From Receptors to Ligands: Fragment-assisted Drug Design for GPCRs Applied to the Discovery of H3 and H4 Receptor Antagonists

Alexander Heifetz1*, Michael P Mazanetz1, Tim James1, Sandeep Pal1, Richard J Law1, Mark Slack2 and Philip C Biggin3

1Evotec (UK) Ltd., 114 Innovation Drive, Milton Park, Abingdon, Oxfordshire OX14 4RZ, UK
2Evotec AG, Manfred Eigen Campus, Essener Bogen 7, 22419 Hamburg, Germany
3Department of Biochemistry, University of Oxford, South Parks Road Oxford OX1 3QU, UK

Abstract

G-Protein Coupled Receptors (GPCRs) have enormous physiological and biomedical importance, being the primary target of a large number of modern drugs. The availability of structural information of the binding site of the targeted GPCR plays a key role in rationalization, efficiency and cost-effectiveness of the drug discovery process. However, obtaining structural information on GPCRs using X-ray crystallography or NMR requires a large investment of time and is technically very challenging. This situation significantly limits the ability of these methods to have an impact in drug discovery for GPCR targets in the short term and hence there is an urgent need for other effective and cost-efficient alternatives. We present here a practical approach that integrates GPCR modelling with fragment based screening to provide structural insights on the H3 and H4 histamine receptor binding sites. This approach creates a cost-efficient new avenue for structure-based drug design (SBDD) against GPCR targets. We report here a success of using this protocol for the discovery of selective and dual H3 and H4 antagonists. Our fragment screen yielded 44 H3, 21 H4 selective and 20 dual fragment hits. These fragments were used to construct high-quality H3 and H4 models followed by binding site exploration and structure based virtual screening (VS). Overall, 172 compounds were purchased for testing based on the virtual screening results. Of the 74 compounds predicted to have dual activity, 33 had activity against one or other of the two receptors (44%), of which 17 had activity against both. Of the 19 compounds predicted to be H3 selective, 13 were active against H3 (68%) and 10 of these also had selectivity over H4. Of the 79 compounds predicted to be H4 selective, 36 were active against H4 (45%) and 2 of these also had selectivity over H3.

Abbreviations: H3: Histamine 3 Receptor; H4: Histamine 4 Receptor; MD: Molecular Dynamics; FADD: Fragment-Assisted Drug Discovery; SBDD: Structure-Based Drug Discovery; SDM: Site Directed Mutagenesis; FBDD: Fragment-Based Drug Discovery; GPCRs: G-Protein Coupled Receptors; 3D: Three-Dimensional; 7TMD: Seven-Transmembrane Domain; TM: Transmembrane Helix; ECL: Extracellular Loop; ICL: Intracellular Loop; β2AR: β2-Adrenergic Receptor; PDB: Protein Data Bank

Introduction

G-Protein Coupled Receptors (GPCRs) have enormous physiological and biomedical importance, being the primary target of a large number of modern drugs. The availability of structural information on the binding site of the targeted GPCR plays a key role in rationalization, efficiency and cost-effectiveness of the drug discovery process. 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. X-ray crystallography and NMR require a large investment of time and is technically very challenging. This situation significantly limits the ability of these methods to impact the drug discovery process for GPCR targets. X-ray crystallography and NMR require a large investment of time and is technically very challenging. This situation significantly limits the ability of these methods to impact the drug discovery process for GPCR targets.

Fragment-assisted drug discovery (FADD) has emerged as a new tool for drug discovery in recent years [2,3] and is typically aimed at a target for which a crystal structure can be determined in order to rationally guide fragment hit expansion. While the majority of historical fragment screens have been focused towards biochemical targets, only few examples, to date, have been published in which this method has been used to identify ligands for GPCRs [4]. This is due to the current infeasibility of regularly crystallizing the GPCR-fragment complexes that are required for further fragment expansion. Recently, there has been several lines of evidence suggesting that FADD might be useful for GPCR hit identification [5]. These publications encouraged us to try this approach in the discovery of GPCR antagonists.

The H3 and the H4 receptor belong to the Class A histamine GPCRs subfamily. The histamine GPCRs signal through Gi/o proteins leading to inhibition of cAMP formation, mobilisation of calcium from intracellular stores and stimulation of MAP kinase in both heterologous expression systems and native immune cells [5,6]. H3 and H4 receptors remain highly attractive targets for drug discovery and medicinal chemistry development programs [7-10]. While the H3 receptor seems to be an interesting target for treatment of certain CNS disorders like cognitive disorders, narcolepsy, ADHD and pain [11,12], the H4 receptor plays an important role in inflammation and allergy [13]. Experimental evidence has demonstrated the high therapeutic potential of modulating both the H3 and H4 receptors in cancer treatment [14,15].

