Recognition of Privileged Structures by G-Protein Coupled Receptors

Kent Bondensgaard,*,‡ Michael Ankersen,⊥ Henning Thøgersen,‡ Birgit S. Hansen,# Birgitte S. Wulff,# and Robert P. Bywater‡

Protein Engineering, Medicinal Chemistry, and Discovery Biology, Novo Nordisk A/S,
*Novo Nordisk Park, DK-2760 Måløv, Denmark*

Privileged structures are ligand substructures that are widely used to generate high-affinity ligands for more than one type of receptor. To explain this, we surmised that there must be some common feature in the target proteins. For a set of class A GPCRs, we found a good correlation between conservation patterns of residues in the ligand binding pocket and the privileged structure fragments in class A GPCR ligands. A major part of interior surface of the common ligand binding pocket of class A receptors, identified in many GPCRs, is lined with variable residues that are responsible for selectivity in ligand recognition, while other regions, typically located deeper into the binding pocket, are more conserved and retain a predominantly hydrophobic and aromatic character. The latter is reflected in the chemical nature of most GPCR privileged structures and is proposed to be the common feature that is recognized by the privileged structures. Further, we find that this subpocket is conserved even in distant orthologs within the class A family. Three pairs of ligands recognizing widely different receptor types were docked into receptor models of their target receptors utilizing available structure-activity relationships and mutagenesis data. For each pair of ligands, the ligand−receptor complexes reveal that the nature of the privileged structure binding pocket is conserved between the two complexes, in support of our hypothesis. Only part of the privileged structures can be accommodated within the conserved subpocket. Some contacts are established between the privileged structure and the nonconserved parts of the binding pocket. This implies that any one particular privileged structure can target only a subset of receptors, those complementary to the full privileged structure. Our hypothesis leads to a valuable novelty in that ligand libraries can be designed without any foreknowledge of the structure of the endogenous ligand, which in turn means that even orphan receptors can in principle now be addressed as potential drug targets.

Introduction

Members of the superfamily of the integral membrane G-protein coupled receptors (GPCRs or 7TM receptors) perform a key information-transfer role in eukaryote biology. They mediate intercellular communication over the cell membrane by recognition of extracellular messengers followed by coupling to intracellular heterotrimeric G proteins which activate secondary-messenger pathways. Primary messenger molecules include endogenous biogenic amines, peptides, amino acids, glycoproteins, lipids, nucleotides, ions, and proteases as well as exogenous messengers such as odor and taste molecules. Activated heterotrimeric G proteins subsequently modulate a variety of intracellular responses that regulate cell function.

The ability to selectively regulate GPCR activity provides means for intervening in many cellular signaling pathways and modulating cellular responses that are associated with disease states. With these as potential drug targets, ligands acting as agonists or antagonists have been developed for a number of receptor systems. Identification of recurring fragments in these ligands for a variety of GPCRs led to the concept of privileged structures[1] (Figure 1). A privileged struc-
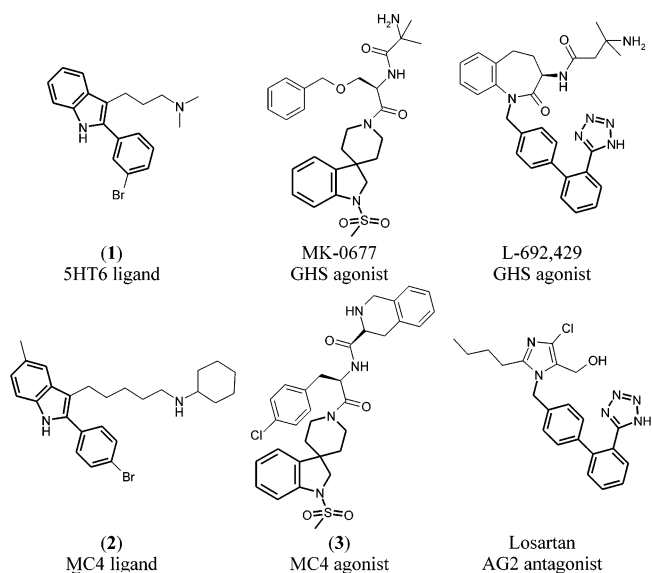
Figure 1. The three pairs of GPCR ligands containing privileged structure fragments used in this study. (Left) The 2-phenylindole derived ligands (**1**) and (**2**) target the 5HT6 and MC4 receptors, respectively. (Middle) The spiro-piperidine-indane derived ligands MK-0677 and (**3**) target the GHS and MC4 receptors, respectively. (Right) The 2-tetrazole-biphenyl ligands Losartan and L-692,429 target the AG2 and GHS receptors, respectively.

ture is defined as a selected substructure that is able to provide high-affinity ligands for more than one type

* To whom correspondence should be addressed. Telephone: +45 4443 4510. Fax: +45 4466 3450. E-mail: kbaa@novonordisk.com.
‡ Protein Engineering.
⊥ Medicinal Chemistry.
Discovery Biology.

Table 1. Affinities of Ligands toward Receptors

receptor	$(1)^a$	(2)	receptor	MK-0677	(3)	receptor	L-692,429	losartan
5HT6-R	0.7 nM		GHS-R	0.4 nM	190 nM	AG2-R	3600 nM	28 nM
MC4-R		612 nM	MC4-R	NB^b	100 nM	GHS-R	2800 nM^c	NB^b

^{a} Data on (1) and (2) are from ref 11. IC$_{50}$ values for the AG2, GHS, and MC4 receptors were determined in competition assays using saralasin,[31] ghrelin,[15] and NDP-α-MSH as tracers (see experimental section). ^{b} No binding detected. ^{c} Previously, a binding affinity of 63 nM[17] and a potency of 60 nM[18] has been determined in a porcine and rat pituitary assay, respectively.

of receptor. Privileged structures have successfully been utilized for design of ligands in receptor systems, for example, a new class of growth hormone secretagogues were identified by screening compounds from other receptor projects.[2] The spiro-piperidine-indane core of these compounds is also found in oxytocin, somatostatin, tachykinines, melanocortin, anaphylatoxin chemotactic receptor ligands, and is a privileged structure according to the definition given above. To date, no straightforward explanation to the concept of privileged structures in GPCR systems exists.

One hypothesis is that the privileged structures recognize a conserved binding pocket in a subset of GCPRs and that this "common" binding domain is complementary to the privileged structure. To investigate such a hypothesis, a thorough analysis of a number of GPCRs and the corresponding ligands is needed. Here we report an analysis of the class A family of GPCRs with emphasis on issues important for ligand recognition. Using sequence analysis tools together with the crystal structure of bovine rhodopsin,[3] a generic ligand binding pocket for the class A family has been analyzed. This pocket corresponds well with a pocket identified earlier in many different GPCRs on the basis of both receptor and ligand mutation studies[4] and has become regarded as a common feature throughout the entire class A family.[5−9] The pocket is located between transmembrane helices (TMs) 3, 5, and 6, and overlaps partially with a mini-core of ∼35 conserved residues. The overlapping subpocket is mainly composed of aromatic and nonpolar residues, which renders them complementary to several privileged structures such as 2-phenyl-indole, spiro-piperidine-indane, and 2-tetrazole-biphenyl. Three pairs of ligands (Figure 1), sharing a privileged structure as a common subfragment while still targeting widely different receptor types, were docked into models of their target receptors. The resulting complexes support the hypothesis that the privileged structures can be accommodated in the conserved subpocket. The existence of such a conserved (or privileged) pocket provides a straightforward explanation to the concept of privileged structures in GPCR systems and could be an additional aid to the design of new ligands for GPCRs.

Results

Degree of Conservation throughout the Receptor. 111 class A human GPCR sequences were aligned and used for calculation of positional entropy with the program AL2CO[10] for each of the 208 residue positions within the seven-transmembrane (7TM) domain. In Figure 2A, residues in the rhodopsin structure are color-coded according to the entropy. Blue and red correspond to the highest and lowest degree of conservation, respectively. The entropy or residue type conservation varies over the 7TM domain with the buried residues in the central part being the most conserved, and the membrane or solvent facing residues toward either the extra- or intracellular face the most variable.

Privileged Structure Ligands. To investigate our hypothesis about a conserved ligand binding pocket ligands containing 2-phenyl-indole, spiro-piperidine-indane, or 2-tetrazole-biphenyl privileged structure fragments (Figure 1) were selected from the literature. In each case, pairs of ligands targeting different receptor types were chosen. The 2-phenyl-indole derived ligands (1) and (2) target the 5-hydroxytryptamine-6 (5HT6) and melanocortin-4 (MC4) receptors, respectively,[11] the spiro-piperidine-indane derived ligands MK-0677[2] and (3)[12] target the growth hormone secretagogue (GHS) and MC4 receptors, respectively, and finally the 2-tetrazole-biphenyl ligands Losartan[13] and L-692,429[14] target the type-1 angiotensin II (AG2) and GHS receptors, respectively. For each ligand, affinities toward its target receptor and toward the second receptor, recognizing the same privileged structure, were determined in binding assays[15,16] (or see Methods) or in the case of (1) and (2) the binding data were taken from the literature[11] (Table 1). With compounds (1) and (2) only affinities toward their target receptors were given;[11] however, it seems reasonable to assume that absence of affinities implies binding affinities of less than 10 μM. For each pair of ligands, the one targeting a given receptor displays a significantly higher affinity for this receptor than the second ligand. L-692,429 only displays an affinity of 2800 nM for the GHS receptor while Losartan does not bind at 100 μM. However, a binding affinity of 63 nM[17] and a potency of 60 nM[18] have previously been determined in porcine and rat pituitary assays, respectively. The discrepancy observed between the two binding assays could be due to different binding epitopes of the two tracers ghrelin and MK0677 as discussed by Hansen et al.[15]

Docking of Privileged Structure Ligands. Each ligand was docked into a receptor model of the 7TM domain of its target receptor. Sequence alignment of the four sequences is shown in Figure 3 together with numbering schemes both for the individual sequences and according to the Ballesteros-Weinstein convention.[19]

