAST: a new network BLAST application for interactive or automated sequence analysis and annotation. *Genome Res.* **1997**, *7*, 649–656.
- (44) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.
- (45) Marti-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291–325.
- (46) Shen, M. Y.; Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **2006**, *15*, 2507–2524.
- (47) Momany, F. A.; Rone, R. Validation of the general purpose QUANTA 3.2/CHARMM force field. *J. Comput. Chem.* **1992**, *13*, 888–900.
- (48) Im, W.; Lee, M. S.; Brooks, C. L. 3rd Generalized born model with a simple smoothing function. *J. Comput. Chem.* **2003**, *24*, 1691–1702.
- (49) von Heijne, G. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* **1992**, *225*, 487–494.
- (50) Luthy, R.; Bowie, J. U.; Eisenberg, D. Assessment of protein models with three-dimensional profiles. *Nature* **1992**, *356*, 83–85.
- (51) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK - a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
- (52) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **2006**, *22*, 195–201.
- (53) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (54) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196.

- (55) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pllard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
- (56) Kontoyianni, M.; Sokol, G. S.; McClellan, L. M. Evaluation of library ranking efficacy in virtual screening. *J. Comput. Chem.* **2005**, *26*, 11–22.
- (57) 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.
- (58) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins* **2003**, *52*, 609–623.
- (59) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425–445.
- (60) Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graphics Modell.* **2003**, *21*, 289–307.
- (61) Krammer, A.; Kirchhoff, P. D.; Jiang, X.; Venkatachalam, C. M.; Waldman, M. LigScore: a novel scoring function for predicting binding affinities. *J. Mol. Graphics Modell.* **2005**, *23*, 395–407.
- (62) Gehlhaar, D. K.; Moerder, K. E.; Zichi, D.; Sherman, C. J.; Ogden, R. C.; Freer, S. T. De novo design of enzyme inhibitors by Monte Carlo ligand generation. *J. Med. Chem.* **1995**, *38*, 466–472.
- (63) Gehlhaar, D. K.; Verkhivker, G. M.; Rejto, P. A.; Sherman, C. J.; Fogel, D. B.; Fogel, L. J.; Freer, S. T. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem. Biol.* **1995**, *2*, 317–324.
- (64) Muegge, I. PMF scoring revisited. *J. Med. Chem.* **2006**, *49*, 5895–5902.
- (65) Mobarec, J. C.; Sanchez, R.; Filizola, M. Modern homology modeling of G-protein coupled receptors: which structural template to use? *J. Med. Chem.* **2009**, *52*, 5207–5216.
- (66) Chothia, C.; Lesk, A. M. The relation between the divergence of sequence and structure in proteins. *EMBO J.* **1986**, *5*, 823–826.
- (67) Rasmussen, S. G.; Jensen, A. D.; Liapakis, G.; Ghanouni, P.; Javitch, J. A.; Gether, U. Mutation of a highly conserved aspartic acid in the beta₂ adrenergic receptor: constitutive activation, structural instability, and conformational rearrangement of transmembrane segment 6. *Mol. Pharmacol.* **1999**, *56*, 175–184.
- (68) Shi, L.; Liapakis, G.; Xu, R.; Guarneri, F.; Ballesteros, J. A.; Javitch, J. A. Beta₂ adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *J. Biol. Chem.* **2002**, *277*, 40989–40996.
- (69) Fritze, O.; Filipek, S.; Kuksa, V.; Palczewski, K.; Hofmann, K. P.; Ernst, O. P. Role of the conserved NPxxY(x)S/F motif in the rhodopsin ground state and during activation. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2290–2295.
- (70) Palczewski, K. G protein-coupled receptor rhodopsin. *Annu. Rev. Biochem.* **2006**, *75*, 743–767.
- (71) Lodowski, D. T.; Angel, T. E.; Palczewski, K. Comparative analysis of GPCR crystal structures. *Photochem. Photobiol.* **2009**, *85*, 425–430.
- (72) Favre, N.; Fanelli, F.; Missotten, M.; Nichols, A.; Wilson, J.; Di Tiani, M.; Rommel, C.; Scheer, A. The DRY motif as a molecular switch of the human oxytocin receptor. *Biochemistry* **2005**, *44*, 9990–10008.
- (73) Ballesteros, J. A.; Weinstein, H.; Stuart, C. S. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In *Methods in Neurosciences*; Academic Press: San Diego, CA, 1995; Vol. 25, pp 366–428.
- (74) Strader, C. D.; Sigal, I. S.; Candelore, M. R.; Rands, E.; Hill, W. S.; Dixon, R. A. F. Conserved aspartic acid residues 79 and 113 of the beta-adrenergic receptor have different roles in receptor function. *J. Biol. Chem.* **1988**, *263*, 10267–10271.
- (75) Strader, C. D.; Gaffney, T.; Sugg, E. E.; Rios Candelore, M.; Keys, R.; Patchett, A. A.; Dixon, R. A. F. Allele-specific activation of genetically engineered receptors. *J. Biol. Chem.* **1991**, *266*, 5–8.
