

Fighting Obesity with a Sugar-Based Library: Discovery of Novel MCH-1R Antagonists by a New Computational–VAST Approach for Exploration of GPCR Binding Sites

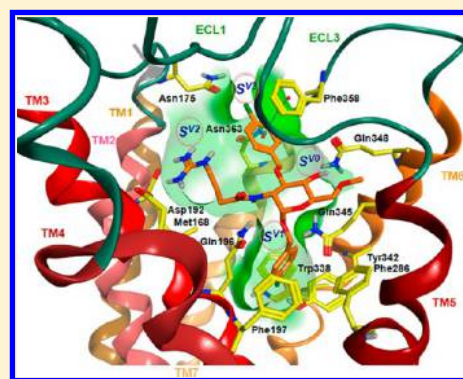
Alexander Heifetz,^{*,†} Oliver Barker,[†] Geraldine Verquin,[‡] Norbert Wimmer,[‡] Wim Meutermans,[‡] Sandeep Pal,[†] Richard J. Law,[†] and Mark Whittaker[†]

[†]Evotec (U.K.), Ltd., 114 Innovation Drive, Milton Park, Abingdon, Oxfordshire OX14 4SD, United Kingdom

[‡]Alchemia, Ltd., Eight Mile Plains, Queensland 4113, Australia

S Supporting Information

ABSTRACT: Obesity is an increasingly common disease. While antagonism of the melanin-concentrating hormone-1 receptor (MCH-1R) has been widely reported as a promising therapeutic avenue for obesity treatment, no MCH-1R antagonists have reached the market. Discovery and optimization of new chemical matter targeting MCH-1R is hindered by reduced HTS success rates and a lack of structural information about the MCH-1R binding site. X-ray crystallography and NMR, the major experimental sources of structural information, are very slow processes for membrane proteins and are not currently feasible for every GPCR or GPCR–ligand complex. This situation significantly limits the ability of these methods to impact the drug discovery process for GPCR targets in “real-time”, and hence, there is an urgent need for other practical and cost-efficient alternatives. We present here a conceptually pioneering approach that integrates GPCR modeling with design, synthesis, and screening of a diverse library of sugar-based compounds from the VAST technology (versatile assembly on stable templates) to provide structural insights on the MCH-1R binding site. This approach creates a cost-efficient new avenue for structure-based drug discovery (SBDD) against GPCR targets. In our work, a primary VAST hit was used to construct a high-quality MCH-1R model. Following model validation, a structure-based virtual screen yielded a 14% hit rate and 10 novel chemotypes of potent MCH-1R antagonists, including EOAI3367472 (IC_{50} = 131 nM) and EOAI3367474 (IC_{50} = 213 nM).



1. INTRODUCTION

Obesity is an increasingly problematic medical condition, particularly in the developed world, associated with an increased risk of developing type-2 diabetes, cardiovascular disease, hypertension, stroke, and cancer.¹ Obesity is characterized by a chronic imbalance between energy expenditure and energy intake.² While many of the mechanisms underlying obesity are far from being fully understood, it has become clear that the condition is influenced by six neuropeptides^{3,4} that are ligands of G-protein-coupled receptors (GPCRs), i.e., the 19-amino-acid polypeptide melanin-concentrating hormone (MCH) of MCH-1 (MCH-1R) and -2 (MCH-2R) receptors,⁵ the ghrelin hormone that binds to the growth hormone secretagogue receptor (GHSR1a), orexin-A and -B peptides of orexin-1 and -2 receptors, respectively, and neuropeptides B/W that partner with the neuropeptide B/W receptors. It was discovered that small organic molecules antagonize obesity-related GPCRs and as such have the potential to be used as anti-obesity drugs.²

Wide experimental evidence has identified MCH receptors as promising therapeutic targets in the treatment of obesity.^{6,7} The MCH-1R and MCH-2R receptors belong to the class A family of GPCRs, which are integral membrane proteins. MCH-1R is

widely distributed in the central nervous system and is particularly concentrated in regions that are involved in rewarding behavior, feeding behavior, and metabolic regulation.² The distribution of the MCH-2R receptor is less clear.⁸ Behavior studies have demonstrated a physiological role of the MCH system in regulating food intake.⁴ Acute central administration of MCH or MCH-1R synthetic agonists leads to a rapid and significant increase in food intake in rodents.⁹ This increase in food consumption is correlated with the functional activity of the agonist for the receptor, suggesting that the MCH-1R directly mediates the appetite stimulant effects of MCH.² The activation of MCH-1R by MCH or by synthetic agonists could be a key factor in the development of obesity, increasing energy intake while also reducing energy expenditure.² Conversely, MCH-1R antagonists could reduce MCH functional activity and as such slow down the development of obesity.²

The first nonpeptide MCH-1R selective antagonist reported was T-226296,¹⁰ which effectively blocked oral food intake in rats. A second nonpeptide antagonist, SNAP7941,² provided

Received: February 5, 2013

Published: April 16, 2013

the first evidence that chronic oral administration of an antagonist can cause sustained reduction in body weight. A number of MCH-1R antagonists including GW-803430¹¹ have been reported to reduce the amount of food intake in rats and cause weight loss in obese mouse models. The chronic administration of other MCH-1R antagonists in rodents led to significant reduction in food intake, body weight and fat gain during the whole treatment period.^{12,13} These promising experimental results maintain the high interest in further development of MCH-1R antagonists as potential drugs for obesity.⁷ More than 2700 MCH-1R antagonists have been discovered, but none of them have been successful in clinical trials so far.⁷

A lack of direct structural information on the binding site of MCH-1R has limited the discovery and optimization of MCH-1R antagonists to a trial and error approach.^{7,14} This observation has resulted in a significant reduction in the reporting of discovery and optimization of new chemotypes of MCH-1R antagonists in recent years.⁶ However, some success has been reported for structure-based virtual screening (VS) campaigns employing theoretical 3D models of MCH-1R, indicating that structural information can benefit the drug discovery process for MCH-1R antagonists.^{14,15}

The observation that MCH-1R binds ligands with a variety of chemotypes and sizes hints at the complex topology and large size of its binding site.¹⁴ X-ray crystallography and NMR, the major experimental sources of structural information, are very slow processes for membrane proteins and are not currently feasible for every GPCR or GPCR–ligand complex.¹⁶ This situation significantly limits the ability of these methods to impact the drug discovery process for GPCR targets in “real-time”, and hence, there is an urgent need for other practical and cost-efficient alternatives. Prediction of GPCR structures in the absence of direct experimentally derived structural information on target receptor is possible but speculative.

Discovery of small drug-like GPCR antagonists is also especially challenging for peptidergic GPCRs like MCH-1R because it depends on correctly transposing known binding elements of the parent peptide onto small drug-like molecules while maintaining bioactivity.¹⁷ The binding of peptides to GPCRs is defined by two key factors: the combination of unique binding elements (motifs) to generate key interactions with specific residues of the receptor and the relative orientation of these binding elements in 3D space (bioactive conformation). The peptide bioactive conformations and motifs can be mimicked by appending substituents onto a rigid organic scaffold in a similar relative orientation as in the original peptide. A significant amount of work has been reported on the applicability of using a monosaccharide scaffold in peptidomimetic design approaches.^{17–25} However, the design of bioactive peptidomimetics, particularly for GPCRs, remains a challenge.

We present here a novel integrated experimental–computational approach that enabled molecular-level exploration of the MCH-1R structure, including prediction of its antagonist binding site, its topology, and the key residues involved in antagonist binding followed by structure-based discovery of 10 new chemotypes of potent MCH-1R antagonists. This practical approach, integrating the VAST technology with molecular modeling, opens an additional avenue for structural exploration of GPCRs and for structure-based drug discovery against GPCR targets. Our method is able to satisfy the immediate need of the drug-discovery process in the structural information

on the target GPCR, particularly when X-ray crystal data is not available. The general workflow for the discovery of novel MCH-1R antagonists from initial screening of a sugar-based (VAST) library and subsequent development of an optimized MCH-1R model is shown in Figure 1a.

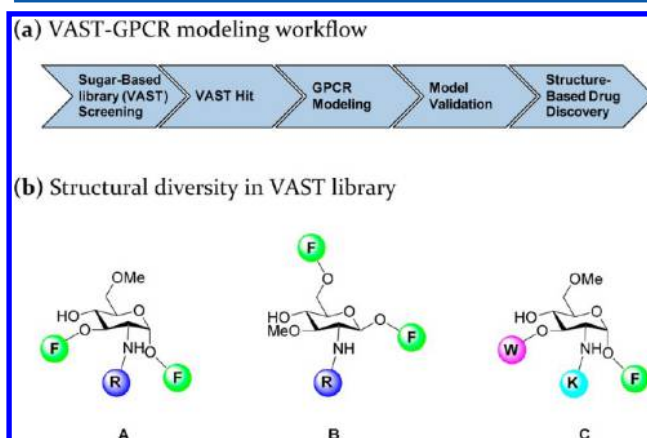


Figure 1. (a) General strategy followed in this work, proceeding from the initial VAST library screening to the discovery of novel chemically unrelated antagonists through the development of an optimized MCH-1R model. (b) Examples of structural diversity in the VAST library scanning the general aromatic–aromatic-positive charge motif. Substituents on the various positions of the sugar scaffold are represented by a letter corresponding to the amino-acid side-chain mimicked (F = phenylalanine, W = tryptophan, R = arginine, K = lysine). Compounds of type A and B share the same motif (FRF) but differ in terms of chemoform (relative orientation of each substituent defined by the sugar scaffold type and the positioning of each substituent on the scaffold). Compounds of type A and C share the same chemoform but differ in terms of motifs (FRF and FKW, respectively).

The VAST technology is ideally suited to the exploration of large and/or topologically complex GPCR binding sites.^{19,21} It is based on the use of a sugar (pyranose) scaffold on which it is possible to append various substituents in a systematic manner to obtain a structurally diverse library. Many GPCR peptides can be minimized to tripeptide molecules and still maintain biological activity.²⁶ We showed that the pyranose scaffold is ideally suited to mimic these tripeptides because of its unique versatility in terms of positioning three substituents in 3D space and the significant overlay of topographical space thus accessible with the corresponding tripeptide conformational space. An example of a diverse VAST library, described in detail in a previous paper, was designed to systematically scan the general aromatic–aromatic-positive charge motif.¹⁷ The compounds in this library each contained three substituents representing various tripeptide sequences of this general motif (more precisely FKW, FKF, WKW, FRW, FRF, or WRW with F = phenylalanine, W = tryptophan, K = lysine, R = arginine side-chain mimetic) in a particular relative orientation referred to as a chemoform, defined by the choice of sugar-based scaffold and substitution pattern on the selected scaffold (Figure 1b). This concept can be extended to a wide range of different motifs; the premise of our approach being that each GPCR will have a unique pattern of motif and chemoform preference. The 490 VAST compounds obtained from this library design were screened against MCH-1R, resulting in the discovery of a potent MCH-1R antagonist, ACL21823

(radioligand binding RLB MCH-1R IC_{50} = 306 nM, aequorin functional assay AEQ MCH-1R EC_{50} = 679 nM; Figure 2).

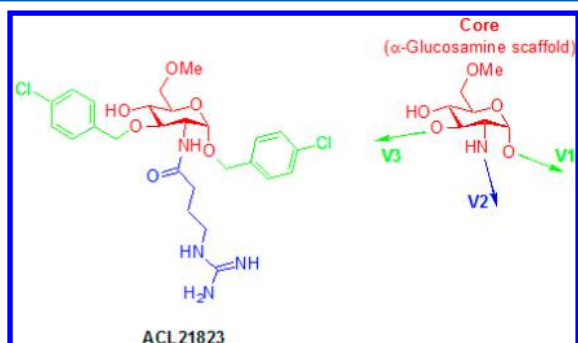


Figure 2. Structure of ACL21823 and schematic representation in terms of vectors of substituent positioning from the pyranose scaffold (vectors V1, V2, and V3 for the substituents at position 1, 2, and 3, respectively, on the sugar scaffold).

The discovery of ACL21823 was a key factor in the construction of a high quality MCH-1R model and the location of its antagonist binding site. The high quality of the MCH-1R model was demonstrated by a virtual enrichment experiment and the model-driven structure-based expansion of ACL21823, which allowed the generation of a list of potential key MCH-1R residues involved in antagonist binding. The utility of our method to drug discovery was demonstrated by a structure-based virtual screen, achieving a hit rate of 14% and yielding 10 new chemotypes of MCH-1R antagonists including EOAI3367472 (IC_{50} = 131 nM) and EOAI3367474 (IC_{50} = 213 nM).

2. RESULTS AND DISCUSSION

2.1. VAST Library Synthesis and Screening. We reported on the method of design, synthesis, and screening of a structurally diverse sugar-based VAST library of 490 compounds.¹⁷ The screening of this library against MCH-1R resulted in the discovery of the potent MCH-1R antagonist ACL21823 (Figure 2 and Table 1).