2-Phenyl-indole Containing Ligands (Figures 1 and 4). In the 5HT6 and MC4 receptors, the indole fragment of the 2-phenyl-indole privileged structure fits well into a pocket between TM 5 and 6 spanned by a group of conserved aromatic residues (PheV.47, PheVI.44, and TrpVI.48). The bromo-phenyl group is placed between TM 3 and 5 pointing toward TM 4. A meta substitution pattern in (1) places the bromine within van der Waals distance of Phe188(V.38) in the top of TM 5 in the 5HT6 receptor, whereas a para substitution pattern in (2) places the bromine close to Phe176(IV.60) in the top of TM 4 in the MC4 receptor. The indole nitrogen forms a hydrogen bond with a backbone

A

B
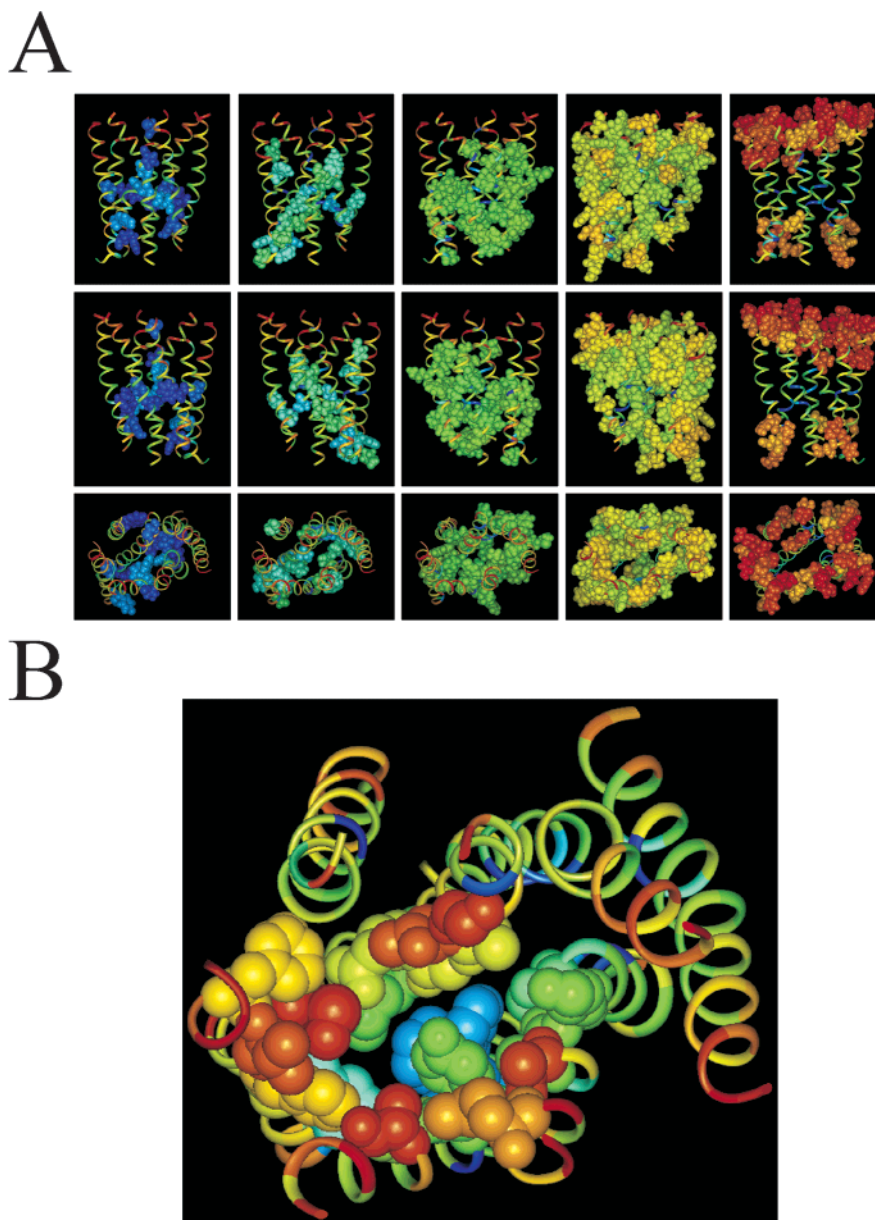
Figure 2. (A) Rhodopsin structure with residues color-coded according to entropy, where blue and red corresponds to the highest and lowest degree of conservation, respectively. Residues are divided into five groups by dividing the entropy range into five equisized regions. In the first column residues in the lowest region of the entropy range are showed as space-filling atoms, in the second column residue in the second lowest region, and etc. Side view of TM 1, 5, 6, and 7 (top), side view of TM 2, 3 and 4 (middle), and top view of extracellular side (bottom). (B) Rhodopsin structure viewed from the extracellular side with residues facing the binding pocket showed as space-filling atoms. Residues are color-coded according to entropy, where blue and red corresponds to the highest and lowest degree of conservation, respectively.

carbonyl oxygen on TM 5 in both cases. This is an important conserved feature of TM 5 since this carbonyl is located in the i-4th position relative to the highly conserved Pro residue in TM 5 and is unable to form a hydrogen bond to the amide nitrogen in that residue. In both receptors, the amino-alkyl group attached to the indole is extended and positioned parallel with TM 3 between TM 3 and 6 with the basic nitrogen within contact distance of Asp106(III.32) and Asp118(III.29) in the 5HT6 and MC4 receptors, respectively.

Spiro-piperidine-indane Containing Ligands (Figures 1 and 5). Mutagenesis data and structure−activity relationships available for MK-0677 were used to guide the docking. The spiro-piperidine-indane privileged structure part of MK-0677 can be substituted with aromatic/lipophilic moieties of a similar size.[20] Polar substituents such as NSO_2CH_3 or CHCOOH are preferred in the spiro-indane benzylic position.[2,21,22] The benzyloxymethyl group in MK-0677 can be substituted with a range of aromatic isosteres as well as 2-cyclohexylethyl without sacrifycing potency.[23] The demands on the amino-alkyl group are quite strict, both chain length and the degree of branching have great impact on potency.[24] Glu124(III.33) in TM 3 acts as a counterion for the amino group, as MK-0677 can activate the E124D mutant but not the E124Q mutant. This pattern is reversed when substituting the amino group in MK-0677 with a hydroxyl group providing strong evidence for an interaction between Glu124(III.33) and the amino group.[25]

In the GHS receptor, the indane fragment of the spiro-piperidine-indane privileged structure docks in the

```
TM 1          I.30                              I.59
bOPSD (35)    WQFSMLAAYMFLLIMLGFPINFLTLYVTVQ
h5H6R (24)    GGSGWVAAALCVVIALTAAANSLLIALICT
hMC4R (34)    EQLFVSPEVFVTLGVISLLENILVIVAIAK
hGHSR (41)    PLLAGVTATCVALFVVGIAGNLLTMLVVSR
hAG2R (26)    YIFVMIPTLYSIIFVVGIFGNSLVVIVIYF

TM 2          II.38                             II.67
bOPSD (71)    PLNYILLNLAVADLFMVFGGFTTTLYTSLH
h5H6R (60)    TSNFFLVSLFTSDLMVGLVVMPPAMLNALY
hMC4R (70)    PMYFFICSLAVADMLVSVSNGSETIIITLL
hGHSR (77)    TTNLYLSSMAFSDLLIFLCMPLDLVRLWQY
hAG2R (62)    VASVFLLNLALADLCFLLTLPLWAVYTAME

TM 3          III.22                           III.5'
bOPSD (107)   PTGCNLEGFFATLGGEIALWSLVVLAIERYVVVCKP
h5H6R (96)    RGLCLLWTAFDVMCCSASILNLCLISLDRYLLILSP
hMC4R (111)   VNIDNVIDSVICSSLLASICSLLSIAVDRYFTIFYA
hGHSR (113)   DLLCKLFQFVSESCTYATVLTITALSVERYFAICFP
hAG2R (98)    NYLCKIASASVSFNLYASVFLLTCLSIDRYLAIVHP

TM 4          IV.38                    IV.61
bOPSD (149)   GENHAIMGVAFTWVMALACAAPPL
h5H6R (139)   TPLRALALVLGAWSLAALASFLPL
hMC4R (154)   TVKRVGIIISCIWAACTVSGILFI
hGHSR (156)   TKGRVKLVIFVIWAVAFCSAGPIF
hAG2R (141)   TMLVAKVTCIIIWLLAGLASLPAI

TM 5          V.34                     V.61
bOPSD (199)   NNESFVIYMFVVHFIIPLIVIFFCYGQL
h5H6R (184)   ASLPFVLVASGLTFFLPSGAICFTYCRI
hMC4R (180)   SDSSAVIICLITMFFTMLALMASLYVHM
hGHSR (208)   GLLTVMVWVSSIFFFLPVFCLTVLYSLI
hAG2R (191)   LPIGLGLTKNILGFLFPFLIILTSYTLI

TM 6          VI.29                            VI.61
bOPSD (246)   AEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTH
h5H6R (262)   KALKASLTLGILLGMFFVTWLPFFVANIVQAVC
hMC4R (231)   ANMKGAITLTILIGVFVVCWAPFFLHLIFYISC
hGHSR (257)   NHKQTVKMLAVVVFAFILCWLPFHVGRYLFSKS
hAG2R (234)   RNDDIFKIIMAIVLFFFFSWIPHQIFTFLDVLI

TM 7          VII.32                           VII.58
bOPSD (285)   PIFMTIPAFFAKTSAVYNPVIYIMMNK
h5H6R (299)   PGLFDVLTWLGYCNSTMNPIIYPLFMR
hMC4R (273)   MSHFNLYLILIMCNSIIDPLIYALRSQ
hGHSR (302)   QYCNLVSFVLFYLSAAINPILYNIMSK
hAG2R (281)   DTAMPITICIAYFNNCLNPLFYGFLGK
```