In the light of this wide clinical evidence of the importance of H3 and H4 receptors as potential drug targets [7-10], the need for exploration of the key structural features and key residues involved in antagonist binding and selectivity cannot be underestimated. Such structural data...
is vital for driving the development of the next generation of new H3 or H4 antagonists. Recently several studies were conducted to address this problem, involving in-silico guided site directed mutagenesis data (SDM) [16,17]. This data shows that key residues, responsible for antagonist binding to H3 receptors are D114, Y151, E206 and for H4 receptors are D94, D182. Mutations of V3.40A, N4.57Y, T4.52A, T5.55 and Q4.1 showed an average of a ~10 fold decrease in potency for a series of clobenpropit derivatives (selective H4 antagonists) to H4 [18].

In our work, a fragment screen yielded 44 H3, 21 H4 selective and 20 dual fragment hits. These fragments were used to construct high quality H3 and H4 models followed by binding site exploration and structure based VS. Overall, 172 compounds were purchased for testing based on the VS results. Of the 74 compounds predicted to have dual activity, 33 had activity against one or other of the two receptors (44%), of which 17 had activity against both. Of the 19 compounds predicted to be H3 selective, 13 were active against H3 (68%) and 10 of these also had selectivity over H4. Of the 79 compounds predicted to be H4 selective, 36 were active against H4 (45%) and 2 of these also had selectivity over H3.

The results reported here are comparable to those from other similar works [19-22] which showed that GPCR modeling [23-31] in the absence of a crystal structure can be a valid replacement [32-39] for structural and functional exploration of GPCR receptors, and for the discovery [21,40-43], VS [44-52] and optimisation [23,53] of their ligands.

Materials and Methods

Computational methods

Residue numbering: The position of each amino acid residue of H3 and H4 receptors was identified both by its sequence number and by its generic number proposed by Ballesteros and Weinstein for class A GPCRs [54]. Briefly, in this numbering scheme, amino acid residues in the 7 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.

H3 and H4 modeling procedure

We performed our hierarchical GPCR modeling protocol [55] in 5 sequential steps: (1) Multiple-Sequence Alignment (2) Template selection (3) Homology Modeling of the receptors (4) Docking procedure and generation of receptor-fragments complex (5) Receptor binding site optimization with “low-mode” molecular dynamics (LowModeMD) simulation.

Multiple-Sequence Alignment is required for selection of the optimum template from the available GPCR crystal structures for further homology modeling of H3 and H4. The amino acid sequences of human H3 (UniProtKB/Swiss-Prot ID Q9YSN1) and human H4 (UniProtKB/Swiss-Prot ID Q9H3N8) were retrieved from the Swiss-Prot database. The sequences of the H3 and H4 receptors were aligned with four published GPCR crystal structures [human dopamine D3 receptor (D3, PDB entry 2PBL), β2-adrenergic receptor (β2AR, PDB entry 2RH1), human A2A adenosine receptor (A2A, PDB entry 3EML) and bovine rhodopsin (PDB entry 1F88) - the more recently published crystal structure of the histamine 1 (H1) receptor [56] was not available at the time of this work, using a multiple-sequence alignment tool implemented in MOE version 2010.10 (Chemical Computing Group). In this approach, originally introduced by Needleman in 1970, [57] alignments were computed by optimizing a function based on residue similarity scores obtained from applying an amino acid substitution matrix (blosum62) [58] 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 of the best template for modeling of any GPCR is usually very challenging [37]. Class A GPCRs share the same arrangement of the seven helices [59] and their 7TMD sequence similarities are relatively high [60-63]. However, even small sequence differences can lead to significant differences in overall structure and particularly in the topology of the ligand binding site [64]. 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 [65-67]. 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 [66]. We ranked the quality of crystal structures as potential templates for homology of H3 and H4 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 H3 and H4 were 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 of the H3 and H4 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) which superposed well onto anchor residues on each side of the insertion area, after the method described by Fechteler et al. [68]. Following the selection of appropriate loop templates, multiple model candidates for each loop were constructed and scored using the OPLS-AA energy function [69,70]. 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. Sidechain data is assembled from an extensive rotamer library generated by systematic clustering of high-resolution PDB data. After all of the backbone segment and sidechain conformations were chosen, the model was minimized using the MMFF94x force-field [71].