2.2. Modeling of MCH-1R Structure. The structure of human MCH-1R was modeled based on a 2.8 Å resolution crystal structure of the β 2-adrenergic receptor (β 2AR, PDB entry 2RH1). Multiple sequence alignment, shown in Figure S1 of the Supporting Information, indicated that β 2AR had a higher sequence similarity with MCH-1R compared with other publicly available experimental GPCR structures, with 69.5% similarity within the 7TMD region and four aligned prolines. The second extracellular loop was modeled using ECL2 from bovine rhodopsin (PDB entry 1F88), attached to the structure and optimized using OPLS-AA energy function. We selected rhodopsin as a template to model the ECL2 loop of MCH-1R due to two reasons. Firstly, the ECL2 loop of MCH-1R has 24.1% sequence similarity with rhodopsin and just 13.8% similarity with the β 2AR ECL2 loop. Secondly, the conformation of the β 2AR ECL2 loop is defined by two disulfide bonds formed between cysteine residues in the ECL2 loop and the 7TM domain. Both MCH-1R and rhodopsin only have one cysteine in the ECL2 loop.

2.3. MCH-1R Binding Site and Complex with ACL21823. With the MCH-1R model and ACL21823 in hand, we next turned to the prediction, refinement, and exploration of the MCH-1R binding site. We docked

Table 1. SAR Analogues of ACL21823 and Corresponding MCH-1R Activity (Radioligand Binding)

Compound	R ₁	R ₂	R ₃	MCH-1R IC_{50} (nM) ^a
8a (ACL21823)	Cl	H	Cl	306
8b	CF ₃	H	Cl	417
8c	F	H	Cl	832
8d	F	F	Cl	196
8e	Cl	H	CF ₃	143
8f	Cl	H	F	8724
8g	F	H	CF ₃	377
8h	F	F	F	3410
8i (ACL22482)	F	F	CF ₃	65
8j	Cl	H	H	N/D ^b

^aEach IC_{50} is the average result of a single experiment (two replicates for each data point). ^bN/D = IC_{50} not determined (no activity detected at 3 μ M).

ACL21823 into MCH-1R using a flexible docking procedure. Analysis of the 10 top-scoring docked poses of ACL21823 revealed that they were all very similar. Heavy atom RMSD between docked poses did not exceed 1.5 Å, indicating that a single conserved pose had been identified. The top scoring pose of ACL21823 was selected for further binding site refinement with “low-mode” molecular dynamics (LowModeMD) simulation.

As was previously published²⁷ the MD simulation step allows refinement of the GPCR models to a degree that is not possible with static homology modeling. The proposed docking pose of ACL21823 in MCH-1R and the location of the binding site are shown in Figure 3a. We observed that the MCH-1R antagonist binding site has a complicated surface topology that allows it to accommodate large molecules such as ACL21823. The predicted binding site of ACL21823 in MCH-1R is located between TM2, 3, 5, 6, and 7 (Figure 3b); it can be divided into four main regions (subpockets), S^{V1} , S^{V2} , S^{V3} , and S^{V0} (Figure 3c), which accommodate the corresponding vectors V1, V2, and V3 and the sugar scaffold (V0; core), shown in Figure 2, of ACL21823.

A deep, narrow subpocket S^{V1} is located between TM3, 5, and 6 and contains the aromatic residues Phe197^{3,37}, Phe286^{5,47}, Trp338^{6,48}, and Tyr342^{6,52}. Additional potentially important residues located in the “bottle neck” of S^{V1} subpocket are Gln196^{3,36} and Gln345^{6,55}. The upper subpocket S^{V3} is located between TMS, 6, and 7 and contains residues Phe358^{ECL3}, Asn175^{ECL1}, and Asn363^{7,36}. The third long subpocket S^{V2} is generated by TM2, 3, and 7 and contains residues Met168^{2,60} and Asp192^{3,32}. The fourth central subpocket S^{V0} is located between subpockets S^{V1} , S^{V2} and S^{V3} , is formed by TMS, 6, and 7, and contains the residue Gln348^{6,58}.

The predicted binding mode of ACL21823 is shown in Figure 3b. ACL21823 occupies all four subpockets, S^{V1} , S^{V2} , S^{V3} and S^{V0} , of the MCH-1R binding site. In subpocket S^{V2} , the guanidine moiety of ACL21823 can form a salt-bridge with

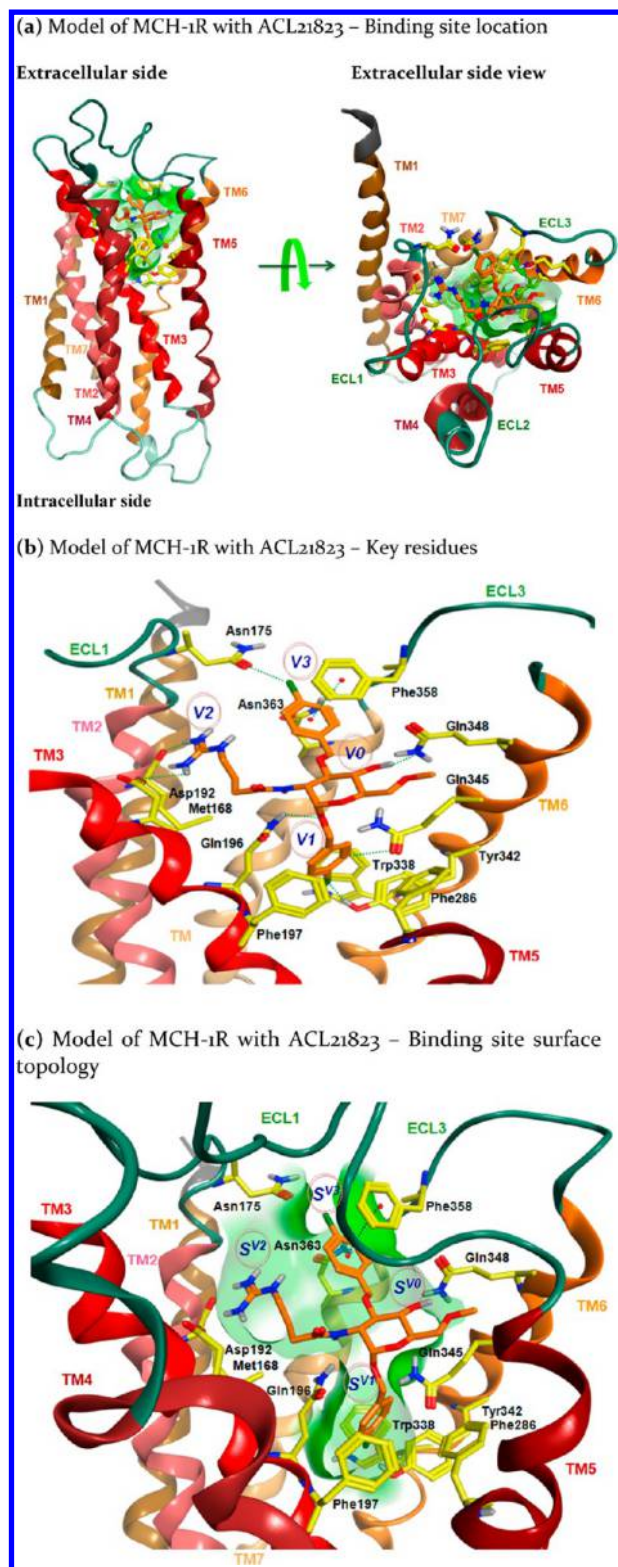


Figure 3. (a), (b), and (c) show the final model of MCH-1R with the predicted docking pose of ACL21823. The ribbon color coding of TMs 1 to 7 is dark orange, pink, red, purple, dark red, orange, and light yellow, respectively, with ICLs and ECLs in dark green. The carbon atoms of ACL21823 are shown in orange and those of MCH-1R in yellow. Nitrogen atoms are shown in blue, oxygen in red, sulfur in yellow, and chlorine in light green. (b) Interactions with key residues are indicated by dotted lines. (c) The surface of MCH-1R binding site is colored in green and clipped to allow a better view of the four subpockets.

Asp192^{3,32} and a hydrogen bond with the backbone carbonyl of Met168^{2,60}. In subpocket S^{V3} the aromatic moiety of ACL21823 can form a π - π interaction with Phe358^{ECL3} and a halogen bond between the chlorine substituent and the side-chain carbonyl of Asn175^{ECL1}. Experimental evidence that halogen bonds can play a key role in binding and recognition in complexes between proteins and ligands has been previously reported.²⁸ In the central subpocket S^{V0} , the sugar core can form a hydrogen bond with Gln348^{6,48}. In subpocket S^{V1} , we find an aromatic cluster including the residues Phe197^{3,37}, Phe286^{5,47}, Trp338^{6,48}, and Tyr342^{6,52}. The aromatic moiety of ACL21823 can form an edge-to-face π - π interaction with Phe197^{3,37}, while the chlorine substituent can form a halogen- π interaction with the Trp338^{6,48} ring system.²⁹ A nonclassical hydrogen bond can be formed between the aromatic hydrogen of ACL21823 and the carbonyl oxygen of Gln345^{6,55,50}. Finally, the side chain of Gln196^{3,36} can form a hydrogen bond with the oxygen linker of ACL21823.

2.4. Validation of MCH-1R Structure by Virtual "Enrichment" Experiment. To validate the model, a virtual "enrichment" experiment was carried out. We docked a diverse set of 126 known MCH-1R ligands, together within a library of 3921 similar decoy compounds. When docked compounds were ranked by AMBER interaction energy, an enrichment factor 5-fold better than random was observed at the point where 50% of the known compounds were identified (Figure 4). Furthermore, 90% of the known MCH-1R antagonists were

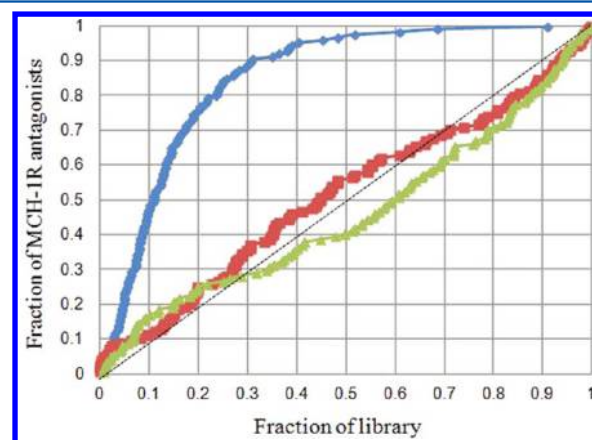
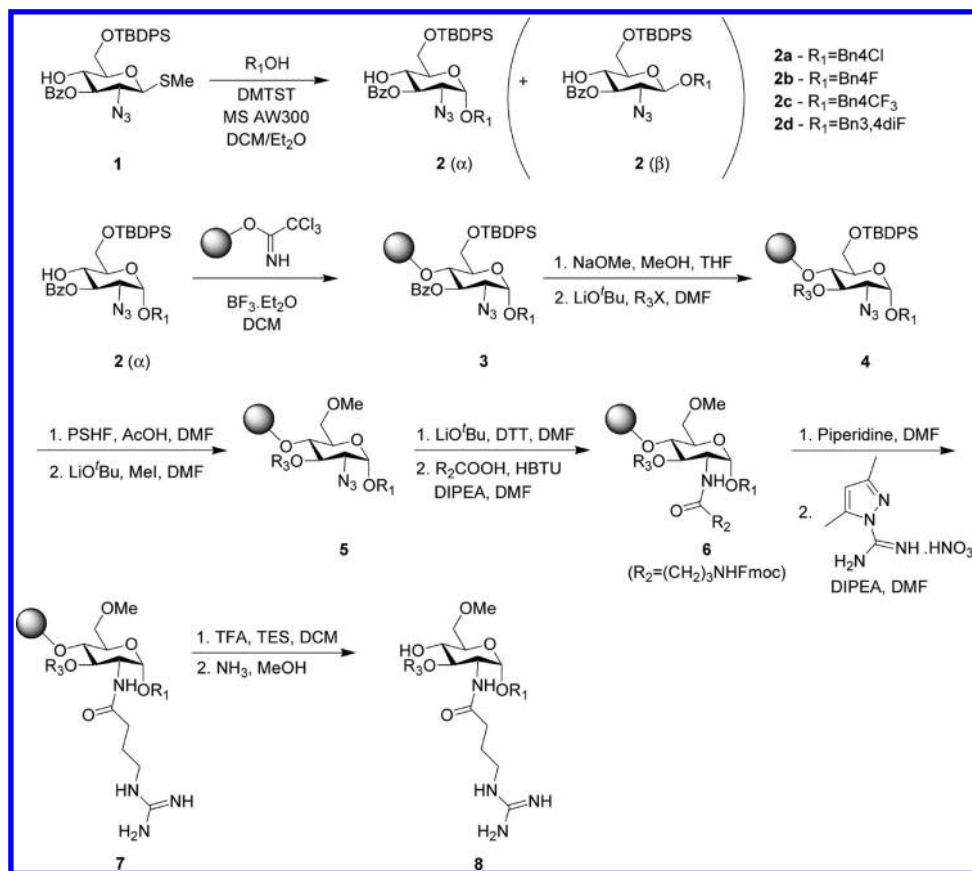


Figure 4. Enrichment plot for the in silico screening of 126 known MCH-1R ligands embedded in a library of 3921 similar compounds with >0.7 Tanimoto similarity after docking into MCH-1R. The fraction of library compounds is ranked along the x axis, and fraction of known MCH-1R antagonists is ranked along the y axis. The blue curve (best enrichment) resulted from ranking by the AMBER protein-ligand interaction energy, while the red and green curves resulted from ligand-based shape similarity (ROCS) searches that used the docked pose of ACL21823 (red) or ACL21823 minimized structure (green) as queries.

ranked among the top 30% of the library based on AMBER interaction energy. This "enrichment" test demonstrated that the MCH-1R model was of high quality, as proven by its ability to discriminate between known binders and very similar decoys. ROCS shape-based similarity searches, using both the docked pose of ACL21823 and its minimized structure as queries, failed to distinguish between MCH-1R binders and similar decoys (Figure 4).