Figure 3. Sequence alignment of bovine rhodopsin (bOPSD), 5-hydroxytryptamine-6 (h5H6R), melanocortin-4 (hMC4R), growth hormone secretagogue (hGHSR), and type-1 angiotensin II (hAG2R) for the 7TM domain. Individual sequence numbers for the first residue in each TM is given in parentheses. Above each TM the generic numbering according to Ballesteros-Weinstein is shown for the first and last residue position.

conserved pocket between TM 5 and 6, spanned by the conserved group of aromatic residues Phe221(V.47), Phe272(VI.44), and Trp276(VI.48), with the NSO$_2$CH$_3$ group facing TM 6 in the vicinity of H280(VI.52). Positioned central in the cavity between TM 3, 5, and 6, the piperidine moiety has contact with several mainly aromatic and lipophilic residues on all three TMs. The branching point between the benzyloxymethyl and amino-alkyl groups is positioned in the upper part of the cavity, and directs the amino group toward Glu124-(III.33) on TM 3, placing the alkyl group between TM 3 and 5, and places the benzyloxymethyl in a pocket between TM 3, 6, and 7. In the MC4 receptor, the spiro compound (**3**) is docked, in the same manner as MK-0677 in the GHS receptor, with the privileged structure in the conserved pocket, the tetrahydro-isoquinoline moiety between TM 3 and 5, and the 4-chloro-phenyl group in the pocket between TM 3, 6, and 7. The tetrahydro-isoquinoline moiety is positioned toward TM

3 where the amino functionality can interact with Asp118(III.29), which is positioned one turn up toward the extracellular side relative to Glu124(III.33) in the GHS receptor. The 4-chloro-phenyl group is placed as the benzyloxomethyl in the GHS receptor between TM 3, 6, and 7.

2-Tetrazole-biphenyl Containing Ligands (Figures 1 and 6). As with MK-0677, mutagenesis data and structure−activity relationships available for L-692,429 and Losartan were used to guide the docking. The salient features of Losartan are an electron-rich substituent on the biphenyl fragment (tetrazole in Losartan), a hydrogen acceptor in the 5-position on the imidazole (MeOH in Losartan), and an alkyl group in the 2-position on the imidazole (butyl in Losartan).[26] Mutagenesis studies on the AG2 receptor and structure−activity relationship on Losartan and derivatives hereof have indicated that regions around Val108(III.32) and Ser109(III.33) on TM 3[27,28] and around Lys199(V.42) and His256(VI.51) on TM 5 and 6, respectively,[29] are implicated in binding to Losartan. The tetrazole moiety presumable binds in the latter region.[29] Studies have shown that Asn294(VII.45) and Asn295(VII.46) also perturb binding, for example, the N295S mutation results in a ~70-fold decrease in binding affinity.[30,31] However, both of these effects are most likely indirect as the N294A mutation results in an impaired receptor, indicating a structural role in stabilizing the active state of the receptor for this residue,[30] and as Asn295(VII.46) is buried and points away from the binding cavity in our model of the AG2 receptor. As in Losartan, a main pharmacophoric element of L-692,429 is a biphenyl fragment with an electron-rich substituent like tetrazole.[32,33] In addition both an amino group[34] and the aromatic moiety in the benzazepione (M. Ankersen, unpublished results) are necessary for high affinity toward the GHS receptor. Similarly to MK-0677, the Glu124(III.33) in TM 3 acts as a counterion for the amino group, as a close derivative of L-692,429 can activate the E124D mutant but not the E124Q mutant.[25]

In the AG2 receptor, the tetrazole moiety of Losartan was placed in the vicinity of Lys199(V.42), Asn200-(V.43), and Gln257(VI.52) with the biphenyl fragment positioned between the conserved aromatic residues on TM 5 and 6 (PheV.47, PheVI.44, and TrpVI.48). The imidazole was placed in the center of the cavity between TM 3, 5, and 6 pointing toward TM 3. Such an arrangement allows the hydroxyl group to interact with Lys199-(V.42) and the unsubstituted imidazole nitrogen to hydrogen bond with Ser105(III.29). The imidazole butyl substituent is situated in a narrow groove between TM 3 and 6 forming contacts with Val108(III.32). In the GHS receptor, the 2-tetrazole-biphenyl moiety of L-692,-429 was docked in the same region as the identical fragment of Losartan in the AG2 receptor. Again the biphenyl fragment can make favorable contacts with the conserved aromatic residues on TM 5 and 6. Lys199-(V.42), Asn200(V.43), and Gln257(VI.52) in the AG2 receptor are replaced with a Val216(V.42), Ser217(V.43), and His280(VI.52) in the GHS receptor, of these, only His280(VI.52) is complementary to the acidic tetrazole moiety. However, Thr260(VI.55) in TM 6 in the AG2 receptor is replaced with Arg283(VI.55) in the GHS receptor, which points toward TM 5 and, as a result,
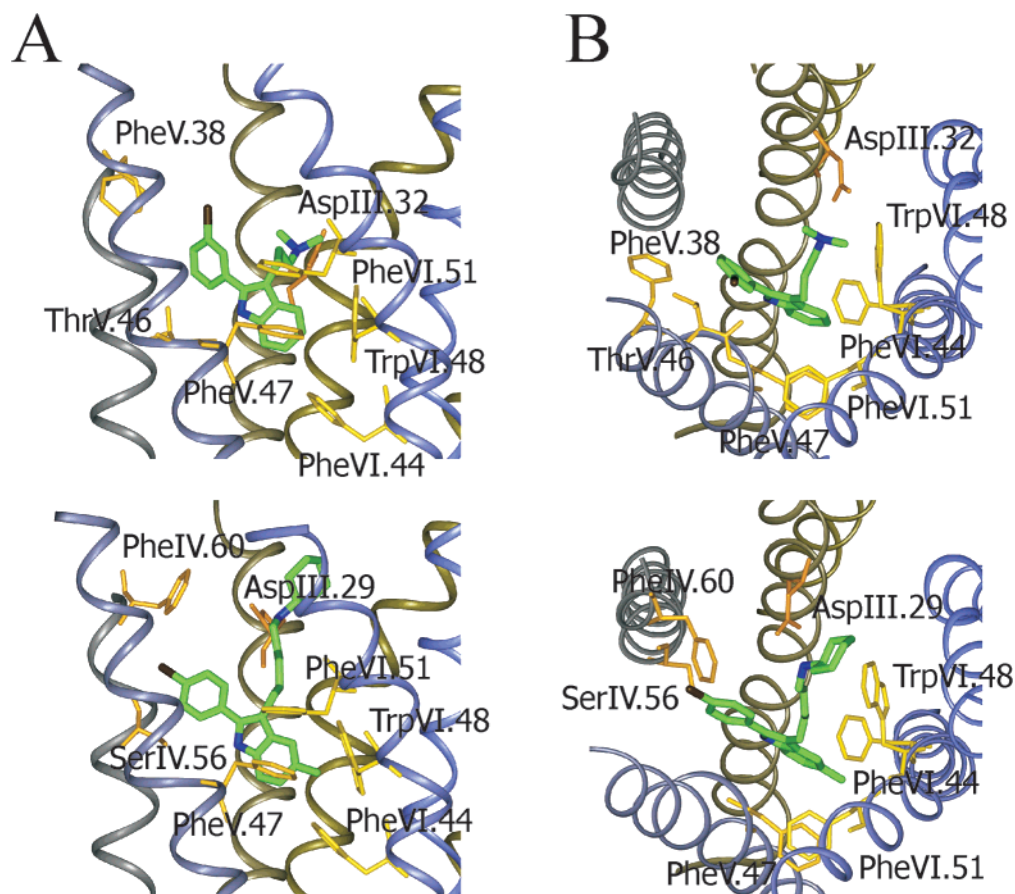
Figure 4. The ligand binding pocket of the (**1**)-5HT6 receptor (top) and (**2**)-MC4 receptor complexes (bottom). (A) Side view from with TM 5, 6, and 7 in front, and (B) top view from the extracellular side. Ligand atoms are color-coded according to atom types: carbon, green; oxygen, red; nitrogen, blue; and bromine, dark green. Receptor atoms are shown in yellow and TMs are shown in shades of brown (TM1) to blue (TM7).

places the positive charge in the same region of the GHS receptor as Lys199(V.42) does in the AG2 receptor. Thus, in both receptors a small cluster of positive ionisable residues is positioned in the same region and can interact favorably with the tetrazole moiety. The benzazepinone of L-692,429 is placed in the central part of the cavity between TM 3, 5, and 6 in the GHS receptor with the benzene ring pointing toward a group of aromatic residues on TM 6 and 7, and the amino group within contact distance of Glu124(III.33).

Root-Mean-Square-Deviation (RMSD). Each pair of ligand−receptor complexes was superimposed and the RMSD between the privileged structure part of the ligands was computed. For the 2-phenyl-indole, spiro-piperidine-indane, and 2-tetrazole-biphenyl privileged structures the RMSD was 0.6, 0.9, and 2.5 Å, respectively. The large RMSD for the 2-tetrazole-biphenyl case is partly due to a large difference between the two complexes in the position of the tetrazole group caused by a rotation around the long axis of the biphenyl. The corresponding value for just the biphenyl part is 1.8 Å.

Discussion

The concept of privileged structures in GPCRs has predominantly been described from a ligand point of view. Collections of ligands targeting a variety of receptors while sharing a common subfragment, a privileged structure, have been compiled.[1,35] These privileged structures have successfully been incorpo-

rated into compound libraries used to screen against other GPCR types.[2,11] However, without an understanding of how these fragments are recognized by GPCRs any rational design approach utilizing privileged structure will be intractable. An appealing explanation to the prevalence of privileged structures in ligands recognizing a range of GPCRs would be the existence of a complementary "conserved" binding domain in the GPCRs.

