J Comput Aided Mol Des
DOI 10.1007/s10822-007-9127-x

Generation of a homology model of the human histamine H_3 receptor for ligand docking and pharmacophore-based screening

Birgit Schlegel · Christian Laggner · Rene Meier · Thierry Langer · David Schnell · Roland Seifert · Holger Stark · Hans-Dieter Höltje · Wolfgang Sippl

Received: 20 November 2006 / Accepted: 29 June 2007

Abstract The human histamine H_3 receptor (hH_3R) is a G-protein coupled receptor (GPCR), which modulates the release of various neurotransmitters in the central and peripheral nervous system and therefore is a potential target in the therapy of numerous diseases. Although ligands addressing this receptor are already known, the discovery of alternative lead structures represents an important goal in drug design. The goal of this work was to study the hH_3R and its antagonists by means of molecular modelling tools. For this purpose, a strategy was pursued in which a homology model of the hH_3R based on the crystal structure of bovine rhodopsin was generated and refined by molecular dynamics simulations in a dipalmitoylphosphatidylcholine (DPPC)/water membrane mimic before the resulting binding pocket was used for high-throughput docking using the program GOLD. Alternatively, a pharmacophore-based procedure was carried out where the alleged bioactive conformations of three different potent hH_3R antagonists were used as templates for the generation of pharmacophore models. A pharmacophore-based screening was then carried out using the program Catalyst. Based upon a database of 418 validated hH_3R antagonists both strategies could be validated in respect of their performance. Seven hits obtained during this screening procedure were commercially purchased, and experimentally tested in a [^{3}H]N$^\alpha$-methylhistamine binding assay. The compounds tested showed affinities at hH_3R with K_i values ranging from 0.079 to 6.3 μM.

Keywords Histamine · H_3 receptor · Docking · Molecular dynamics · GOLD · Catalyst · Pharmacophore · Rhodopsin

B. Schlegel · H.-D. Höltje
Institute of Pharmaceutical Chemistry,
Heinrich-Heine-Universität Düsseldorf,
Universitätsstr. 1, 40197 Dusseldorf, Germany

C. Laggner · T. Langer
Institute of Pharmaceutical Chemistry,
Leopold-Franzens-Universität Innsbruck,
Innrain 52, Innsbruck 6020, Austria

R. Meier · W. Sippl (✉)
Department of Pharmaceutical Chemistry,
Martin-Luther Universität Halle-Wittenberg,
Wolfgang-Langenbeckstr. 4, 06120 Halle/Saale, Germany
e-mail: wolfgang.sippl@pharmazie.uni-halle.de

D. Schnell · R. Seifert
Department of Pharmacology and Toxicology,
Institute of Pharmacy, University of Regensburg,
93040 Regensburg, Germany

H. Stark
Institute of Pharmaceutical Chemistry, Biozentrum,
ZAFES/CMP, Johann Wolfgang Goethe-Universität,
60438 Frankfurt, Germany

Introduction

The histamine H_3 receptor was discovered in 1983 by Arrang and co-workers [1] and has been the focus of intense research for over more than 20 years since then. Recently, several review articles have been published on the histamine H_3 receptor, [2] H_3R isoforms, [3, 4] on H_3R antagonists [5, 6] and agonists, [7] which summarise the current knowledge on this receptor. Briefly, the hH_3R is a GPCR protein expressed presynaptically in several regions of the central and peripheral nervous system where it functions either as an autoreceptor regulating the release of histamine from histaminergic neurons or as an heteroreceptor regulating the release of several other

neurotransmitters. Due to this regulatory function it is expected that the hH_3R could be exploited as a potential target for several therapeutic applications including obesity, cognitive disorders and insomnia.

In recent years, the number of hH_3R ligands has rapidly increased due to the combined effort of university research groups and pharmaceutical companies. In order to find new hH_3R ligands, initially, derivatives of the intrinsic ligand histamine were generated leading to the class of imidazole-containing compounds [for a review see 6, 7]. Due to several potential drawbacks of the imidazole-moiety (interaction with P450 enzymes, substrate of the inactivating enzyme, histamine N^{τ}-methyl transferase, and low CNS penetration[1]) great effort was put into the replacement of this moiety, resulting in the nowadays heterogeneous class of non-imidazole ligands containing mostly piperidine, pyrrolidine or structurally related groups [10, for a review see 5, 7]. In most cases new compounds were designed from scratch or by variation of hits found during *in vitro* screening of large compound libraries. Examples of hH_3R antagonists are given in Fig. 1.

Recently, the successful application of in silico screening tools such as ligand docking or pharmacophore based screening for retrieving GPCR antagonists was shown, [11–15]. Several in silico studies have already been carried out on the hH_3R focusing however mainly on the placement of ligands in the binding pocket and on the derivation of a putative binding site for reasoning the design of new compounds rather than on an automatic screening for new compounds. In 2000 De Esch et al. [16, 17] published a study on imidazole-containing hH_3R ligands and proposed a pharmacophore model consisting of a common anchor site for the imidazole moiety which was expected to interact with E5.46 in helix 5, and two lipophilic pockets. At the same time branched compounds published by Schering in the patent application WO00/53596 confirmed the existence of two lipophilic pockets.

The first H_3R homology model based on the crystal structure of bovine rhodopsin was published in 2001 by Sippl et al. [18] with the aim of explaining the striking species differences observed for some antagonists on the rat versus the human hH_3R [19, 20]. In the complexes studied, the imidazole moiety of antagonists was—in analogy with the imidazole moiety of histamine—assumed to interact with E5.46[2]. In 2002, Uveges and co-workers studied the natural agonist histamine in an hH_3R homology model [22]. Histamine was manually placed such as to contact E5.46 with its imidazole moiety and D3.32 with its primary amine functionality. In order to simultaneously establish an interaction of histamine with D3.32 and E5.46, manual adjustments of helix 5 were required. Interestingly, Uveges further reported that the mutation E5.46A had only minor effects on the binding of the protean ligand [^{125}I]iodoproxyfan, which was later supported also by Jacobsen et al. who analysed binding of iodoproxyfan to an E5.46Q hH_3R mutant [23]. In the same study of Jacobsen and coworkers it was concluded that a common feature of antagonists, which were most affected by the E5.46Q mutation, i.e. iodophenpropit, clobenpropit and NNC-0038–1035, was the presence of structural groups in their side-chain, which could make interactions with the carboxylic acid in E5.46. Implicitly, it was thus suggested that the imidazole moiety of these antagonists was in contact with D3.32.

In 2003, Yao and co-workers [24] further attempted to explain species differences observed for the binding of antagonists such as A-304121, and showed that by carrying out the point mutations A3.37T and V3.40A in the rat H_3R, the binding profile of the human H_3R was restored. Different to the model of Sippl et al. which suggested an indirect influence of amino-acids varying between species, in the model of Yao antagonists made a direct contact to those residues, resulting in a ligand placement extending from D3.32 orthogonal to the membrane plane down to residue D2.50 [25, 26]. In 2005, a model of the rat H_3R was published by Lorenzi et al. which was used to guide the successful design of further imidazole-containing H_3R compounds [27]. Antagonists were placed into the homology model starting from the hypothesis that their imidazole ring interacted with E5.46. Very recently, a further modelling study on the hH_3R was published by Axe et al. [28] in which complexes of the hH_3R with bi-cationic antagonists were studied by means of MD simulation in a continuum dielectric membrane model. The compounds were manually docked such as to contact both D3.32 and E5.46. The question of how mono-cationic compounds would be oriented in the binding site had not been addressed.

In the present work, two well established tools for in silico screening, namely molecular docking into a rhodopsin-based homology model and a pharmacophore based search, were carried out for the target hH_3R. The hH_3R ligand dataset is extremely challenging in this respect as the ligands are in most cases highly flexible and different inverse agonist classes interact with different sets of receptor site points. Furthermore, the low number of mutational studies published for the hH_3R merely indicated the approximate position of the ligands in the hH_3R model rather than giving a detailed view on the amino acid side chains involved in ligand binding. Still,

[1] low CNS penetration can also be an advantage in situations where a peripheral application of hH3 ligands is pursued, such as the design of nasal decongestants by Schering or cytoprotective agents. [8, 9]

[2] numbering scheme according to Ballesteros et al. [21]: the most conserved residue in each transmembrane segment is assigned position 50. The first number refers to the helical segment.