Scheme 1. Synthesis of ACL21823 Analogues on Solid Phase



2.5. Validation of MCH-1R Structure with ACL21823 Analogues. Analogues of the primary hit ACL21823 were synthesized for SAR exploration and validation of MCH-1R. We focused on modifications of the substituents R_1 , R_2 (located in subpocket S^{V1}), and R_3 (located in subpocket S^{V3}), as shown in Table 1.

The compounds listed in Table 1 were synthesized as part of a SAR library of 80 compounds, focused on the structure of the initial hit with variation of the substituents on the various positions of the sugar scaffold. These ACL21823 analogues were prepared on solid phase (Scheme 1) using standard procedures for glycosylation, resin loading, library production, and cleavage off-resin described previously.^{17,31} On the basis of the docking pose of ACL21823 in MCH-1R, the R_1 , R_2 , and R_3 substituents were predicted to be particularly significant for the binding of ACL21823 to MCH-1R. They were predicted to affect binding either through direct interaction with key MCH-1R residues or via nonspecific contacts with the surface of the MCH-1R binding site or through electronic effects exerted on the aromatic rings in the ligand that modulate the strength of the π - π interactions formed by these groups. As the IC_{50} s were determined from a single experiment, and due to the inherent uncertainty of homology modeling, the model was only used to rationalize large shifts in affinity (>10 fold) compared to the reference compound ACL21823 (8a).

In compounds 8b and 8c, the single-point substitution of chlorine in the R_1 position with either CF_3 or fluorine, respectively, resulted in a similar potency range to the reference ACL21823 (8a). Further substitution in positions R_2 and R_3 , in compounds 8d and 8e did not cause a significant variation in potency.

In compound 8f, the single-point replacement of the R_3 chlorine with fluorine resulted in a 28-fold reduction in MCH-1R potency compared to ACL21823. The model suggested several possible explanations for this. Firstly, a potential halogen bond was observed formed between the R_3 chlorine and $\text{Asn175}^{\text{ECL1}}$. Fluorine has not been observed to form halogen bonds,²⁸ so the fluorine substitution in 8f would result in the loss of this interaction. Secondly, the electronic effects of the R_3 substituent may also affect the predicted π - π interaction with $\text{Phe358}^{\text{ECL3}}$. While fluorine is more electronegative than chlorine, suggesting that it should have a greater electron-withdrawing effect, resonance effects mean that this is not the case in practice, and indeed fluorine has been reported to have a weak electron-donating effect.³² Aromatic rings with electron-withdrawing groups favor face-to-face π - π interactions, while rings with electron-donating groups tend toward edge-to-face π - π configurations. Therefore, the R_3 fluorine substituent may also affect affinity by reducing the strength of the proposed π - π interaction with $\text{Phe358}^{\text{ECL3}}$.

Compound 8g is a hybrid of compounds 8c and 8e, with the chlorines in positions R_1 and R_3 replaced by fluorine and CF_3 , respectively. This compound retains similar potency range to the original ACL21823.

In compound 8h, as with compound 8f, a significant loss of binding is caused by the replacement of chlorine with fluorine in position R_3 . The R_1 and/or R_2 fluorine substituents appear to ameliorate this somewhat, as there is only a 11-fold loss of binding for compound 8h (compared to ACL21823), as opposed to the 28-fold loss observed for 8f.

Compound 8i (ACL22482) provided the highest potency measurement of all of the VAST compounds tested.

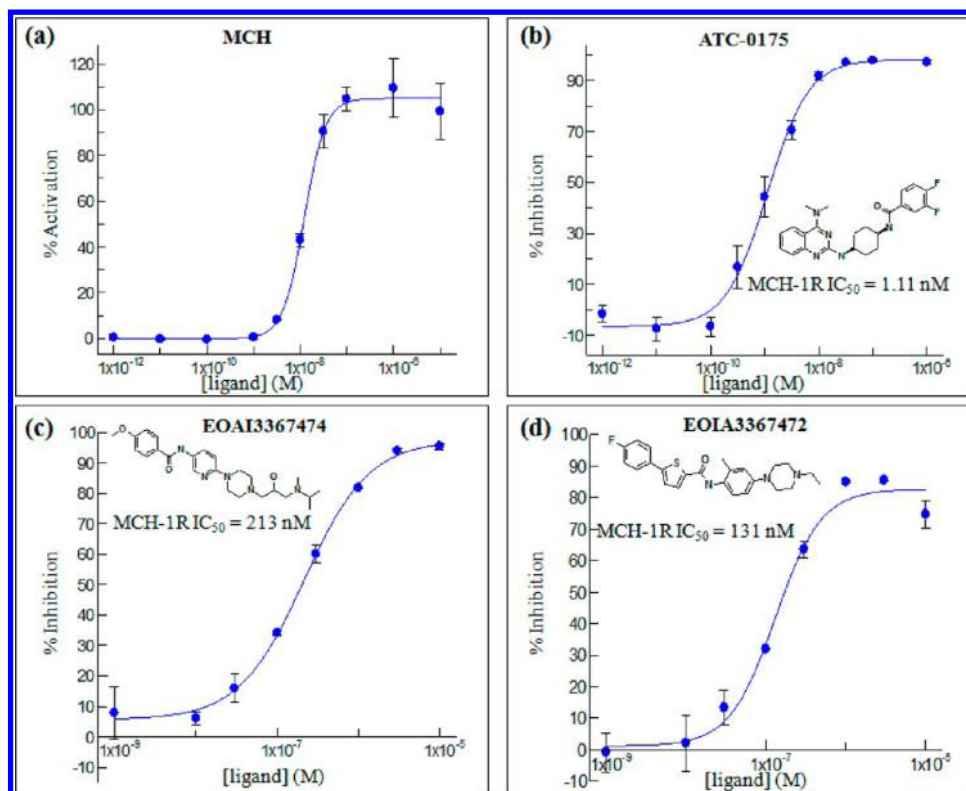


Figure 5. MCH-1R dose–response curves. (a) MCH and (b) ATC-0175 used as controls, (c) EOAI3367474 ($IC_{50} = 213$ nM), and (d) EOAI3367472 ($IC_{50} = 131$ nM). Points represent mean \pm SEM values from single experiments with two replicates.

Modification of the R_1 , R_2 , and R_3 substituents to F, F, and CF_3 , respectively, resulted in a measured IC_{50} of 65 nM.

In compound **8j**, the single-point replacement of the R_3 chlorine with hydrogen resulted in a substantial decrease in MCH-1R binding compared to ACL21823. As with compound **8f**, the model suggests two potential contributors to this. Firstly, a potential halogen bond with Asn175^{ECL1} is lost. Secondly, as discussed previously, aromatic rings with electron-withdrawing groups, such as chlorine in the original ACL21823 hit, favor face-to-face π – π interactions. The R_3 hydrogen substituent may affect binding affinity by reducing the strength of the π – π interaction with Phe358^{ECL3}. Tolerance of R_3 CF_3 substituents in compounds **8e**, **8g**, and **8i** would appear to support this latter hypothesis as CF_3 is, like Cl, an electron withdrawing group.

As predicted by the MCH-1R model, the nature of R_1 , R_2 , and in particular R_3 substituents appears to affect MCH-1R binding of the VAST hits. The independent and additive effects of some of these substituents suggests that they interact with MCH-1R via separate subpockets, supporting the predicted MCH-1R binding site topology and modeled binding mode in which the substituents occupy “geographically” distant subpockets (S^{V1} and S^{V3}).

This study also identified a list of residues predicted to be involved in ligand binding to MCH-1R: Asn175^{ECL1}, Asp192^{3.32}, Gln196^{3.36}, Phe197^{3.37}, Thr200^{3.40}, Phe358^{ECL3}, Phe286^{5.47}, Gln348^{6.58}, Trp338^{6.48}, Tyr342^{6.52}, Gln345^{6.55}, and Asn363^{7.36}. Site-directed mutagenesis studies targeting these residues (particularly Asn175^{ECL1} and Phe358^{ECL3}) will provide further experimental validation of the proposed binding mode, as will biological testing of additional VAST compounds designed to strengthen and weaken the predicted interactions.

2.6. Structure-Based Virtual Screening. A virtual screening (VS) library of 45,000 compounds with ≤ 0.7 Tanimoto similarity to the 126 known MCH-1R antagonists was extracted from a commercial vendor library of 16 M compounds. This was done to ensure that only MCH-1R binders with novel chemotypes would be identified. The same GOLD docking procedure and calculation of AMBER interaction energy as used in the “enrichment” validation test was applied to docking, scoring, and ranking of these 45,000 compounds. The top 70 compounds from the ranked list, which showed AMBER interaction energies in the range of the top 10% of ranked known MCH-1R antagonists, were selected for biochemical screening. These 70 compounds were purchased from vendors and experimentally tested in an aequorin-based inhibition assay for MCH-1R. This in vitro assay confirmed that 10 of the 70 compounds tested were hits with % inhibition $>80\%$ at 10 μ M, reflecting a 14% hit rate. Dose–response curves were obtained for these 10 hits. MCH (Figure 5a) and ATC-0175 (Figure 5b) were used as controls. The 10 VS hits showed low 2D Tanimoto similarity (<0.5) to the known MCH-1R antagonists and hence represent novel chemotypes of MCH-1R antagonists. EOAI3367472 ($IC_{50} = 131$ nM; Figure 5d) and EOAI3367474 ($IC_{50} = 213$ nM; Figure 5c) were the most potent VS hits identified.

2.7. Prediction of Key Structural Features Involved in Binding Small Molecule MCH-1R Antagonists. Next, we used the model to generate a prediction of the key structural features of MCH-1R involved in the binding of small organic antagonists. Information of this sort can be useful for guiding the discovery, hit to lead, and lead optimization of novel MCH-1R antagonists. In order to obtain this information, we generated a training set that consisted of docking poses of (1) VS hits EOAI3367472 and EOAI3367474, (2) a

representative set of 32 known MCH-1R antagonists (the most potent representative per cluster from 126 clusters of known MCH-1R antagonists with K_s or IC_{50} s < 160 nM (details are in Table S1, Supporting Information), and (3) a potent literature hit ATC-0175³³ (IC_{50} = 1.11 nM; Figure 5b). Ten docked poses were produced for each molecule and ranked by AMBER interaction energy. The top-ranked pose was retained for further study; other poses were also inspected visually and were found to be similar to the top-ranked pose or to have fewer interactions with the protein.

Docked poses for the VS hits EOAI3367472 and EOAI3367474 are shown in Figure 6a and b, respectively. Both compounds were predicted to have distinct binding modes that access three of the four subpockets occupied by the VAST hits. The docked pose of the potent hit ATC-0175 is also illustrated here for the purpose of comparison with the VS hits (Figure 6c).

EOAI3367472 is predicted to occupy subpockets S^{V0} , S^{V1} , and S^{V3} (Figure 6a); S^{V2} does not appear to be accessed by this molecule. In subpocket S^{V3} , the docked pose shows the aromatic moiety of EOAI3367472 forming a face-to-face π - π interaction with Phe358^{ECL3}. Hydrogen bonds are predicted to occur between the basic nitrogen of the piperazine moiety and Asn175^{ECL1} and also between Asn363^{7.36} and the nonbasic piperazine nitrogen. In the central subpocket S^{V0} , the carbonyl oxygen of the amide linker can accept a hydrogen bond from Gln348^{6.58}. In subpocket S^{V1} , the aromatic moiety of EOAI3367472 can form an edge-to-face π - π interaction with Trp338^{6.48} and Phe197^{3.37}, while the fluorine substituent potentially forms a weak hydrogen bond with the OH group of the Tyr342^{6.52} side chain. It is not clear from our model what types of interactions might exist between the thiophene moiety of EOAI3367472 and Gln196^{3.36} and Gln345^{6.55} residues.

EOAI3367474 is predicted to have a different binding mode to EOAI3367472 (Figure 6b). It does not appear to access S^{V1} , but seems to occupy subpockets S^{V0} and S^{V2} . In the case of S^{V2} , docking shows EOAI3367474 occupying an extension to the subpocket between TM2 and TM3 that is left unfilled by the VAST compounds. It also partially accesses S^{V3} . In subpocket S^{V0} , the aromatic moiety of EOAI3367474 can form an edge-to-face π - π interaction with the S^{V3} residue Phe358^{ECL3}. The carbonyl oxygen of the amide linker can accept a hydrogen bond from the S^{V3} residue Asn363^{7.36}, and the NH of the amide linker can form a hydrogen bond with Gln345^{6.55}. In subpocket S^{V2} , a hydrogen bond is potentially formed between the pyridine nitrogen of EOAI3367474 and Gln196^{3.36}. There is a potential for a salt-bridge between the basic piperazine nitrogen and Asp192^{3.32}; however, the compound contains multiple ionizable centers with similar pKas, and it is likely that multiple microspecies will be present. Not all of them will have this nitrogen protonated, reducing the significance of this interaction. A further hydrogen bond can be generated between the side chain oxygen of Tyr370^{7.43} and the OH group of EOAI3367474. Finally, the isopropyl portion of EOAI3367474 appears to pack against the side chain of Phe161^{2.53}, making either a generic hydrophobic contact or potentially a CH- π interaction.