Structural Characteristics of GPCRs. All GPCRs share the common feature of seven transmembrane-spanning α-helices connected by alternating intra- and extracellular loops with the amino terminus located on the extracellular side and the carboxy terminus on the intracellular side, respectively. On the basis of sequence similarity the family can be divided into several sub-families with no sequence similarity between subfami-lies.[36] The class A family is by far the largest and includes receptors related to the rhodopsin and β-adre-nergic receptors. Despite a highly diverse family with a low overall sequence homology (∼20−30%), a number of key residues are highly conserved indicating that the overall topology of the 7TM bundle is highly conserved between class A receptors.

The only receptor structure determined to date is the crystal structure of bovine rhodopsin,[3] which has pro-vided detailed knowledge about the structural organiza-tion of the 7TM bundle including loop regions. Helix−helix packing contacts have been unambiguously identi-
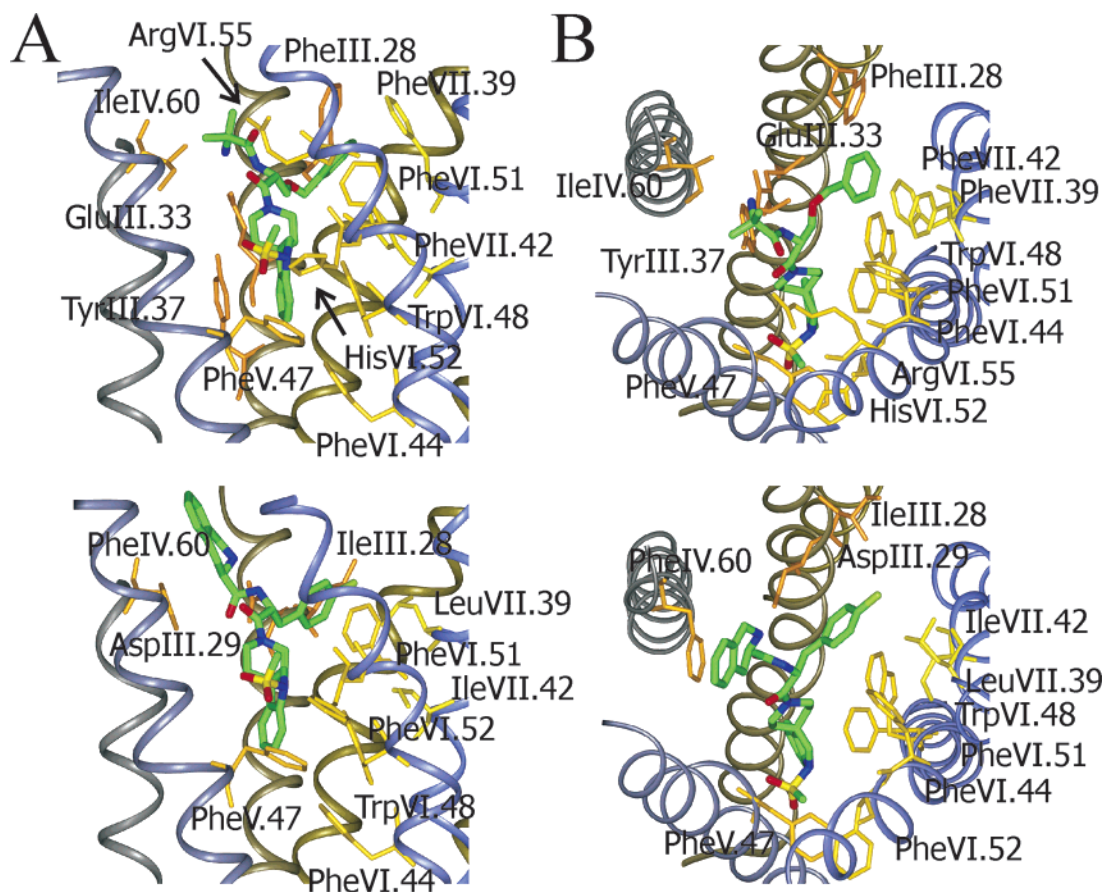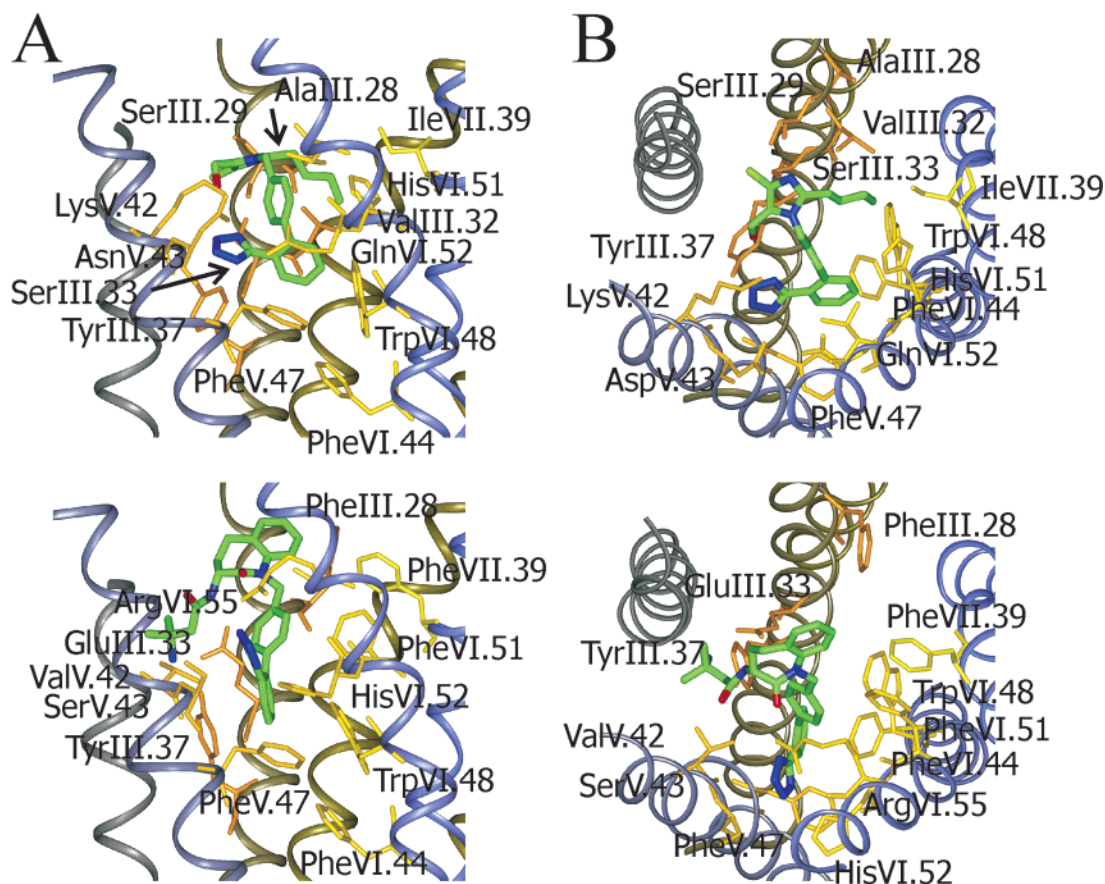
Figure 5. The ligand binding pocket of the MK-0677-GHS receptor (top) and (**3**)-MC4 receptor complexes (bottom). (A) Side view from with TM 5, 6, and 7 in front, and (B) top view from the extracellular side. Ligand atoms are color-coded according to atom types: carbon, green; oxygen, red; nitrogen, blue; chloride, light yellow; and sulfur, yellow. Receptor atoms are shown in yellow and TMs are shown in shades of brown (TM1) to blue (TM7).

fied together with helix–membrane contacts. Further details such as helix–helix crossing angles and deviations from ideal α-helices are also evident from the structure. The endogenous covalently attached ligand retinal–the light absorbing chromophore of rhodopsin–is bound in a cavity within the 7TM bundle, which is located between TM 3, 4, 5, 6, and 7.

The question of whether this structure constitutes an acceptable template for construction of homology models of other rhodopsin-like (class A) GPCRs has been thoroughly examined[37] with the conclusion that rhodopsin is indeed an excellent template, at least for antagonist studies. This conclusion emanates from the fact that the rhodopsin crystal structure was determined for the dark (i.e., inactive) form of the receptor, and it is known that certain conformational changes take place during activation.[38] For our purposes, while we do not wish to neglect this dynamical feature of GPCR structure, in the absence of a crystal structure of a representative GPCR, we prefer not to let our results become biased by a purported structure that may turn out to be inaccurate. For example, in the work in which the homology models were "activated",[37] there was judged to be a need to expand the size of the ligand binding pocket to accommodate agonists. This may be so, but there is no evidence to support this decision. In any case, it has been shown that while there are indeed certain conformational adjustments to the structure, both in the backbone and in side chains, there are at most only very few changes[38] to the residues that actually participate in ligand recognition. This further justifies our choice of a knowledge-based docking approach where experimental facts concerning the ligand moieties and the residues involved are used, rather than a method based on a molecular mechanical force field. The latter requires an accurate structure for the receptor and not one that has been adjusted according to purely qualitative criteria.

Sequence/Structure Conservation. The degree of residue type conservation varies a great deal over the 7TM domain (Figure 2A). In general, residues facing the 7TM bundle (inward facing) are more conserved than those facing the membrane (outward facing). This is in accordance with previous observations and is rationalized by the structural constraints on buried residues for optimal packing with surrounding residues.[39] Among the inward facing residues, there are also large variations in the degree of conservation. A number of residues such as the GN motif in TM 1 and the NPXXY motif (where X is any amino acid) in TM 7 are highly conserved. These key residues most likely play an essential role for structural and/or functional integrity of the receptors and tend to cluster in the central part of the 7TM domain when mapped onto the rhodopsin structure (Figure 2A). The most variable or less conserved residues are predominately found at the extracellular face of the receptor, lining what is generally considered to be the ligand binding site in many GPCRs. Some highly variable residues are also found at the intracellular face of the receptor (Figure 2A). Not

Figure 6. The ligand binding pocket of the Losartan-AG2 receptor (top) and L-692,429−GHS receptor complexes (bottom). (A) Side view from with TM 5, 6, and 7 in front, and (B) top view from the extracellular side. Ligand atoms are color-coded according to atom types: carbon, green; oxygen, red; nitrogen, blue; and chloride, light yellow. Receptor atoms are shown in yellow and TMs are shown in shades of brown (TM1) to blue (TM7).

included in the analysis are both intra- and extracellular loops together with C- and N-termini. These regions could not be included in the analysis as both length and sequences are nonconserved, precluding any meaningful sequence alignment.