Docking procedure and generation of receptor-fragments complex: In this stage we docked our fragment hits into the binding sites of H3 and H4 using a flexible docking procedure. In all docking experiments described in this manuscript we used the GOLD docking package (Cambridge Crystallographic Data Centre, version 5.0), followed by re-scoring and re-ranking procedures. When docking fragments a flexible GOLD docking procedure was used, as fragments were being docked into an unrefined binding site. The rotamer library of GOLD [72] (Cambridge Crystallographic Data Centre) was used to
add flexibility to the key residues found by SDM. The docking itself was performed once for each molecule, with the 10 top ranked docking poses scored by the GOLD default scoring function [73] retained. We then re-scored and re-ranked these 10 docked poses using AMBER interaction energy. We used the MM-PBSA/GBSA approach [74] to calculate the AMBER interaction energy between protein and ligand. As was recently published [75,76] 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 re-score and to rank docking poses. The 10 top-ranked docking poses, according to the AMBER interaction energy, was taken for further analysis. The top-ranked poses of fragments, according to the AMBER interaction energy and SDM, was selected for further binding site refinement.

Receptor binding site optimization: In the final modeling stage, we aimed to optimize the H3 and H4 binding sites by applying a “low-mode” molecular dynamics simulation (LowModeMD [77]) to docked fragments pose. The LowModeMD protocol is a stochastic conformation generation protocol implemented in MOE. For the LowModeMD refinement the H3 and H4 binding sites were defined by residues within a radius of 7.0 Å of the docked fragments. Flexibility was permitted for all atoms within this radius when the rest of the atoms of H3 and H4 and of fragments were restrained. The radius of 7.0 Å was selected to ensure that all the atoms of helical fraction that include close to fragment residues will be also included in refinement. The dielectric constant was set to 3. A LowModeMD constant temperature MD simulation was performed at 300K, using the Berendsen thermostat [78] and the velocity Verlet algorithm. The default value of the energy minimisation gradient (0.001 kcal.mol⁻¹.Å⁻¹) 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.

Structure-based virtual screen

We performed structure-based VS in 2 sequential steps: (1) Similarity search followed by (2) Filtering by docking. Figure 1 gives a schematic overview of the strategy that was used starting from GPCR modelling.

Similarity search: We used a shape-pharmacophore superposition method, as it is implemented in OMEGA and ROCS tools of the OpenEye software package (version 3.0.0) [79] to rank the similarity between fragment hits and in-house vendor screening library of 4.8 M compounds. We used OMEGA [80] to generate potential bioactive and energetically accessible conformations for each molecule in the vendor screening library and ROCS for shape-based superposition. A Gaussian description of molecular shape is used to compare the shapes of any two molecules by maximization of their volume intersection [79]. We used docking poses of our fragment hits as templates for superposing with our vendor screening library. Molecules were superposed by a solid-body optimization process that maximizes the overlap volume between them. We used the ROCS default scoring function Combo Score to rank the quantity of the overlap between the fragments and screening library. The Combo Score evaluates the shape and pharmacophore properties (like donor, acceptor, hydrophobe, cation, anion, and ring) overlap between a pair of molecules. It was observed by Hawkins et al. [81] that adding to the shape the score for the appropriate overlap of pharmacophore properties, and then ranking on this summed score, improved VS performance considerably. The top compounds from the screening library, which showed overlap with Combo Score > 1 to our fragment hits, were selected for further docking and filtering.

Filtering by docking: These top ranked molecules from screening library were docked into the H3 and H4 receptors using rigid docking procedure. We used the same docking and scoring protocol as for docking fragments, however no flexibility was allowed for the backbone and sidechains of the receptors. The top ranked compounds according to AMBER interaction energy were manually inspected and selected for binding-functional assays.

Binding and functional assays

Reagents: Histamine dihydrochloride (53300) was purchased from Fluka, IBMX (410957) was from Calbiochem, HTRF-cAMP dynamic 2 assay kit (62AM4PEB) was from CisBio Bioassays, and FLIPR Calcium 3 Assay Kit, Express (R8108) was from MDS Analytical Technologies. The reference antagonists Thioperamide maleate (0444) and Clobenpropit dihydrobromide (0752) were both purchased from Tocris. HEPES (1 M, 15630-056) and Hanks’ Balanced Salt Solution (HBBS 10x, 14065-049) were obtained from Invitrogen while PBS (D8537), Probenecid (P-8761) and Forskolin (F6886) were purchased from Sigma.

Cell lines: The recombinant CHO-K1 cell line expressing, Galphal6 with the human H4 (cat no. ES-393-A) and the recombinant CHO-K1 cell line expressing Galphal6 with the human H3 (cat no. ES-393-F) were purchased from Euroscreen (Brussels, Belgium). All cell lines were cultured in Nutrient Mixture F-12 Ham (Sigma, N4888) supplemented with 10% (vol/vol) heat-inactivated fetal calf serum (Sigma, F9665), 100 U/mL penicillin (PAA, P11- 010), 100 µg/mL streptomycin (PAA, P11-010), 250 µg/mL Zeocin (Invitrogen, R250-01), and 500 µg/mL G418 (PAA, P31-011).

