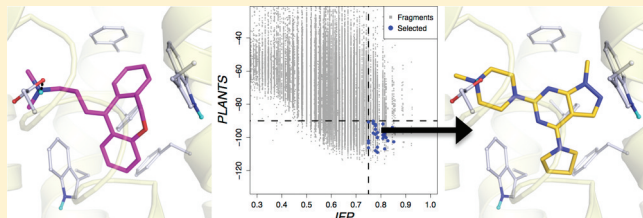


Crystal Structure-Based Virtual Screening for Fragment-like Ligands of the Human Histamine H₁ ReceptorChris de Graaf,^{†,●} Albert J. Kooistra,^{†,●} Henry F. Vischer,[†] Vsevolod Katritch,[‡] Martien Kuijer,[†] Mitsunori Shiroishi,^{§,||} So Iwata,^{§,⊥, #, ×} Tatsuro Shimamura,^{§, ×} Raymond C. Stevens,[‡] Iwan J. P. de Esch,[†] and Rob Leurs^{*,†}[†]Leiden/Amsterdam Center for Drug Research (LACDR), Division of Medicinal Chemistry, Faculty of Science, VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands[‡]Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, GAC-1200, La Jolla, California 92037, United States[§]Human Receptor Crystallography Project, ERATO, Japan Science and Technology Agency, Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8501, Japan^{||}Graduate School of Pharmaceutical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan[⊥]Division of Molecular Biosciences, Membrane Protein Crystallography Group, Imperial College, London SW7 2AZ, U.K.[#]Diamond Light Source, Harwell Science and Innovation Campus, Chilton, Didcot, Oxfordshire OX11 0DE, U.K.[×]Department of Cell Biology, Graduate School of Medicine, Kyoto University, Yoshidakonoe-cho, Sakyo-Ku, Kyoto 606-8501, Japan

S Supporting Information

ABSTRACT: The recent crystal structure determinations of druggable class A G protein-coupled receptors (GPCRs) have opened up excellent opportunities in structure-based ligand discovery for this pharmaceutically important protein family. We have developed and validated a customized structure-based virtual fragment screening protocol against the recently determined human histamine H₁ receptor (H₁R) crystal structure. The method combines molecular docking simulations with a protein–ligand interaction fingerprint (IFP) scoring method. The optimized in silico screening approach was successfully applied to identify a chemically diverse set of novel fragment-like (≤ 22 heavy atoms) H₁R ligands with an exceptionally high hit rate of 73%. Of the 26 tested fragments, 19 compounds had affinities ranging from 10 μ M to 6 nM. The current study shows the potential of in silico screening against GPCR crystal structures to explore novel, fragment-like GPCR ligand space.



■ INTRODUCTION

G protein-coupled receptors (GPCRs) comprise the largest family of transmembrane proteins, mediating the intracellular signals of a wide array of signaling molecules and playing an essential role in numerous cellular and physiological effects.¹ Knowledge of the three-dimensional structure of GPCRs provides valuable insights into receptor function and receptor–ligand interactions^{2,3} and is key for the in silico discovery of new bioactive molecules that can target this family of pharmaceutically relevant drug targets.^{4,5} The human histamine H₁ receptor (hH₁R) is a key player in allergic responses, and so-called “antihistamines” are widely used to relieve the symptoms of allergic rhinitis by inhibiting the constitutive activity of H₁R, as well as antagonizing histamine binding to the H₁R.⁶ Recently the first inverse agonist bound H₁R crystal structure was solved,⁷ opening up new possibilities in the rational design and in silico discovery of novel H₁R ligands. GPCR homology models have already been used successfully to identify new ligands,^{5,8–19} but the increasing

number of GPCR X-ray structures solved in the past few years²⁰ offers unique opportunities to push the limits of structure-based virtual screening (SBVS).^{4,21–24} With more and detailed structural information, one should be able to increase hit rates of SBVS campaigns and to specifically apply the in silico approach in the field of fragment-based drug discovery (FBDD).^{25,26} FBDD is a new paradigm in drug discovery that utilizes small molecules (≤ 22 heavy atoms) as starting points for efficient hit optimization.²⁵ While previous GPCR SBVS campaigns have mainly identified larger molecules,⁵ the aim of the current study was to overcome the challenges of structure-based virtual fragment screening, the in silico discovery of smaller, fragment-like molecules, based on the recently elucidated crystal structure of the H₁R. Although more than 70% of H₁R ligands have a heavy atom count higher than 22 (Figure 1A), doxepin, the cocrystallized high affinity inverse

Received: August 31, 2011

Published: October 18, 2011

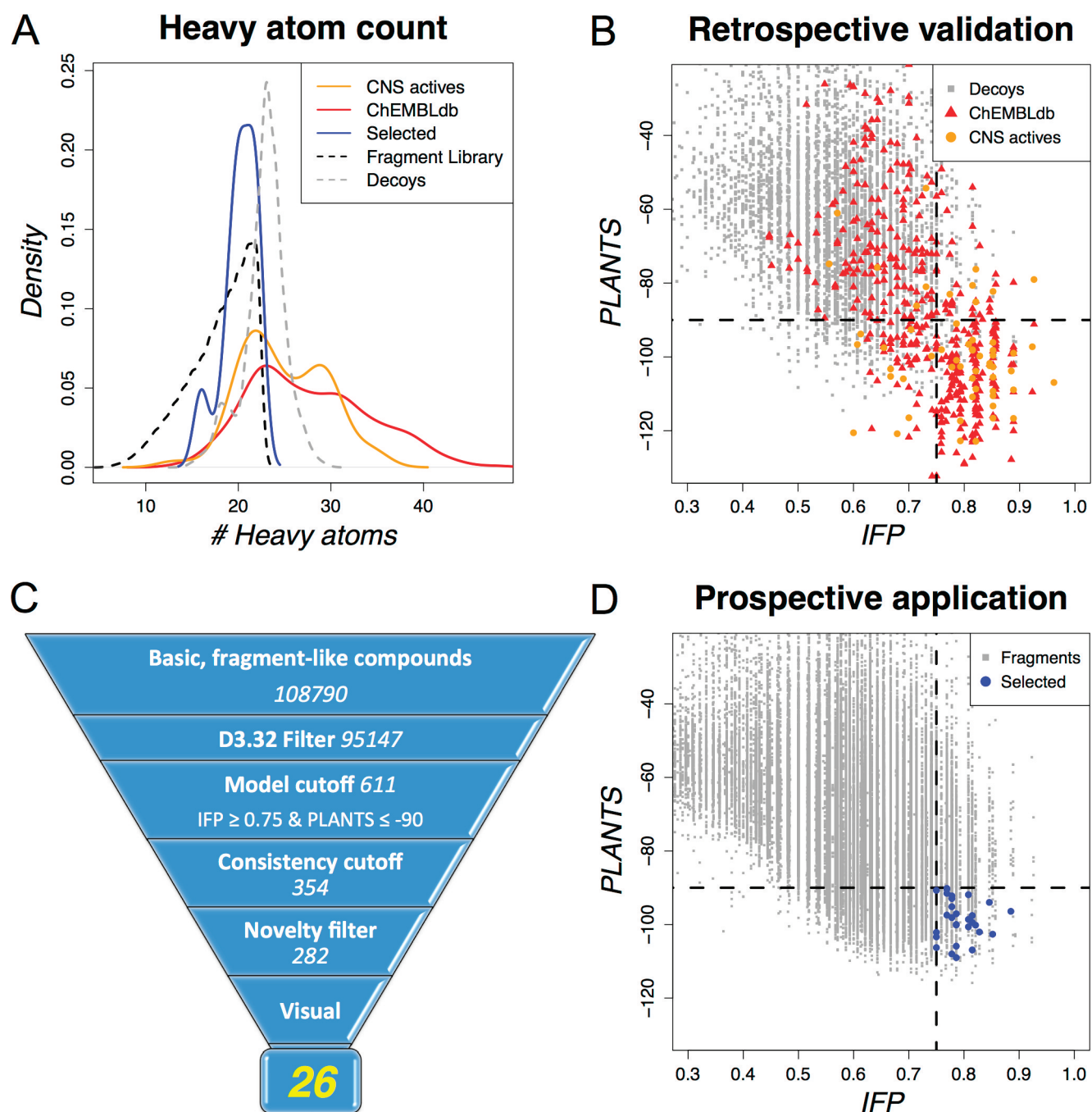


Figure 1. Four-panel overview of the preparation, validation, selection, and virtual screening process. (A) Distribution of the heavy atom count for known H_1R ligands (ChEMBLdb (red) and CNS active drugs³⁰ (orange)), decoys used for retrospective validation (gray), fragment-like compounds from ZINC used for prospective virtual screening (black), and in silico hits selected by our structure-based virtual screening method (blue) is shown. (B) Scatter plot of PLANTS scores versus IFP scores for known actives from the ChEMBLdb (red) and CNS active drugs³⁰ (orange) and physicochemically similar decoys (gray). (C) Overview of the structure-based virtual screening postprocessing steps of 108 790 fragment-like, basic compounds, which resulted in final selection of 26 fragment-like compounds: D^{3.32} filter, docking poses making an ionic interaction with D107^{3,32}; model cutoff, compounds for which docking poses are generated with IFP $T_c \geq 0.75$ and PLANTS ≤ -90 (not necessarily the same pose); consistency cutoff, only compounds with an IFP score of ≥ 0.7 according to the best PLANTS pose as well as a PLANTS score of ≤ -75 according to the best IFP pose were selected; novelty filter, ECFP-4 Tanimoto similarity of <0.40 to any known H_1R ligand; visual inspection, close analogues with highest IFP score are kept and compounds for which buried polar groups are placed in hydrophobic parts of the binding site in all filtered docking poses are discarded). The number indicates the number of compounds present at each step. (D) Scatter plot of the PLANTS scores versus the IFP scores for the fragment screening data set (gray) with the selected compounds (blue). The dotted lines in parts B and D indicate the selected model cutoffs (IFP $T_c \geq 0.75$ and PLANTS ≤ -90).