Overall the 7TM domain consists of a conserved central part joined with two more variable regions—one at the extracellular face and the other at the intracellular face of the 7TM domain. A simplified view of a GPCR receptor would be two variable interfaces—one interacting with various endogenous ligands in the extracellular environment and one interacting with intracellular G-proteins—joined by a conserved machinery responsible for transducing a signal from the extracellular to the intracellular interface.

Ligand Binding Site. A ligand binding pocket located between TM 3, 5, 6, and 7 has been identified in many GPCRs systems by mutagenesis studies.[4−8] The presence of such a ligand binding pocket has become regarded as a common feature throughout the entire GPCR superfamily, even including members of class B[40] which bind their endogenous ligand at mainly extracellular sites. The rhodopsin structure revealed that the retinal binding site indeed also is found in this part of the receptor in accordance with earlier studies. In Figure 2B, residues lining this binding pocket are shown as space filling atoms. In the rhodopsin structure, the E2 loop forms a lid on the binding cavity and consequently several of these residues contact the retinal. A pair of conserved cysteines in the E2 loop and the top

of TM 3, respectively, has led to the general consensus that such an arrangement of the E2 loop is general for most class A GPCRs (with some exceptions, notably, the melanocortin family). However, the loop involvement in ligand binding could not be analyzed due to lack of homology within the loop regions.

Most residues lining the ligand binding site are moderately to highly variable, which fits the need for recognition of a wide range of ligands. Among the highly variable are residues on TM 3 and 5, including those involved in binding biogenic transmitter molecules as indicated in mutagenesis studies.[41] For example, residue AspIII.32 conserved throughout the biogenic transmitter subfamily (but still highly variable within the entire class A family) forms a salt-bridge with the amino group present in all these transmitter molecules, and SerV.43 and SerV.46 in the adrenergic receptors form hydrogen bonds with the hydroxyls on the catechol ring of epinephrine.[8,9] Within the pocket the trend is that residues toward the lower part of the pocket are more conserved. In particular, the subpocket spanned by residues V.47, VI.44, and VI.48 between TM 5 and 6 is highly conserved. Residues V.47 and VI.44 are either Phe or Tyr in 91 and 98% of the receptors, respectively, and residue VI.48 is Trp in 90% of the receptors. Thus, this cluster of aromatic residues is conserved in the vast majority of receptors and seems to mark the borderline between the more conserved central part of the 7TM domain and the more variable part toward the extracellular face. In the rhodopsin structure, the β-ionone

Table 2. Contact Residues for the Privileged Structure Part of Each Ligand[a]

residue position[b]	2-phenyl-indole		spiro-piperidine-indane		2-tetrazole-biphenyl	
	5HT6-R	MC4-R	GHS-R	MC4-R	AG2-R	GHS−R
III.33	V107	C122				
III.36	C110	L126	T127	L125	L112	T127
III.37	S111	L125	Y128	L126	Y113	Y128
III.40	I114	I129	V131	I129		
IV.56	A157	S172				
IV.60		F176				
V.38	F188					
V.39						
V.42	A192	C188	V216	C188	K199	V216
V.43			S217	L189	N200	S217
V.46	T196	M192				
V.47	F197	F193	F221	F193	F204	F221
VI.44	F277	F246	F272	F246	F249	F272
VI.48	W281	W250	W276	W250	W253	W276
VI.51			F279	F253	H256	F279
VI.52	F285	F254	H280	F254	Q257	H280
VI.55			R283	L257	T260	R283

[a] Residues within a 4 Å radius of the privileged structure were selected. [b] Numbering according to Ballesteros-Weinstein.

ring of retinal packs within this subpocket of conserved aromatic residues.

Docked Privileged Structure Ligands. The sequence analysis revealed that, while there is considerable variability in the ligand binding pocket, a part of the pocket is highly conserved. Such a conserved subpocket is an attractive binding site candidate for the privileged structures as this part of the pocket would remain conserved even between distant orthologues within the class A family. To test this hypothesis, three pairs of ligands (Figure 1) were docked into receptor models of their target receptors. Each pair targets two different receptor types belonging to the class A family and shares a common subfragment, the privileged structure, but are otherwise different. Ideally, the privileged structure part of each pair should dock in the same region, whereas the remainder—the nonconserved part of the ligands—should target distinct features in the corresponding target receptors and consequently lead to receptor specificity of the particular ligand.

2-Phenyl-indole Containing Ligands (Figures 1 and 4). With an RMSD of only 0.6 Å between the 2-phenyl-indole part of the two superimposed complexes, the privileged structure part of the ligand is docked in near identical positions in the two complexes. The indole moiety is positioned in the subpocket spanned by the three conserved aromatic residues PheV.47, PheVI.44, and TrpVI.48 as well as PheVI.52, conserved between the 5HT6 and MC4 receptors, making favorable aromatic—aromatic interactions. The bromo-phenyl part of the privileged structure is placed between TM 3 and 5, a nonpolar region in both receptors except for Ser111-(III.37) and Thr196(V.46) in the 5HT6 receptor and Ser172(IV.56) the MC4 receptor. Thus, the privileged structure makes similar contacts in both receptors (Table 2). The main difference between the two ligands is the linker connecting the amino group with the indole. In (**1**) and (**2**), propylene and pentylene, respectively, are used as linkers leading to differences in the ligand—receptor complexes. Both linkers are positioned in a groove between TM 3 and 6, but the longer pentylene

linker in (**2**) can reach longer into the groove than the shorter propylene linker. However, this correlates very well with the position of the Asp in TM 3 acting as a counterion for the amino group in both complexes. In the 5HT6 receptor, the Asp106(III.32) is postioned three turns from the extracellular end of TM 3, whereas Asp118(III.29) in the MC4 receptor is positioned almost a full turn closer to the extracellular end. This means that the distance from the indole to the Asp is longer in the MC4 receptor than in the 5HT6 receptor correlating nicely with the need for a longer linker in the MC4 receptor ligand (**2**).

Spiro-piperidine-indane Containing Ligands (Figures 1 and 5). Again the privileged structure part of the two ligands is in very similar positions in the two complexes with an RMSD of only 0.9 Å. As with 2-phenyl-indole, the indane part of the spiro-piperidine-indane docks in the subpocket formed by the cluster of three conserved aromatic residues PheV.47, PheVI.44, and TrpVI.48 along with Tyr128(III.37) and His280-(VI.52) in the GHS receptor, and Phe254(VI.52) in the MC4 receptor (Table 2). In both cases, the NSO_2CH_3 group faces TM 6. Neighboring residues His280(VI.52) and Arg283(VI.55) in the GHS receptor might also interact favorably with the NSO_2CH_3 group explaining the preference for polar groups in this position. In both the GHS and MC4 receptors, the piperidine ring is centered in the cavity between TM 3, 5, and 6 having contacts to several lipophilic residues on all three TMs (Table 2). Besides the privileged structure the two ligands also share a dipeptide scaffold resulting in an almost identical placement of the attached groups in the two ligand—receptor complexes. In both ligands, the two attached groups contain an amino group and an aromatic group, respectively. The amino-carrying group is 2-amino-isobutyl in MK-0677 and tetrahydro-isoquinoline in (**3**), one aliphatic and the other more aromatic in nature. This is mirrored in the receptors where Ile178(IV.60) in the GHS receptor is substituted for Phe176(IV.60) in the MC4 receptor. The amino group forms a salt-bridge with an Asp or Glu on TM 3. Even though these acidic residues are positioned one turn apart in the two receptors, the amino group in the ligands is docked almost in identical positions between the two residues. The third group attached to the dipeptide scaffold, besides the privileged structure and the amino-carrying group, is benzyloxymethyl and 4-chlorobenzyl in MK-0677 and (**3**), respectively. In the GHS receptor—ligand complex the benzene ring of benzyloxymethyl in MK-0677 is placed in a pocket formed by four aromatic residues Phe119(III.28), Phe279-(VI.51), Phe309(VII.39), and Phe312(VII.42). Three of these Phe119(III.28), Phe309(VII.39), and Phe312-(VII.42) are substituted with aliphatic residues Ile117-(III.28), Leu280(VII.39), and Ile283(VII.42), respectively, in the MC4 receptor. This correlates with the position of the benzene ring of the 4-chlorobenzyl group in (**3**), which in the MC4 receptor is positioned closer to TM 6 next to the last remaining aromatic residue Phe253-(VI.51). This is due to a shorter linker between the dipeptide scaffold and the benzene ring in (**3**) compared to MK-0677. So even though the ligand—receptor complexes overall are very similar as is also reflected by

the similar affinities of (**3**) toward the GHS and MC4 receptors, small differences do prevail.