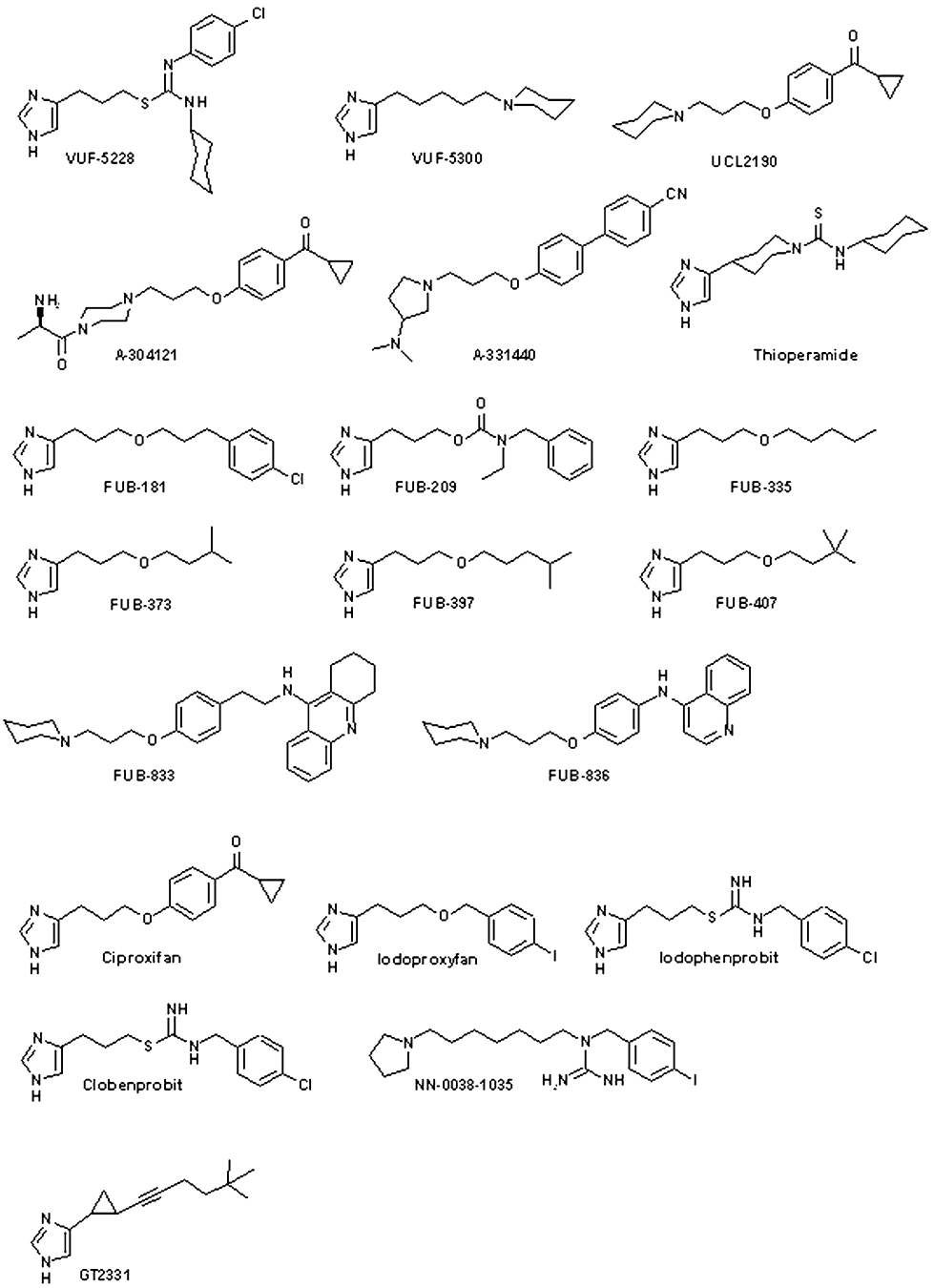

Fig. 1 Structures of hH_3R antagonists mentioned in the text

the existence of a large dataset of experimentally tested hH_3R inverse agonists aided the generation and the subsequent validation of hH_3R homology- and ligand-based pharmacophore-models.

In order to obtain a binding site suitable for subsequent ligand docking, complexes of the hH_3R with antagonists were simulated in a DPPC/water environment. The protonated headgroup of the ligands was thereby oriented such as to contact D3.32. The resulting receptor binding site and the pharmacophore models were then applied in a virtual screening experiment using a validated data set of known hH_3R ligands. A significant portion of validated actives could be retrieved by applying either method indicating that both the generated receptor binding site and the pharmacophore models are suitable for virtual screening. In order to test the predictive value of the generated model, seven hits obtained during the screening procedures were purchase from the Maybridge Database (MDB) and tested for their affinity in a $[^3H]N^{\alpha}$-methylhistamine binding experiment.

Materials and methods

Generation of homology models of the hH_3R and hH_3R ligand complexes

A model of the human histamine H_3 receptor was generated based on the crystal structure 1HZX of bovine rhodopsin [29]. The initial sequence-structure alignment was based on multiple sequence alignments, the prediction of secondary structure, transmembrane helices and highly conserved residues identified by Ballesteros et al. [21] and resulted to be identical to the alignment shown by Mor and coworkers [27]. After truncating the 3rd intracellular loop to a comparable length as present in the template structure bovine rhodopsin by excising the stretch A240-Q346 from the hH_3R sequence, amino acid side chain conformations were added using program SCWRL3.0 [30]. One internal water molecule was included in the hH_3R receptor model, which was located in proximity to D2.50 and linked helices 2, 3 and 7 at a comparable position as water molecule 1b in the structure of bovine rhodopsin [26].

For generating a FUB836/hH_3R complex, FUB836 was flexibly docked into a set of alternative hH_3R binding sites generated by assuming alternative rotamers for several amino acid residues lining the binding pocket (see Table 1) using the program GOLD version 2.3 [31]. During this docking procedure, a distance constraint was applied between the piperidyl-nitrogen of FUB836 and D3.32. The resulting complexes were ranked according to the obtained GoldScores and the potential energy of the FUB836 conformation within these complexes was calculated. The highest ranked FUB836/hH_3R complex was used as a starting conformation for MD simulation. The orientation of FUB836 in the hH_3R binding pocket is shown in Fig. 2.

Table 1 Residues used in the approach of inverse docking

TM	Residues
1	–
2	V2.53, C2.57, I2.58, **Y2.61(3)**
3	**W3.28(3)**, L3.29, **D3.32(3)**, **Y3.33(3)**, L3.35, **C3.36(2)**, **T3.37(3)**
4	**Y4.57(2)**
5	L5.39, A5.42, S5.43, **E5.46(3)**, F5.47
6	W6.48, Y6.51, **T6.52(3)**, M6.55, I6,56
7	F7.39, W7.40, L7.42, **W7.43(5)**, S7.46
E2	A5.30

Residues lining the hH3R binding pocket, for which various rotamers had been considered (number of alternative rotamers given in parentheses) or which had been included as additional constraints for docking FUB836. Amino acids in contact with the ligands are made in boldface

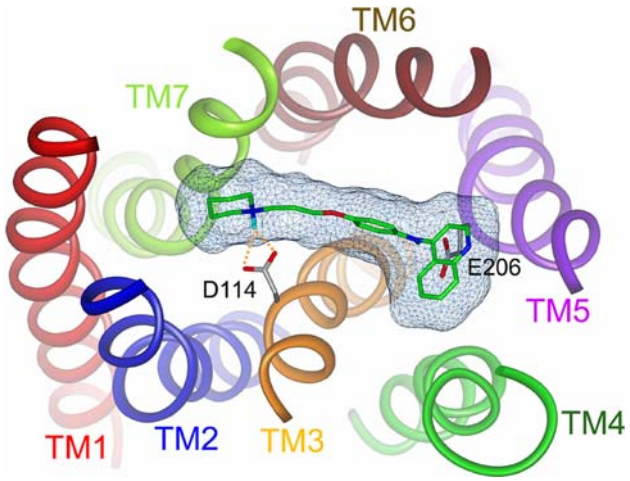

Fig. 2 Orientation of FUB836 in the hH3R binding pocket

Calculation of pK_a-shifts in the hH_3R binding pocket

Calculation of pK_a-shifts was carried out using the program UHBD [32] with default settings. Calculation of pK_a-shifts were carried out for the uncomplexed hH_3R model comprising one internal water molecule, and for a complex of VUF5300/hH_3R, in which the piperidyl-moiety of VUF5300 was interacting with D3.32 and the imidazole group was interacting with E5.46.