agonist in the H_1R X-ray structure,⁷ can be considered as a large fragment-like compound containing 21 heavy atoms.²⁵ We have developed and validated a target-customized, docking-based

virtual screening protocol that combines molecular docking with a protein–ligand interaction scoring method.²⁷ This optimized SBVS method was successfully applied to identify

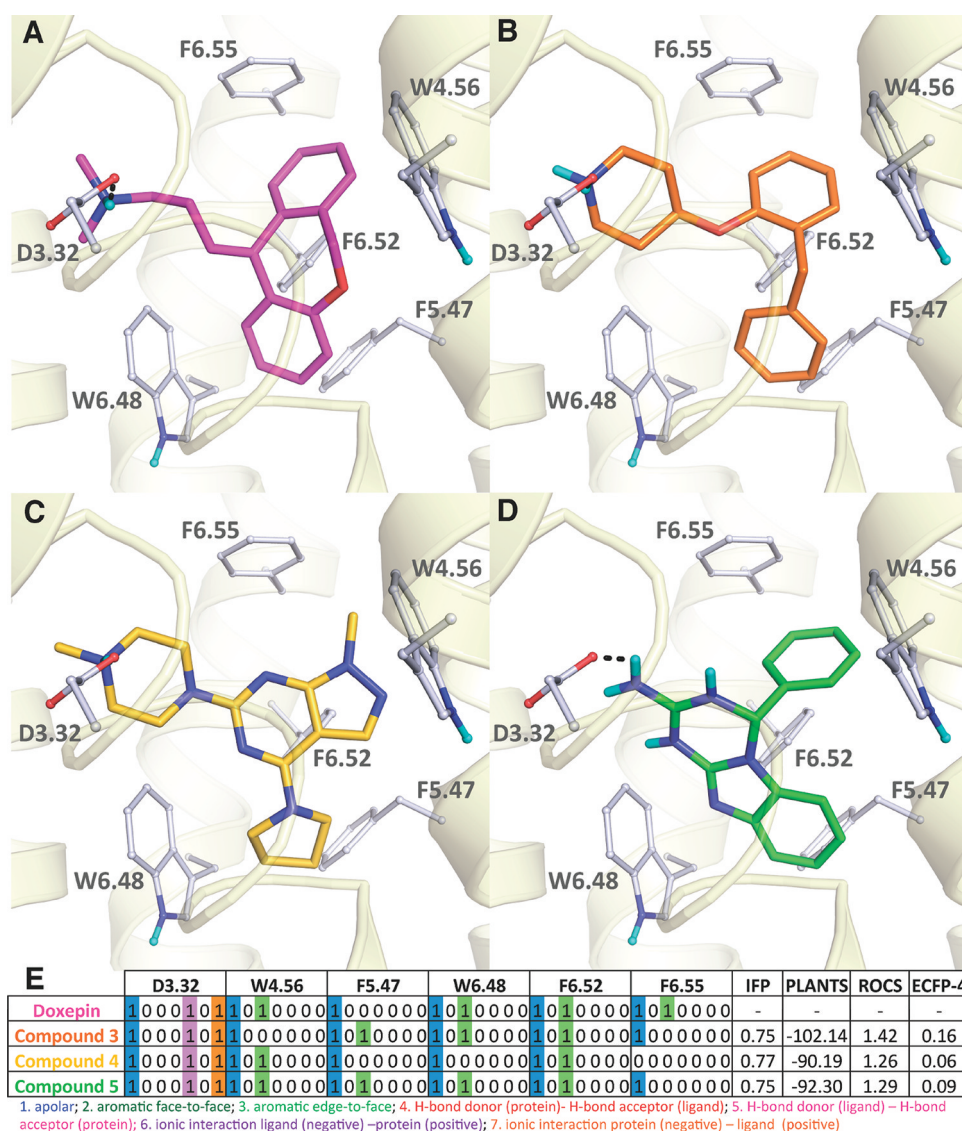


Figure 2. Binding pose of doxepin (magenta carbon atoms, A) in the H₁R structure (PDB code 3RZE) and the predicted binding poses of the new fragment-like H₁R ligands identified by prospective structure-based virtual screening: 3 (orange, B), 4 (gold, C), and 5 (green, D). The IFPs corresponding to the compounds in the displayed pose are partially presented (E). Parts of the backbone of transmembrane (TM) helices 3, 4, 5, 6, and 7 are represented by transparent light yellow ribbons. Important binding residues are depicted as ball-and-sticks with gray carbon atoms. Oxygen, nitrogen, and hydrogen atoms are colored red, blue, and cyan, respectively. H-Bonds described in the text are depicted by black dots. The IFP bit strings of the docking poses of the hits 3–5 (B–D) are compared to the reference IFP of doxepin 1 (A) in panel E, encoding different interaction types with each residue in the binding site. For reasons of clarity, the bit strings of only 6 residues (out of 33) are shown as an example.

fragment-like H₁R ligands with an exceptionally high hit rate. The current study shows the potential of *in silico* screening against GPCR crystal structures to explore novel fragment-like ligand space and to investigate the fine atomic details of molecular recognition by this pharmaceutically relevant family of protein targets.

RESULTS AND DISCUSSION

Development and Validation of a Customized Structure-Based Virtual Screening Approach. We performed docking experiments with PLANTS,²⁸ using the H₁R crystal structure on 543 known H₁R ligands from the ChEMBL database²⁹ ($K_i \leq 10 \mu\text{M}$) and 59 CNS active drugs acting as inverse agonists on H₁R,³⁰ as well as 7088 decoys with physicochemical properties similar to the ChEMBLdb actives (Figure 1A, Supporting Information Table 1). PLANTS

combines an ant colony optimization algorithm with an empirical scoring function³¹ for the prediction and scoring of binding poses in a protein structure. The resulting docking poses were postprocessed using molecular interaction fingerprints (IFPs).³² The IFP scoring method determines ligand binding mode similarity to experimentally supported ligand binding poses (Figure 2). IFPs have been used as an efficient alternative postprocessing method of docking poses^{27,32} to overcome target dependent scoring problems.³³ Seven different interaction types (negatively charged, positively charged, H-bond acceptor, H-bond donor, aromatic face-to-edge, aromatic face-to-face, and hydrophobic interactions) were used to define the IFP. A Tanimoto coefficient (Tc) measuring IFP similarity with the reference IFP pose in the H₁R crystal structure (Figure 2) was used to score the docking poses of known actives and decoys.³² The scatter plot in Figure 1B shows that

Table 1. Retrospective Virtual Screening Accuracy of Different SBVFS Scoring Methods^a

	combination IFP (≥ 0.75) and PLANTS (≤ -90) filters		IFP (≥ 0.75) filter only		PLANTS (≤ -90) filter only	
	ChEMBLdb	CNS drugs	ChEMBLdb	CNS drugs	ChEMBLdb	CNS drugs
actives (%) ^b	39.3	57.6	48.0	69.5	50.7	76.3
decoys (%) ^c	1.0	1.0	3.4	3.4	6.4	6.4
enrichment ^d	39.3	57.6	14.1	20.4	7.9	11.9

^aPercentage of known H₁R ligands (from ChEMBLdb and CNS drugs³⁰ sets) and decoys in the regions defined by IFP-Tc (≥ 0.75), PLANTS docking score (≤ -90), and combined IFP-Tc and PLANTS docking score cutoffs (Figure 1B,D). ^bActives in the ChEMBLdb set are defined by $pK_i \geq 5$ and in the CNS actives set by $pIC_{50} \geq 5$. ^cThe decoy set consists of ~ 7000 compounds, selected based on the physicochemical properties of the ChEMBLdb data set. ^dThe enrichment reported here is the percentage of retrieved true positives (TP) divided by the percentage of retrieved false positives, TP(%) / FP(%), at the specified cutoffs.

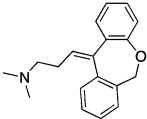
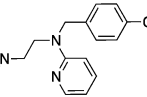
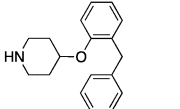
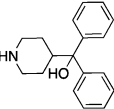
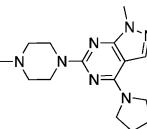
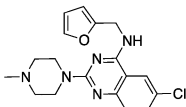
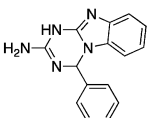
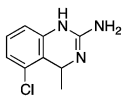
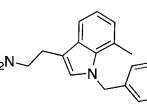
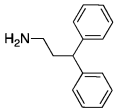
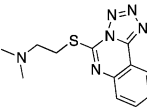
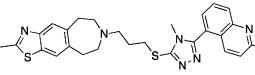
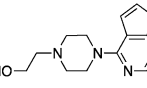
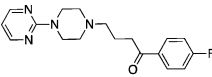
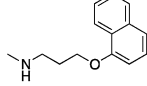
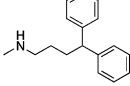
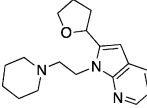
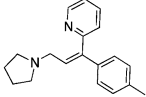
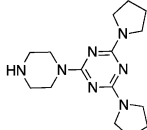
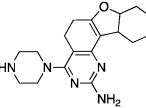
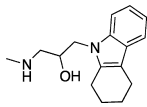
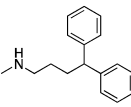
active compounds can be discriminated from decoys by considering both the best IFP score and best PLANTS score for each compound, and the Kernel density plot for the ChEMBLdb actives in Supporting Information Figure 2 clearly demonstrates that most actives can obtain binding modes in the H₁R binding site that (i) are similar to the binding mode of doxepin (indicated by high IFP Tanimoto similarity scores) and (ii) are energetically favorable (high (negative) PLANTS docking scores). On the basis of this analysis, we determined IFP (Tc ≥ 0.75) and PLANTS (Total_Score ≤ -90) cutoffs to discriminate H₁R ligands from decoys (Table 1). This combination of binding mode similarity and docking scores resulted in high retrospective virtual screening enrichments³⁴ of known H₁R ligands from two independent test sets, the ChEMBLdb (39-fold enrichment of over random picking) and CNS drugs (58-fold enrichment) test sets over decoys with similar physicochemical properties (Table 1).