2-Tetrazole-biphenyl Containing Ligands (Figures 1 and 6). The biphenyl is positioned in a subpocket, spanned by the three conserved residues PheV.47, PheVI.44, and TrpVI.48 similar to the 2-phenyl-indole and spiro-piperidine-indane privileged structures. Other contacts includes TyrIII.37 conserved between the AG2 and GHS receptors, and His256(VI.51) and Phe279-(VI.51) in the AG2 and GHS receptor, respectively, leading to numerous aromatic−aromatic interactions (Table 2). The tetrazole group is placed between TM 5 and 6, where Lys199(V.42), Asn200(V.43), and Gln257-(VI.52) in the AG2 receptor can stabilize a negative charge on the tetrazole. In the GHS receptor, Asn200-(V.43) and Gln257(VI.52) are substituted with Ser217-(V.43) and His280(VI.52), respectively, and the role of Lys199(V.42) is taken over by Arg283(VI.55) on TM 6 as described above (Table 2). In both cases, the combination of the conserved cluster of aromatic residues together with a group of positive ionizable residues on TM 5 and/or 6 serves well as a binding pocket for the 2-tetrazole-biphenyl privileged structure. An RMSD of 2.5 Å between the 2-tetrazole-biphenyl part of the ligands is fairly large compared to the previous two cases. However, this is partly due to a different orientation of the tetrazole resulting from differences in position of the positive ionisable residues stabilizing the negative charge on the tetrazole. The RMSD for just the biphenyl part is only 1.8 Å. Another explanation may be that both ligands were docked into homology models based on the rhodopsin structure regardless of ligand type (agonist or antagonist). These models represent an inactive state of class A receptors and may diverge somewhat from the active state recognized by agonists. The imidazole and benzazepinone of Losartan and L-692,492, respectively, are both roughly positioned in the same region in the cavity between TM 3, 5, and 6. Substituents on these core fragments are the butyl and hydroxymethyl substituents on the imidazole of Losartan, and the benzene and amino-akyl substituents on the azepinone ring of L-692,429. Even though the directionality of the substituents on the two core fragments do not match exactly, the substituents pair up so the butyl and benzene groups, and the hydroxymethyl and aminoalkyl groups occupy the same space in the AG2 and GHS receptors. The butyl and benzene groups are positioned between TM 3, 6, and 7. In the GHS receptor, a cluster of aromatic residues Phe119(III.28), Phe279(VI.51), and Phe309(VII.39) forms aromatic−aromatic interactions with the bezene ring. Two of these residues Phe119(III.28) and Phe309(VII.39) are, in contrast, substituted with aliphatic residues Ala104-(III.28) and Ile288(VII.39) in the AG2 receptor explaining the preference for an aliphatic over an aromatic group in the AG2 receptor compared with the GHS receptor. The hydroxymethyl and amino-alkyl groups are found between TM 3, 4, and 5, where the amino group of L-692,492 interacts with Glu124(III.33) in the GHS receptor as shown by mutagenesis. In the AG2 receptor the hydroxyl is pointing toward TM 5 and is fairly close to Lys199(V.42) which also interacts with the tetrazole. Proximity of the hydroxyl group and Lys199(V.42) also explains why the carboxylate deriva-
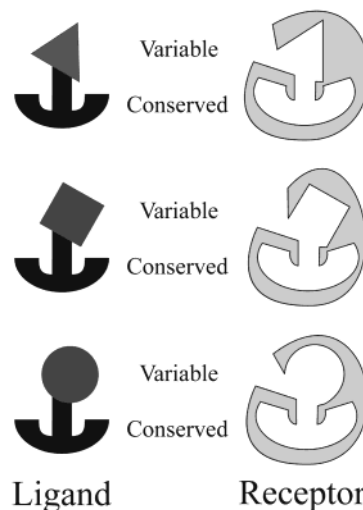
Figure 7. Schematic illustration of a class of ligands composed by a conserved privileged structure part and a variable part together with the corresponding receptor binding pockets.

tive of Losartan has a higher affinity toward the AG2 receptor.[42] In this respect, it is also interesting that the Glu124(III.33) in the GHS receptor is substituted with Ser109(III.33) in the AG2 receptor avoiding charge−charge repulsion between the two carboxylate groups.

Summary. For each pair of ligands the privileged structure part could be docked in the same region of the target receptors as measured by RMSD values between the privileged structure part of the superimposed ligand−receptor complexes. The nature of the privileged structure binding pocket was conserved between receptors even in the case of 2-tetrazole-biphenyl where the residue acting as counterion for the tetrazole was positioned on different TMs in the two complexes (Table 2). Differences between the nonprivileged parts of the ligands could to some degree be rationalized from differences in the receptors leading to a preference for one pharmacophoric element over another, and thus responsible for the observed selectivity displayed by the receptors toward the ligands. Ligand−receptor interactions can therefore be divided into two categories. One category being those interactions between receptor and privileged structure, which are conserved among complexes containing the same privileged structure, and the second category being those interactions between receptor and the remainder of the molecule, which may be highly variable (Figure 7).

Privileged Structures and the Privileged Receptor Site. All three examples of privileged structures dock in the same region of the receptors—the subpocket deep within the 7TM domain spanned by three conserved aromatic residues PheV.47, PheVI.44, and TrpVI.48 on TM 5 and 6. The existence of such a conserved (or privileged) pocket would seem to suggest a straightforward explanation to the concept of privileged structures in GPCR systems. However, with all three privileged structure this subpocket can only accommodate part of the privileged structure. Some contacts are established between the privileged structure and the nonconserved parts of the binding pocket (Table 2). This would imply that any one particular privileged structure can target only a subset of receptors, those complementary to the full privileged structure. A good example of

this is the 2-tetrazole-biphenyl privileged structure which requires one or more positive ionisable residues to stabilize the negative charge on the tetrazole. Such residues exist in the GHS and AG2 receptors but not in the MC4 receptor indicating that the latter receptor will fail to accommodate the 2-tetrazole-biphenyl privileged structure. So even though privileged structures can target widely different receptor subtypes, it seems unlikely that an universal privileged structure exists for all members of the class A family.

Typically, ligand recognition presupposes a match between ligand and binding pocket not only of "shape" but also of physicochemical interactions such as charge, dipolar interactions, and polarizability. Hydrogen bonds, for example, are important in this regard. They not only represent an important contribution to binding enthalpy, but provide directionality as well. But the residue types and the corresponding privileged structures are conspicuously aromatic in character, giving rise to a more or less nonpolar type of interaction. It is well-known that aromatic−aromatic interactions involve more than just the entropic contribution provided by the desolvation of nonpolar surfaces, but there is an enthalpic contribution too, due to $\pi-\pi$ interactions. These can be of comparable strength to hydrogen bonds, and likewise they are directional in character.[43] Thus, there would seem to be a natural explanation for the conservation of aromatic residues in the deeper regions of the binding pocket and the "conservation" of recurrent aromatic groups in the privileged structures. The problem with this conclusion is that the receptors have not evolved under the selection pressure of these ligands. These are modern in origin. Rather, we must resort to other lines of reasoning to explain this apparent compatibility between privileged structures and the corresponding residues which they so readily bind to. The three aromatic residues, spanning the pocket, overlap with a conserved mini-core of ∼35 residues (Figure 2A) and may have to do with an essential part of the machinery of the receptor, the transfer between active and inactive states, for example. They just happen to be partly exposed in the bottom of the ligand binding pocket. Further, they may be differently exposed in the active or inactive structures, such that small differences in privileged ligand structures, or in groups attached to them, may help to stabilize one or another state (active or inactive) of the receptor,[44] but the machinery itself is likely to be very similar in all GPCRs judged by the high degree of conservation in the central part of the receptors. This may be the explanation for the observation that privileged structures can be engineered with only small changes, to generate agonists or antagonists.[44] It is likely as improved GPCR models will appear in the future that the hypothesis here described may be an additional aid to design new ligands for GPCRs. We see an advantage in such a design since no endogenous ligand is required as starting point, and many orphan receptors may thus be amenable to ligand design simply by testing whether a given privileged structure binding site is present in a orphan receptor or not. To illustrate this, the binding site for the 2-tetrazole-biphenyl privileged structure in GHS receptor was compared with the corresponding binding site in other class A GPCRs having a positively charged

Table 3. Comparison of the 2-Tetrazole-Biphenyl Binding Site in the GHS Receptor with Other Class A GPCRs[a]

	receptor				
residue position[b]	GHS-R	NT1-R	APJ-R	CC6-R	ET1-R
III.36	T	T	M	F	V
III.37	Y	Y	Y	Y	G
V.42	V	N	S	Q	L
V.43	S	T	T	M	F
V.47	F	F	F	F	F
VI.44	F	F	F	F	F
VI.48	W	W	W	Q	W
VI.51	F	Y	Y	F	L
VI.52	H	H	H	N	H
VI.55	R	R	K	K	R

[a] Residues within a 4 Å radius of the privileged structure were selected. [b] Numbering according to Ballesteros-Weinstein.

amino acid in position VI.55 (Table 3). Evidently, the NT1 receptor has a binding site quite similar to the one in the GHS receptor, and a literature search also revealed a patent with compounds containing 2-tetrazole-biphenyl directed against neurotensin.[45] To our knowledge, the other three receptors in Table 3 do not have any known ligands with 2-tetrazole-biphenyl groups.

Experimental Procedures

Sequence Alignment and Conservation Entropy. Aligned subfamilies were downloaded from the GPCR database.[46] The multiple sequence alignments were submitted to PHDhtm[47] to predict the location and topology of transmembrane segments. Subsequently, the transmembrane part of the subgroups were aligned relative to each other using the predicted transmembrane segments and the conserved residues within each transmembrane helix. The exact location of the seven transmembrane segments were taken as the ends of the transmembrane helices in the rhodopsin structure[3] determined using the DSSP routine in WHAT IF.[48] As a control of the alignment, the program Perscan[39] was used to analyze the sequence alignment of the subgroups. Perscan analyses a multiple sequence alignment for conserved or variable positions based on the hypothesis that membrane-facing residues are less conserved than residues facing other helices in the 7TM bundle. In the case of an ideal α-helix, this should lead to a cosine type pattern of conserved/variable residues with a period of 3.6 residues. A very good agreement between different subgroups for all seven transmembrane helices validates the alignment.

For each position in the 7TM domain the degree of conservation in the multiple sequence alignment was evaluated using an entropy-based measure implemented in the program AL2CO.[10] Weighted amino acid frequencies were used to decrease the impact of large groups of similar sequences.[10] 111 class A human GPCR sequences were included in the analysis.