Molecular dynamics simulations of hH_3R models

All MD simulations were carried out using the program GROMACS and the *ffG43a1* force filed [33, 34]. For testing the influence of alternative rotamer conformations for specific residues on the resulting binding pocket geometry, MD simulations of uncomplexed hH_3R models in a CCl_4/water membrane mimic were carried out. A CCl_4/solvent box of the dimensions $8.56 \times 6.45 \times 9.01$ nm was generated and the hH_3R models were simulated without applying any constraints on the model. The generation and equilibration of the CCl_4/solvent box, the insertion of the receptor into this box and the subsequent simulation of the receptor were carried out in analogy to the simulation of bovine rhodospin in a CCl_4/solvent box (see [35] for details).

For the simulation of the FUB836/hH_3R ligand receptor complex, a DPPC/water box was used comprising 92 DPPC molecules, 7085 solvent molecules, 11 sodium and 27 chlorine atoms. The insertion of the ligand receptor complex into this membrane mimic, details on equilibration and simulations are described in reference [35] for the analogue simulation of bovine rhodopsin. During the simulation of the FUB836/hH_3R model the following interhelical hydrogen bond contacts were restrained by applying the following

distance constraints: N1.50:ND2-S7.46:O, N2.45:ND2-N3.42:OD1, N2.45:OD1-W4.50:HE1, L7.55:O-R7.61:NH1, D2.50:OD2-H2O:OW, S3.39:OG-H2O:HW2, N7.49:OD1-H2O:HW1 and Y4.57:HH-E5.46:OE2.

Ligand docking into the hH_3R binding site

Docking was carried out using a inhouse dataset of 418 experimentally tested hH_3R antagonists with a range of pK_i from 5.29 to 10.04. The structural diversity/similarity of the 418 hH_3R ligands was analyzed by carrying out a cluster analyis in MOE2006.08 (Chemical Computing Group, Montreal, Canada) using MACCS and graph-3-point-pharmacophore fingerprints. Considering a conservative Tanimoto cutoff of 0.8, 198 (MACCS) and 275 (graph-3-point-pharmacophore) individual clusters were obtained, respectively. Histograms showing the individual cluster populations are given in Fig. 3. Due to the fact that the 418 ligand data set represents the result from a medicinal chemistry guided optimisation strategy, individual clusters show higher population including structurally related analogs.

Ligand docking was carried out using the program GOLD version 2.3 [31] and default parameters except when otherwise indicated. For ligand docking, all compounds were treated as being in their natural protonation state under physiological conditions and all imidazole groups were considered in their protonated form. During the approach of inverse docking of FUB836 into various hH_3R binding pockets and for the validation experiments, a "2-times accelerated" genetic algorithm was used and a distance constraint was set between D3.32 and the protonated head group in order to guarantee the establishment of this ionic interaction. When screening against the unfocused library, the correction term $1/\sqrt{N}$, where N is the number of non-hydrogen atoms [36], was applied to the resulting GoldScores in order to reduce the bias of docking programs towards higher molecular weight compounds. For screening WDI and the MDB, in a first step, ligands comprising a secondary or tertiary amine moiety were selected (compounds comprising primary amines were excluded due their unfavourable physicochemical properties) and a molecular weight cut-off of 600Da was applied resulting in 13,524 compounds. For docking this larger number of compounds, the default parameters for a "library screening" genetic algorithm were applied.

The "receiver operating characteristic" (ROC) curves were calculated to assess the accuracy of the used virtual screening procedure (for details of the method see [37]). ROC curves are obtained by plotting the sensitivity versus the specificity of a virtual screening experiment. Sensitivity is the percentage of truly active compound being selected from the virtual screening workflow and is calculated by dividing the number of true positives by the sum of true positives and false negatives. Specificity, on the other hand is the percentage of truly inactive compounds being correctly identified by the virtual screening experiment. It is calculated by dividing the number of true negative results by the sum of true negatives and false positives. Thus, in ROC curves, the activity signal (i.e. % actives) is plotted versus the detected noise (% inactives) at all possible detection thresholds.

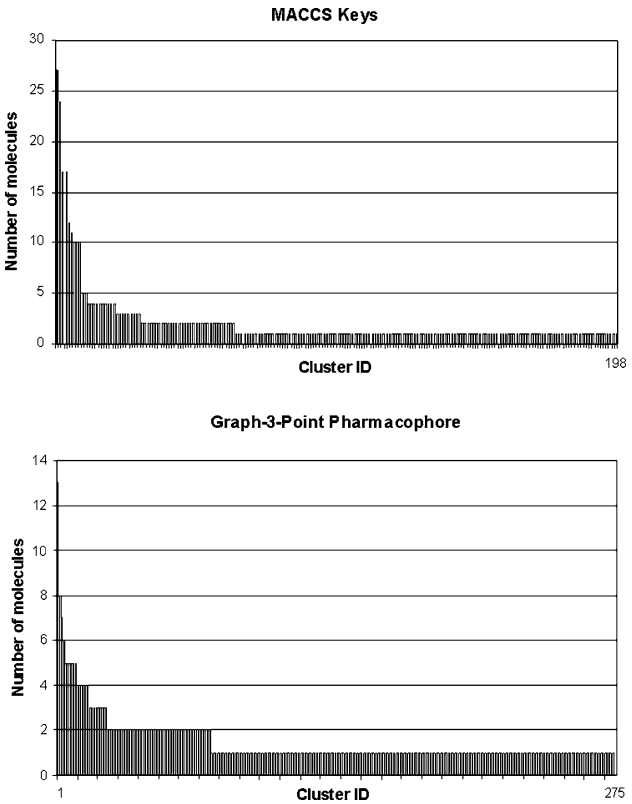

Fig. 3 Cluster analysis of the 418 hH_3R ligand data set. The structural diversity of the studied hH3R ligand data set was analyzed using MACCS keys and graph-3-point pharmacophore fingerprint as similarity metrices. Using a Tanimoto coefficient of 0.8 198 individual clusters were obtained using MACCS keys (Top), whereas the graph-3-point pharmacophore fingerprint yielded 275 clusters (Bottom)

Generation of a focused library

For generating a focused library, the strategy described by Verdonk et al. was followed [38]. Thus, in a first step the distances D(i, j) between all pairs of 138 active hH_3R antagonists with a binding affinity of $K_i < 10$ nM was calculated using formula (1), which takes into account the 1D properties (i) number of hydrogen-bond donors (N_D), (ii) number of hydrogen-bond acceptors (N_A) and (iii) number of nonpolar atoms (N_{NP}).

$$D(i,j) = \sqrt{(N_D(i)-N_D(j))^2+(N_A(i)-N_A(j))^2+(N_{NP}(i)-N_{NP}(j))^2} \quad (1)$$

Then the average distance $D_{\min}$ was calculated as an average of the individual distances D(i, j) over all active compounds, which resulted in $D_{\min} = 0.55$ for the set of 138 inverse hH_3R agonists. A focused library was then generated by chosing ligands from the WDI, which lay within the distance $D_{\min}$ to at least one of the 138 highly active hH_3R compounds and—as a further constraint—contained a secondary or tertiary amine moiety or an imidazole group. From the 3298 unitary WDI compounds, which fulfilled these requirements, 473 structures were randomly chosen for the subsequent validation experiment.