Structure-Based Identification of Fragment-like H₁R Ligands with Exceptionally High Virtual Screening Hit Rates. The validated SBVFS protocol was subsequently used to identify new fragment-like H₁R ligands from a subset of 108 790 molecules extracted from 13 million commercially available compounds (Figure 1C) stored in the ZINC database³⁵ that (i) obey fragment rules based on rule-of-three^{36,37} (number of heavy atoms, ≤ 22 ; log $P < 3$; number of H-bond donors, ≤ 3 ; number of H-bond acceptors, ≤ 3 ; number of rotatable bonds, ≤ 5 ; number of rings, ≥ 1) and (ii) contain a basic moiety (to enable ionic interactions with the essential D107^{3,32} residue³⁸ in the H₁R binding pocket (Figure 2)). The docking poses of 354 fragments were capable of making an ionic interaction with D107^{3,32} (ref 38) and complied with the optimized IFP and PLANTS score cutoffs as well as the consistency cutoffs (see Methods, Figure 1C, and Supporting Information Figure S3). Interestingly, this set contained nine compounds with known activity on the H₁ receptor ($K_i \leq 10 \mu\text{M}$ in ChEMBLdb²⁹), of which seven compounds are FDA approved drugs (epinastine, mianserin, triprolidine, promazine, amoxapine, imipramine, and desipramine; see Supporting Information Table 2). In fact, 282 of the 354 compounds were chemically *dissimilar* to any known H₁R ligand (ECFP-4 Tanimoto similarity of < 0.40 ³⁹). This set of fragments was visually clustered, and for each cluster the fragment with highest IFP and/or PLANTS score was selected. Fragments for which buried polar groups were placed in hydrophobic parts of the H₁R binding site in all filtered docking poses were discarded based on visual inspection. This resulted in a final selection of 30 compounds, of which 26 were actually available and experimentally tested for their binding affinity to H₁R (Figure 1C). Of the 26 experimentally tested compounds 19 had affinities ranging from $10 \mu\text{M}$ to 6 nM for H₁R

(Figure 2, Table 2). Seven fragments have submicromolar affinity (3–9, Figure 3A,B), and fragment 3 ($K_i = 6 \text{ nM}$) has one of the highest affinities reported for a GPCR ligand identified by structure-based virtual screening.^{12,40} Nine of the 19 H₁R binders (3–8, 10, 13, and 15) were characterized as inverse agonists in histamine stimulated inositol phosphate (InsP) accumulation assays, while fragment 9 was determined to be a very weak partial agonist (intrinsic activity of 0.07). When tested against histamine ($0.1 \mu\text{M}$), doxepin (1), mepyramine (2), and fragment 3 were able to completely inhibit the histamine-induced effects. Also, the other fragment hits with submicromolar affinities were able to interfere with the agonist response (Supporting Information Figure S6). Similar to the doxepin binding mode in the H₁R crystal structure⁷ (Figure 2A), the predicted docking poses of the new inverse agonists in H₁R (Figure 2B–D) make extensive hydrophobic interactions with W^{6,48}, a highly conserved key residue in GPCR activation.⁴¹ This observation is in line with the hypothesis that inverse agonists of H₁R should be able to lock W^{6,48} in an inactive conformation to reduce H₁R basal activity.^{7,42} In light of the recently solved agonist bound GPCR crystal structures, it is interesting to discover a partial agonist with an inverse agonist bound H₁R structure. However, it is emphasized that the efficacy of 9 is very weak compared to histamine. Moreover, GPCR structure-based virtual screening studies have shown that agonists can be retrieved by using receptor models based on inactive crystal structure templates,^{13,14,18} and vice versa, antagonists have been found using agonist-biased receptor models.^{16,17} Selective virtual screening for ligands with a specific function (e.g., inverse agonist/antagonist, partial/full agonist) probably requires a customized modeling protocol, as recently reported in retrospective virtual screening studies against modeled agonist bound conformations of ADRB2^{27,43} (prior to the release of the first agonist bound ADRB2 crystal structure).⁴⁴

The hit rate of 73% (19 out of 26 tested compounds) of our SBVS study is exceptionally high compared to other prospective studies.⁴⁵ In fact, to the best of our knowledge, this is the highest hit rate reported for any prospective SBVS campaign reported for GPCRs⁵ as shown in Figure 4. It should be noted that our *in silico* screening approach combines a docking scoring function with a molecular interaction fingerprint scoring method using knowledge of the cocrystallized ligand binding mode, while many of the other recent SBVS runs against GPCR crystal structures used only energy-based scoring functions to rank the docking poses.^{21–23} However, like in our study, all these virtual screening exercises also included additional selection criteria (e.g., complementarity to the binding site, clustering/novelty, polarity) to select $\sim 10\%$ of the top hit list. Interestingly, not only high hit rates have been obtained based on docking studies against GPCR crystal structures but also successful SBVS studies with high hit rates ($> 20\%$) have been reported based

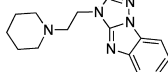
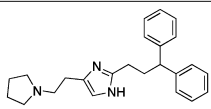
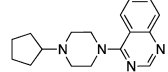
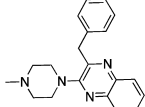
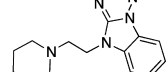
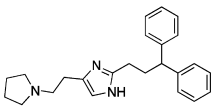
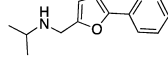
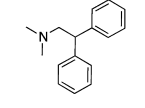
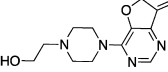
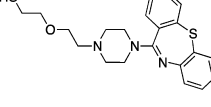
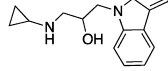
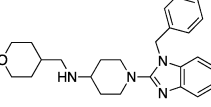
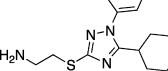
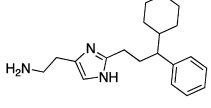
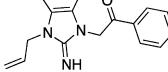
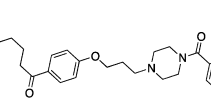
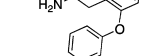
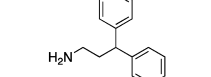
Table 2. Histamine H₁ Receptor Binding Affinities of Validated Fragment-like Hits ($K_i \leq 10 \mu\text{M}$) Selected by an Optimized Structure-Based Virtual Screening Protocol (Figure 2) against the H₁R Crystal Structureⁱ

cpd		pK _i H ₁ R ^{a,b}	IFP (rank) ^c	PLANTS (rank) ^d	ROCS _{DOX} (rank) ^e	ECFP-4 _{DOX} (rank) ^f	ECFP ₄ ^g	closest known H ₁ R ligand ^f
1		9.75 ± 0.14	1.00 (1) ^h	-115.92 (1) ^h	-	-	-	-
2		8.68 ± 0.03	0.85 (31) ^h	-106.39 (106) ^h	-	-	-	-
3		8.20 ± 0.10	0.75 (1836)	-102.14 (429)	1.416 (146)	0.16 (8677)	0.34	
4		7.21 ± 0.03	0.77 (1326)	-90.19 (6205)	1.263 (3601)	0.06 (102773)	0.34	
5		6.37 ± 0.06	0.75 (1926)	-92.3 (4270)	1.287 (2819)	0.09 (68145)	0.25	
6		6.27 ± 0.07	0.79 (500)	-109.00 (31)	1.305 (2240)	0.11 (45055)	0.23	
7		6.15 ± 0.08	0.78 (754)	-92.21 (4346)	1.379 (699)	0.21 (1206)	0.28	
8		6.15 ± 0.10	0.77 (1319)	-90.44 (5944)	1.406 (399)	0.07 (17047)	0.32	
9		6.10 ± 0.28	0.77 (1229)	-97.43 (1416)	1.374 (707)	0.14 (17047)	0.38	
10		5.75 ± 0.04	0.81 (312)	-98.61 (1095)	1.311 (2061)	0.13 (27979)	0.25	
11		5.72 ± 0.09	0.79 (518)	-100.13 (757)	1.125 (17913)	0.04 (108215)	0.27	
12		5.64 ± 0.03	0.77 (1294)	-91.51 (4934)	1.402 (425)	0.14 (16574)	0.24	

on (bovine rhodopsin based) GPCR homology models.^{8,12,18,46} Moreover, in a recent comparative virtual screening, comparable high hit rates were obtained for prospective virtual screening runs against the recently solved dopamine D3 receptor (DRD3) crystal structure and a previously constructed DRD3 homology model.⁴⁶

Our analysis furthermore shows that of 16 previously published GPCR SBVS campaigns,^{8–19,21–23,46} our experimentally validated hit set also has the lowest average heavy count of all GPCR SBVS studies (Figure 4). This makes the high virtual screening hit rate even more remarkable because of the typically low-affinity

Table 2. continued

cpd		pK _i H ₁ R ^{a,b}	IFP (rank) ^c	PLANTS (rank) ^d	ROCS _{DOX} (rank) ^e	ECFP-4 _{DOX} (rank) ^f	ECFP ₄ ^g	closest known H ₁ R ligand ^f
13		5.58 ± 0.14	0.78 (689)	-98.13 (1240)	1.178 (10007)	0.13 (29867)	0.32	
14		5.49 ± 0.04	0.82 (183)	-97.56 (1419)	1.356 (972)	0.11 (47280)	0.30	
15		5.38 ± 0.03	0.85 (83)	-93.98 (3060)	1.325 (1605)	0.11 (40224)	0.34	
16		5.34 ± 0.14	0.77 (1278)	-92.53 (4082)	1.356 (972)	0.16 (7054)	0.26	
17		5.27 ± 0.04	0.78 (744)	-92.93 (3775)	1.106 (22077)	0.12 (39372)	0.31	
18		5.20 ± 0.13	0.81 (308)	-100.63 (675)	1.296 (2535)	0.16 (9372)	0.28	
19		5.09 ± 0.07	0.83 (115)	-102.04 (446)	1.375 (690)	0.14 (17769)	0.29	
20		4.97 ± 0.10	0.85 (43)	-102.63 (372)	1.301 (2368)	0.15 (11737)	0.24	
21		4.96 ± 0.02	0.78 (719)	-95.17 (2397)	1.016 (47198)	0.13 (11736)	0.34	