The literature antagonist ATC-0175 adopts a similar docked pose to EOAI3367472, occupying subpockets S^{V0} , S^{V1} , and S^{V3} (Figure 6c). In the central subpocket S^{V0} , the cyclohexane scaffold of ATC-0175 is predicted to form several nonspecific hydrophobic contacts with MCH-1R, and a hydrogen bond can be formed between the NH linker and Gln348^{6.58}. In subpocket

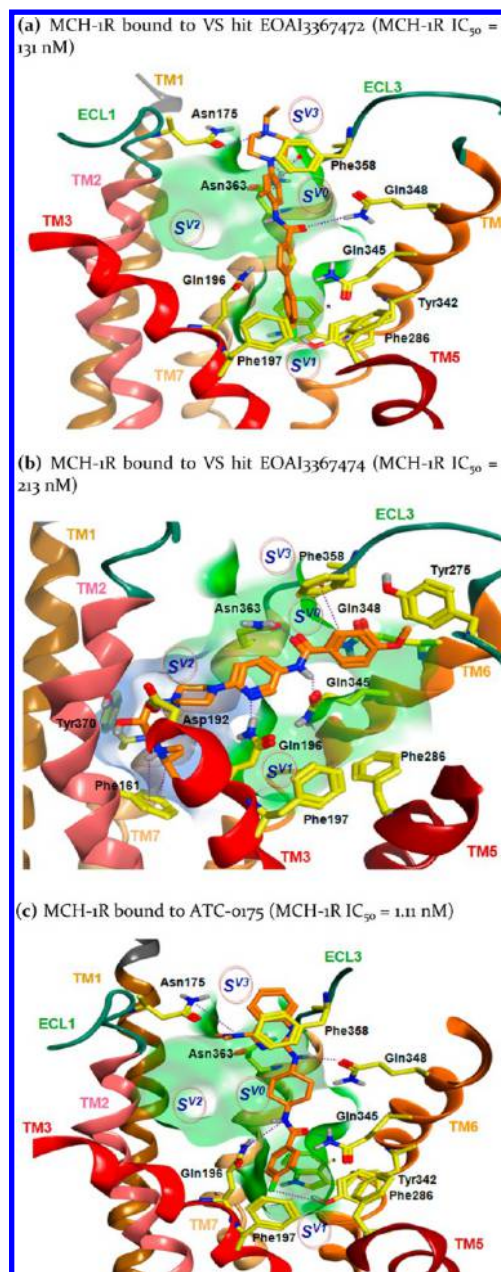


Figure 6. Docked poses of MCH-1R with EOAI3367472 (a), EOAI3367474 (b), and ATC-0175 (c). The ribbon color coding of TMs 1 to 7 is dark orange, pink, red, purple, dark red, orange, and light yellow respectively, with ICLs and ECLs in dark green. The ligands carbon atoms are shown in orange and those of MCH-1R in yellow. Nitrogen atoms are shown in blue, oxygen in red, sulfur in yellow, and fluorine in light green. Interactions with key residues are indicated by dotted lines, and the surface of MCH-1R binding site is colored in green and clipped to allow a better view of the four subpockets. In (b), the surface of the extra space in the S^{V2} pocket is colored in blue. Trp338 is marked as (*) in (a) and (c) to avoid obscuring the image.

S^{V1} , the aromatic moiety of ATC-0175 can form an edge-to-face π - π interaction with Trp338^{6.48}, while one fluorine substituent can form a hydrogen bond with Tyr342^{6.52}. Two hydrogen bonds can be formed by the amide linker of ATC-0175 with Gln196^{3.36} and Gln345^{6.55}. In subpocket S^{V3} , the quinazoline moiety of ATC-0175 can form a face-to-face π - π interaction with Phe358^{ECL3} and a hydrogen bond with Asn175^{ECL1}. While ATC-0175 and EOAI3367472 adopt similar docked poses, they

exhibit a 100-fold difference in potencies. The model suggests a possible rationale for this difference, as ATC-0175 is predicted to form five hydrogen bonds with MCH-1R, while EOAI3357472 would form only three.

While the VAST hits are predicted to occupy all four subpockets of the MCH-1R antagonist site, docked poses for ATC-0175, EOAI3367472, and EOAI3367474 only accessed three of them. In addition, EOAI3367474 occupied a different subset of the subpockets to ATC-0175 and EOAI336742.

A further 32 published and potent MCH-1R antagonists were docked into the model in order to investigate how frequently each subpocket was predicted to be involved in ligand binding. As illustrated in Figure 7, this analysis indicates

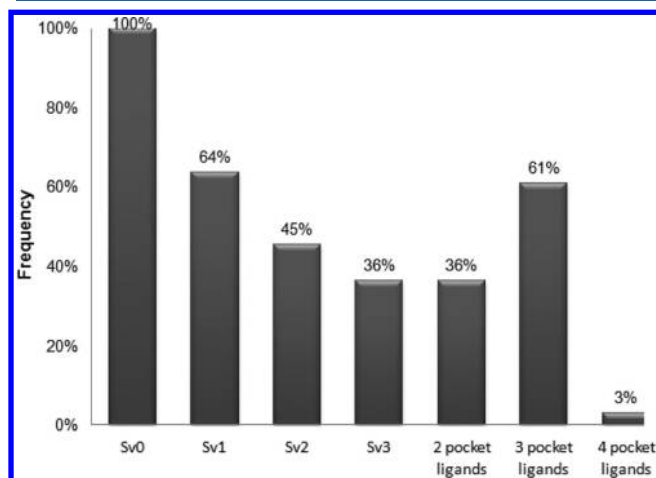


Figure 7. Plot based on analysis of docking poses for 33 highly potent MCH-1R antagonists to determine subpocket occupancy.

that 61% of the 33 antagonists (ATC-0175 + 32 other literature compounds) are predicted to occupy three of the four subpockets. In 36% of cases, the antagonists only occupied two subpockets. Potent small molecule antagonists that accessed all four subpockets were, according to the model, very rare, only accounting for a single example. The S^{V0} subpocket was occupied by the docked poses of all 33 antagonists. The central position of the S^{V0} subpocket makes it an important junction allowing access to neighboring subpockets. Because of the relatively small number of data points, and as the reported potencies for the literature compounds were taken from different assays, it was not possible to provide a clear prediction as to which of the S^{V1} , S^{V2} , or S^{V3} subpockets are most significant with regards to binding potency, but it does suggest that potent small molecule MCH-1R antagonists do not need to occupy all four subpockets occupied by the VAST compounds.

We also calculated the frequency of interaction of each of the binding site residues with docked small molecule ligands. This exercise was performed to identify a list of residues predicted to be involved in small ligand binding, which could be used to prioritize future site-directed mutagenesis studies and to suggest further development of EOAI3367472 and EOAI3367474. The same training set of 33 MCH-1R potent binders detailed above was used. In subpocket S^{V3} , 100% of the potent antagonists that occupied the S^{V3} subpocket were predicted to form a π - π interaction with Phe358^{ECL3}. Phe358^{ECL3} was also involved in π - π interactions with 11% of the antagonists predicted to occupy subpockets other than

the S^{V3} subpocket. Asn175^{ECL1} was predicted to be involved in a hydrogen or halogen bond with 53% of the antagonists with docked poses that occupied S^{V3} . Asn363^{7.36} is located at the border between subpockets S^{V3} and S^{V0} and observed to form hydrogen bonds with 61% of antagonists that dock in S^{V3} and also with 22% of antagonists where the pose did not access S^{V3} . In the S^{V0} subpocket, the residue Gln348^{6.58} was predicted to form a hydrogen bond with antagonists in 35% of cases. Involvement of Tyr275^{ECL3} in binding of MCH-1R antagonists was also predicted in 8% of cases. The S^{V1} subpocket contains the aromatic residues Phe197^{3.37}, Phe286^{5.47}, Trp338^{6.48}, and Tyr342^{6.52}, which formed interactions with all docked poses of small MCH-1R antagonists that occupied the S^{V1} subpocket. Unfortunately, the MCH-1R model could not separate the individual contributions of each of these residues. The residues Gln196^{3.36} and Gln345^{6.55}, located on the border between subpockets S^{V0} and S^{V1} , were observed to be involved in hydrogen bonding in 71% and 64% of cases, respectively, when docked poses accessed the S^{V1} subpocket, and in 32% and 15% of cases, respectively, when S^{V1} was not occupied. In the S^{V2} subpocket, the analysis suggested that potent MCH-1R antagonists predicted to occupy this pocket would mainly interact with residues Asp192^{3.32} (72% frequency), Met168^{2.60} (35% frequency), Phe161^{2.53} (25% frequency), and Tyr370^{7.43} (32% frequency).

This exercise provided a shortlist of potential key residues, identified from the modeling of the MCH-1R structure with literature antagonists, that were predicted to be involved in small molecule ligand binding to MCH-1R: Asn175^{ECL1}, Asp192^{3.32}, Gln196^{3.36}, Phe197^{3.37}, Thr200^{3.40}, Phe358^{ECL3}, Phe286^{5.47}, Gln348^{6.58}, Trp338^{6.48}, Tyr342^{6.52}, Gln345^{6.55}, and Asn363^{7.36}. The fact that MCH-1R binds ligands of many chemotypes and sizes can be explained by the observation that MCH-1R antagonists have a large selection of key residues to bind to and the freedom to occupy multiple combinations of subpockets S^{V1} , S^{V2} , or S^{V3} .

3. CONCLUSION

Obesity is a chronic global epidemic disease and MCH-1R receptor antagonism remains a viable approach for obesity treatment.⁷ However, the progress in discovery of MCH-1R antagonists depends on the availability of molecular-level insights on the MCH-1R binding site. Here, we present a new practical approach that opens an alternative avenue for the structural exploration of GPCRs and structure-based GPCR drug discovery. The practicality and efficiency of this approach is enhanced by generation of structural information on the binding site of a targeted GPCR to satisfy the immediate need of the drug discovery process for such information and to drive structure-based drug design. This conceptually pioneering approach was applied to exploration of the MCH-1R structure and resulted in the discovery of 10 new antagonist chemotypes, obviating the need for running an expensive and time-consuming HTS. The usefulness of this approach to drug discovery, demonstrated in the complicated case of MCH-1R, should now allow SBDD even for some of the most challenging GPCR targets. The results presented here are comparable to those from reports that demonstrate that GPCR modeling supported by experimental data in the absence of a crystal structure can be a valid replacement for structural and functional exploration of GPCRs and for structure-based drug discovery.^{34–38}

Ligand SAR data, when combined with modeling, can provide a useful source of structural information on GPCR binding sites. The VAST compounds are designed to be highly efficient in exploring GPCR binding sites, and the experimental information they provide allows the construction of high quality GPCR models. The large size, complexity, and rigidity of the VAST hits improves the prediction of binding site location, topology, and key residues by limiting the number of feasible locations and orientations in which these hits can be docked into the receptor. In the case of MCH-1R, only one feasible binding site and binding mode was found for the VAST hits. Optimization of the GPCR model using VAST hits does not bias the GPCR binding site toward a single ligand chemotype, a common problem observed when performing the optimization with known ligands.¹⁴ For this reason, we obtained high enrichment for all chemotypes of known MCH-1R antagonists over chemically similar decoys and high performance in virtual screening.

We have demonstrated that the peptide's bioactive conformation and motifs can be mimicked by appending substituents onto a pyranose–sugar scaffold in a way that positions these elements in a similar relative orientation. These molecules are ideal for exploring the structural and chemical features required for compounds to interact with GPCR targets—particularly those GPCRs with large endogenous ligands, e.g., peptides or hormones—and for the discovery of ligands with novel chemotypes. The work described here was based on the original VAST library of 490 compounds that was biased toward specific motifs (represented by the general motif aromatic–aromatic–positive charge), targeting selected receptors including MCH-1R. A larger VAST library that is systematically diverse in motif and chemoform is now available (>14,000 compounds), thus providing an exciting opportunity to broaden the applicability of this approach to any other GPCR.

In this work, we aimed to use homology modeling, refined by VAST ligand directed modeling, to identify the MCH-1R structural features involved in the binding of small organic antagonists. In spite of the inherent limitations of modeling, such information can be usefully applied to the future discovery and lead optimization of MCH-1R antagonists and of GPCR ligands in general. On the basis of our analysis, it seems that MCH-1R antagonists can be classified not only based on their chemotype but also on the subpockets that they occupy and the unique interactions that they form with the MCH-1R binding site residues. While the model suggests that there are four distinct subpockets in the MCH-1R binding site, modeling of known antagonists indicates that it is not necessary for compounds to occupy all of these four subpockets. Two or three subpockets are sufficient for potent antagonism, allowing the design of smaller compounds that occupy a more drug-like property space in comparison with larger molecules. Design and optimization of novel MCH-1R antagonists that already access two or three subpockets can therefore be focused on optimizing interactions within those subpockets rather than expansion into additional parts of the site.