Calcium flux assay: The functionality of each histamine receptor was assessed using a standard calcium flux assay and the corresponding stable expressing cell line. Histamine H3 or H4 receptor expressing CHO-K1 cells (adherent cultures) were seeded into, tissue culture-treated, 384-well, black clear-bottom plates (CellBind Corning 7086); at either 15,000 (H4) or 20,000 (H3) cells/well in culture medium and maintained in an incubator (5% CO2 at 37°C) overnight prior to performing the calcium mobilization assays. Cells were washed once with assay buffer (HBSS 1x, 20 mM HEPES pH 7.4) using an automated cell washer leaving the cells in 25 µl/well assay buffer. 25 µl of the calcium assay kit solution (containing 5 mM probenecid) was added into each well. The final liquid volume before antagonist treatment was calculated from the binding site optimization, which showed overlap with Combo Score > 1 to our fragment hits.
Results and Discussion

Fragments screening

A collection of five compound plates representing 1,708 fragments from Evotec’s fragment library (consisting of 30,000 highly diverse fragments), were screened at 2 and 20 µM in the functional calcium flux assays on the H3 and H4 receptors to identify antagonists. The compounds were tested at n=4, at each concentration, on each receptor, to provide a rate of confirmed hits. The simultaneous testing on the different receptors allowed an initial assessment of specificity and selectivity of the identified hit compounds. Table 1 summarizes the results of the fragment screening.

The obtained mean Z’ values of 0.83 (H3) and 0.59 (H4) indicate a good assay performance. Based on negative sample control wells, thresholds for hit selection between 15% and 32% inhibition were calculated. As expected the hit rates were higher at the higher compound concentration whereby the majority of the hits identified at 2 µM were also identified at 20 µM (see hit overlap). The highest hit rate (64 hits or 3.7% at 20 µM) was obtained for the H3 whereas H4 showed a significantly lower hit rate (21 hits or 1.2% at 20 µM). Out of 64 H3 fragment hits 20 were dual active and 44 were H3 selective, whilst all 21 of the H4 fragment hits were selective. The potency and selectivity of the fragment hits were tested at various concentrations on the three cell lines expressing the histamine receptors H3 and H4 in the corresponding calcium flux assays. Finally, our fragment screen yielded 44 H3 selective, 21 H4 selective and 20 dual active hits. Finally our fragment screen yielded 44 H3 selective, 21 H4 selective and 20 dual fragment hits. In Figure 2 we show three examples of these fragment hits that we used to demonstrate our method.

Modelling of human H3 and H4 receptors

The structure of the human H3 and H4 receptors were modelled based on a 2.8 Å resolution crystal structure of the β2-adrenergic receptor (β2AR, PDB entry 2RH1 [82]). Multiple sequence alignment, shown in Figure 3, indicated that the β2AR had a higher sequence similarity with the H3 and H4 receptors compared with other publically-available in the literature [86]. A post factum analysis suggest that the β2AR is similarly as good as H1 as a template for the modelling of H3 and H4. The sequence identity and similarity between β2AR and H3 (27% identity, 52% similarity) or H4 (28% identity, 52% similarity) receptors are almost identical to those between H1 and these receptors (H1:H3 32% identity, 58% similarity & H1:H4 29% identity, 54% similarity). As described previously [2,83-85]. With the models of H3 and H4, SDM data and fragment hits in hand, we next turned to the docking, optimization and exploration of receptors binding sites followed by VS.

As outlined above, we used the β2AR rather than H1 [56] as a template because a crystal structure of H1 was not publically available at the time that this work was performed. The use of β2AR as a template for modelling of the H3 and H4 receptors is widely reported in the literature [86]. A post factum analysis suggest that the β2AR is similarly as good as H1 as a template for the modelling of H3 and H4. The sequence identity and similarity between β2AR and H3 (27% identity, 52% similarity) or H4 (28% identity, 52% similarity) receptors are almost identical to those between H1 and these receptors (H1:H3 32% identity, 58% similarity & H1:H4 29% identity, 54% similarity). As described before [55,87], our modeling method is less dependent on the initial template due to the fact that we apply post homology modeling optimisation procedures. The LowModeMD used by us allowed refinement of the GPCR models to a degree that is not possible with static homology modelling alone [1]. The structural insights gained from this process are critical for the discovery of a new generation of H3 and H4 antagonists.

Binding pocket of H3

The binding site of H3 (Figure 5A) is located between TM3, 5, 6 and 7 respectively. These differences in the binding site topology between H3 and H4 are highly important for the design and optimization of selective or dual H3 and H4 ligands. For the design of H3 selective compounds we hypothesize that occupation of the S1 and S2 subpockets is required, and compounds with dual potency would not occupy these subpockets. We also noticed that the conserved F1835.47 is positioned further inside the H4 binding site, compared to its position in the H3 binding site, making it more accessible for interactions with ligands. This observation can help the design of H4 selective antagonists, as shown previously [2,83-85]. With the models of H3 and H4, SDM data and fragment hits in hand, we next turned to the docking, optimization and exploration of receptors binding sites followed by VS.