GPCR Receptor Binding Assays. With compounds **1** and **2**, affinities were taken from Willoughby et al.[11] Ligand affinities toward the GHS receptor were determined as described in Hansen et al.[15] Affinities toward to the AG2 receptor were determined by Cerep (Celle l'Evescault, France) as described in Bergsma et al.[16] The MC4 receptor binding assay is described below.

MC4 Receptor Binding Assay. BKH cells expressing MC4 receptors were made by transfecting MC4 receptor cDNA into BHK570 cells, and stable clones were selected and grown in DMEM media containing 1 mg/mL G418, 10% FCS, and 1% MTX. Membranes were prepared by homogenization (3×) in 20 mM Hepes pH 7.1, 5 mM MgCl$_2$, bacitracin 1 mg/mL, and centrifuged at 15 000 rpm at 4 °C, 10 min in a Sorvall RC 5B plus, SS-34 rotor. After the third centrifugation the pellet was dissolved in binding buffer; 25 mM HEPES pH 7.0, 1 mM CaCl$_2$, 1 mM MgSO$_4$, 1 mM EGTA, 0.02% Bacitracin 0.005% tween 20, and 0.1% HSA. The binding assay was performed in 96-well microtiter plates, (Millipore MADVN 6550) in a total

volume of 200 μL by mixing the membranes (10 μg) with 80 pM ^{125}NDP-α-MSH and test compound in varying concentrations, all dissolved in binding buffer. After addition of all assay components, the samples are incubated for 2 h at 25 °C. The incubation is terminated by placing the 96-well filter plates on a vacuum-manifold, and sucking the supernatant through the filters. The wells are washed twice with 100 μL ice-cold 0.9% NaCl. The plates are then air-dried and the filters are punched into counting-vials. The radioactivity retained on the filters was counted using a Cobra II auto gamma counter. IC$_{50}$ values are calculated by nonlinear regression analysis of binding curves using the windows program GraphPad Prism, GraphPad software, USA.

Receptor Models and Docking. Models of the 7TM domain for the human 5HT6, MC4, AG2, and GHS receptors were constructed using Modeler.[49] Sequence alignment of the four sequences together with bovine rhodopsin is shown in Figure 3. The bovine rhodopsin crystal structure[3] was used as template structure. Side chain conformations were fixed as in the crystal structure for residues satisfying the following two conditions: (a) identical in bovine rhodopsin and in the sequence used for the model and (b) conserved/identical in more than 80% of the class A receptor sequences including bovine rhodopsin. These constraints were included based on the assumption that such highly conserved residues most likely also have a conserved structural or functional role that demands a specific conformation, the one observed in the bovine rhodopsin structure. Loop regions as well as the C- and N-terminus were not included into the models as these regions are highly variable and for practical purposes impossible to model accurately.

Docking was done by hand followed by minimization with MMFF[50] in Sybyl6.8 (Tripos, St. Louis). The amino group present in all ligands except Losartan was protonated and assigned a formal charge of +1. Likewise, the tetrazole group of L-692,429 and Losartan were deprotonated and assigned a formal charge of -1. Side chain orientations were adjusted in a few cases to avoid steric clashes between receptor and ligand atoms. When available mutagenesis data and structure-activity relationships were used to guide the dockings. The RMSD between heavy atoms in the privileged structure part of the ligand-receptor complexes was computed for each pair of complexes by superimposing the complexes using protein backbone atoms.

References

(1) Evans, B. E.; Rittle, K. E.; Bock, M. G.; Dipardo, R. M.; Freidinger, R. M.; Whitter, W. L.; Lundell, G. F.; Veber, D. F.; Anderson, P. S.; Chang, R. S. L.; Lotti, V. J.; Cerino, D. J.; Chen, T. B.; Kling, P. J.; Kunkel, K. A.; Springer, J. P.; Hirshfield, J. Methods for Drug Discovery − Development of Potent, Selective, Orally Effective Cholecystokinin Antagonists. *J. Med. Chem.* **1988**, *31*, 2235−2246.

(2) Patchett, A. A.; Nargund, R. P.; Tata, J. R.; Chen, M. H.; Barakat, K. J.; Johnston, D. B. R.; Cheng, K.; Chan, W. W. S.; Butler, B.; Hickey, G.; Jacks, T.; Schleim, K.; Pong, S. S.; Chaung, L. Y. P.; Chen, H. Y.; Frazier, E.; Leung, K. H.; Chiu, S. H. L.; Smith, R. G. Design and Biological-Activities of L-163,-191 (Mk-0677) − A Potent, Orally-Active Growth-Hormone Secretagogue. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 7001−7005.

(3) 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.

(4) Labrou, N. E.; Bhogal, N.; Hurrell, C. R.; Findlay, J. B. C. Interaction of Met(297) in the Seventh Transmembrane Segment of the Tachykinin NK2 Receptor With Neurokinin A. *J. Biol. Chem.* **2001**, *276*, 37944−37949.

(5) Renzetti, A. R.; Catalioto, R. M.; Criscuoli, M.; Cucchi, P.; Ferrer, C.; Giolitti, A.; Guelfi, M.; Rotondaro, L.; Warner, F. J.; Maggi, C. A. Relevance of Aromatic Residues in Transmembrane Segments V to VII for Binding of Peptide and Nonpeptide Antagonists to the Human Tachykinin NK2 Receptor. *J. Pharmacol. Exp. Ther.* **1999**, *290*, 487−495.

(6) Spalding, T. A.; Burstein, E. S.; Henderson, S. C.; Ducote, K. R.; Brann, M. R. Identification of a Ligand-Dependent Switch Within a Muscarinic Receptor. *J. Biol. Chem.* **1998**, *273*, 21563−21568.

(7) Fong, T. M.; Huang, R. R. C.; Strader, C. D. Localization of Agonist and Antagonist Binding Domains of the Human Neurokinin-1 Receptor. *J. Biol. Chem.* **1992**, *267*, 25664−25667.

(8) Strader, C. D.; Candelore, M. R.; Hill, W. S.; Sigal, I. S.; Dixon, R. A. F. Identification of 2 Serine Residues Involved in Agonist Activation of the Beta-Adrenergic-Receptor. *J. Biol. Chem.* **1989**, *264*, 13572−13578.

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

(10) Pei, J. M.; Grishin, N. V. AL2CO: Calculation of Positional Conservation in a Protein Sequence Alignment. *Bioinformatics* **2001**, *17*, 700−712.

(11) Willoughby, C. A.; Hutchins, S. M.; Rosauer, K. G.; Dhar, M. J.; Chapman, K. T.; Chicchi, G. G.; Sadowski, S.; Weinberg, D. H.; Patel, S.; Malkowitz, L.; Di Salvo, J.; Pacholok, S. G.; Cheng, K. Combinatorial Synthesis of 3-(Amidoalkyl) and 3-(Aminoalkyl)-2- Arylindole Derivatives: Discovery of Potent Ligands for a Variety of G-Protein Coupled Receptors. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 93−96.

(12) Nargund, R.; Ye, Z.; Palucki, B. L., Bakshi, R. K., Patchett, A. A., Van der Ploeg, L. H. T. 1999, Patent No. WO 9964002.

(13) Chiu, A. T.; McCall, D. E.; Price, W. A.; Wong, P. C.; Carini, D. J.; Duncia, J. V.; Wexler, R. R.; Yoo, S. E.; Johnson, A. L.; Timmermans, P. B. M. W. Nonpeptide Angiotensin-Ii Receptor Antagonists 0.7. Cellular and Biochemical Pharmacology of Dup-753, An Orally Active Antihypertensive Agent. *J. Pharmacol. Exp. Ther.* **1990**, *252*, 711−718.

(14) Schoen, W. R.; Pisano, J. M.; Prendergast, T. K.; Wyvratt, M. J.; Fisher, M. H.; Cheng, K.; Chan, W. W. S.; Butler, B.; Smith, R. G.; Ball, R. G. A Novel 3-Substituted Benzazepinone Growth-Hormone Secretagogue (L-692,429). *J. Med. Chem.* **1994**, *37*, 897−906.

(15) Hansen, B. S.; Raun, K.; Nielsen, K. K.; Johansen, P. B.; Mansen, T. K.; Peschke, B.; Lau, J.; Andersen, P. H.; Ankersen, M. Pharmacological Characterisation of a New Oral GH Secretagogue, NN703. *Eur. J. Endocrinol.* **1999**, *141*, 180−189.

(16) Bergsma, D. J.; Ellis, C.; Kumar, C.; Nuthulaganti, P.; Kersten, H.; Elshourbagy, N.; Griffin, E.; Stadel, J. M.; Aiyar, N. Cloning and Characterization of A Human Angiotensin-Ii Type-1 Receptor. *Biochem. Biophys. Res. Commun.* **1992**, *183*, 989−995.

(17) Dean, D. C.; Nargund, R. P.; Pong, S. S.; Chaung, L. Y. P.; Griffin, P.; Melillo, D. G.; Ellsworth, R. L.; VanDerPloeg, L. H. T.; Patchett, A. A.; Smith, R. G. Development of a High Specific Activity Sulfur-35-Labeled Sulfonamide Radioligand That Allowed the Identification of a New Growth Hormone Secretagogue Receptor. *J. Med. Chem.* **1996**, *39*, 1767−1770.

(18) Devita, R. J.; Bochis, R.; Frontier, A. J.; Kotliar, A.; Fisher, M. H.; Schoen, W. R.; Wyvratt, M. J.; Cheng, K.; Chan, W. W. S.; Butler, B.; Jacks, T. M.; Hickey, G. J.; Schleim, K. D.; Leung, K.; Chen, Z. S.; Chiu, S. H. L.; Feeney, W. P.; Cunningham, P. K.; Smith, R. G. A Potent, Orally Bioavailable Benzazepinone Growth Hormone Secretagogue. *J. Med. Chem.* **1998**, *41*, 1716−1728.