EXPERIMENTAL SECTION

Computational Methods. MCH-1R Residue Numbering. The position of each amino acid residue of MCH-1R was identified both by its sequence number and by its generic number proposed by Ballesteros and Weinstein.³⁹ Briefly, in this numbering system, amino acid residues in the seven trans-

membrane domain (7TMD) are given two numbers; the first corresponds to the trans-membrane helix (TM) number (1 to 7), while the second indicates the residue position relative to a highly conserved residue in class A GPCRs in that TM, which is arbitrarily assigned to 50. The numbering of the loops is done in a similar manner, for example, extracellular loop 2 (ECL2) is labeled 45 to indicate its location between helices 4 and 5, and the conserved cysteine (thought to be part of a disulfide bond) is given the index number 45.50. The residues within the ECL2 loop are then numbered relative to this position.

MCH-1R Modeling Procedure. We performed our hierarchical GPCR modeling protocol²⁷ in five sequential steps: (1) multiple sequence alignment, (2) template selection, (3) homology modeling of MCH-1R, (4) docking procedure and generation of ACL21823 complex, and (5) MCH-1R binding site optimization with “low-mode” molecular dynamics (LowModeMD) simulation.

Multiple Sequence Alignment. Multiple sequence alignment is required for selection of the optimum template from the available GPCR crystal structures for further homology modeling of MCH-1R. The amino acid sequence for human MCH-1R (Q99705) was retrieved from the SwissProt database (Figure 8). The sequence of MCH-1R was aligned with four

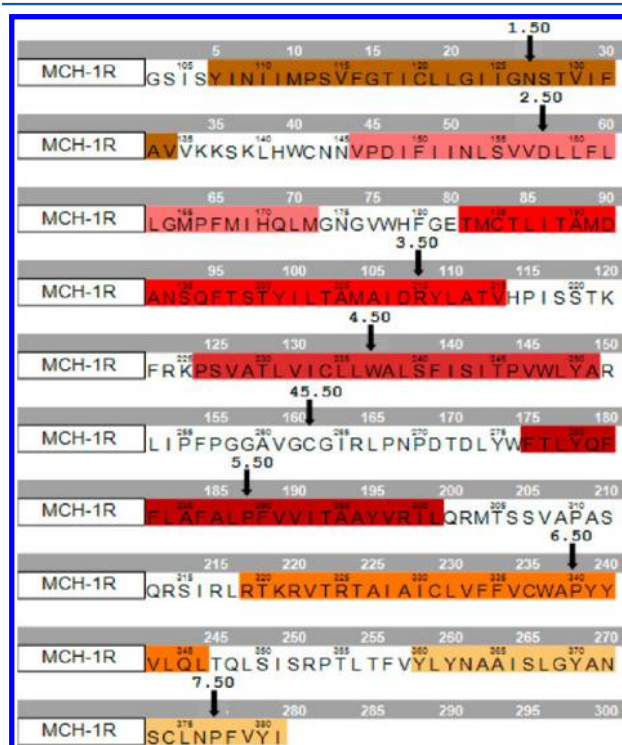


Figure 8. Amino-acid sequence of human MCH-1R, the color coding for TM 1 to 7 is dark orange, pink, red, purple, dark red, orange, and light yellow, respectively. The conserved residues in each TM that are assigned to 50 by the Ballesteros–Weinstein numbering scheme are shown by black arrows.

published GPCR crystal structures [human dopamine D₃ receptor (D₃, PDB entry 3PBL), β 2-adrenergic receptor (β 2AR, PDB entry 2RH1), human A2A adenosine receptor (A2A, PDB entry 3EML), and bovine rhodopsin (PDB entry 1F88)], using a multiple sequence alignment tool implemented in MOE version 2010.10 (Chemical Computing Group). In this approach, originally introduced by Needleman in 1970,⁴⁰ alignments were computed by optimizing a function based on

residue similarity scores obtained from applying an amino acid substitution matrix (blosum62)⁴¹ to pairs of aligned residues and gap penalties. Penalties were imposed for introducing and extending gaps in one sequence with respect to another. The gap start penalty was set at 7, and the penalty for gap extension was set at 1. The conserved residues and conserved GPCR motifs were constrained to ensure their proper alignment. The position of each amino acid was identified by its sequence number and by the generic number proposed by Ballesteros and Weinstein.

Template Selection. Template selection of the best template for modeling of any GPCR is usually very challenging.⁴² Class A GPCRs share the same arrangement of the seven helices,⁴³ and their 7TMD sequence similarities are relatively high.^{44–47} However, even small sequence differences can lead to significant differences in overall structure and particularly in the topology of the ligand binding site.⁴⁸ This renders each GPCR unique to its exclusive biological function. The most critical sequence difference in GPCRs is the difference in the positions of proline residues.^{49–51} The prolines force kinks in TM secondary structure, and as a result, even the smallest difference in the positions of prolines in the sequence alignment of the modeled GPCR and the template can result in a significant decrease in the accuracy of the model.⁵⁰ We ranked the quality of crystal structures as potential templates for homology of MCH-1R based on the maximum number of correctly aligned prolines in the 7TMD. The crystal structure that had the highest number of aligned prolines was chosen as a template for further homology modeling.

Homology Modeling of MCH-1R. Modeling of MCH-1R was performed using the homology modeling tool as implemented in the MOE software package (Chemical Computing Group, version 2010.10), based on the template selected in the previous stage. Due to insertions in the MCH-1R sequences with respect to the template, some residues, particularly in the loop regions, did not have assigned backbone geometries based on the template. These insertions were modeled from segments of high-resolution chains from the protein data bank (PDB) that superposed well onto anchor residues on each side of the insertion area, after the method described by Fechteler et al.⁵² Following the selection of appropriate loop templates, multiple model candidates for each loop were constructed and scored using the OPLS-AA energy function.^{53,54} The coordinates of the top-ranked loop model were added to the global model. After all of the loops had been added, the side chains were modeled. Side chain data is assembled from an extensive rotamer library generated by systematic clustering of high-resolution PDB data. After all of the backbone segment and side chain conformations were chosen, the model was minimized using the OPLS-AA^{53,54} force field.⁵⁵

Docking Procedure and Generation of ACL21823 Complex. In all docking experiments described in this manuscript, we used the GOLD docking package (Cambridge Crystallographic Data Centre, version 5.0), followed by a consistent rescoring and reranking procedure. The docking itself was performed once for each molecule, with the 10 top-ranked docking poses scored by the GOLD default scoring function⁵⁶ retained. We then rescored and reranked these 10 docked poses using AMBER interaction energy. We used the MM_PBSA/GBSA approach⁵⁷ to calculate the AMBER interaction energy (see details in the Supporting Information) in order to take the entropy and solvation energy change upon binding into

account, as well as the binding enthalpy between protein and ligand. As was recently published,^{58,59} the AMBER interaction energy, while subject to the same limitations as all force field based methods, was able to accurately predict relative binding affinities between ligand and protein and was therefore selected as a reliable method to rescore and to rank docking poses. The top-ranked docking pose, according to the AMBER interaction energy, was taken for further analysis. Except when docking ACL21823, a rigid docking procedure was used, with the backbone and side chains of MCH-1R restrained during docking.

When docking ACL21823, a flexible GOLD docking procedure was used, as ACL21823 was being docked into an unrefined binding site. We assigned flexibility to all aromatic residues using the GOLD rotamer library to ensure that the docking procedure could adjust the site in order to accommodate large molecules such as ACL21823. Because the exact location of the MCH-1R binding site is not known, we initially defined a large region toward the extracellular end of the MCH-1R 7TMD as a potential binding site. Analysis of the 10 top-ranked docking poses of ACL21823 revealed that they all share high similarity; pairwise heavy-atom RMSD was less than 1.5 Å for all pairs of poses. The top-ranked pose of ACL21823, according to the AMBER interaction energy, was selected for further binding site refinement.

MCH-1R Binding Site Optimization. In the final MCH-1R modeling stage, we aimed to optimize the MCH-1R binding site of MCH-1R by applying a “low-mode” molecular dynamics simulation (LowModeMD⁶⁰) to docked ACL21823 pose. The LowModeMD protocol is a stochastic conformation generation protocol implemented in MOE. For the LowModeMD refinement, the MCH-1R binding site was defined by residues within a radius of 7.0 Å of the docked ACL21823. Flexibility was permitted for all atoms within this radius when the rest of the atoms of MCH-1R and of ACL21823 were restrained. The radius of 7.0 Å was selected to ensure that all the atoms of helical fraction that include close to ACL21823 residues will be also included in refinement. The OPLS-AA force field was used for the conformational search as the MOE AMBER implementation did not contain appropriate atom types for the ligand. The dielectric constant was set to 3. A LowModeMD constant temperature MD simulation was performed at 300 K, using the Berendsen thermostat⁶¹ and the velocity Verlet algorithm. The default value of the energy minimization gradient (0.001 kcal mol⁻¹ Å⁻²) was used. The LowModeMD and stochastic searches were terminated after 200 failed attempts to generate a new conformation, with a maximum of 10,000 iterations. To ensure that the TM helices did not “unwind” during the optimization, simple harmonic distance constraints were applied to mimic the α -helical $i, i + 4$ carbonyl–amine hydrogen bonds.

Validation of the MCH-1R Model. Virtual “Enrichment” Experiment. For primary validation of the MCH-1R model, a virtual enrichment experiment was carried out. This step was designed to test the ability of the MCH-1R model to separate known MCH-1R antagonists from a library of chemically similar analogues and to determine if the quality of the MCH-1R model was high enough to be used in structure-based virtual screening. To generate a representative library of MCH-1R antagonists, a total of 2723 potent MCH-1R antagonists were extracted from the Linceptor database⁶² and clustered using a spectral clustering method.⁶³ One representative from each of the 126 clusters identified was retained, providing exhaustive

coverage of all available MCH-1R antagonist chemotypes. For the decoys, a set of 3921 compounds was extracted from commercial vendor libraries using Tripos UNITY fingerprints,⁶⁴ with a Tanimoto similarity cutoff to the known 126 MCH-1R binders >0.7. The use of decoys similar to the known compounds provided a challenging validation exercise, success in which would indicate a high-quality MCH-1R model. The library of 4047 compounds (126 known binders and 3921 decoys) was docked into the MCH-1R model with the binding site refined based on ACL21823. The docked poses were rescored and ranked by protein–ligand interaction energies calculated with the AMBER force field as described in previous sections. For comparison, compounds were also ranked with ROCS, a shape-based similarity method (see details in the Supporting Information), using both the docked pose of ACL21823 and its minimized structure as queries.

Validation of MCH-1R Structure with ACL21823 Analogues. To further validate the model, we synthesized ACL21823 analogues with various R₁, R₂, and R₃ substituents (Table 1), predicted to be highly significant for MCH-1R binding. Synthesis of these analogues is depicted in Scheme 1 and described below.

Structure-Based Virtual Screen. This step aimed to discover a novel chemotype for potent MCH-1R ligands. A virtual library of 45,000 compounds with Tanimoto similarity ≤0.7 to the 126 known MCH-1R antagonists was extracted from commercial vendor catalogues. We used the same docking and scoring procedure as the enrichment study to rank the library compounds and select a shortlist of highly ranked virtual hits for testing. These 70 compounds were purchased and tested in an aequorin-based inhibition assay as described below.

Synthesis of VAST Compounds. The 490 VAST library compounds as well as the ACL21823 analogues were prepared following previously published solid-phase synthetic methods.^{17,31} The synthetic route for the preparation of the ACL21823 analogues is depicted in Scheme 1.

General Procedure for Glycosylation Reaction. DMTST (Dimethyl(methylthio)sulfonium triflate) Preparation. Methylidisulfide (30 mL, 0.34 mol) was diluted in 300 mL dry DCM under N₂. Methyl triflate (50 g, 0.30 mol) was added dropwise at RT to the stirred methylidisulfide solution, over 15 min. The resulting solution was stirred at RT overnight under N₂. The reaction mixture was then poured into 1.4 L dry diethyl ether; the mixture was stirred well and then left to stand for 30 min. The resulting white precipitate was collected by filtration, washed with dry diethyl ether, and dried under high vacuum overnight (74.4 g of white solid, 95%).

Glycosylation Reaction. Orthogonally protected building block **1** (Glc-SMe) and alcohol R₁OH (5 equiv) were dissolved in dry DCM (6 mL per gram of **1**), and the mixture was stirred at RT over molecular sieves (AW300, previously dried at 300 °C) under N₂. After 1 h, a solution of DMTST (0.5 M, 2.5 equiv) in dry DCM was added to the sugar and alcohol mixture, and the resulting solution was stirred under N₂ at RT. Evolution of reaction was checked by TLC and LC/MS. After complete reaction of starting material (1–3 h), the reaction mixture was filtered on Celite, and the Celite was washed twice with DCM. The organic phase was washed three times with a saturated NaHCO₃ solution. The aqueous phase was re-extracted once with DCM. The combined organic phases were dried over MgSO₄ and, after filtration, evaporated under reduced pressure to afford a crude residue, which contained desired α/β anomers in both TBDPS-protected and

unprotected form. To reprotect the 6-position with the TBDPS group, the crude material was treated with imidazole (7 equiv to starting building block **1**) and TBDPS-Cl (4.5 equiv to starting building block **1**) in dry DMF (7 mL per gram of starting material). The mixture was stirred under N₂ at RT, and evolution of reaction was checked by TLC and LC/MS. After completion (typically 2 h), the reaction mixture was diluted with DCM, and the solution was washed with distilled water. The water layer was re-extracted three times with DCM, and all organic layers were combined, dried over MgSO₄, and after filtration, evaporated under reduced pressure. The α- and β-anomers were separated and purified from the recovered residue by column chromatography.