Pharmacophore based screening

Pharmacophore-based screening was carried out using the program Catalyst (Accelrys Inc.: San Diego, CA, 2002). Conformational models were generated for a database of 418 validated hH_3R antagonists and for all compounds of the WDI and MDB by using the default routine of the program. An energy cut-off of 20 kcalmol^{-1} from each energetic minimum structure was set in order to avoid highly energetic structures. Three pharmacophore models were then defined based on the template molecules FUB836, FUB833 and FUB209 in their supposedly bioactive conformation. In more detail, the generation of the pharmacophore model based on FUB836 shall be described. In order to obtain the allegedly bioactive conformation of FUB836, first a conformational analysis was carried out in which each torsion angle was rotated in 15° increments. In the energetic global minimum conformation of the propyloxy linker of FUB836, the protonated piperidyl-nitrogen was pointing towards the aromatic ether atom (C1–N–C2–C3: -159/68/-59; C2 and C3 form part of the propyloxy linker), thereby impeding that the piperidyl-nitrogen could interact with D3.32, when placing this conformation into the hH_3R binding site. Furthermore, in its global minimum conformation, the piperidinopropyl-oxy-fragment of FUB836 could not be overlaid with more rigid hH_3R antagonists such as 1S,2S-GT2331, when common interaction sites were assumed for both compounds. For these reasons, an extended conformation was assumed for the propyloxy-linker (C1–N–C2–C3: -75/-174/-179°), which deviated only 2.4 kJ/mol from the global minimum structure in solution. For the two torsion angles in the spacious aromatic system in the side chain of FUB836, favourable torsion angles were calculated to be within the range [–30–30°] for the bond between the phenyl ring and the secondary amine, while four energetic minima were observed for the bond between the secondary amine group and the quinoline system at –150, –30, 30, and 150°, which were separated by low energetic barriers. These data were in good agreement with the CCD-structure VOTFIT (amodiaquine hydroxide dihydrochloride: 28° and 165°, respectively).

Pharmacophoric features were then directly defined upon this alleged bioactive conformation of FUB836. In a first step, three spheres (see Fig. 4) were defined whereby the red sphere represents a volume in which positively charged moieties and imidazole groups of test compounds have to be accommodated in order to fulfil this pharmacophoric feature; the orange sphere represents linker groups observed in hH_3R compounds (ethers, thioethers, aliphatic un/saturated hydrocarbon chains, cyclopropyl moieties or aromatic ring systems and hydrophobic groups as internally defined by Catalyst); and the cyan sphere represents π-electron rich systems such as aromatic ring systems (predefined in Catalyst), carbamate, ester, urea, and thiourea groups and additionally *t*-butyl moieties. In a next step, the van-der-Waals volume of FUB836 was included as a further constraint. For defining this shape query, default parameters of Catalyst were used, except for the value of similarity tolerance, which was adjusted to a minimum value of 0.45 instead of 0.5 in order to further increase the number of hH_3R antagonists retrieved by this model. Finally, also forbidden volumes (black spheres) were defined in order to account for the fact that some ligands extending into these areas were inactive although resembling other active compounds. An additional

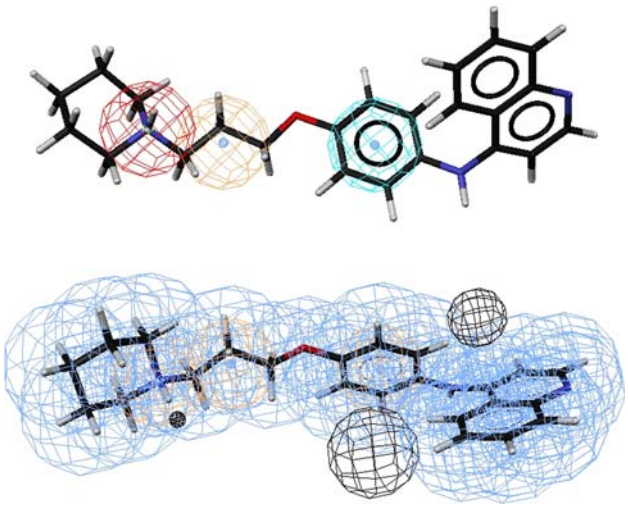

Fig. 4 Top: Pharmacophoric features defined upon the alleged bioactive conformation of FUB836 (see text for interpretation). Bottom: The complete pharmacophore model based on FUB836 additionally including a shape query (blue spheres) and forbidden volumes (black spheres)

forbidden volume was defined in proximity to the space occupied by the positive head-group of hH_3R antagonists in order to avoid larger substituents at this site, which would - if the pharmacophore model was seen in its context with the binding site - produce a clash with D3.32 (see Fig. 4). No pharmacophoric features were defined upon the 4-aminoquinoline moiety as a high degree of chemical diversity was observed in active hH_3R ligands within this region. Any restriction regarding chemical features was thus avoided in the first instance.

For a more stringent screening, a leave-one-out (LOO) filter was defined on the pharmacophoric features of FUB836 as depicted in Fig. 5. The FUB836-LOO model consisted of a combination of five individual pharmacophore models each lacking one pharmacophoric feature found in FUB836 at a time, with the exception of the positive ionisable group and the spacer moiety which were required in all models.

[^{3}H]N$^\alpha$-methylhistamine binding experiments

Competition binding experiments were carried out with Sf9 cell membranes co-expressing the hH_3R, $G\alpha_o$ and $G\beta_1\gamma_2$ complex. Briefly, membranes were thawed and sedimented by a 15-min centrifugation at 4 °C and 15,000*g* to remove residual endogenous guanine nucleotides as far as possible and then resuspended in binding buffer (12.5 mM $MgCl_2$, 1 mM EDTA, and 75 mM Tris/HCl, pH 7.4). For ligand-competition, membranes (15–40 µg of protein per tube, depending on the expression level), 3 nM [^{3}H]NAMH (PerkinElmer Life and Analytical Sciences, Boston, MA) and test compounds (Maybridge, Trevillet, UK) at various concentrations were used. The total volume of the binding reaction was 250 µL. Incubations were performed for 60 min at 25 °C and shaking at 250 rpm. Bound [^{3}H]NAMH was separated from free [^{3}H]NAMH by filtration through 0.3% polyethyleneimine-pretreated GF/C filters using a 48-well brandel harvester (model M-48R, Brandel, Gaithersburg, MD, USA), followed by three washes with 2 mL of binding buffer (4 °C). Filter-bound radioactivity was determined by liquid scintillation counting using Rotiszint® eco plus cocktail (Carl Roth GmbH + Co. KG, Karlsruhe, Germany). The experimental conditions chosen ensured that not more than 10% of the total amount of [^{3}H]NAMH added to binding tubes was bound to filters. All analyses of experimental data were performed with the Prism 4 program (GraphPad Software, San Diego, CA). K_i values were calculated using the Cheng and Prusoff equation [39] and a K_D of 1,081 nM for [^{3}H]NAMH

Results

Generation of a homology model of the hH_3R

A model of the human histamine H_3 receptor missing the stretch A240-Q346 was generated based on the backbone coordinates of the crystal structure 1HZX of bovine rhodopsin [29]. Favourable side chain conformations were added using the program SCWRL3.0 [30]. At two sites, where small residues observed in the structure of bovine rhodopsin were mutated to sterically more demanding residues, steric clashes persisted involving

- residues Y3.33 and Y4.57, and
- residues Y2.61, W3.28, W7.40 and W7.43

In order to find reasonable placements for these residues, two strategies were followed:

For residues Y3.33 and Y4.57, molecular dynamics (MD) simulations of uncomplexed hH_3R models were carried out in a CCl_4/water membrane mimic, using alternative start conformations for residue Y4.57. Based on a better overall structural preservation of the model and a more reasonable hydrogen-bond pattern evolving between residues T3.37, Y4.57 and E5.46, the placement of Y4.57 into the binding site was favoured. Such a placement is also in accordance with the observation that for residue 4.57 an involvement in ligand binding or receptor activation has been reported for other GPCRs [40–42].