- (76) Cascieri, M. A.; Fong, T. M.; Strader, C. D. Molecular characterization of a common binding site for small molecules within the transmembrane domain of G-protein coupled receptors. *J. Pharmacol. Toxicol. Methods* **1995**, *33*, 179–185.
- (77) Liu, S.; Crider, A. M.; Tang, C.; Ho, B.; Ankersen, M.; Stidsen, C. E. 2-pyridylthioureas: novel nonpeptide somatostatin agonists with SST4 selectivity. *Curr. Pharm. Des.* **1999**, *5*, 255–263.
- (78) Isaacs, R. C.; Solinsky, M. G.; Cutrona, K. J.; Newton, C. L.; Naylor-Olsen, A. M.; Krueger, J. A.; Lewis, S. D.; Lucas, B. J. Structure-based design of novel groups for use in the P1 position of thrombin inhibitor scaffolds. Part 1: Weakly basic azoles. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 338–342.
- (79) Tomperi, J.; Engstrom, M.; Wurster, S. Sulfonylamino-peptidomimetics active on the somatostatin receptor subtypes 4(sstr4) and 1 (sstr1). WO 2005033124A, 2004.
- (80) Liu, S.; Tang, C.; Ho, B.; Ankersen, M.; Stidsen, C. E.; Crider, A. M. Nonpeptide somatostatin agonists with sst4 selectivity: synthesis and structure-activity relationships of thioureas. *J. Med. Chem.* **1998**, *41*, 4693–4705.
- (81) Crider, A. M.; Liu, S.; Li, T.; Mahajan, S.; Ankersen, M.; Stidsen, C. E. Somatostatin Receptor Subtype 4 (sst 4) Ligands: Synthesis and Evaluation of Indol-3-yl- and 2-Pyridyl-thioureas. *Lett. Drug Des. Discov.* **2004**, *1*, 84–87.
- (82) Ankersen, M.; Dorwald, F. Z.; Stidsen, C. E.; Crider, A. M. Constrained Somatostatin Agonists. U.S. Patent 6,020,349, 2000.
- (83) Chianelli, D.; Kim, Y. C.; Lvovsky, D.; Webb, T. R. Application of a novel design paradigm to generate general nonpeptide combinatorial scaffolds mimicking beta turns: synthesis of ligands for somatostatin receptors. *Bioorg. Med. Chem.* **2003**, *11*, 5059–5068.
- (84) Feytens, D.; Cescato, R.; Reubi, J. C.; Tourwe, D. New sst4/S-selective somatostatin peptidomimetics based on a constrained tryptophan scaffold. *J. Med. Chem.* **2007**, *50*, 3397–3401.
- (85) Prasad, V.; Birzin, E. T.; McVaugh, C. T.; Van Rijn, R. D.; Rohrer, S. P.; Chicchi, G.; Underwood, D. J.; Thornton, E. R.; Smith, A. B. III; Hirschmann, R. Effects of heterocyclic aromatic substituents on binding affinities at two distinct sites of somatostatin receptors. Correlation with the electrostatic potential of the substituents. *J. Med. Chem.* **2003**, *46*, 1858–1869.
- (86) Rohrer, S. P.; Birzin, E. T.; Mosley, R. T.; Berk, S. C.; Hutchins, S. M.; Shen, D.-M.; Xiong, Y.; Hayes, E. C.; Parmar, R. M.; Foor, F.; Mitra, S. W.; Degrado, S. J.; Shu, M.; Koop, J. M.; Cai, S.-J.; Blake, A.; Chan, W. W. S.; Pasternak, A.; Yang, L.; Patchett, A. A.; Smith, R. G.; Chapman, K. T.; Schaeffer, J. M. Rapid identification of subtype-selective agonists of the somatostatin receptor through combinatorial chemistry. *Science* **1998**, *282*, 737–740.
- (87) Hirschmann, R.; Hynes, J. Jr.; Cichy-Knight, M. A.; van Rijn, R. D.; Sprengeler, P. A.; Spoors, P. G.; Shakespeare, W. C.; Pietranico-Cole, S.; Barbosa, J.; Liu, J.; Yao, W.; Rohrer, S.; Smith, A. B. III. Modulation of receptor and receptor subtype affinities using diastereomeric and enantiomeric monosaccharide scaffolds as a means to structural and biological diversity. A new route to ether synthesis. *J. Med. Chem.* **1998**, *41*, 1382–1391.
- (88) Mowery, B. P.; Prasad, V.; Kenesky, C. S.; Angeles, A. R.; Taylor, L. L.; Feng, J.-J.; Chen, W.-L.; Lin, A.; Cheng, F.-C.; Smith, A. B. III; Hirschmann, R. Catechol: A minimal scaffold for non-peptide peptidomimetics of the i + 1 and i + 2 positions of the beta-turn of somatostatin. *Org. Lett.* **2006**, *8*, 4397–4400.
- (89) Wang, X. M. Synthesis of Thiohydantoins as Nonpeptide Somatostatin Receptor Ligands. M.S. Thesis, University of Louisiana at Monroe, Monroe, LA, 2003.
- (90) Crider, A. M. Unpublished work, 2011.
- (91) Li, T. Synthesis of Nonpeptide Somatostatin Ligands. M.S. Thesis, University of Louisiana at Monroe, Monroe, LA, 2000.