(4'-Chlorobenzyl)-2-azido-3-O-benzoyl-6-O-(tert-butylidiphenylsilyl)-2-deoxy-1-α-D-glucopyranoside (2a). GlcSMe **1** (5.0 g, 8.7 mmol) and 4-chlorobenzyl alcohol (6.2 g, 43.3 mmol) were used in the glycosylation reaction as described above. The crude material was purified by column chromatography using PE/EtOAc (95/5) as eluent to give the α-anomer **2a** as a white solid (1.45 g, 25%): ¹H NMR (400 MHz, CDCl₃) δ 1.08 (s, 9H, CH₃), 2.87 (d, 1H, J_{H4,OH} = 3.5 Hz, OH), 3.40 (dd, 1H, J_{H1,H2} = 3.6 Hz, J_{H2,H3} = 10.6 Hz, H-2), 3.81–3.89 (m, 2H, H-4, H-5), 3.90–3.97 (m, 2H, H-6a, H-6b), 4.55 (d, 1H, J_{A,B} = 12.2 Hz, OCH₂(A)), 4.72 (d, 1H, OCH₂(B)), 5.04 (d, 1H, J_{H1,H2} = 3.5 Hz, H-1), 5.60 (dd, 1H, J_{H2,H3} = 10.4 Hz, J_{H3,H4} = 8.4 Hz, H-3), 7.28–8.12 (m, 19H, H_{Ar}); MS (ES) *m/z* 694.14 [M+Na]⁺.

(4'-Fluorobenzyl)-2-azido-3-O-benzoyl-6-O-(tert-butylidiphenylsilyl)-2-deoxy-1-α-D-glucopyranoside (2b). GlcSMe **1** (5.1 g, 8.8 mmol) and 4-fluorobenzyl alcohol (4.8 mL, 44.5 mmol) were used in the glycosylation reaction as described above. The crude material was purified by column chromatography using PE/EtOAc (92/8) as eluent to give the α-anomer **2b** as a white solid (1.44 g, 25%): ¹H NMR (400 MHz, CDCl₃) δ 1.08 (s, 9H, CH₃), 2.89 (d, 1H, J_{H4,OH} = 3.8 Hz, OH), 3.39 (dd, 1H, J_{H1,H2} = 3.6 Hz, J_{H2,H3} = 10.6 Hz, H-2), 3.81–3.86 (m, 2H, H-4, H-5), 3.92–3.94 (m, 2H, H-6a, H-6b), 4.55 (d, 1H, J_{A,B} = 11.9 Hz, OCH₂(A)), 4.72 (d, 1H, OCH₂(B)), 5.04 (d, 1H, J_{H1,H2} = 3.5 Hz, H-1), 5.60 (dd, 1H, J_{H2,H3} = 10.4 Hz, J_{H3,H4} = 8.8 Hz, H-3), 7.01–7.06 (m, 2H, H_{Ar}), 7.31–8.11 (m, 17H, H_{Ar}); MS (ES) *m/z* 678.13 [M+Na]⁺.

(4'-Trifluoromethylbenzyl)-2-azido-3-O-benzoyl-6-O-(tert-butylidiphenylsilyl)-2-deoxy-1-α-D-glucopyranoside (2c). GlcSMe **1** (5.1 g, 8.9 mmol) and 4-trifluoromethylbenzyl alcohol (5.0 g, 28.4 mmol, 3.2 equiv) were used in a modified glycosylation procedure. Et₂O (80 mL) was used as the solvent instead of DCM and DMTST (2 equiv) was added as a solid. After overnight reaction at RT, the reaction mixture was diluted with DCM and worked up as described in the general glycosylation procedure. TBDPS reprotection was not necessary in this case. The crude material was purified by column chromatography using PE/EtOAc (95/5 to 90/10) as eluent to give the α-anomer **2c** as a white solid (2.82 g, 45%): ¹H NMR (400 MHz, CDCl₃) δ 1.08 (s, 9H, CH₃), 2.87 (d, 1H, J_{H4,OH} = 3.5 Hz, OH), 3.44 (dd, 1H, J_{H1,H2} = 3.2 Hz, J_{H2,H3} = 10.4 Hz, H-2), 3.81–3.90 (m, 2H, H-4, H-5), 3.91–3.95 (m, 2H, H-6a, H-6b), 4.64 (d, 1H, J_{A,B} = 12.6 Hz, OCH₂(A)), 4.82 (d, 1H, OCH₂(B)), 5.07 (d, 1H, J_{H1,H2} = 3.5 Hz, H-1), 5.61 (dd, 1H, J_{H2,H3} = 10.4 Hz, J_{H3,H4} = 8.4 Hz, H-3), 7.36–8.12 (m, 19H, H_{Ar}); MS (ES) *m/z* 728.10 [M+Na]⁺.

(3',4'-Difluorobenzyl)-2-azido-3-O-benzoyl-6-O-(tert-butylidiphenylsilyl)-2-deoxy-1-α-D-glucopyranoside (2d).

GlcSMe **1** (5.0 g, 8.7 mmol) and 3,4-difluorobenzyl alcohol (3.65 mL, 31.9 mmol, 3.7 equiv) were used in a modified glycosylation procedure. Et₂O/DCM (2:1) was used as the solvent and DMTST (2 equiv) was added as a solid. After overnight reaction at 0 °C, the reaction mixture was diluted with EtOAc and worked up as described in the general glycosylation procedure. TBDPS reprotection was not necessary in this case. The crude material was purified by column chromatography using PE/EtOAc (90/10 to 0/100) as eluent to give the α -anomer **2d** as a white solid (1.98 g, 34%): ¹H NMR (400 MHz, CDCl₃) δ 1.08 (s, 9H, CH₃), 2.88 (d, 1H, $J_{H_4,OH}$ = 3.6 Hz, OH), 3.42 (dd, 1H, J_{H_1,H_2} = 3.2 Hz, J_{H_2,H_3} = 10.4 Hz, H-2), 3.80–3.88 (m, 2H, H-4, H-5), 3.91–3.98 (m, 2H, H-6a, H-6b), 4.51 (d, 1H, $J_{A,B}$ = 12.0 Hz, OCH₂(A)), 4.69 (d, 1H, OCH₂(B)), 5.03 (d, 1H, J_{H_1,H_2} = 3.5 Hz, H-1), 5.59 (dd, 1H, J_{H_2,H_3} = 10.4 Hz, J_{H_3,H_4} = 8.0 Hz, H-3), 7.08–7.21 (m, 3H, H_{Ar}), 7.38–8.11 (m, 15H, H_{Ar}); MS (ES) m/z 696.14 [M + Na]⁺.

General Solid-Phase Procedures. Resin Loading. Commercial TCA Wang resin (loading = 1.08 mmol/g) and glycoside **2** were dried under high vacuum over KOH overnight before use. The resin was washed with dry THF ($\times 1$) and dry DCM ($\times 2$) under N₂. The sugar (2 equiv) was dissolved in dry DCM (8 mL per gram of resin) under N₂, and the resulting solution was added to the resin, which was shaken for 2 min at RT. BF₃·Et₂O (0.3 equiv) was then added to the resin, which was shaken for 10 min. The addition of BF₃·Et₂O (0.3 equiv) was repeated once, followed by shaking for 10 min at RT. After draining, the resin was finally washed with dry DCM ($\times 5$) to recover the glycoside. The filtrate with unreacted glycoside was washed with saturated NaHCO₃ solution in water ($\times 3$). The organic phase was dried over MgSO₄, and after filtration, evaporated under reduced pressure to give crude recovered sugar **2**, which was dried under high vacuum. Resin **3** was further washed with dry THF ($\times 4$) and dry DCM ($\times 4$) and then dried under high vacuum.

Resin Loading Determination. The glycoside resin **3** was dried under high vacuum for a minimum of 16 h. A mixture of anhydrous DCM/TES/TFA (7:2:1) was prepared under N₂. About 70 mg of the dried loaded resin was weighed precisely, and the cleavage solution (1 mL) was added. The resulting mixture was left at RT for 3 h. The solution was then drained and the resin was washed with DCM (4 \times 2 mL). The filtrates were combined and diluted to a known volume. The concentration of the corresponding glycoside was then assessed using HPLC/UV chromatography comparison against a standard curve obtained from five known glycoside concentrations (ranging from 0.06 to 1.2 mg/mL), and the resulting resin loading value was derived. Typical resin loading values were in the range 0.2–0.4 mmol/g.

Library Synthesis. The library compounds were synthesized following the “split-and-pool” technique, using the IRORI directed sorting process and the Synthesis Manager software. Each individual resin was split in the required number of Kan microreactors. Around 80 mg of resin was introduced in each Kan, using the IRORI dry resin filler. Each microreactor contained a unique tag identifier (ceramic 2D-coded cap). By pooling microreactors into common building block groups, the directed sorting process made the synthesis efficient, while enabling the synthesis of one discrete compound in each microreactor. Quality control (QC) Kans were added to monitor evolution of each reaction along the library production process.

Benzoate Removal at O3. A solution of NaOMe 240 mM in dry MeOH was prepared under N₂ from Na metal. The Kans were treated with a mixture dry THF/NaOMe solution in MeOH (83:17, 2.5 mL per Kan, final NaOMe concentration = 40 mM) and were shaken overnight at RT. The solution was drained and the Kans washed with MeOH, THF, MeOH, and DCM ($\times 3$ each). The resin was dried under high vacuum overnight.

Alkylation at O3. A solution of LiO^tBu in dry DMF (0.25 M) was prepared under N₂, as well as a solution of alkylating agent in dry DMF (0.25 M). The Kans were washed once with dry DMF and then treated with the LiO^tBu solution (2.5 mL per Kan) for 10 min at RT. The LiO^tBu solution was removed, and the alkylating agent solution (2.5 mL per Kan) was added. The Kans were shaken for 20 min at RT and then the solution was removed. This LiO^tBu/alkylating agent cycle was performed three times, using fresh LiO^tBu solution and fresh alkylating agent solution. The Kans were finally washed with DMF, MeOH, and DCM ($\times 3$ each), and resin **4** was dried under high vacuum overnight.

Preparation of PSHF Reagent. To a solution of proton sponge (1,8-bis-(dimethylamino)-naphthalene, 50 g, 0.233 mol) in dry THF (300 mL) was added HF-Pyridine (HF 70%, Pyridine 30%, 12.5 mL, 2 equiv), in a plastic vessel, dropwise over 10 min under N₂. The mixture was stirred for 15 min before filtration of the resulting white precipitate. The precipitate was washed with dry THF (4 \times 50 mL) and dry Et₂O (4 \times 100 mL) and then dried under high vacuum overnight (44 g, 81%).

TBDPS Deprotection at O6. A solution of 0.5 M PSHF and AcOH 5% in dry DMF was added to the Kans (2.5 mL per Kan). The mixture was stirred and heated to 65 °C overnight. After cooling down, the solution was aspirated and the Kans washed with DMF, THF, and DCM ($\times 3$ each). The resulting resin was dried under high vacuum overnight.

Methylation at O6. This reaction was performed following the procedure described for alkylation at O3 (see above), using methyl iodide as the alkylating agent, to give resin **5**.

Azide Reduction. A solution of 0.2 M LiO^tBu and 0.2 M DTT in dry DMF was added to the Kans (2.5 mL per Kan), which were shaken overnight at RT. The solution was aspirated, the Kans washed with DMF, MeOH, and DCM ($\times 3$ each), and the resin dried under high vacuum overnight.

Acylation at N2. A solution of HBTU 0.5 M in dry DMF was prepared. The acid was dissolved in dry DMF (0.5 M) and treated with HBTU (0.5 M, same volume) and DIPEA (final concentration 0.5 M). The resulting mixture was swirled, allowed to stand for 10 min, and then added to the Kans (2.5 mL per Kan), which were shaken at RT overnight. The solution was then aspirated, the Kans washed with DMF, MeOH, and DCM ($\times 3$ each), and resulting resin **6** dried under high vacuum overnight.

Fmoc Deprotection. The Kans were treated with a solution of 20% piperidine in dry DMF (2.5 mL per Kan), and the resulting mixture was shaken for 2 h at RT. The solution was aspirated, the Kans washed with DMF, MeOH, and DCM ($\times 3$ each), and the resin dried under high vacuum overnight.

Guanidine Formation. The Kans were treated with a solution of DMPFN (3,5-dimethyl-1-pyrazolylformamidinium nitrate, 0.25 M) and DIPEA (1 M) in dry DMF (2.5 mL per Kan), and the resulting mixture was stirred and heated to 65 °C overnight. After cooling, the solution was aspirated, the Kans

washed with DMF, MeOH, and DCM ($\times 3$ each), and resin 7 was dried under high vacuum overnight.

Cleavage Off-Resin. Each Kan of the library was placed in a tube of a Bohdan block and treated with 2 mL of cleavage solution (TFA/TES/dry DCM 10/20/70) for 3 h at RT. The cleavage solution was drained in the corresponding collecting tube and the Kan washed with THF, DCM, and THF (1 mL each). The solvents were evaporated to obtain crude cleavage products.