Table 1: Statistics of fragment screening results.

| Number of fragments | H3 | H4 |
|---------------------|----|----|
| Concentration       | 2µM| 20µM| 2µM| 20µM|
| Compounds           | 1708| 1708| 1708| 1708|
| Mean Z’ value       | 0.83| 0.59| 0.83| 0.59|
| Hit Threshold       | 19.1%| 16.7%| 32%| 31.4%|
| Confirmed Hits      | 19| 44| 9| 21|
| Selective fragments | 12| 40| 3| 10|
| Hit Overlap         | 17| 5| 17| 5|

Figure 2: Examples of fragment hits used for H3 and H4 structure modeling.
Figure 3: Multiple sequence alignment of H3 and H4 receptors aligned with the template (β2AR, PDB entry 2RH1), the colour coding for TM 1 to 7 is dark orange, pink, red, purple, dark red, orange and light yellow, respectively.

Figure 4: (A) and (B) show the final models of H3 and H4. 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 and the surface of the receptor binding site in pink for H3 and green for H4. (B) The ICLs and ECLs are omitted to ease the view.
7. The selective fragment f1 (Figure 2) is predicted to have two distinct binding poses with the H3 receptor as illustrated in Figure 5A. In the first docking pose the central core of f1 forms a salt bridge with D1144.52 and in the second docking pose, a salt bridge with E2066.46. In both docking poses the saturated ring of f1 occupies a hydrophobic subpocket (Figure 4B): in pose 1 it is S1 and in pose 2 it is S2. In the first docking pose the aromatic moiety of f1 forms a face-to-edge-stacking interaction with W3716.48 and in the second docking pose the same moiety forms a face-to-face π-stacking interaction with W3716.48. The interaction with conserved W3716.48 can have a strong impact on the functionality of H3 due to the fact that W6.48 is involved as a "transmission switch". The mechanism and structural changes associated with the functionality of GPCRs remain unclear and many experimental and computational methods have been applied to investigate this [87,88]. These methods provided an explanation of receptor functionality based on the action of so-called "molecular switches" buried in the receptor structure. The "transmission switch" accounts for the location of conserved residues W6.48 and F9.44 towards P6.50 and is possibly the most common switch among Class A GPCRs. The difference in energy interactions between poses 1 and 2 of f1 was negligible at < 1.0 kcal/mol, likely due to the fact that in both poses f1 forms two identical types of interaction with H3 (one salt bridge and one π-stack). Based on this observation, it was inconclusive which pose was the more likely, and therefore both were used in further VS.

The binding pose of dual fragment f3 (Figure 2) with H3 is illustrated in Figure 5B. The hydrogen donor of the f3 imidazole moiety forms a hydrogen bond with D1143.32 and the hydrogen donor of the f3 amide linker can potentially form a second hydrogen bond with E2065.46.

**Binding pocket of H4**

The binding site of H4 (Figure 5C) is located between TM3, 5, 6 and 7. The binding pose of the selective fragment f2 (Figure 2) with H4 is illustrated in Figure 5C. The basic centre of f2 forms a salt bridge with D941.32 and the hydrogen donor of the indole core of f2 forms a hydrogen bond with E1825.46. The indole core of f2 occupies a hydrophobic subpocket located between TM5 and 6 and forms two face-to-face stacking interactions with W3166.48 and F1835.47. As mentioned for H3 the interaction of f3 with transmission switch residue W3166.48 can have a strong impact on the functionality of the H4 receptor.

The docking pose of the dual fragment f3 (Figure 2) with H4 is illustrated in Figure 5D. The amide linker hydrogen donor of f3 forms a hydrogen bond with E1825.46, whilst the hydrogen donor of the f3 imidazole moiety potentially forms a hydrogen bond with D941.32. The aromatic moiety of f3 occupies a hydrophobic subpocket located between TM5 and 6 of the H4 binding site and forms face-to-face stacking interaction with F1835.47.

**Virtual screening**

An in-house vendor screening library of 4.8M compounds was overlapped with the docking poses of 44 (H3 selective), 21 (H4 selective) and 20 (dual active) fragment hits using the OMEGA and ROCS tools. The ComboScore top ranked 7,500 (H3), 2,000 (H4) and 1,500 (dual) compounds were selected for docking, manual inspection and SDM (pharmacophore) filtering. A total of 172 virtual hits (19 compounds predicted as H3 selective, 79 as H4 selective and 74 as dual active) were purchased for further biological testing. While not all of the virtual hits contained the core scaffold of their "parent" fragment, they preserved the pharmacophore, key interactions and other binding features of the original fragment hit. The results and the statistics of the VS are summarized in Table 2.