^a pK_i values are calculated from at least three independent measurements as the mean ± SEM. ^bMeasured by displacement of [³H]mepyramine binding using membranes of HEK293T cells transiently expressing the human H₁R. ^cIFP Tanimoto similarity with doxepin pose in the H₁R crystal structure. Optimized IFP score cutoff of ≥0.75. IFP ranking is given in parentheses. ^dScore and rank according to PLANTS scoring function.³¹ Optimized PLANTS score cutoff of ≤−90. PLANTS ranking is given in parentheses. ^eROCS shape-based 3D similarity to doxepin based on Comboscore.⁵⁴ ROCS ranking is given in parentheses. ^fECFP-4 2D Tanimoto similarity to doxepin. A similarity higher than 0.40 is considered as significant.³⁹ ECFP-4 ranking is given in parentheses. ^gECFP-4 circular fingerprint Tanimoto similarity to closest known H₁R active in ChEMBLdb. A similarity higher than 0.40 is considered as significant.³⁹ ^hThe rankings indicated for reference compounds 1 and 2 (doxepin and mepyramine) were determined as if they were included in the screening library. ⁱPLANTS docking scores, interaction fingerprint similarities (IFP), and 2D (ECFP-4) and 3D (ROCS) chemical similarities to doxepin (binding mode), as well as closest chemical similarity to any known H₁R ligand are given for each validated hit.

fragment–protein interactions. Apparently the H₁R binding site, combining a strong polar ligand recognition site (e.g., D107^{3,32} (refs 38 and 47)) with a tight hydrophobic binding site surrounded by several aromatic residues (W158^{4,56}/F199^{5,47}/W428^{6,48}/F432^{6,52}/F435^{6,55} (refs 38, 47, 48)), is particularly suitable for the identification of (small) ligands using a customized docking-based virtual screening protocol that combines a knowledge-based scoring function (IFP) to determine ligand binding mode similarity and a docking scoring function (PLANTS) to assess ligand–protein complementarity.

While IFP rankings are in most cases higher than PLANTS score ranking (Table 2), our optimized screening approach enabled the selection of new H₁R ligands from a small subset of fragments (Figure 1D). Moreover, most of the validated hits would not have been selected using only one of the two SBVS techniques applied. The top 300 compounds of the IFP and PLANTS ranking lists (of the 2274 and 6416 fragments passing the IFP and PLANTS filters) contain only four and one of the validated fragments, respectively (see Figure 1C, Table 2, and Supporting Information Figure 3). Compound 4, for example, one of the seven validated hits with submicromolar affinity

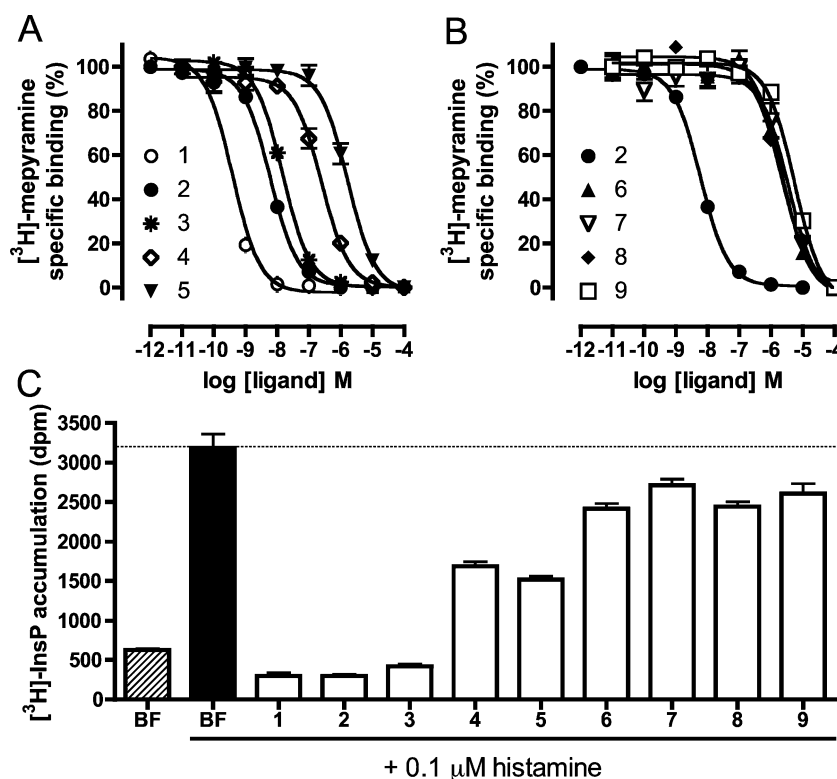


Figure 3. Radioligand displacement by (A, B) and functional effects of (C) compounds 1–9. (A, B) Displacement curves of [³H]mepyramine in HEK293T cells transiently transfected with hH₁R (*n* = 3, each performed in triplicate). (C) Inhibition of histamine-stimulated inositol phosphate accumulation assay in HEK293T cells transiently transfected with hH₁R by compounds 1–9 (*n* = 2, each performed in triplicate).

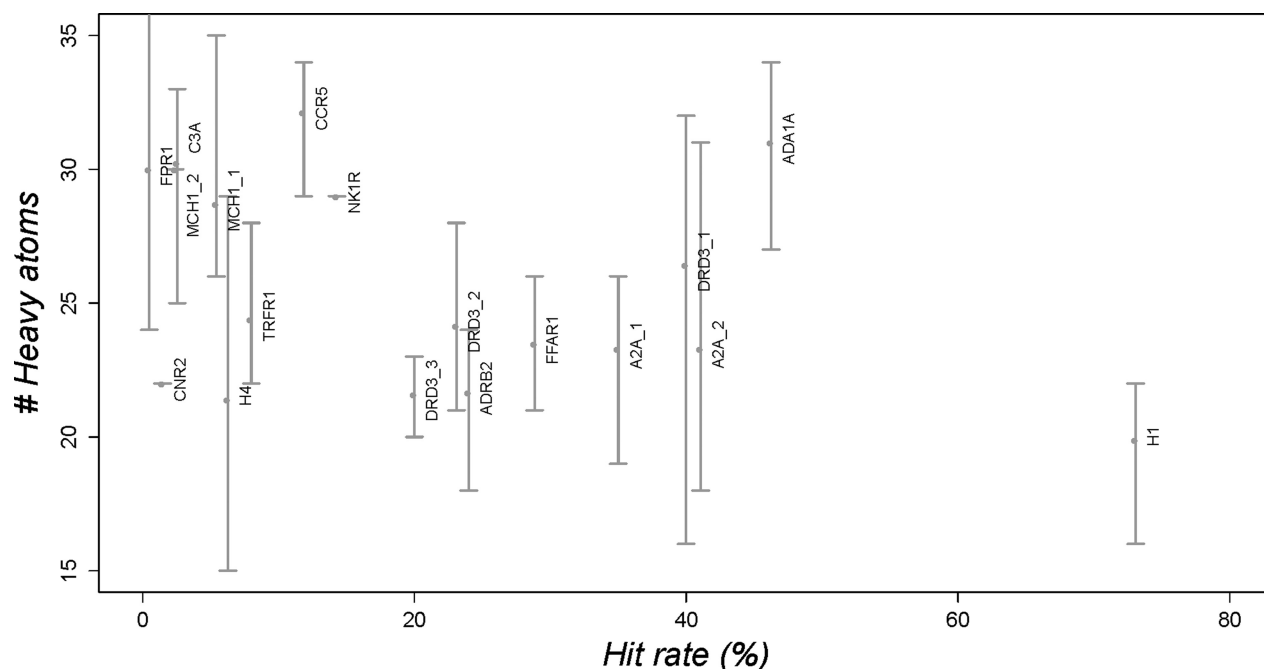


Figure 4. Hit rate and size of hits identified in prospective structure-based virtual screening studies against GPCR crystal structures^{21–23} and homology models.^{8–19} The bars shown for the heavy atom count indicate the minimum and maximum heavy atom count for all hits of each SBVS. The labels indicate the screening on the following receptors adenosine α -2a receptor (A2A_1²² and A2A_2²³), adrenergic α -1a receptor (ADA1A¹⁰), adrenergic β -2 receptor (ADRB2²¹), complement component 3a receptor 1 (C3A¹⁹), C–C chemokine receptor type 5 (CCR5¹⁴), cannabinoid receptor 2 (CNR2¹³), dopamine receptor D3 (DRD3_1⁸ and DRD3_2⁴⁶ homology models and DRD3_3 crystal structure⁴⁶), free fatty acid receptor 1 (FFAR1¹⁸), formyl peptide receptor 1 (FPR1¹¹), histamine receptor H₁ (H1), histamine receptor H₄ (H4¹⁷), melanin-concentrating hormone receptor 1 (MCH1_1¹⁵ and MCH1_2⁹), neurokinin 1 receptor (NK1R¹²), and transferrin receptor 1 (TRFR1¹⁶). The maximum heavy atom count of FPR1 (41) is not shown for clarity purposes. Only hits (i) for which binding affinity ($K_i \leq 15 \mu\text{M}$) or potency ($\text{EC}_{50} \leq 15 \mu\text{M}$) was experimentally determined and (ii) for which a molecular structure was reported are included in the analysis.

for H₁R (K_i = 62 nM), is ranked only 1326th and 6205th in the IFP and PLANTS hits lists, respectively, but is one of the few (354) fragments (corresponding to 0.325% of the initial fragment data set) which passed the combined IFP and PLANTS filter (see Figure 1C and Supporting Information Figure 3). Apparently the combination of binding mode similarity with the experimentally supported H₁R-doxepin pose (determined by IFP) and an energetically favorable H₁R–ligand configuration (assessed by the PLANTS docking score) was required to obtain this high hit rate of novel *in silico* predicted fragment-like H₁R ligands. Further research is needed to determine whether this customized virtual screening approach is generally applicable to other targets as well. Although it is beyond the scope of the current study, it would furthermore be particularly interesting to test whether novel H₁R ligands can be retrieved from alternative hit sets (e.g., hits only selected by PLANTS or only selected by IFP) and/or the omission of specific virtual screening filters (e.g., essential polar interaction with D^{3.32}). While general energy-based scoring functions have been successfully applied in recent virtual screening studies against GPCR crystal structures,^{21–23} the recent community wide GPCR DOCK 2010 challenge showed that a customized and experimentally supported modeling methodology can improve the prediction of GPCR–ligand interactions.⁴⁹ Moreover, consensus scoring strategies^{50,51} and target-customized scoring strategies^{5,52} have been successfully applied in structure-based virtual screening exercises (and in GPCR-based *in silico* screening studies^{8–19} in particular).