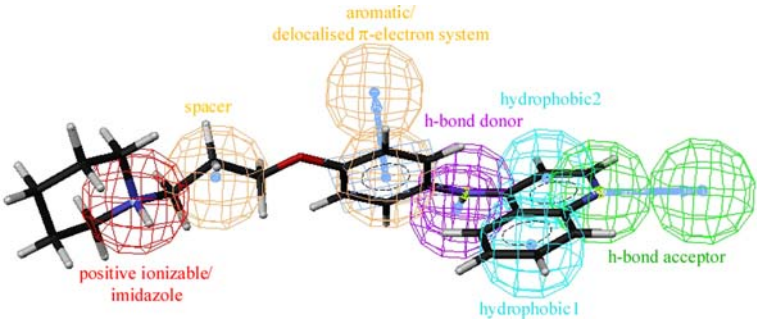

Fig. 5 Features in FUB836 used for the definition of a leave-one-out pharmacophore model. While the positive ionisable moiety/imidazole group and the spacer moiety were required in all models, of all other features each was allowed to be missed in a combinatorial way resulting in a LOO filter embracing five individual filters

For the clash involving Y2.61, W3.28, W7.40, and W7.43 too many placements and corresponding start conformations for MD-simulations would have resulted when following this strategy. Thus, in order to find in an objective way a reasonable conformation for this part of the binding site, a docking procedure of antagonist FUB836 (see Fig. 1) was carried out. Several rotamers were considered for each residue, for which no definite placement could be obtained by applying the SCWRL algorithm and which was likely participating in the binding site. The approximate position of antagonists in the hH_3R binding pocket was thereby known from mutational studies, which showed that D3.32 and E5.46 were the major sites of interaction [22]. During docking FUB836 into all alternative binding sites, a distance constraint between the piperidyl-nitrogen of FUB836 and D3.32 was applied. The resulting complexes were then ranked according to the obtained GoldScore and the potential energy of the FUB836 conformation within this complex. The binding site geometry, for which FUB836 obtained the highest docking score, was considered to be the most likely geometry and thus further used in subsequent MD simulations. Most strikingly, W7.43 was predicted to point into the cleft between helices 1 and 7.

One internal water molecule was included in the receptor model, which was located in proximity to D2.50 and linked helices 2, 3 and 7 at a comparable position as water molecule 1b in the structure of bovine rhodopsin [26].

The final model was submitted to pK_a -shift calculations using the program UHBD [32], which suggested that residue D2.50 was – in analogy to D2.50 in bovine rhodopsin—in its protonated state [25].

Generation of a FUB836-hH_3R complex

By applying the "inverse" docking of FUB836 into the hH_3R binding site, not only a decision on the placement of residues Y2.61, W3.28, W7.40 and W7.43 could be taken, but also a FUB836/hH_3R complex was generated.

Characteristics of this complex were:

- an ionic interactions between the piperidyl-nitrogen and D3.32
- a contact of the exocyclic nitrogen of the 4-aminoquinoline moiety with E5.46
- the sterically demanding quinoline system of FUB836 occupied the gap between helices 3, 4 and 5
- a hydrogen bond interaction of the endocyclic nitrogen of the quinoline system with Y5.34
- accommodation of the propyloxy-linker in an extended conformation in a cleft formed between helices 3, 6 and 7 in proximity to the voluminous leucine residue 7.42.
- cation-π-interactions between the protonated piperidyl-moiety interacting with D3.32 and residues W3.28, F7.39 and W7.40
- T-shaped interactions between the aromatic ring linked to the propyloxy moiety in FUB836 with Y6.51 and
- a parallel-displaced interaction between the aromatic ring linked to the propyloxy moiety in FUB836 with Y5.29 from the second extracellular loop

In order to assess if a pK_a -shift occurred in proximity to E5.46, the complex VUF5300/hH_3R, in which VUF5300 was interacting with its piperidyl-moiety with D3.32 and in which the imidazole moiety was located in proximity to E5.46, was submitted to calculations with the UHBD program. For the imidazole moiety a significant pK_a -shift from 6.5 to 8.6 was predicted in proximity to E5.46, indicating that the interaction will be of electrostatic nature.

When superimposing VUF5300 onto FUB836, the imidazole moiety of VUF5300 (pK_a (imidazole) ~ 6.5, K_i = 8.05 nM) [43] could be superimposed onto the 4-aminoquinoline group of FUB836 (expected pK_a -value based upon the similarity to the compound amodiaquine = 7.53 [44], K_i = 10.04 nM [45]). As an ionic interaction was predicted between the less basic imidazole moiety of VUF5300 and E5.46, the same was assumed for the interaction between the more basic 4-aminoquinoline moiety of FUB836 with E5.46.

MD-simulations of hH_3R models

MD-simulations of an uncomplexed and ligand-complexed hH_3R model were carried out in a DPPC/water environment. During all simulations conserved interhelical hydrogen bond contacts were included as distance constraints. During the simulation of an uncomplexed hH_3R model, residue W7.43, which had been predicted to point into the cleft between helices 1 and 7 by the approach of inverse docking, switched back into a conformation pointing into the binding pocket. As such a placement is not compatible with subsequent ligand docking, the simulation was not further prolonged.

During the simulation of the FUB836/hH_3R complex, the placement of W7.43 in the cleft between helices 1 and 7 was preserved and prompted helix 7 to adopt an idealised helical conformation in proximity to W7.43. Further overall structural adaptations consisted in a slight outward shift of helix 4 and the adoption of an idealised helical conformation of transmembrane segment 1. In proximity to the binding site no significant changes from the start geometry were observed after 1ns of MD simulation. Figure 6 shows the course of root mean square deviation (RMSD) from the start structure during the MD simulation.

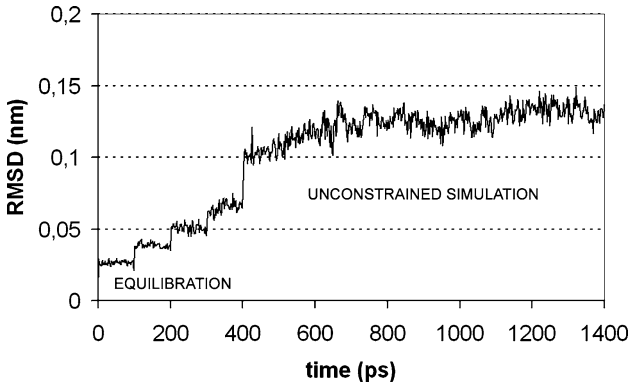

Fig. 6 Course of RMSD during the simulation of a FUB836/hH_3R complex within the backbone of the transmembrane region. The simulation protocol included a stepwise reduction of tether forces in 100 ps time scales from 1000 to 500 to 200 to 100 $kJmol^{-1}nm^{-2}$ (equilibration) before the tethers on the backbone were completely removed (unconstrained simulation)

Validation of the hH_3R binding pocket by screening against a random and focused library

After carrying out an MD-simulation of the FUB836/hH_3R complex, the ligand was removed, and the binding pocket model was validated for its capability to discriminate between a database comprising 418 active hH_3R antagonists with a range of pK_i from 5.29 to 10.04 and a database comprising either 473 randomly chosen compounds from the WDI (screening against a random library) or 473 compounds chosen from the WDI based on their 1D properties, which resembled the 1D properties of active hH_3R antagonists (screening against a focused library). All ligands, hH_3R antagonists and WDI compounds, were docked using the program GOLD and ranked according to their GoldScore. In case of the screening procedure against the random library, the resulting scores were multiplied by the correction term $1/\sqrt{N}$, where N represents the number of non-hydrogen atoms, before the complexes were ranked [36]. Exemplary ligand placements obtained during this docking procedure for an imidazole- and piperidine-containing inverse hH_3R agonist are shown in Fig. 7 for FUB181, and Fig. 8 for UCL2190, respectively.

When screening was carried out against the random library, 11.4% WDI ligands scored among the 80% top ranked hH_3R antagonists reflecting a good discrimination between validated actives and randomly chosen compounds [37]. (Fig. 9, top) When screening was carried out against the focused library, 23% WDI ligands ranked among the top 80% scored hH_3R antagonists, indicating that the discrimination of hH_3R ligands from WDI compounds with similar 1D physicochemical property was significantly more difficult (Fig. 9, bottom).