Overall, 172 compounds were purchased for testing based on the virtual screening results. Of the 74 compounds predicted to have dual activity, 33 had activity against one or other of the two receptors (44%), of which 17 had activity against both (51%). Of the 19 compounds predicted to be H3 selective, 13 were active against H3 (68%) and 10 of these also had selectivity over H4 (76%). Of the 79 compounds predicted to be H4 selective, 36 were active against H4 (45%) and 2 of these also had selectivity over H3 (5%).

A few examples of H3 virtual hits are shown in Figure 6A and 6B. The docking pose of H3 selective virtual hit 29 (Figure 6A) is illustrated in Figure 7A. The basic centre of 29 forms a salt bridge with D1143.32 and the aromatic moiety and piperazine ring of 29 form face-to-face and face-to-face π-stacking with W3716.48. The cyclopentane ring of 29 occupies hydrophobic pocket S1 (Figure 4B). The docking pose of the H3 virtual hit 166 (Figure 6B) is illustrated in Figure 7B. The basic centre of 166 forms a salt bridge with E2065.46, with the benzimidazole core forming hydrogen bond and face-to-face π-stacking interactions with W3716.48. The hydrogen donor of the 166 benzimidazole core forms a hydrogen bond with D11413.32. The linker of 166 forms hydrophobic contacts with W3716.48.

**Conclusion**

We present here a practical approach that integrates GPCRmodelling with fragment based screening to provide structural insight into the H3 and H4 histamine receptor binding sites, followed by VS. This FADD-modelling approach creates a cost-efficient new avenue for SBDD against GPCR targets. This method extends the boundaries of traditional FBDD where the practicality and cost-efficiency of this approach is enhanced by the generation of structural information on the binding site of a GPCR, to satisfy the immediate need of the
Conventional fragment-based lead generation would typically envisage iterative rounds of fragment expansion through cataloguing combined with medicinal chemistry efforts. In an effort to accelerate and streamline this process, we have successfully achieved hit expansion via the application of VS against an optimized GPCR model, immediately after fragment screening. The identified VS hits themselves can now be used as tools for the further refinement of the existing GPCR models. The LowModeMD simulation protocol used in our work [1, 91], followed by flexible docking, has gone beyond the use of static models and allowed for a more detailed exploration of the H3 and H4 receptors. The structural insights gained from this process are critical for the discovery of a new generation of H3 and H4 antagonists.

**References**

1. Heifetz A, Barker O, Verquin G, Wimmer N, Meutermans W, et al. (2013) 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. J Chem Inf Model 53: 1084-1099.
2. Sirci F, Istyastono EP, Vischer HF, Kooistra AJ, Nijmeijer S, et al. (2012) Virtual fragment screening: discovery of histamine H3 receptor ligands using ligand-based and protein-based molecular fingerprints. J Chem Inf Model 52: 3308-3324.
3. Law R, Barker O, Barker JJ, Hesterkamp T, Godemann R, et al. (2009) The multiple roles of computational chemistry in fragment-based drug design. J Comput Aided Mol Des 23: 459-473.
4. Albert JS, Blomberg N, Breeze AL, Brown AJ, Burrows JN, et al. (2007) An integrated approach to fragment-based lead generation: philosophy, strategy and case studies from AstraZeneca’s drug discovery programmes. Curr Top Med Chem 7: 1600-1629.
5. Gbahou F, Vincent L, Humbert-Claude M, Tardivel-Lacombe J, Chabret C, et al. (2008) Compared pharmacology of human histamine H3 and H4 receptors: structure-activity relationships of histamine derivatives. Br J Pharmacol 147: 744-754.
6. Hofstra CL, Desai PJ, Thurmond RL, Fung-Leung WP (2003) Histamine H4 receptor mediates chemotaxis and calcium mobilization of mast cells. J Pharmacol Exp Ther 305: 1212-1221.
7. Cole DC, Gross JL, Comery TA, Aschmies S, Hirst WD, et al. (2004) Benzimidazole and indole-substituted 1,3-bipyrrolidine benzamides as histamine H3 receptor antagonists. Bioorg. Med Chem Lett 14: 2689-2693.
8. Letavic MA, Aluisio L, Atack JR, Bonaventure P, Carruthers N, et al. (2010) J Pre-clinical characterization of aryloxypyridine amides as histamine H3 receptor antagonists: identification of candidates for clinical development. Bioorg Med Chem Lett 20: 4210-4214.
9. Cramp S, Dyke HJ, Higgs C, Clark DE, Gill M, et al. (2010) Identification and hit-to-lead exploration of a novel series of histamine H4 receptor inverse agonists. Bioorg Med Chem Lett 20: 2516-2519.
10. Engelhardt H, Smits RA, Leurs R, Haaksma E, de Esch IJ (2009) A new generation of anti-histamines: Histamine H4 receptor antagonists on their way to the clinic. Curr Opin Drug Discov Devel 12: 628-643.
11. Gernkow MJ, Davenport AJ, Harich S, Ellenbroek BA, Cesura A, et al. (2009) The histamine H3 receptor as a therapeutic drug target for CNS disorders. Drug Discov Today 14: 509-515.
12. Leurs R, Bakker RA, Timmerman H, de Esch IJ (2005) The histamine H3 receptor: from gene cloning to H3 receptor drugs. Nat Rev Drug Discov 4: 107-120.
13. Thurmond RL, Gefand EW, Dunford PJ (2008) The role of histamine H1 and H4 receptors in allergic inflammation: the search for new antihistamines. Nat Rev Drug Discov 7: 41-53.
14. Martineli Lamas DJ, Croci M, Carabajal E, Crescenti EJ, Sambuclo L, et al.
34. Worth CL, Kleinau G, Krause G (2009) Comparative sequence and structural analyses of G-protein-coupled receptor crystal structures and implications for molecular models. PLoS One 4: e7011.