Interaction Fingerprint (IFP) Based Virtual Screening Explores Novel Fragment-like Ligand Space. Interestingly, 18 out of the 19 experimentally validated hits do not rank within the top 300 (a selection comparable to the 282 fragments we clustered and visually inspected) of 2D topological (ECFP-4)⁵³ or 3D shape-based (ROCS)⁵⁴ similarity searches of the fragment library against doxepin, the cocrystallized ligand in the H₁R X-ray structure (Table 2). A combination of previously defined, minimal ECFP-4 (Tanimoto similarity of ≥ 0.26 ⁵⁵ and ROCS Comboscore score of ≥ 1.20 ⁵⁶) cutoffs yields a large hit list of 1536 compounds (1.5% of the initial fragment data set), containing none of the validated hits (Supporting Information Figure 4). Only the high affinity compound 3 (K_i = 6 nM) is ranked relatively high (146th) in the ROCS list with a Comboscore of 1.416 (but is ranked 8677th in the ECFP-4 list). This indicates that 3 has a similar three-dimensional pharmacophore and can adopt a similar shape as doxepin in H₁R (Figure 2A). This is confirmed by the docking pose presented in Figure 2B. Compound 3, containing a piperidine ring and a linear chain connecting two benzene rings, is chemically dissimilar from doxepin (and any other known H₁R ligand), which combines a linear amine head with a tricyclic ring system (Table 2). Both ligands, however, make the same ionic and aromatic interactions with D107^{3,32} (refs 38 and 47) and W158^{4,56}/F199^{5,47}/W428^{6,48}/F432^{6,52}/F435^{6,55} (refs 38, 47, 48), respectively, by placing their amine group and benzene rings in almost exactly the same locations in the H₁R binding pocket (Figure 2A,B). Another high affinity hit, compound 4 (K_i = 62 nM), combines a basic piperazine group with a pyrazolo[3,4-*d*]pyrimidine ring system and has a low 2D and 3D similarity with doxepin or any other known H₁R ligand (Table 2, Figure 2C). Data mining of the ChEMBLdb indicates that the chemically complex pyrazolo[3,4-*d*]pyrimidine scaffold is included in adenosine receptor⁵⁷ and corticotropin receptor (CRFR1)⁵⁸ ligands but has so far

not been incorporated in ligands of bioaminergic GPCRs. In the docking pose of 4, the pyrazolopyrimidine group and its pyrrolidine substituent mimic the benzene rings of doxepin by binding in the same aromatic cavity between TM helices 3, 4, 5, and 6 (Figure 2C). Chemically complex bicyclic or tricyclic aromatic ring systems that have not yet been included in histamine H₁R ligands are also present in validated virtual screening hits 5 (dihydrobenzoimidazotriazine), 7 (tetrazoloquinazoline), 8 (tienopyrimidine), 9 (pyrolopyridine), 13 (benzoamidazotriazole), 15 (triazoloindazole), and 17 (benzofuropyrimidine) and play the same role in aromatic π stacking with W158^{4,56}/F199^{5,47}/W428^{6,48}/F432^{6,52}/F435^{6,55}. While noncyclic amine groups (6, 7, 9, 12, 16, 18–21), piperidine (3, 10, 13, 15), and piperazine (4, 8, 11, 14, 17) are typical basic groups in bioaminergic GPCR ligands, compound 5 forms an ionic link with the negatively charged carboxylate group of the conserved D^{3.32} residue^{59,60} via a dihydrobenzoimidazotriazineamine group (Figure 2D). This demonstrates the potential of the IFP scoring method to identify novel ligands with alternative chemical groups that allow formation of the same protein–ligand interactions.⁶¹ This scaffold hopping potential of our structure-based virtual fragment screening approach is further emphasized by the fact that none of the hits are chemically similar to any known H₁R ligand (ECFP-4 Tanimoto similarity of < 0.40 ,³⁹ Table 2). Scaffold diversity analysis⁶² indicated that the validated fragment hits cover the complexity vs cyclicity space of known fragment-like H₁R ligands (Figure 5). Many of the hits have a high complexity and

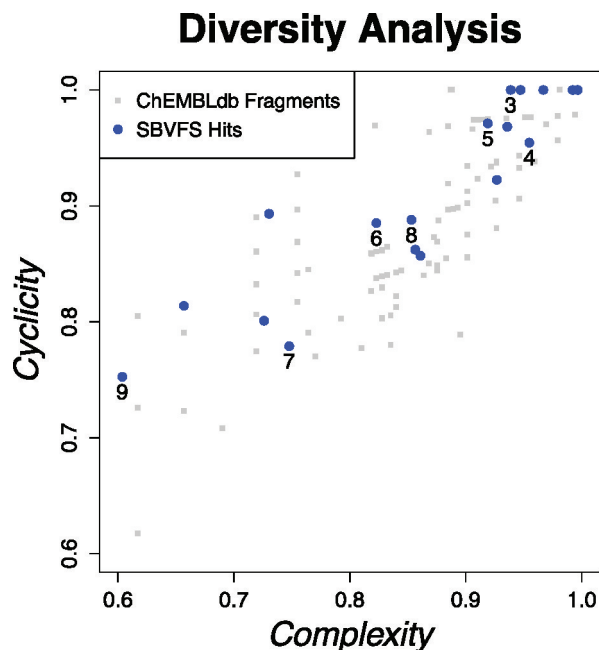


Figure 5. Scatter plot showing the distribution of experimentally validated structure-based virtual screening hits of H₁R (blue dots) in the chemical space covered by previously fragment-like H₁R ligands in the ChEMBLdb. The positions of all submicromolar affinity hits are indicated by the numbers (see Table 2).

cyclicity score, including high affinity hits 3–5, while submicromolar hit 9 has relatively low complexity and cyclicity scores compared to other fragment-like H₁R ligands. Together with the ECFP-4 score reported in Table 2, this clearly indicates that our SBVFS approach is able to successfully retrieve a diverse set of ligands for the H₁R receptor.

Two of the H₁R hits, compounds **13** and **14**, have affinity for H₃R, with K_i values of 0.6 and 0.1 μ M, respectively (Supporting Information Table 3). Two other H₁R hits, **4** and **18**, have medium affinity for H₄R, (K_i values of 2.9 and 2.4 μ M, respectively). On one hand this shows that our structure-based virtual screening protocol yields mainly histamine H₁R subtype selective fragment-like ligands. Many known H₃R ligands contain piperidines and substituted piperazines,⁶³ a scaffold present in H₃R binders **13** and **14**, respectively. Compounds **10** and **15**, however, demonstrate that piperidine ligands (without H₃R affinity) can also bind H₁R selectively. The discovery of fragments **4** and **18** offer new opportunities to develop a dual H₁R–H₄R ligand with synergistic anti-inflammatory properties.⁶⁴ For example, the pyrazolomethyl group of fragment **4** could be used as a handle to grow the ligand toward the extracellular region of the binding site between TMS and TM6 (Figure 2C) to fine-tune H₁R and H₄R affinity by introducing groups matching previously identified interaction hot spots for these receptors (e.g., K191^{5,39} in H₁R, and L175^{5,39} in H₄R).^{47,65,66} Moreover, experimentally determined (or modeled) binding modes of more selective H₁R ligands than doxepin, like the zwitterion ceterizine that is expected to bind to K191^{5,39} in the “anion” binding pocket,^{7,47} might serve as IFP²⁷ references for future structure-based virtual screening studies.

CONCLUSION

We have developed a structure-based virtual fragment screening protocol with which we efficiently identified new fragment-like H₁R ligands with an exceptionally high hit rate of 73%, the highest reported for GPCRs. Many of the identified fragments are both promising and challenging new starting points for structure-based ligand optimization. The current study shows the potential of in silico screening against GPCR crystal structures to explore novel, fragment-like GPCR ligand space.

METHODS

Residue Numbering and Nomenclature. The Ballesteros–Weinstein residue numbering scheme⁶⁷ was used throughout this manuscript. For explicitly numbered residues in specific receptors, the UniProt⁶⁸ residue number is given before the Ballesteros–Weinstein residue number in superscript (e.g., D107^{3,32} in H₁R).