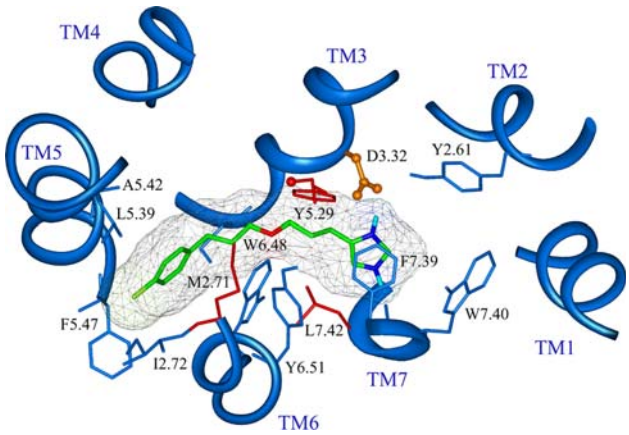

Fig. 7 FUB181 [10] in the hH_3R binding site. Amino acids varying between the hH_3R and hH_4R are shown in red

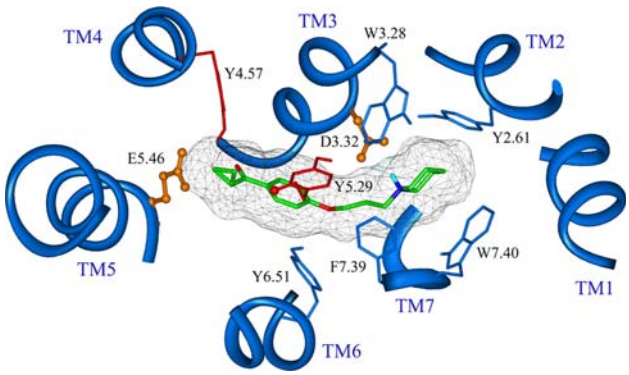

Fig. 8 UCL2190 [10] in the hH_3R binding site. The carbonyl moiety is located in hydrogen bonding distance to Y4.57, which is in turn interacting with T3.37 (not shown) and E5.46. Amino acids varying between the hH_3R and hH_4R are shown in red

Receptor-based virtual screening using the hH_3R binding site

After the ability of the hH_3R binding site to discriminate between actives and non-actives had been verified, the hH_3R binding site was used for virtual screening of WDI and MDB. For this purpose, ligands comprising a secondary or tertiary amine moiety were selected in a first step and a molecular weight cut-off of 600Da was applied resulting in 13,524 compounds. The compounds were docked using the program GOLD and ranked according to their GoldScores. Figure 10 shows a histogram comparing the GoldScores obtained for docking WDI and MDB compounds to the scores obtained when docking the 418 active hH_3R antagonists using the same parameters. As can be seen from Fig. 10, the mean docking score for the hH_3R actives lay in the cluster of [40, 50] and was thus significantly shifted by a value of 20 to higher docking scores, when compared to the mean value of the distribution of

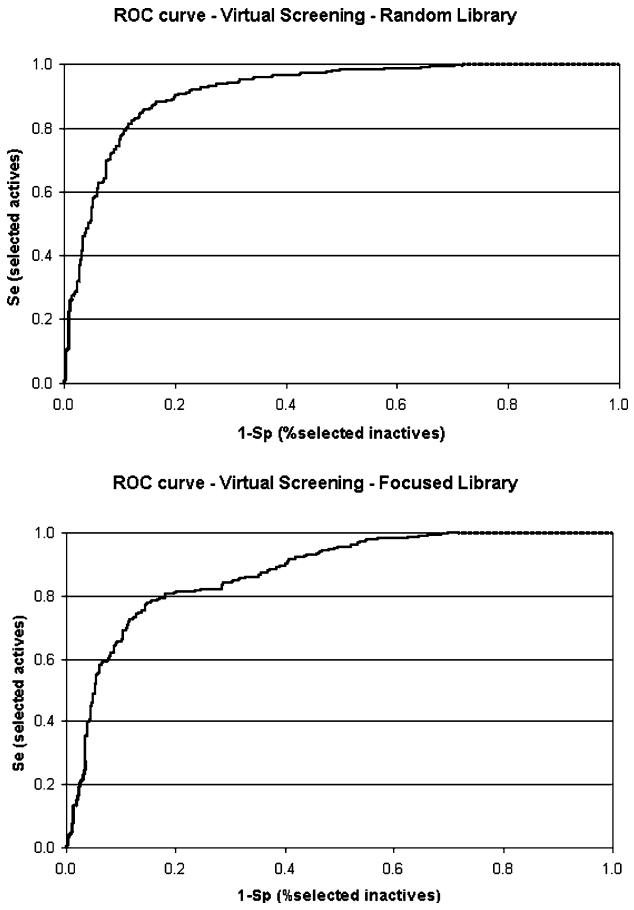

Fig. 9 Top: ROC curve obtained when carrying out GOLD docking of hH_3R actives against a non-focused library of 473 randomly selected compounds from the WDI. Bottom: ROC curve obtained when carrying our docking of 418 hH_3R actives against a focused library of 473 ligands. The corrected GoldScores were used as scoring values

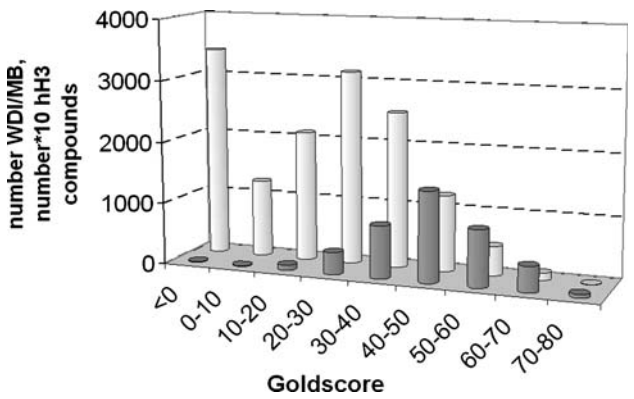

Fig. 10 Comparison of GoldScores obtained when docking WDI/MDB compounds (grey columns) and hH_3R compounds (black columns) into the hH_3R binding site. The distribution of hH_3R compounds is scaled by a factor of 10 in order to facilitate inspection

WDI and MDB compounds ([20, 30]), thereby indicating that by docking into the hH_3R binding site, significant higher scores were in average obtained for validated hH_3R ligands. At the arbitrary chosen GoldScore cut-off of 40, at which 66.5% of the validated hH_3R antagonists would have been retrieved, 87% of the WDI and MDB compounds were filtered out, resulting in 1720 structures, which were further analysed by visual inspection.

Pharmacophore-based virtual screening

Three pharmacophore models were defined based on the template molecules FUB836, FUB833 and FUB209 in their supposedly bioactive conformation. The pharmacophore models were then used to screen MDB and WDI compounds, for which conformational models had been generated. For validation purposes, again the database of 418 active hH_3R antagonists was simultaneously processed in order to assess the quality of the pharmacophore models in terms of retrieving known active hH_3R ligands. Application of the pharmacophore model based on FUB836 shown in Fig. 4 resulted in the retrieval of 316/428 ligands from the hH_3R database comprising 418 active hH_3R antagonists. In order to further increase these percentages, additional pharmacophore models based upon compounds FUB833 (K_i = 0.33nM [45]) and FUB209 (K_i = 69nM [46]). By combining these three models, 369 of 398 (93%) hH_3R ligands with a $pK_i > 7$ could be obtained, while 2668 compounds (2.5% of the entire databases) were retrieved as hits when screening WDI and MDB. Within this subset of 2668 WDI and MDB hits, the more stringent leave-one-out (LOO) filter based on FUB836 (see Fig. 5) was applied, reducing the number of hits to 320.

In order to assess whether all of these 320 compounds selected were also compatible with the hH_3R binding site, the compounds selected via the pharmacophore search were additionally docked using GOLD. After ranking and clustering the candidates into a histogram similar to that shown in Fig. 10, the distribution of candidate compounds and validated hH_3R antagonists interestingly showed the same maximum GoldScores indicating that the pre-screening with Catalyst was successful in selecting compounds that later resulted in a high docking score. By visual inspection we manually selected seven structurally diverse top-ranked compounds from the Maybridge Database fullfilling the pharmacophore requirements for experimental testing (Fig. 11).