NH₃ Treatment. Each crude residue from the cleavage step was treated with freshly prepared saturated ammonia solution in MeOH (2 mL). The solution was gently shaken and left to stand for 2 h at RT. The solution was evaporated, and crude samples 8 were purified by preparative LC/MS on a standard C18 column using gradient programs of acetonitrile and water (containing 0.01% TFA). The purified samples were reanalysed by LC/MS with parallel ELSD detection to determine purity (column Zorbax C18 (4.6 \times 50 mm); solvent = H₂O (0.01% TFA)/ACN (0.01% TFA); gradient = 5 (0–1), 100 (7–12); flow = 2 mL/min).

(4'-Chlorobenzyl)-2-deoxy-2-(4'-guanidinobutrylamido)-3-O-(4'-chlorobenzyl)-6-O-methyl- α -D-glucopyranoside (8a). Resin loaded α -glycoside 2a (80 mg resin, 0.44 mmol/g) was carried through the solid phase synthetic process. The title product was isolated by HPLC purification and MS detection as a white solid (13.1 mg, 65%): C18 HPLC/ELSD *Rt* 5.14 min (99.82%); MS (ES) *m/z* 568.99 [M+H]⁺.

(4'-Trifluoromethylbenzyl)-2-deoxy-2-(4'-guanidinobutrylamido)-3-O-(4'-chlorobenzyl)-6-O-methyl- α -D-glucopyranoside (8b). Resin loaded α -glycoside 2c (80 mg resin, 0.40 mmol/g) was carried through the solid phase synthetic process. The title product was isolated by HPLC purification and MS detection as a white solid (13.3 mg, 69%): C18 HPLC/ELSD *Rt* 5.33 min (99.99%), MS (ES) *m/z* 603.02 [M+H]⁺.

(4'-Fluorobenzyl)-2-deoxy-2-(4'-guanidinobutrylamido)-3-O-(4'-chlorobenzyl)-6-O-methyl- α -D-glucopyranoside (8c). Resin loaded α -glycoside 2b (80 mg resin, 0.44 mmol/g) was carried through the solid phase synthetic process. The title product was isolated by HPLC purification and MS detection as a white solid (10.3 mg, 53%): C18 HPLC/ELSD *Rt* 5.05 min (99.37%), MS (ES) *m/z* 553.02 [M+H]⁺.

(3',4'-Difluorobenzyl)-2-deoxy-2-(4'-guanidinobutrylamido)-3-O-(4'-chlorobenzyl)-6-O-methyl- α -D-glucopyranoside (8d). Resin loaded α -glycoside 2d (80 mg resin, 0.43 mmol/g) was carried through the solid phase synthetic process. The title product was isolated by HPLC purification and MS detection as a white solid (12.4 mg, 63%): C18 HPLC/ELSD *Rt* 5.05 min (100%), MS (ES) *m/z* 571.09 [M+H]⁺.

(4'-Chlorobenzyl)-2-deoxy-2-(4'-guanidinobutrylamido)-3-O-(4'-trifluoromethylbenzyl)-6-O-methyl- α -D-glucopyranoside (8e). Resin loaded α -glycoside 2a (80 mg resin, 0.44 mmol/g) was carried through the solid phase synthetic process. The title product was isolated by HPLC purification and MS detection as a white solid (11.7 mg, 55%): C18 HPLC/ELSD *Rt* 5.25 min (98.44%), MS (ES) *m/z* 603.02 [M+H]⁺.

(4'-Chlorobenzyl)-2-deoxy-2-(4'-guanidinobutrylamido)-3-O-(4'-fluorobenzyl)-6-O-methyl- α -D-glucopyranoside (8f). Resin loaded α -glycoside 2a (80 mg resin, 0.44 mmol/g) was carried through the solid phase synthetic process. The title product was isolated by HPLC purification and MS

detection as a white solid (11.4 mg, 59%): C18 HPLC/ELSD *Rt* 5.02 min (98.03%), MS (ES) *m/z* 553.02 [M+H]⁺.

(4'-Fluorobenzyl)-2-deoxy-2-(4'-guanidinobutrylamido)-3-O-(4'-trifluoromethylbenzyl)-6-O-methyl- α -D-glucopyranoside (8g). Resin loaded α -glycoside 2b (80 mg resin, 0.44 mmol/g) was carried through the solid phase synthetic process. The title product was isolated by HPLC purification and MS detection as a white solid (12.5 mg, 61%): C18 HPLC/ELSD *Rt* 5.14 min (97.34%), MS (ES) *m/z* 587.06 [M+H]⁺.

(3',4'-Difluorobenzyl)-2-deoxy-2-(4'-guanidinobutrylamido)-3-O-(4'-fluorobenzyl)-6-O-methyl- α -D-glucopyranoside (8h). Resin loaded α -glycoside 2d (80 mg resin, 0.43 mmol/g) was carried through the solid phase synthetic process. The title product was isolated by HPLC purification and MS detection as a white solid (13.8 mg, 72%): C18 HPLC/ELSD *Rt* 4.93 min (99.77%), MS (ES) *m/z* 555.00 [M+H]⁺.

(3',4'-Difluorobenzyl)-2-deoxy-2-(4'-guanidinobutrylamido)-3-O-(4'-trifluoromethylbenzyl)-6-O-methyl- α -D-glucopyranoside (8i). Resin loaded α -glycoside 2d (80 mg resin, 0.43 mmol/g) was carried through the solid phase synthetic process. The title product was isolated by HPLC purification and MS detection as a white solid (11.8 mg, 57%): C18 HPLC/ELSD *Rt* 5.08 min (99.2%), MS (ES) *m/z* 605.13 [M+H]⁺.

(4'-Chlorobenzyl)-2-deoxy-2-(4'-guanidinobutrylamido)-3-O-benzyl-6-O-methyl- α -D-glucopyranoside (8j). Resin loaded α -glycoside 2a (80 mg resin, 0.44 mmol/g) was carried through the solid phase synthetic process. The title product was isolated by HPLC purification and MS detection as a white solid (11.1 mg, 59%): C18 HPLC/ELSD *Rt* 4.71 min (100%), MS (ES) *m/z* 535.20 [M+H]⁺.

Competition Binding Assay on MCH-1R. The competition binding assay was performed using membranes prepared from CHO cells stably and selectively expressing the cloned human MCH-1R. The assay was carried out in an Optiplat (Perkin-Elmer, 60052) wherein test compound (50 μ L, used at increasing concentrations) was combined with successive additions of 25 μ L of radioligand ([¹²⁵I]-MCH, Perkin-Elmer, NEX 375, 2200 Ci/mMol) diluted in assay buffer (25 mM HEPES pH 7.4, 1 mM CaCl₂, 5 mM MgCl₂, 0.2% protease free BSA) and 25 μ L of a premix of membrane/beads (PVT-WGA beads, Amersham, RPNQ0001, 0.25 mg/well). Plates were incubated at room temperature for 2 h before counting for 1 min/well in a TopCount (Packard). Phe-(D)Tyr-MCH (Bachem, H-2218) was used as the reference ligand. IC₅₀ values were calculated by nonlinear regression analysis using a four parameter sigmoidal dose–response model (XLfit, IDBS Software).

Functional Aequorin-Based Assay on MCH-1R. The aequorin-based assay was performed in CHO cells stably and selectively expressing the MCH₁ receptor. Cells in mid-log phase, grown in media without antibiotic 18 h prior to the test, were detached by gentle flushing with PBS-EDTA (5 mM EDTA), recovered by centrifugation, and resuspended in “BSA medium” (DMEM/HAMs F12 Gibco Cat#11039-021, with HEPES, without phenol red +0.1% BSA). Cells were then counted, centrifuged, and resuspended in a 15 mL Falcon tube at a concentration of 1.7×10^6 cells/mL. Coelenterazine h (Promega, Cat n° X300X, stock solution: 500 μ M in methanol) was added to the cells in suspension at a final concentration of 5 μ M. The Falcon tube, wrapped in aluminum foil, was then placed on a vertical rotating wheel (6 rpm) and incubated at

room temperature (temperature should be maintained below 22 °C) overnight in order to reconstitute active aequorin. Cells were counted, diluted in “BSA-medium”, and incubated for 60 min prior to experimentation. Test and reference compounds were similarly prepared and distributed in a 384 well plate (30 μ L/well) prior to injection of cells (5000 cells/30 μ L). Cells were injected, and agonist activity was measured by recording the emission of light (FDSS 6000 Hamamatsu) for a period of 90 s. Antagonist activity was tested on the same preparation following the addition and 3 min incubation with 30 μ L of the reference agonist (MCH, Bachem, H1482) at a final concentration equivalent to the EC₈₀ determined for the same experimental day (using the reader injection system). All values were normalized relative to background (defined as 0%), and the maximal response to MCH (defined as 100%). EC₅₀ values were calculated by nonlinear regression analysis using a four parameter sigmoidal dose–response model (XLfit, IDBS Software).

■ ASSOCIATED CONTENT

■ Supporting Information

Methods used to calculate AMBER MCH-1R ligand interaction energy. Shape comparison and statistical comparison of subpocket occupancy effects, multiple amino-acid sequence alignment results of MCH-1R sequence aligned with four GPCR crystal structures [human dopamine D3 receptor (D3, PDB entry 3PBL), β 2-adrenergic receptor (β 2AR, PDB entry 2RH1), human A2A adenosine receptor (A2A, PDB entry 3EML), and bovine rhodopsin (PDB entry 1U19)]. Table of 32 molecules used as training set for statistical analysis. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Alexander.Heifetz@Evotec.com.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge Euroscreen, in particular Drs. Graeme Fraser and Sébastien Blanc, for the determination of MCH-1R IC₅₀ and EC₅₀ values for the SAR analogues and Dr. Daniel Warner for his help in statistical analysis of MCH-1R antagonists docking poses. We are also grateful to the Royal Society for an Industry Award (A.H.) and to the Queensland Government Smart State Fund for a NIRAP grant.

■ ABBREVIATIONS

MCH-1R, melanin-concentrating hormone receptor 1; MCH-2R, melanin-concentrating hormone receptor 2; GPCR, G-protein-coupled receptor; MCH, melanin-concentrating hormone; VAST, versatile assembly on stable templates; 7TMD, seven transmembrane domain; TM, transmembrane helix; ECL, extracellular loop; PDB, Protein Data Bank; SAR, structure–activity relationship; VS, virtual screening; SBDD, structure-based drug discovery

■ REFERENCES

- (1) Forte, V.; Pandey, A.; Abdelmessih, R.; Forte, G.; Whaley-Connell, A.; Sowers, J. R.; McFarlane, S. I. Obesity, diabetes, the cardiorenal syndrome, and risk for cancer. *Cardiorenal. Med.* **2012**, *2*, 143–162.
- (2) Saito, Y.; Maruyama, K. Identification of melanin-concentrating hormone receptor and its impact on drug discovery. *J. Exp. Zool., Part A: Comp. Exp. Biol.* **2006**, *305*, 761–768.
- (3) Xu, Y. L.; Jackson, V. R.; Civelli, O. Orphan G protein-coupled receptors and obesity. *Eur. J. Pharmacol.* **2004**, *500*, 243–253.
- (4) Schioth, H. B. G protein-coupled receptors in regulation of body weight. *CNS. Neurol. Disord. Drug Targets* **2006**, *5*, 241–249.
- (5) Bednarek, M. A. Peptide ligands for the melanin-concentrating hormone (MCH) receptor 1. *Curr. Top. Med. Chem.* **2007**, *7*, 1425–1432.
- (6) Kowalski, T. J.; McBriar, M. D. Therapeutic potential of melanin-concentrating hormone-1 receptor antagonists for the treatment of obesity. *Expert. Opin. Investig. Drugs* **2004**, *13*, 1113–1122.
- (7) Hogberg, T.; Frimurer, T. M.; Sasmal, P. K. Melanin concentrating hormone receptor 1 (MCHR1) antagonists—Still a viable approach for obesity treatment? *Bioorg. Med. Chem. Lett.* **2012**, *22*, 6039–6047.
- (8) Miyamoto-Matsubara, M.; Chung, S.; Saito, Y. Functional interaction of regulator of G protein signaling-2 with melanin-concentrating hormone receptor 1. *Ann. N.Y. Acad. Sci.* **2010**, *1200*, 112–119.
- (9) Della-Zuana, O.; Presse, F.; Ortola, C.; Duhault, J.; Nahon, J. L.; Levens, N. Acute and chronic administration of melanin-concentrating hormone enhances food intake and body weight in Wistar and Sprague-Dawley rats. *Int. J. Obes. Relat. Metab. Disord.* **2002**, *26*, 1289–1295.
- (10) Takekawa, S.; Asami, A.; Ishihara, Y.; Terauchi, J.; Kato, K.; Shimomura, Y.; Mori, M.; Murakoshi, H.; Kato, K.; Suzuki, N.; Nishimura, O.; Fujino, M. T-226296: A novel, orally active and selective melanin-concentrating hormone receptor antagonist. *Eur. J. Pharmacol.* **2002**, *438*, 129–135.
- (11) Cipitelli, A.; Karlsson, C.; Shaw, J. L.; Thorsell, A.; Gehlert, D. R.; Heilig, M. Suppression of alcohol self-administration and reinstatement of alcohol seeking by melanin-concentrating hormone receptor 1 (MCH1-R) antagonism in Wistar rats. *Psychopharmacology (Berl)* **2010**, *211*, 367–375.
- (12) Borowsky, B.; Durkin, M. M.; Ogozalek, K.; Marzabadi, M. R.; DeLeon, J.; Lagu, B.; Heurich, R.; Lichtblau, H.; Shaposhnik, Z.; Daniewska, I.; Blackburn, T. P.; Branchek, T. A.; Gerald, C.; Vaysse, P. J.; Forray, C. Antidepressant, anxiolytic and anorectic effects of a melanin-concentrating hormone-1 receptor antagonist. *Nat. Med.* **2002**, *8*, 825–830.
- (13) Shearman, L. P.; Camacho, R. E.; Sloan, S. D.; Zhou, D.; Bednarek, M. A.; Hreniuk, D. L.; Feighner, S. D.; Tan, C. P.; Howard, A. D.; Van der Ploeg, L. H.; Macintyre, D. E.; Hickey, G. J.; Strack, A. M. Chronic MCH-1 receptor modulation alters appetite, body weight and adiposity in rats. *Eur. J. Pharmacol.* **2003**, *475*, 37–47.
- (14) Cavasotto, C. N.; Orry, A. J.; Murgolo, N. J.; Czarniecki, M. F.; Kocsi, S. A.; Hawes, B. E.; O'Neill, K. A.; Hine, H.; Burton, M. S.; Voigt, J. H.; Abagyan, R. A.; Bayne, M. L.; Monsma, F. J., Jr. Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. *J. Med. Chem.* **2008**, *51*, 581–588.
- (15) Clark, D. E.; Higgs, C.; Wren, S. P.; Dyke, H. J.; Wong, M.; Norman, D.; Lockey, P. M.; Roach, A. G. A virtual screening approach to finding novel and potent antagonists at the melanin-concentrating hormone 1 receptor. *J. Med. Chem.* **2004**, *47*, 3962–3971.
- (16) Tate, C. G. A crystal clear solution for determining G-protein-coupled receptor structures. *Trends Biochem. Sci.* **2012**, *37*, 343–352.
- (17) Abbenante, G.; Becker, B.; Blanc, S.; Clark, C.; Condie, G.; Fraser, G.; Grathwohl, M.; Halliday, J.; Henderson, S.; Lam, A.; Liu, L.; Mann, M.; Muldoon, C.; Pearson, A.; Premraj, R.; Ramsdale, T.; Rossetti, T.; Schafer, K.; Le, T. G.; Tometzki, G.; Vari, F.; Verquin, G.; Waanders, J.; West, M.; Wimmer, N.; Yau, A.; Zuegg, J.; Meutermans, W. Biological diversity from a structurally diverse library: Systematically scanning conformational space using a pyranose scaffold. *J. Med. Chem.* **2010**, *53*, 5576–5586.