35. Michino M, Abola E: GPCR Dock 2008 participants, Brooks CL 3rd, Dixon JS, Moul J, et al. (2009) Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. Nat Rev Drug Discov 8: 455-463.

36. Fanelli F, De Benedetti PG, Raimondi F, Seebeger M (2009) Computational modeling of intramolecular and intermolecular communication in GPCRs. Curr Protein Pept Sci 10: 173-185.

37. Fanelli F, De Benedetti PG (2005) Computational modeling approaches to structure-function analysis of G protein-coupled receptors. Chem Rev 105: 3297-3351.

38. Katritch V, Abagyan R (2011) GPCR agonist binding revealed by modeling and crystallography. Trends Pharmacol Sci 32: 637-643.

39. Kufareva I, Rueda M, Katritch V, Stevens RC, Abagyan R: GPCR Dock 2010 participants (2011) Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. Structure 19: 1108-1126.

40. de Graaf C, Reif C, Pwnica D, Giordanetto F, Rognan D (2011) Structure-based discovery of allosteric modulators of two related class B G-protein-coupled receptors. ChemMedChem 6: 2159-2169.

41. Hubbard RE (2011) Structure-based drug discovery and protein targets in the CNS. Neuropharmacology 60: 7-23.

42. Sela I, Golan G, Strajbl M, Rivenzon-Segal D, Bar-Haim S, et al. (2010) G protein coupled receptors in silico drug discovery and design. Curr Top Med Chem 10: 638-656.

43. Millar RP, Newton CL (2010) The year in G-protein coupled receptor research. Mol Endocrinol 24: 261-274.

44. Schneider M, Wolf S, Schlitter J, Gwerlt K (2011) The structure of active opioid as a basis for identification of GPCR agonists by dynamic homology modelling and virtual screening assays. FEBS Lett 585: 3587-3592.

45. Sage C, Wang R, Jones G (2011) G-protein coupled receptors virtual screening using genetic algorithm focused chemical space. J Chem Inf Model 51: 1754-1761.

46. Vilas S, Ferino G, Phatak SS, Berk B, Cavasotto CN, et al. (2011) Docking-based virtual screening for ligands of G protein-coupled receptors: not only crystal structures but also in silico models. J Mol Graph Model 29: 614-623.

47. Jacob L, Hoffmann B, Steven V, Vert JP (2008) Virtual screening of GPCRs: an in silico chemogenomics approach. BMC Bioinformatics 9: 363.

48. Radestock S, Weit T, Renner S (2008) Homology model-based virtual screening for GPCR ligands using docking and target-biased scoring. J Chem Inf Model 48: 1104-1117.

49. Engel S, Skoumbourdis AP, Childress J, Neumann S, Deschamps JR, et al. (2008) A virtual screen for diverse ligands: discovery of selective G protein-coupled receptor antagonists. J Am Chem Soc 130: 5115-5123.

50. Cavasotto CN, Orry AJ, Murgolo NJ, Czarniecki MF, Kossi SA, et al. (2008) Discovery of novel chemotypes to a G protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. J Med Chem 51: 581-588.

51. Chen JZ, Wang J, Xie XQ (2007) GPCR structure-based virtual screening approach for CB2 antagonist search. J Chem Inf Model 47: 1626-1637.

52. Evers A, Klabunde T (2005) Structure-based drug discovery using GPCR homology modeling: successful virtual screening for antagonists of the alpha 1A adrenergic receptor. J Med Chem 48: 1088-1097.

53. Tautermann CS (2011) The use of G-protein coupled receptor models in lead optimization. Future Med Chem 3: 799-821.

54. Ballesteros JA, Weinstein H (1995) Integrated methods for construction three dimensional models and computational probing of structure-function relations in G protein-coupled receptors. Methods Neurosci 25: 366-428.