(19) Each residue is numbered according to the Ballesteros-Weinstein two digit representation, whereby the first digit is the transmembrane helix number (I to VII) under consideration followed by the sequential order of the residue position in that helix relative to the most conserved residue in the helix which is given number 50. With residues in one particular receptor the sequence number is given followed by the Ballesteros-Weinstein numbering in parentheses.

(20) Yang, L. H.; Morriello, G.; Prendergast, K.; Cheng, K.; Jacks, T.; Chan, W. W. S.; Schleim, K. D.; Smith, R. G.; Patchett, A. A. Potent 3-Spiropiperidine Growth Hormone Secretagogues. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 107−112.

(21) Tata, J. R.; Lu, Z. J.; Jacks, T. M.; Schleim, K. D.; Cheng, K.; Wei, L. T.; Chan, W. W. S.; Butler, B.; Tsou, N.; Leung, K.; Chiu, S. H. L.; Hickey, G.; Smith, R. G.; Patchett, A. A. The Design and Synthesis of Orally Active Short Duration Spiroindane Growth Hormone Secretagogues. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2319−2324.

(22) Tata, J. R.; Nargund, R. P.; Murphy, M. M.; Johnston, D. B. R.; Patchett, A. A.; Cheng, K.; Wei, L.; Chan, W. W. S.; Butler, B.; Jacks, T. M.; Hickey, G.; Smith, R. The Synthesis and Activity of Spiroindane Growth Hormone Secretagogues. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 663−668.

(23) Chen, M. H.; Pollard, P. P.; Patchett, A. A. Synthesis and Biological Activities of Spiroheterocyclic Growth Hormone Secretagogues. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1261−1266.

(24) Chen, M. H.; Steiner, M. G.; Patchett, A. A.; Cheng, K.; Wei, L. T.; Chan, W. W. S.; Butler, B.; Jacks, T. M.; Smith, R. G. Analogues of the Orally Active Growth Hormone Secretagogue L- 162,752. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2163−2168.

(25) Feighner, S. D.; Howard, A. D.; Prendergast, K.; Palyha, O. C.; Hreniuk, D. L.; Nargund, R.; Underwood, D.; Tata, J. R.; Dean, D. C.; Tan, C. P.; McKee, K. K.; Woods, J. W.; Patchett, A. A.;

Smith, R. G.; Van der Ploeg, L. H. T. Structural Requirements for the Activation of the Human Growth Hormone Secretagogue Receptor by Peptide and Nonpeptide Secretagogues. *Mol. Endocrinol.* **1998**, *12*, 137−145.

(26) Yanagisawa, H.; Amemiya, Y.; Kanazaki, T.; Shimoji, Y.; Fujimoto, K.; Kitahara, Y.; Sada, T.; Mizuno, M.; Ikeda, M.; Miyamoto, S.; Furukawa, Y.; Koike, H. Nonpeptide Angiotensin II Receptor Antagonists: Synthesis, Biological Activities, and Structure − Activity Relationships of Imidazole-5-Carboxylic Acids Bearing Alkyl, Alkenyl, and Hydroxyalkyl Substituents at the 4-Position and Their Related Compounds. *J. Med. Chem.* **1996**, *39*, 323−338.

(27) Dascal, D.; Nirula, V.; Lawus, K.; Yoo, S. E.; Walsh, T. F.; Sandberg, K. Shared Determinants of Receptor Binding for Subtype Selective, and Dual Endothelin-Angiotensin Antagonists on the AT(1) Angiotensin II Receptor. *FEBS Lett.* **1998**, *423*, 15−18.

(28) Ji, H.; Leung, M.; Zhang, Y.; Catt, K. J.; Sandberg, K. Differential Structural Requirements for Specific Binding of Nonpeptide and Peptide Antagonists to the At(1) Angiotensin Receptor − Identification of Amino-Acid-Residues That Determine Binding of the Antihypertensive Drug Losartan. *J. Biol. Chem.* **1994**, *269*, 16533−16536.

(29) Noda, K.; Saad, Y.; Kinoshita, A.; Boyle, T. P.; Graham, R. M.; Husain, A.; Karnik, S. S. Tetrazole and Carboxylate Groups of Angiotensin Receptor Antagonists Bind to the Same Subsite by Different Mechanisms. *J. Biol. Chem.* **1995**, *270*, 2284−2289.

(30) Hunyady, L.; Ji, H.; Jagadeesh, G.; Zhang, M.; Gaborik, Z.; Mihalik, B.; Catt, K. J. Dependence of at(1) Angiotensin Receptor Function on Adjacent Asparagine Residues in the 7th Transmembrane Helix. *Mol. Pharm.* **1998**, *54*, 427−434.

(31) Schambye, H. T.; Hjorth, S. A.; Bergsma, D. J.; Sathe, G.; Schwartz, T. W. Differentiation Between Binding-Sites for Angiotensin-Ii and Nonpeptide Antagonists on the Angiotensin-Ii Type-1 Receptors. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7046−7050.

(32) Devita, R. J.; Schoen, W. R.; Fisher, M. H.; Frontier, A. J.; Pisano, J. M.; Wyvratt, M. J.; Cheng, K.; Chan, W. W. S.; Butler, B. S.; Hickey, G. J.; Jacks, T. M.; Smith, R. G. Benzolactam Growth-Hormone Secretagogues − Carboxamides As Replacements for the 2′-Tetrazole Moiety of L-692,429. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2249−2254.

(33) Ankersen, M.; Peschke, B.; Hansen, B. S.; Hansen, T. K. Investigation of Bioisosters of the Growth Hormone Secretagogue L-692,429. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1293−1298.

(34) Schoen, W. R.; Ok, D.; Devita, R. J.; Pisano, J. M.; Hodges, P.; Cheng, K.; Chan, W. W. S.; Butler, B. S.; Smith, R. G.; Wyvratt, M. J.; Fisher, M. H. Structure−Activity-Relationships in the Amino-Acid Side-Chain of L-692,429. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1117−1122.

(35) Mason, J. S.; Morize, I.; Menard, P. R.; Cheney, D. L.; Hulme, C.; Labaudiniere, R. F. New 4-Point Pharmacophore Method for Molecular Similarity and Diversity Applications: Overview of the Method and Applications, Including a Novel Approach to the Design of Combinatorial Libraries Containing Privileged Substructures. *J. Med. Chem.* **1999**, *42*, 3251−3264.

(36) Gether, U. Uncovering Molecular Mechanisms Involved in Activation of G Protein-Coupled Receptors. *Endocr. Rev.* **2000**, *21*, 90−113.

(37) 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-Struct. Funct. Genet.* **2003**, *50*, 5−25.

(38) Gouldson, P.; Kidley, N. J.; Bywater, R. P.; Psaroudakis, G.; Brooks, H. D.; Diaz, C.; Shire, D.; Reynolds, C. A. Towards the Active Conformations of Rhopdopsin and the Beta2-Adrenergic Receptor. *Proteins-Struct. Funct. Genet.* **2004**, in press.

(39) Donnelly, D.; Overington, J. P.; Ruffle, S. V.; Nugent, J. H. A.; Blundell, T. L. Modeling Alpha-Helical Transmembrane Domains − the Calculation and Use of Substitution Tables for Lipid-Facing Residues. *Protein Sci.* **1993**, *2*, 55−70.

(40) Sheikh, S. P.; Vilardarga, J. P.; Baranski, T. J.; Lichtarge, O.; Iiri, T.; Meng, E. C.; Nissenson, R. A.; Bourne, H. R. Similar Structures and Shared Switch Mechanisms of the Beta(2)-Adrenoceptor and the Parathyroid Hormone Receptor − Zn(II) Bridges Between Helices III and VI Block Activation. *J. Biol. Chem.* **1999**, *274*, 17033−17041.

(41) Shi, L.; Javitch, J. A. The Binding Site of Aminergic G Protein-Coupled Receptors: The Transmembrane Segments and Second Extracellular Loop. *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 437−467.

(42) Wong, P. C.; Price, W. A., Jr; Chiu, A. T.; Duncia, J. V.; Carini, D. J.; Wexler, R. R.; Johnson, A. L.; Timmermans, P. B. M. W. Nonpeptide Angiotensin II Receptor Antagonists. XI. Pharmacology of EXP3174: An Active Metabolite of DuP 753, an Orally Active Antihypertensive Agent. *J. Pharmacol. Exp. Ther.* **1990**, *255*, 211−217.

(43) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. Estimates of the Ab Initio Limit for Pi-Pi Interactions: The Benzene Dimer. *J. Am. Chem. Soc.* **2002**, *124*, 10887−10893.

(44) Perlman, S.; CostaNeto, C. M.; Miyakawa, A. A.; Schambye, H. T.; Hjorth, S. A.; Paiva, A. C. M.; Rivero, R. A.; Greenlee, W. J.; Schwartz, T. W. Dual Agonistic and Antagonistic Property of Nonpeptide Angiotensin AT(1) Ligands: Susceptibility to Receptor Mutations. *Mol. Pharm.* **1997**, *51*, 301−311.

(45) Chakravarty, P. K.; Naylor, E. M.; Ransom, R. W. 1993, Patent No. US5204354.

(46) Horn, F.; Vriend, G.; Cohen, F. E. Collecting and Harvesting Biological Data: the GPCRDB and NucleaRDB Information Systems. *Nucleic Acids Res.* **2001**, *29*, 346−349.

(47) Rost, B.; Casadio, R.; Fariselli, P.; Sander, C. Transmembrane Helices Predicted at 95-Percent Accuracy. *Protein Sci.* **1995**, *4*, 521−533.

(48) Vriend, G. What If − A Molecular Modeling and Drug Design Program. *J. Mol. Graphics* **1990**, *8*, 52.

(49) Sali, A.; Blundell, T. L. Comparative Protein Modeling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.

(50) Halgren, T. A. Maximally Diagonal Force-Constants in Dependent Angle-Bending Coordinates 0.2. Implications for the Design of Empirical Force- Fields. *J. Am. Chem. Soc.* **1990**, *112*, 4710−4723.