- (18) Meutermans, W. D.; Bourne, G. T.; Golding, S. W.; Horton, D. A.; Campitelli, M. R.; Craik, D.; Scanlon, M.; Smythe, M. L. Difficult macrocyclizations: New strategies for synthesizing highly strained cyclic tetrapeptides. *Org. Lett.* **2003**, *5*, 2711–2714.
- (19) Becker, B.; Condie, G. C.; Le, G. T.; Meutermans, W. Carbohydrate-based scaffolds in drug discovery. *Mini-Rev. Med. Chem.* **2006**, *6*, 1299–1309.
- (20) Meutermans, W.; Le, G. T.; Becker, B. Carbohydrates as scaffolds in drug discovery. *ChemMedChem*. **2006**, *1*, 1164–1194.
- (21) Halliday, J.; McKeveney, D.; Muldoon, C.; Rajaratnam, P.; Meutermans, W. Targeting the forgotten transglycosylases. *Biochem. Pharmacol.* **2006**, *71*, 957–967.
- (22) Tran, T. T.; McKie, J.; Meutermans, W. D.; Bourne, G. T.; Andrews, P. R.; Smythe, M. L. Topological side-chain classification of beta-turns: Ideal motifs for peptidomimetic development. *J. Comput.-Aided Mol. Des.* **2005**, *19*, 551–566.
- (23) Le, G. T.; Abbenante, G.; Becker, B.; Grathwohl, M.; Halliday, J.; Tometzki, G.; Zuegg, J.; Meutermans, W. Molecular diversity through sugar scaffolds. *Drug Discovery Today* **2003**, *8*, 701–709.
- (24) Gruner, S. A.; Locardi, E.; Lohof, E.; Kessler, H. Carbohydrate-based mimetics in drug design: Sugar amino acids and carbohydrate scaffolds. *Chem. Rev.* **2002**, *102*, 491–514.
- (25) Gruner, S. A.; Keri, G.; Schwab, R.; Venetianer, A.; Kessler, H. Sugar amino acid containing somatostatin analogues that induce apoptosis in both drug-sensitive and multidrug-resistant tumor cells. *Org. Lett.* **2001**, *3*, 3723–3725.
- (26) Beinborn, M. Class B GPCRs: A hidden agonist within? *Mol. Pharmacol.* **2006**, *70*, 1–4.
- (27) Heifetz, A.; Morris, G. B.; Biggin, P. C.; Barker, O.; Fryatt, T.; Bentley, J.; Hallett, D.; Manikowski, D.; Pal, S.; Reifegerste, R.; Slack, M.; Law, R. Study of human orexin-1 and -2 G-protein-coupled receptors with novel and published antagonists by modeling, molecular dynamics simulations, and site-directed mutagenesis. *Biochemistry* **2012**, *51*, 3178–3197.
- (28) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. Halogen bonds in biological molecules. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16789–16794.
- (29) Yumi, N. I.; Yoshihisa, I.; Isao, N.; Kazuo, K. Cl-pi interactions in protein–ligand complexes. *Protein Sci.* **2008**, *17*, 1129–1137.
- (30) Anand, S.; Anbarasu, A.; Sethumadhavan, R. Non-classical hydrogen bonds in interleukins: The role of C-H...O interactions. *Int. J. Biol. Macromol.* **2008**, *43*, 468–473.
- (31) Thanh, G. L.; Abbenante, G.; Adamson, G.; Becker, B.; Clark, C.; Condie, G.; Falzun, T.; Grathwohl, M.; Gupta, P.; Hanson, M.; Huynh, N.; Katavic, P.; Kuipers, K.; Lam, A.; Liu, L.; Mann, M.; Mason, J.; McKeveney, D.; Muldoon, C.; Pearson, A.; Rajaratnam, P.; Ryan, S.; Tometzki, G.; Verquin, G.; Waanders, J.; West, M.; Wilcox, N.; Wimmer, N.; Yau, A.; Zuegg, J.; Meutermans, W. A versatile synthetic approach toward diversity libraries using monosaccharide scaffolds. *J. Org. Chem.* **2010**, *75*, 197–203.
- (32) Carroll, T. X.; Thomas, T. D.; Bergersen, H.; Borge, K. J.; Saethre, L. J. Fluorine as a pi donor. Carbon 1s photoelectron spectroscopy and proton affinities of fluorobenzenes. *J. Org. Chem.* **2006**, *71*, 1961–1968.
- (33) Chaki, S.; Funakoshi, T.; Hirota-Okuno, S.; Nishiguchi, M.; Shimazaki, T.; Iijima, M.; Grottick, A. J.; Kanuma, K.; Omodera, K.; Sekiguchi, Y.; Okuyama, S.; Tran, T. A.; Semple, G.; Thomsen, W. Anxiolytic- and antidepressant-like profile of ATC0065 and ATC0175: Nonpeptidic and orally active melanin-concentrating hormone receptor 1 antagonists. *J. Pharmacol. Exp. Ther.* **2005**, *313*, 831–839.
- (34) Shoichet, B. K.; Kobilka, B. K. Structure-based drug screening for G-protein-coupled receptors. *Trends Pharmacol. Sci.* **2012**, *33*, 268–272.
- (35) Congreve, M.; Langmead, C.; Marshall, F. H. The use of GPCR structures in drug design. *Adv. Pharmacol.* **2011**, *62*, 1–36.
- (36) Congreve, M.; Marshall, F. The impact of GPCR structures on pharmacology and structure-based drug design. *Br. J. Pharmacol.* **2010**, *159*, 986–996.
- (37) Congreve, M.; Langmead, C. J.; Mason, J. S.; Marshall, F. H. Progress in structure based drug design for G protein-coupled receptors. *J. Med. Chem.* **2011**, *54*, 4283–4311.
- (38) Carlsson, J.; Coleman, R. G.; Setola, V.; Irwin, J. J.; Fan, H.; Schlessinger, A.; Sali, A.; Roth, B. L.; Shoichet, B. K. Ligand discovery from a dopamine D3 receptor homology model and crystal structure. *Nat. Chem Biol.* **2011**, *7*, 769–778.
- (39) Ballesteros, J. A.; Weinstein, H. Integrated methods for construction three dimensional models and computational probing of structure–function relations in G protein-coupled receptors. *Methods Neurosci.* **1995**, *25*, 366–428.
- (40) Needleman, S. B.; Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **1970**, *48*, 443–453.
- (41) Styczynski, M. P.; Jensen, K. L.; Rigoutsos, I.; Stephanopoulos, G. BLOSUM62 miscalculations improve search performance. *Nat. Biotechnol.* **2008**, *26*, 274–275.
- (42) Fanelli, F.; De Benedetti, P. G. Computational modeling approaches to structure–function analysis of G protein-coupled receptors. *Chem. Rev.* **2005**, *105*, 3297–3351.
- (43) Baldwin, J. M. The probable arrangement of the helices in G protein-coupled receptors. *EMBO J.* **1993**, *12*, 1693–1703.
- (44) Cobanoglu, M. C.; Saygin, Y.; Sezerman, U. Classification of GPCRs using family specific motifs. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2011**, *8*, 1495–1508.
- (45) Gloriam, D. E.; Foord, S. M.; Blaney, F. E.; Garland, S. L. Definition of the G protein-coupled receptor transmembrane bundle binding pocket and calculation of receptor similarities for drug design. *J. Med. Chem.* **2009**, *52*, 4429–4442.
- (46) Gao, Q. B.; Wang, Z. Z. Classification of G-protein coupled receptors at four levels. *Protein Eng., Des. Sel.* **2006**, *19*, 511–516.
- (47) Wistrand, M.; Kall, L.; Sonnhammer, E. L. A general model of G protein-coupled receptor sequences and its application to detect remote homologs. *Protein Sci.* **2006**, *15*, 509–521.
- (48) Bettinelli, I.; Graziani, D.; Marconi, C.; Pedretti, A.; Vistoli, G. The approach of conformational chimeras to model the role of proline-containing helices on GPCR mobility: the fertile case of Cys-LTR1. *ChemMedChem*. **2011**, *6*, 1217–1227.
- (49) Langelaan, D. N.; Wiczorek, M.; Blouin, C.; Rainey, J. K. Improved helix and kink characterization in membrane proteins allows evaluation of kink sequence predictors. *J. Chem. Inf. Model.* **2010**, *50*, 2213–2220.
- (50) Hall, S. E.; Roberts, K.; Vaidehi, N. Position of helical kinks in membrane protein crystal structures and the accuracy of computational prediction. *J. Mol. Graph. Model.* **2009**, *27*, 944–950.
- (51) Yohannan, S.; Faham, S.; Yang, D.; Whitelegge, J. P.; Bowie, J. U. The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 959–963.
- (52) Fechteler, T.; Dengler, U.; Schomburg, D. Prediction of protein three-dimensional structures in insertion and deletion regions: a procedure for searching data bases of representative protein fragments using geometric scoring criteria. *J. Mol. Biol.* **1995**, *253*, 114–131.
- (53) Jorgensen, W.; Maxwell, L. D. S.; Tirado, R. J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (54) Kaminski, G. A.; Friesner, J.; Tirado, R. J.; Jorgensen, W. Evaluation and reparameterization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B.* **2001**, *105*, 6474–6487.
- (55) Courcot, B.; Bridgeman, A. J. Modeling the interactions between polyoxometalates and their environment. *J. Comput. Chem.* **2011**, *32*, 3143–3153.
- (56) Liebeschuetz, J. W.; Cole, J. C.; Korb, O. Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 737–748.
- (57) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak,

P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897.

(58) Homeyer, N.; Gohlke, H. FEW: A workflow tool for free energy calculations of ligand binding. *J. Comput. Chem.* **2013**, *34*, 965–973.

(59) Hoffer, L.; Horvath, D. S4MPLE—Sampler for multiple protein–ligand entities: Simultaneous docking of several entities. *J. Chem Inf. Model.* **2013**, *53*, 88–102.

(60) Labute, P. LowModeMD: Implicit low-mode velocity filtering applied to conformational search of macrocycles and protein loops. *J. Chem Inf. Model.* **2010**, *50*, 792–800.

(61) Berendsen, H. J. C.; Postma, J. P. M.; Van-Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(62) Tiikkainen, P.; Franke, L. Analysis of commercial and public bioactivity databases. *J. Chem Inf. Model.* **2012**, *52*, 319–326.

(63) Brewer, M. L. Development of a spectral clustering method for the analysis of molecular data sets. *J. Chem Inf. Model.* **2007**, *47*, 1727–1733.

(64) Ma, C.; Wang, L.; Xie, X. Q. GPU accelerated chemical similarity calculation for compound library comparison. *J. Chem Inf. Model.* **2011**, *51*, 1521–1527.