Preparation of Retrospective Validation Databases. For the retrospective validation of our structure-based virtual screening protocol, two independent H₁R ligand test sets were prepared: 543 known active H₁R compounds from the ChEMBL database with $K_i \leq 10$ μ M (test set 1) and 59 CNS drugs with inverse agonist activity on H₁R³⁰ (test set 2). In order to avoid biasing virtual screening results, caution was given to select 7088 decoys from the BioInfo database⁶⁹ covering similar property ranges (molecular weight, number of rotatable bonds, number of rings, hydrogen bond donor/acceptor counts, at least one positively charged atom) as H₁R ligand test set 1 (Supporting Information Table S1). SMILES (available as Supporting Information) data were retrieved from the ChEMBLdb, and plausible tautomers and protonation states were computed for these compounds with ChemAxon's Calculator⁷⁰ and converted into Mol2 format with Molecular Networks' Corina.⁷¹

Preparation of Prospective Virtual Screening Database. From 15 vendors we downloaded their commercial compound data sets in SMILES format from the ZINC Web site (~13 million compounds). With use of Openeye's filter (version 2.1.1),⁷² only fragment-like compounds were selected (757,728 compounds). Plausible tautomers and protonation states were computed for these compounds with Tautothor (version 1.4.90) and Blabber (version 1.4.90), respectively (both part of MolDiscovery's MoKa package⁷³). A second filter was applied to select only compounds with a formal charge of at least +1;

this selection ensures that all selected compounds have the possibility for an ionic bond with key residue D107^{3,32} in the pocket (108 790 compounds).

Automated Docking. All virtual screenings were performed by docking program PLANTS (version 1.1).²⁸ PLANTS combines an ant colony optimization algorithm with an empirical scoring function³¹ for the prediction and scoring of binding poses in a protein structure. For each compound, 25 poses were calculated and scored by the chemplp scoring function at a speed setting 2. The binding pocket of H₁R was defined by the coordinates of the center of cocrystallized doxepin in the 3RZE structure and a radius of 10.8 Å (which is the maximum distance from the center defined by a 5 Å radius around doxepin). All other options of PLANTS were left at their default setting.

IFP Postprocessing. The doxepin binding mode in the original H₁R X-ray structure⁷ was used to generate reference interaction fingerprints (IFPs) as previously described.³² Seven different interaction types (negatively charged, positively charged, H-bond acceptor, H-bond donor, aromatic face-to-edge, aromatic face-to-face, and hydrophobic interactions) were used to define the IFP. The cavity used for the IFP analysis consisted of the same set of 33 residues used in a previous retrospective structure-based virtual screening study²⁶ (30 residues earlier proposed to define a consensus TM binding pocket⁶⁰ plus three additional residues at positions 3.37, 5.47, and 7.40): L^{1.35}, L^{1.39}, I^{1.42}, T^{1.46}, V^{2.57}, M^{2.58}, N^{2.61}, L^{2.65}, W^{3.28}, L^{3.29}, D^{3.32}, Y^{3.33}, S^{3.36}, T^{3.37}, I^{3.40}, W^{4.56}, I^{4.60}, F^{5.38}, K^{5.39}, T^{5.42}, A^{5.43}, N^{5.46}, F^{5.47}, F^{6.44}, W^{6.48}, Y^{6.51}, F^{6.52}, F^{6.55}, H^{7.35}, I^{7.39}, W^{7.40}, Y^{7.43}, N^{7.45}. Note that for each PLANTS docking pose, a unique subset of protein coordinates with rotated hydroxyl hydrogen atoms was used to define the IFP. Standard IFP scoring parameters and a Tanimoto coefficient (Tc) measuring IFP similarity with the reference molecule pose (doxepin) in the H₁R crystal structure (Figure 2A) were used to filter and rank the docking poses of 543 known active H₁R compounds from the ChEMBLdb, 59 CNS drugs with inverse agonist activity on H₁R, 7088 decoys, and the focused database of 108 790 fragment-like molecules (only poses forming an H-bond and ionic interaction with D107^{3,32} are considered). The reference IFP bit string is available as Supporting Information.

Retrospective Virtual Screening Analysis. H₁R ligand test set 1 (543 known H₁R ligands from ChEMBLdb), test set 2 (59 CNS active drugs with inverse agonist activity on H₁R³⁰), and the focused database of 7088 similar decoys were docked into H₁R and scored with PLANTS and IFP. On the basis of optimal virtual screening enrichment of test set 1 (true positives) against the decoy set (false positives), IFP (Tanimoto similarity to doxepin of ≥ 0.75) and PLANTS (≤ -90) score cutoffs were defined. In an independent retrospective virtual screening study, the enrichment of test set 2 against the decoy set was determined at these IFP and PLANTS cutoff values.

Prospective Virtual Screening. The screening database was docked with the same PLANTS protocol used for the retrospective validation. After postprocessing of the results using IFPs and filtering for the ionic and H-bond interaction with D107^{3,32}, the previously mentioned cutoffs (IFP Tc ≥ 0.75 and PLANTS ≤ -90) were applied (see Figure 1C,D). To further focus the data set, we selected only compounds that had a consistently high PLANTS and IFP score for the best poses according to PLANTS and IFP; only compounds with an IFP score of ≥ 0.7 according to the best PLANTS pose as well as a PLANTS score of ≤ -75 according to the best IFP pose were selected (see Figure 1C). To assess the novelty of the selected compounds, all remaining 354 compounds were compared using Pipeline Pilot's ECFP-4⁷⁴ to the known active compounds and only compounds with an ECFP-4 score below 0.40 were selected (282 compounds).

ROCS 3D Similarity Search. The conformer database was generated using standard settings OMEGA⁷⁵ and searched with ROCS⁷⁶ using standard settings as well. The conformations of doxepin found in the H₁R X-ray structure were used as query molecules for independent ROCS runs. Compounds were ranked by decreasing Comboscore⁷⁶ (combination of shape Tanimoto and the normalized color score in this optimized overlay).

ECFP-4 2D Similarity Search. Two-dimensional similarity searches were carried out using ECFP-4 (extended connectivity

fingerprints⁵³) descriptors available in Pipeline Pilot⁷⁴ and compared using the Tanimoto coefficient.

Scaffold Diversity Analysis. Scaffold diversity⁶² of the novel fragment hits in known H₁R fragment-like ligand space was determined using the publicly available sca.svl script in MOE.⁷⁷ In this analysis, fragments are indexed by two parameters: cyclicity and complexity. Cyclicity is the ratio between ring atoms and side chain atoms (thus, if all the atoms of the molecule belong to the ring structure, cyclicity equals 1). In addition, the complexity was calculated as a descriptor of the size and shape of the scaffold, taking into account the smallest set of smallest rings, the number of heavy atoms, the number of bonds between the heavy atoms, and the sum of heavy atoms atomic number.⁶²

Compounds Selected by Virtual Screening. The compounds selected by virtual screening were purchased from available screening collections of six vendors (Supporting Information Table 4): Asinex (www.asinex.com), Chembridge (www.Hit2Lead.com), Enamine (www.enamine.com), IBScreen (www.ibscreen.com), Matrix Scientific (www.matrixscientific.com), and Vitas-M (www.vitasmlab.com). The purity of all compounds was verified by liquid chromatography–mass spectrometry (LC–MS). All 19 experimentally validated hits had a purity of 95% or higher (see Supporting Information Table 5), except compound 11 which in our hands had a purity of 76% (reported to be 95% pure according to the supplier Asinex).

Plasmids. Human H₁R cDNA was kindly provided by Dr. H. Fukui (Japan).⁷⁸ Human H₃R and H₄R cDNA were obtained from the Missouri S&T cDNA Resource Center (www.cdna.org).

Cell Culture and Transfection. HEK293T cells were cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum, 50 IU/mL penicillin, and 50 µg/mL streptomycin at 37 °C and 5% CO₂. Approximately 4 × 10⁶ cells in 10 cm dishes were transiently transfected with 0.5 or 5 µg of receptor DNA using 25 kDa linear polyethylenimine (PEI; Polysciences, Warrington, PA, U.S.) as transfection reagent (1:4 DNA/PEI ratio), for inositol phosphate (InsP) accumulation or radioligand displacement assays, respectively.

Radioligand Displacement Assay. Cells were harvested 2 days after transfection and homogenized in 50 mM Tris-HCl binding buffer (pH 7.4). Cell homogenates were co-incubated with indicated concentrations of fragment-like ligands and ~3 nM [³H]mepyramine (H₁R), ~1 nM [³H]N-α-methylhistamine (H₃R), or ~10 nM [³H]histamine (H₄R) in a total volume of 100 µL/well. The reaction mixtures were for 1–1.5 h at 25 °C on a microtiter shaker (750 rpm). Incubations were terminated by rapid filtration through Unifilter glass fiber C plates (PerkinElmer Life Sciences) that were presoaked in 0.3% polyethylenimine and subsequently washed three times with ice-cold binding buffer (pH 7.4 at 4 °C). Retained radioactivity was measured by liquid scintillation using a MicroBeta Trilux (PerkinElmer Life Sciences). Nonlinear curve fitting was performed using GraphPad Prism, version 4.03, software. The K_i values were calculated using the Cheng–Prusoff equation $K_i = IC_{50}/(1 + [radioligand]/K_d)$.⁷⁹

Inositol Phosphate Accumulation Assay. Twenty-four hours after transfection, cells were collected and seeded in Earle's inositol-free minimal essential medium (Invitrogen) supplemented with 10% fetal bovine serum, 50 IU/mL penicillin, 50 µg/mL streptomycin, and 1 µCi/ml myo-[2-³H]inositol in poly-L-lysine-coated 48-well plates. The next day, cells were washed with DMEM supplemented with 25 mM HEPES, pH 7.4, and 20 mM LiCl and subsequently incubated with the indicated concentrations of fragment-like ligands in the absence or presence of 0.1 µM histamine for 1 h at 37 °C. Incubations were terminated by replacing the assay buffer with 10 mM formic acid. Next, accumulated inositol phosphates were isolated using anion exchange chromatography (Dowex AG1-X8 columns, Bio-Rad) and counted by liquid scintillation.