55. Heftzet A, Morris GB, Biggin PC, Barker O, Frysty T, et al. (2012) 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 51: 3178-3197.

56. Shimamura T, Shiraiishi M, Weyand S, Tsujimoto H, Winter G, et al. (2011) Structure of the human histamine H1 receptor complex with doxepin. Nature 475: 65-70.

57. Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 48: 443-453.
65. Gloriam DE, Fechteler T, Baldini PN, Brown DR, Langelaan DN, Kobilka TS, Horvath D (2013) The approach to conformational search of macrocycles and protein loops. J Chem Inf Model 53: 88-92.

66. Dickinson SB, Kall L, Knecht R, Sonnhammer EL (2006) A general model of G protein-coupled receptor sequences and its application to detect remote homologs. J Mol Biol 359: 926-945.

67. Kall L, Knecht R, Sonnhammer EL (2006) A general model of G protein-coupled receptor sequences and its application to detect remote homologs. J Mol Biol 359: 926-945.

68. Jorgensen W, Maxwell DS, Tirado RJ (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc 118: 11225-11236.

69. Kaminski GA, Friesner J, Jorgensen W (2001) Evaluation and reparameterization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J Phys Chem B 105: 6474-6487.

70. Courcot B, Bridgeman AJ (2011) Modeling the interactions between polyoxometalates and their environment. J Comput Chem 32: 3143-3153.

71. Jones G, Willett P, Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol 245: 43-53.

72. Liebeschuetz JW, Cole JC, Korb O (2012) Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. J Comput Aided Mol Des 26: 737-748.

73. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, et al. (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33: 889-897.

74. Homeyer N, Gohlke H (2013) FFEA: a workflow tool for free energy calculations of ligand binding. J Comput Chem 34: 965-973.

75. Hoffer L, Horvath D (2013) S4MABLE--sample for multiple protein-ligand entities: simultaneous docking of several entities. J Chem Inf Model 53: 88-102.

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

77. Berendsen HJC, Postma JPM, Van Gunsteren WF, Dinola A, Haak JR (1984) Molecular-Dynamics with Coupling to an External Bath. J Chem Phys 81: 3684-3690.

78. Grant JA, Pickup BT, Sykes MJ, Kitchen CA, Nicholls A (2007) The Gaussian Generalized Born model: application to small molecules. Phys Chem Chem Phys 9: 4913-4922.

79. Hawkins PC, Nicholls A (2012) Conformer generation with OMEGA: learning from the data set and the analysis of failures. J Chem Inf Model 52: 2919-2936.

80. Hawkins PC, Skillman AG, Nicholls A (2007) Comparison of shape-matching and docking as virtual screening tools. J Med Chem 50: 74-82.

81. Rasmussen SG, Choi JH, Rosenbaum DM, Kobilka TS, Thian FS, et al. (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. Nature 450: 383-387.

82. Istyastono EP, de Graaf C, Esch IJ, Leurs R (2011) Molecular determinants of selective agonist and antagonist binding to the histamine H4 receptor. Curr Top Med Chem 11: 661-679.

83. Smits RA, Adami M, Istyastono EP, Zuiderveld OP, van Dam CM, et al. (2010) Synthesis and QSAR of quinazoline sulfonamides as highly potent human histamine H4 receptor inverse agonists. J Med Chem 53: 2390-2400.

84. Lin HD, de Graaf C, Jiang W, Sadek P, McGovern PM, et al. (2010) Molecular determinants of ligand binding to H4R species variants. Mol Pharmacol 77: 743-747.

85. Wijtmans MG, de GC, de KG, Istyastono EP, Smits J, et al. (2011) Triazole ligands reveal distinct molecular features that induce histamine H4 receptor affinity and subtly govern H4/H3 subtype selectivity. J Med Chem 54: 1693-1703.

86. Carlen DL, Altenbach C, Yang K, Hubbell WL, Khorana HG (1996) Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. Science 274: 768-770.

87. Gether U, Lin S, Ghanouni P, Ballesteros JA, Weinstein H, et al. (1997) Agonists induce conformational changes in transmembrane domains III and VI of the beta2 adrenergic receptor. EMBO J 16: 6737-6747.

88. Shoichet BK, Kobilka BK (2012) Structure-based drug screening for G-protein-coupled receptors. Trends Pharmacol Sci 33: 268-272.

89. Congreve M, Langmead CJ, Mason JS, Marshall FH (2011) Progress in structure based drug design for G-protein coupled receptors. J Med Chem 54: 4283-4311.

90. Heifetz A, Barker O, Morris GB, Law RJ, Slack M, et al. (2013) Toward an understanding of agonist binding to human orexin-1 and orexin-2 receptors with protein-coupled receptor modeling and site-directed mutagenesis. Biochemistry 52: 8246-8260.