Experimental testing of the identified hits

The seven compounds depicted in Fig. 11 were purchased from Maybridge and experimentally tested for binding to the hH_3R in a competition binding experiment as described

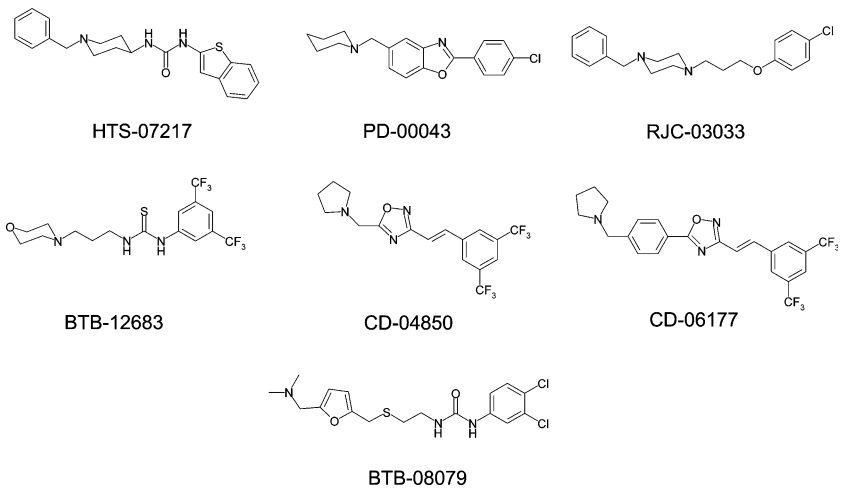

Fig. 11 Hits obtained by screening the MDB with a pharmacophore model based upon FUB836 and subsequently docking the compounds into the hH_3R model

Table 2 Binding properties of compounds shown in Fig. 9 at hH_3R

Compound	K_i (nM)	GoldScore
HTS-07217	2459 (1.510–4.004)	72.89
PD-00043	1024 (599–1749)	79.21
RJC-03033	383 (249–589)	81.10
BTB-12683	3655 (2266–5896)	65.97
CD-04850	6258 (3775–10370)	60.13
CD-06177	2958 (1940–4510)	82.21
BTB-08079	79 (47–131)	87.89

The ligands were tested as described under Materials and Methods. Data shown are the means of two experiments performed in duplicate. Numbers in parentheses represent the 95% confidence intervals. In addition, the GoldScore of the top-ranked docking solution is included

Table 3 Structural similarity between the seven hits and the original 418 hH3R ligand data set

Compound	Graph-3-point pharmacophore	MACCS keys
HTS-07217	0.45	0.65
PD-00043	0.55	0.73
RJC-03033	0.59	0.83
BTB-12683	0.51	0.73
CD-04850	0.47	0.60
CD-06177	0.56	0.63
BTB-08079	0.45	0.57

Tanimoto coefficients between the seven hit structures and the most similar analogs were calculated on the basis of MACCS keys and graph-3-point pharmacophore fingerprints

in the Materials and Methods section. Results are shown in Table 2. All seven compounds are active in the range between 0.079 and 6.3 μM. Two compounds, BTB-08079 and RJC-03033, are active in the nanomolar range. In order to determine the structural similarity between the seven retrieved Maybridge compounds and the 418 hH_3R ligands we calculated similarity indices on the basis of MACCS keys and graph-3-point pharmacophore fingerprints in MOE. In Table 3 the Tanimoto coefficients of the closest neighbour in the 428 ligand dataset are listed. Using the graph-3-point pharmacophore fingerprint the similarity cutoffs are all below 0.60 indicating the low structural similarity between the seven hits and the original hH_3R ligand structures. Using the MACCS key, we found several piperidine derivatives in the original dataset as analogs of RJC-03033. Interestingly, for the most potent hit BTB-08079 (79 nM), the lowest similarity with the original hH3R ligand structures was observed. The dimethylaminofuran fragment, which is already known from the potent histamine H_2 receptor antagonist Ranitidine, has not been reported so far as structural element of potent H_3R antagonists.

In order to analyze the binding orientation of the seven Maybridge compounds, we docked them again in the hH_3R binding pocket using GOLD and standard default docking settings (highest docking accuracy). No constraints were used for this redocking. The obtained GoldScores are shown in Table 2. As known from many docking studies, only a moderate correlation ($r^2 = 0.65$) between the docking scores and the K_i values was observed for the seven compounds. However, the most potent inhibitor BTB-08079 yielded the highest GoldScore value. The interaction between BTB-08079 and hH_3R is shown exemplarily in Fig. 12. The dimethylaminofuran group interacts with the residues of the aromatic cage nearby D3.32 (Y2.61, F7.39, and W7.40). In addition the dimethylamino group makes a hydrogen bond to D3.32. The general orientation and conformation of BTB-08079 is similar to the ones observed for the other investigated hH_3R antagonists, e.g. UCL2190 (Fig. 8).

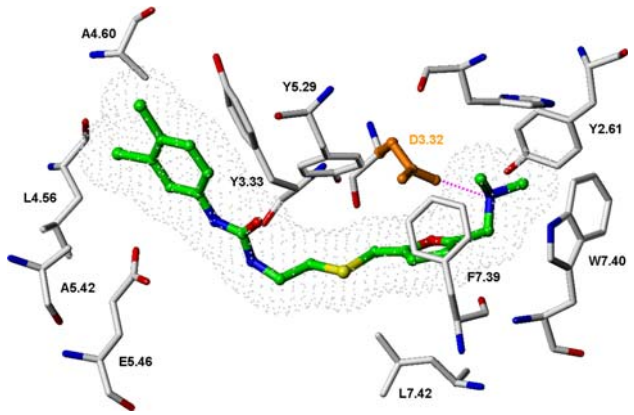

Fig. 12 Docked BTB-08079 in the hH_3R binding site. The key amino acid D3.32 is shown in orange and hydrogen bonds between ligand and receptor are colored magenta. The van-der-Waals volume of the ligand is displayed

Discussion and conclusions

In the present work, two well established tools for in silico screening, namely molecular docking into a rhodopsin-based homology model and a pharmacophore based search, were carried out for the hH_3R. In contrast to the hH_3R models published so far, the model of our work was relaxed by molecular dynamics simulations in an explicit DPPC/water environment, which allowed carrying out the simulation without tether forces on the model backbone, thereby permitting adjustments of helix geometries and topology to take place. In order to avoid model deterioration, which is frequently observed during completely unconstrained MD simulation, [47] conserved interhelical hydrogen bond contacts, which have been previously analysed in simulations of a model of bovine rhodopsin [35], were preserved via the incorporation of distance constraints during the simulations.

Generation of a homology model of the hH_3R and a FUB836/hH_3R complex

The basis for the prior approach was the generation of a suitable homology model of the hH_3R. The main difficulty during the model generation was the generation of a ligand-compatible binding site geometry, which was complicated due to the relatively low number of mutation data and the fact that at several sites small residues in the rhodopsin reference structure had been mutated to more voluminous amino acids in the hH_3R sequence. Thus, after the initial model generation, the binding site was blocked by amino acid side chains and did not allow the automated docking of inverse hH_3R agonists. In order to still allow an objective generation of ligand-receptor complexes, ligand information was included in the placement of amino acids in an "inverse docking" approach, which resulted in a binding site geometry capable of accommodating sterically demanding inverse hH_3R agonists. Different to a "normal'' docking approach, where a binding pocket conformation is used as a filter in screening structure databases for compatible ligands, here, a ligand was used for retrieving the most suitable binding site. FUB836 is a high affinity ligand with a pK_i value of 10.04 [46]. Consequently, the binding conformation of FUB836 should be near to the energetic minimum structure and a good fit between the ligand and the binding site can be expected, which should be reflected in a large docking score. Based on this assumption, FUB836 was flexibly docked into alternative binding site geometries, which varied in the placement of amino acid side chains, for which clashes have been observed after adding side chains using the program SCWRL.