■ ASSOCIATED CONTENT

Supporting Information

Additional analysis results of the retrospective and prospective virtual screening studies, H₁R, H₃R, and H₄R radioligand

displacement curves, and InsP accumulation barplots; a zipped file containing molecular structures of ligand test set (SMILES) and receptor mol2 coordinates (H₁R crystal structure), PLANTS docking configuration file, reference IFP bit-strings, and cavity coordinates used for IFP calculations. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +31 20 5987600. Fax: +31 20 5987610. E-mail: r.leurs@vu.nl.

Author Contributions

• These authors contributed equally to this work.

■ ACKNOWLEDGMENTS

The authors thank Herman D. Lim for technical assistance with the H₄R binding assays and Sebastiaan Kuhne for carefully reading the manuscript and performing the LC–MS analyses. This research was financially supported by The Netherlands Organization for Scientific Research (NWO) through a VENI grant (Grant 700.59.408 to C.d.G.), by TI-Pharma through Grant D1-105 (GPCR Forum to A.J.K.), and a NIGMS PSI:Biological grant (Grant U54 GM094618 to R.C.S. and V.K.).

■ ABBREVIATIONS USED

A2A, adenosine α-2A receptor; ADA1A, adrenergic α-1A receptor; ADRB2, adrenergic β-2 receptor; C3A, complement component 3A receptor; CCR5, C–C chemokine receptor 5; CRFR1, corticotropin releasing factor receptor 1; CNR2, cannabinoid receptor 2; DRD3, dopamine receptor D₃; FBDD, fragment-based drug design; FFAR1, free fatty acid receptor 1; FPR1, formyl peptide receptor 1; GPCR, G protein-coupled receptor; (h)H₁R, (human) histamine H₁ receptor; H₃R, histamine H₃ receptor; H₄R, histamine H₄ receptor; IFP, interaction fingerprint; InsP, inositol phosphate; MCH1, melanin-concentrating hormone receptor 1; NK1R, neurokinin 1 receptor; SBVS, structure-based virtual screening; SBVFS, structure-based virtual fragment screening; TRFR1, transferrin receptor 1

■ REFERENCES

- (1) Lagerstrom, M. C.; Schioth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discovery* **2008**, *7*, 339–357.
- (2) Rosenbaum, D. M.; Rasmussen, S. G.; Kobilka, B. K. The structure and function of G protein-coupled receptors. *Nature* **2009**, *459*, 356–363.
- (3) Topiol, S.; Sabio, M. X-ray structure breakthroughs in the GPCR transmembrane region. *Biochem. Pharmacol.* **2009**, *78*, 11–20.
- (4) 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.
- (5) de Graaf, C.; Rognan, D. Customizing G protein-coupled receptor models for structure-based virtual screening. *Curr. Pharm. Des.* **2009**, *15*, 4026–4048.
- (6) Leurs, R.; Church, M. K.; Tagliatela, M. H₁-Antihistamines: inverse agonism, anti-inflammatory actions and cardiac effects. *Clin. Exp. Allergy* **2002**, *32*, 489–498.
- (7) Shimamura, T.; Shiroishi, M.; Weyand, S.; Tsujimoto, H.; Winter, G.; Katritch, V.; Abagyan, R.; Cherezov, V.; Liu, W.; Han, G. W.; Kobayashi, T.; Stevens, R. C.; Iwata, S. Structure of the human histamine H1 receptor complex with doxepin. *Nature* **2011**, *475*, 65–70.

- (8) Varady, J.; Wu, X.; Fang, X.; Min, J.; Hu, Z.; Levant, B.; Wang, S. Molecular modeling of the three-dimensional structure of dopamine 3 (D3) subtype receptor: discovery of novel and potent D3 ligands through a hybrid pharmacophore- and structure-based database searching approach. *J. Med. Chem.* **2003**, *46*, 4377–4392.
- (9) 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.
- (10) Evers, A.; Klebe, G. Successful virtual screening for a submicromolar antagonist of the neurokinin-1 receptor based on a ligand-supported homology model. *J. Med. Chem.* **2004**, *47*, 5381–5392.
- (11) Edwards, B. S.; Bologa, C.; Young, S. M.; Balakin, K. V.; Prossnitz, E. R.; Savchuck, N. P.; Sklar, L. A.; Oprea, T. I. Integration of virtual screening with high-throughput flow cytometry to identify novel small molecule formylpeptide receptor antagonists. *Mol. Pharmacol.* **2005**, *68*, 1301–1310.
- (12) Evers, A.; Klabunde, T. Structure-based drug discovery using GPCR homology modeling: successful virtual screening for antagonists of the alpha1A adrenergic receptor. *J. Med. Chem.* **2005**, *48*, 1088–1097.
- (13) Salo, O. M.; Raitio, K. H.; Savinainen, J. R.; Nevalainen, T.; Lahtela-Kakkonen, M.; Laitinen, J. T.; Jarvinen, T.; Poso, A. Virtual screening of novel CB2 ligands using a comparative model of the human cannabinoid CB2 receptor. *J. Med. Chem.* **2005**, *48*, 7166–7171.
- (14) Kellenberger, E.; Springael, J. Y.; Parmentier, M.; Hachet-Haas, M.; Galzi, J. L.; Rognan, D. Identification of nonpeptide CCR5 receptor agonists by structure-based virtual screening. *J. Med. Chem.* **2007**, *50*, 1294–1303.
- (15) 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.
- (16) Engel, S.; Skoumbourdis, A. P.; Childress, J.; Neumann, S.; Deschamps, J. R.; Thomas, A. C.; Colson, A. O.; Costanzi, S.; Gershengorn, M. C. A virtual screen for diverse ligands: discovery of selective G protein-coupled receptor antagonists. *J. Am. Chem. Soc.* **2008**, *130*, 5115–5123.
- (17) Kiss, R.; Kiss, B.; Konczol, A.; Szalai, F.; Jelinek, I.; Laszlo, V.; Noszal, B.; Falus, A.; Keseru, G. M. Discovery of novel human histamine H4 receptor ligands by large-scale structure-based virtual screening. *J. Med. Chem.* **2008**, *51*, 3145–3153.
- (18) Tikhonova, I. G.; Sum, C. S.; Neumann, S.; Engel, S.; Raaka, B. M.; Costanzi, S.; Gershengorn, M. C. Discovery of novel agonists and antagonists of the free fatty acid receptor 1 (FFAR1) using virtual screening. *J. Med. Chem.* **2008**, *51*, 625–633.
- (19) Klabunde, T.; Giegerich, C.; Evers, A. Sequence-derived three-dimensional pharmacophore models for G protein-coupled receptors and their application in virtual screening. *J. Med. Chem.* **2009**, *52*, 2923–2932.
- (20) Salon, J. A.; Lodowski, D. T.; Palczewski, K. The significance of G protein-coupled receptor crystallography for drug discovery. *Pharmacol. Rev.* **2011**, *63*, 901–937.
- (21) Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K. Structure-based discovery of beta2-adrenergic receptor ligands. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6843–6848.
- (22) Carlsson, J.; Yoo, L.; Gao, Z. G.; Irwin, J. J.; Shoichet, B. K.; Jacobson, K. A. Structure-based discovery of A2A adenosine receptor ligands. *J. Med. Chem.* **2010**, *53*, 3748–3755.
- (23) Katritch, V.; Jaakola, V. P.; Lane, J. R.; Lin, J.; Ijzerman, A. P.; Yeager, M.; Kufareva, I.; Stevens, R. C.; Abagyan, R. Structure-based discovery of novel chemotypes for adenosine A(2A) receptor antagonists. *J. Med. Chem.* **2010**, *53*, 1799–1809.
- (24) Sabio, M.; Jones, K.; Topiol, S. Use of the X-ray structure of the beta2-adrenergic receptor for drug discovery. Part 2: Identification of active compounds. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5391–5395.
- (25) de Kloe, G. E.; Bailey, D.; Leurs, R.; de Esch, I. J. Transforming fragments into candidates: small becomes big in medicinal chemistry. *Drug Discovery Today* **2009**, *14*, 630–646.
- (26) Rognan, D. Fragment-based approaches and computer-aided drug discovery. *Top. Curr. Chem.* **2011**, *1*–22.
- (27) de Graaf, C.; Rognan, D. Selective structure-based virtual screening for full and partial agonists of the beta2 adrenergic receptor. *J. Med. Chem.* **2008**, *51*, 4978–4985.
- (28) Korb, O.; Stützle, T.; Exner, T. E. An ant colony optimization approach to flexible protein–ligand docking. *Swarm Intell.* **2007**, *1*, 115–134.
- (29) Overington, J. ChEMBL, an interview with John Overington, team leader, chemogenomics at the European Bioinformatics Institute Outstation of the European Molecular Biology Laboratory (EMBL-EBI). Interview by Wendy A. Warr. *J. Comput.-Aided Mol. Des.* **2009**, *23*, 195–198.
- (30) Bakker, R. A.; Nicholas, M. W.; Smith, T. T.; Burstein, E. S.; Hacksell, U.; Timmerman, H.; Leurs, R.; Brann, M. R.; Weiner, D. M. In vitro pharmacology of clinically used central nervous system-active drugs as inverse H(1) receptor agonists. *J. Pharmacol. Exp. Ther.* **2007**, *322*, 172–179.
- (31) Korb, O.; Stützle, T.; Exner, T. E. Empirical scoring functions for advanced protein–ligand docking with PLANTS. *J. Chem. Inf. Model.* **2009**, *49*, 84–96.
- (32) Marcou, G.; Rognan, D. Optimizing fragment and scaffold docking by use of molecular interaction fingerprints. *J. Chem. Inf. Model.* **2007**, *47*, 195–207.
- (33) Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L. 3rd. Assessing scoring functions for protein–ligand interactions. *J. Med. Chem.* **2004**, *47*, 3032–3047.
- (34) Jain, A. N.; Nicholls, A. Recommendations for evaluation of computational methods. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 133–139.
- (35) Irwin, J. J.; Shoichet, B. K. ZINC—a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177–182.
- (36) Verheij, M. H.; de Graaf, C.; de Kloe, G. E.; Nijmeijer, S.; Vischer, H. F.; Smits, R. A.; Zuiderveld, O. P.; Hulscher, S.; Silvestri, L.; Thompson, A. J.; van Muijlwijk-Koezen, J. E.; Lummis, S. C.; Leurs, R.; de Esch, I. J. Fragment library screening reveals remarkable similarities between the G protein-coupled receptor histamine H(4) and the ion channel serotonin 5-HT(3A). *Bioorg. Med. Chem. Lett.* **2011**, *5460*–5464.
- (37) Siegal, G.; Ab, E.; Schultz, J. Integration of fragment screening and library design. *Drug Discovery Today* **2007**, *12*, 1032–1039.
- (38) Bruysters, M.; Pertz, H. H.; Teunissen, A.; Bakker, R. A.; Gillard, M.; Chatelain, P.; Schunack, W.; Timmerman, H.; Leurs, R. Mutational analysis of the histamine H1-receptor binding pocket of histaprodifens. *Eur. J. Pharmacol.* **2004**, *487*, 55–63.
- (39) Wawer, M.; Bajorath, J. Similarity-potency trees: a method to search for SAR information in compound data sets and derive SAR rules. *J. Chem. Inf. Model.* **2010**, *50*, 1395–1409.
- (40) Becker, O. M.; Dhanoa, D. S.; Marantz, Y.; Chen, D.; Shacham, S.; Cheruku, S.; Heifetz, A.; Mohanty, P.; Fichman, M.; Sharadendu, A.; Nudelman, R.; Kauffman, M.; Noiman, S. An integrated in silico 3D model-driven discovery of a novel, potent, and selective amidosulfonamide 5-HT1A agonist (PRX-00023) for the treatment of anxiety and depression. *J. Med. Chem.* **2006**, *49*, 3116–3135.
- (41) Holst, B.; Nygaard, R.; Valentini-Hansen, L.; Bach, A.; Engelstoft, M. S.; Petersen, P. S.; Frimurer, T. M.; Schwartz, T. W. A conserved aromatic lock for the tryptophan rotameric switch in TM-VI of seven-transmembrane receptors. *J. Biol. Chem.* **2010**, *285*, 3973–3985.
- (42) Jongejan, A.; Bruysters, M.; Ballesteros, J. A.; Haaksma, E.; Bakker, R. A.; Pardo, L.; Leurs, R. Linking agonist binding to histamine H1 receptor activation. *Nat. Chem. Biol.* **2005**, *1*, 98–103.

- (43) Katritch, V.; Abagyan, R. GPCR agonist binding revealed by modeling and crystallography. *Trends Pharmacol. Sci.* **2011**, *32*, 637–643.
- (44) Rasmussen, S. G.; Choi, H. J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; Devree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S.; Schnapp, A.; Konetzki, I.; Sunahara, R. K.; Gellman, S. H.; Pautsch, A.; Steyaert, J.; Weis, W. I.; Kobilka, B. K. Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature* **2011**, *469*, 175–180.
- (45) Rognan, D. Docking Methods for Virtual Screening: Principles and Recent Advances. In *Virtual Screening*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2011; pp 153–176.
- (46) 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.
- (47) Wieland, K.; Laak, A. M.; Smit, M. J.; Kuhne, R.; Timmerman, H.; Leurs, R. Mutational analysis of the antagonist-binding site of the histamine H(1) receptor. *J. Biol. Chem.* **1999**, *274*, 29994–30000.
- (48) Ohta, K.; Hayashi, H.; Mizuguchi, H.; Kagamiyama, H.; Fujimoto, K.; Fukui, H. Site-directed mutagenesis of the histamine H1 receptor: roles of aspartic acid107, asparagine198 and threonine194. *Biochem. Biophys. Res. Commun.* **1994**, *203*, 1096–1101.
- (49) Kufareva, I.; Rueda, M.; Katritch, V.; Stevens, R. C.; Abagyan, R. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* **2011**, *19*, 1108–1126.
- (50) O'Boyle, N. M.; Liebeschuetz, J. W.; Cole, J. C. Testing assumptions and hypotheses for rescoring success in protein–ligand docking. *J. Chem. Inf. Model.* **2009**, *49*, 1871–1878.
- (51) Kellenberger, E.; Foata, N.; Rognan, D. Ranking targets in structure-based virtual screening of three-dimensional protein libraries: methods and problems. *J. Chem. Inf. Model.* **2008**, *48*, 1014–1025.
- (52) Knox, A. J.; Meegan, M. J.; Sobolev, V.; Frost, D.; Zisterer, D. M.; Williams, D. C.; Lloyd, D. G. Target specific virtual screening: optimization of an estrogen receptor screening platform. *J. Med. Chem.* **2007**, *50*, 5301–5310.
- (53) Rogers, D.; Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754.
- (54) Grant, J. A.; Gallardo, M. A.; Pickup, B. T. A fast method of molecular shape comparison: a simple application of a Gaussian description of molecular shape. *J. Comput. Chem.* **1996**, *17*, 1653–1666.
- (55) Steffen, A.; Kogej, T.; Tyrchan, C.; Engkvist, O. Comparison of molecular fingerprint methods on the basis of biological profile data. *J. Chem. Inf. Model.* **2009**, *49*, 338–347.
- (56) Fan, Y.; Lai, M. H.; Sullivan, K.; Popielek, M.; Andree, T. H.; Dollings, P.; Pausch, M. H. The identification of neurotensin NTS1 receptor partial agonists through a ligand-based virtual screening approach. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5789–5791.
- (57) Gillespie, R. J.; Cliffe, I. A.; Dawson, C. E.; Dourish, C. T.; Gaur, S.; Jordan, A. M.; Knight, A. R.; Lerpiniere, J.; Misra, A.; Pratt, R. M.; Roffey, J.; Stratton, G. C.; Upton, R.; Weiss, S. M.; Williamson, D. S. Antagonists of the human adenosine A2A receptor. Part 3: Design and synthesis of pyrazolo[3,4-*d*]pyrimidines, pyrrolo[2,3-*d*]pyrimidines and 6-arylpurines. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2924–2929.
- (58) Miller, D. C.; Klute, W.; Calabrese, A.; Brown, A. D. Optimising metabolic stability in lipophilic chemical space: the identification of a metabolically stable pyrazolopyrimidine CRF-1 receptor antagonist. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6144–6147.
- (59) 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.
- (60) Surgand, J. S.; Rodrigo, J.; Kellenberger, E.; Rognan, D. A chemogenomic analysis of the transmembrane binding cavity of human G protein-coupled receptors. *Proteins* **2006**, *62*, 509–538.
- (61) Venhorst, J.; Nunez, S.; Terpstra, J. W.; Kruse, C. G. Assessment of scaffold hopping efficiency by use of molecular interaction fingerprints. *J. Med. Chem.* **2008**, *51*, 3222–3229.
- (62) Xu, J. A new approach to finding natural chemical structure classes. *J. Med. Chem.* **2002**, *45*, 5311–5320.
- (63) Celanire, S.; Wijnmans, M.; Talaga, P.; Leurs, R.; de Esch, I. J. Keynote review: histamine H3 receptor antagonists reach out for the clinic. *Drug Discovery Today* **2005**, *10*, 1613–1627.
- (64) Westly, E. Nothing to sneeze at. *Nat. Med.* **2010**, *16*, 1063–1065.
- (65) Lim, H. D.; de Graaf, C.; Jiang, W.; Sadek, P.; McGovern, P. M.; Istyastono, E. P.; Bakker, R. A.; de Esch, I. J.; Thurmond, R. L.; Leurs, R. Molecular determinants of ligand binding to H4R species variants. *Mol. Pharmacol.* **2010**, *77*, 734–743.
- (66) Gillard, M.; Van Der Perren, C.; Moguilevsky, N.; Massingham, R.; Chatelain, P. Binding characteristics of cetirizine and levocetirizine to human H(1) histamine receptors: contribution of Lys(191) and Thr(194). *Mol. Pharmacol.* **2002**, *61*, 391–399.
- (67) Ballesteros, J. A.; Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure–function relations in G protein-coupled receptors. *Methods Neurosci.* **1995**, *25*, 366–428.
- (68) Wu, C. H.; Apweiler, R.; Bairoch, A.; Natale, D. A.; Barker, W. C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M.; Martin, M. J.; Mazumder, R.; O'Donovan, C.; Redaschi, N.; Suzek, B. The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.* **2006**, *34*, D187–D191.
- (69) Bioinfo 11.1. <http://bioinfo-pharma.u-strasbg.fr/bioinfo> (accessed June 6, 2011).
- (70) Calculator, version 5.1.4; ChemAxon: Budapest, Hungary.
- (71) Corina, version 3.46; Molecular Networks GmbH: Erlangen, Germany.
- (72) FILTER, version 2.1.1; OpenEye Scientific Software: Santa Fe, NM.
- (73) MoKa, version 1.1.0; Molecular Discovery Ltd.: Pinner, Middlesex, U.K.
- (74) Pipeline Pilot, version 6.1.5; Accelrys: San Diego, CA.
- (75) Omega, version 2.3.2; OpenEye Scientific Software: Santa Fe, NM.
- (76) ROCS, version 2.3.1; OpenEye Scientific Software: Santa Fe, NM.
- (77) MOE, version 2010.10; Chemical Computing Group, Inc.: Montreal, Canada.
- (78) Fukui, H.; Fujimoto, K.; Mizuguchi, H.; Sakamoto, K.; Horio, Y.; Takai, S.; Yamada, K.; Ito, S. Molecular cloning of the human histamine H1 receptor gene. *Biochem. Biophys. Res. Commun.* **1994**, *201*, 894–901.
- (79) Cheng, Y.; Prusoff, W. H. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50% inhibition (I₅₀) of an enzymatic reaction. *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.