Although a distance constraint between D3.32 and the piperidyl-nitrogen of FUB836 was included during this docking procedure, this constraint merely served to ensure that FUB836 was placed inside the binding pocket and not on the surface of the receptor. The orientation of FUB836 within the binding pocket, i.e. the piperidyl-nitrogen interacted with D3.32, while the aminoquinoline-system interacted with E5.46 was not biased by the inclusion of such a constraint, as the inverse orientation (i.e. piperidyl-nitrogen interacting with E5.46 and aminoquinoline-system interacting with D3.32) was sterically not possible. The binding site geometry of the FUB836/hH_3R complex, which obtained the highest GoldScore and which simultaneously accommodated FUB836 in an energetically favourable conformation (especially in respect to the conformation of the aromatic system in the side chain of FUB836) was chosen for deciding on a placement of conflicting amino acid side chains and as a start conformation for MD-simulation of the FUB836/hH_3R complex.

The incorporation of ligand information into the generation of the binding site has been recently discussed also by Evers et al. for the NK1 receptor and was shown to significantly improve the quality of the obtained binding site [11]. Similar to the approach described here, Evers and co-workers generated 100 preliminary homology models and used a docking approach to choose a suitable binding site geometry. The approach herein described differs in that a flexible docking of the antagonist was carried out and that the ranking of the obtained complexes was based solely on the docking scores rather than the establishment of required contacts known from mutational studies. Although computationally more demanding, this strategy was given preference as no detailed mutational data were available to *a priori* exclude or favour any receptor-ligand complex.

MD-simulations of hH_3R models

In order to relax the energetically still not favourable models, MD simulations of an uncomplexed hH3R model and a FUB836/hH3R complex were carried out in a DPPC/water membran mimic. As an alternative to completely unconstrained MD simulations, which were shown not to be suitable for homology model refinement [47], a number of constraints was included between conserved residues, which ensured that important contacts within the helix bundle were preserved during the simulation. At the same time, structural adaptations of the hH_3R model such as a slight outward shift of helix 4 and the adoption of an idealised helical conformation of transmembrane segment 1 of the hH_3R model were permitted.

Although MD simulations of uncomplexed models would allow for a more efficient sampling of the binding site conformation, residue W7.43 showed the tendency to adopt a rotamer, which was not compatible with the available ligand data. A possible explanation could be the interdependence of backbone coordinates and amino acid side chain placements [30]. Adoption of the backbone geometry of the reference structure bovine rhodopsin would thus prompt all amino acid side chains to adopt a rotamer consistent with these backbone conformations. K7.43 in bovine rhodopsin is involved in the Schiff-base linkage and thus points into the binding site. Therefore, also W7.43 in the hH_3R model would be triggered to adopt a conformation pointing into the binding site independent on the start-conformation imposed. The difficulty of finding a suitable conformation for helix 7 in GPCR homology models has been discussed also by Konvicka [48]. Although this work focused mainly on the kink in proximity to P7.50, Monte Carlo analysis suggested an idealized helical conformation in proximity of residue 7.43 of the $5HT_{2a}$ receptor. An analogous conformation was obtained from the simulation of a complexed hH_3R model, where the explicit consideration of a ligand impeded W7.43 from switching back into the binding pocket.

On the other hand, the difficulties in finding a consistent placement for W7.43 could however reflect more than just shortcomings of a model built on a relatively distant reference structure with an altered backbone conformation. When analysing the SAR of hH_3R agonists and antagonists, it became apparent that the binding site geometry in proximity to D3.32 significantly varied dependent on which compound was binding. While for agonists the binding site in proximity to D3.32 was shown to be sterically quite demanding, the same site appeared to easily accept the more voluminous groups of antagonists [7]. Given the importance of the comparable residue K7.43 in bovine rhodopsin, one could speculate that also W7.43 in the hH_3R could be involved in receptor activation. In such a model, antagonists could trigger W7.43 to adopt an alternative conformation thereby increasing the free volume around D3.32. If such a mechanism held true, again the simulation of complexed models would be more goal-oriented than the simulation of uncomplexed models. Another residue showing a similar "unstable" behaviour was F5.47, which adopted a rotamer pointing into the binding pocket during the simulation of uncomplexed hH_3R models. For the F5.47A variant a significant drop in potency was observed suggesting that this residue was involved in upholding the receptor structure or in receptor activation [22]. In simulations of antagonist/hH_3R complexes the conformational switch of F5.47 towards the inside of the binding pocket was inhibited due to the presence of the ligand. If one assumes a role of F5.47 in activation, the transition from partial agonism to inverse agonism caused by slight structural changes, as for example observed in the series FUB373, FUB335, FUB407 and FUB397 [49] could be correlated to the conformational changes of F5.47. Antagonists FUB335 and FUB397 would thus block the conformational switch of F5.47 due to the structurally more demanding imidazole side chain.

For the goal of obtaining ligand-compatible binding site geometries, simulations of antagonist-hH_3R complexes were thus given preference, although here the sampling efficiency was significantly reduced due to the presence of the ligand. The time of MD-simulations was restricted to 1ns in order to avoid that the resulting binding site would be over fitted to the ligand which had been used in the simulation. The course of RMSD during the simulation of FUB836/ hH_3R is depicted in Fig. 6. Although the RMSD is only an imprecise measure for the quality of a simulation, it allows to assess if an equilibration of a model has occurred, which is indicated by a plateau of the RMSD curve as observed within the transmembrane region in Fig. 6. Main interactions between FUB836 and the hH_3R binding site are listed in the Result section. Another residue in the hH_3R, yet not interacting with FUB836, was methionine 6.55, which stabilised the second aromatic ring in biphenylic systems, such as A-331440 (results not shown) [50].

Validation of the hH_3R binding pocket by screening against a random and focused library

The resulting binding pocket was then validated with respect to its ability to accommodate hH_3R antagonists not used during the MD simulation and to discriminate between validated actives and other randomly chosen compounds. For this purpose, inverse hH_3R agonists and 473 compounds either chosen randomly from the WDI or via the average distance D_{min} (see Methods and Material section) were docked using the program GOLD. During

this docking procedure a distance constraint was set between the polar headgroup of hH_3R antagonists and D3.32 thereby forcing all antagonists into an orientation, where the polar headgroup is located in proximity to D3.32.

Although in most studies so far published [16, 18, 27] the imidazole-moiety of antagonists was assumed to interact with E5.46, an inverse placement was assumed in this work for antagonists containing no further basic moieties in their side-chain, resulting in a placement where the imidazole moiety interacted with D3.32. In our opinion, this hypothesis is supported by the binding affinities measured for the antagonists ciproxifan, thioperamide, clobenpropit and NNC-0038–1035 described by Jacobsen et al. [23]. Thus, when assuming that all imidazole moieties of this set of antagonists would interact with E5.46, all five imidazole-containing compounds should be negatively affected by the mutation E5.46Q. Yet, for ciproxifan an increase in affinity was observed and the affinity of thioperamide changed only slightly. The dramatic loss in affinity observed for clobenpropit and NNC-0038-1035 was explained by Jacobsen et al. by a potentially less favourable interaction of the side chain moieties of these compounds with E5.46. As it had been outlined in this study that several antagonists could interact with D3.32 and E5.46, the imidazole moieties were thus implicitly suggested to interact with D3.32 rather than E5.46. Although it is also possible, that depending on the specific structure of an antagonist, different orientations in the binding pocket could be adopted, the explanation given by Jacobsen et al. i.e. that the imidazole moiety of an antagonist might as well interact with D3.32 appears to be a more sound interpretation of this data set. Furthermore, accommodation of sterically demanding ligands such as FUB836 or FUB833 in an orientation, where the piperidyl-moiety is interacting with E5.46 is sterically not possible due to the extended aromatic system in the side chain of these compounds, which can not be accommodated in proximity to D3.32 without resulting in significant structural distortions of the model during MD-simulations (results not shown).

It thus appeared likely that although the imidazole moiety in histamine interacted with E5.46, antagonists containing no further basic moieties in their side-chain could contact D3.32. The existence of different imidazole binding environments was further supported by the observation that species differences only affected antagonists while agonists showed almost the same affinity at the rat and the human receptors. Both residues responsible for species differences (A3.37T and V3.40A) are located in proximity to E5.46, which is known to interact with the imidazole moiety of histamine and other hH_3R agonists. If the imidazole moiety of antagonist interacted with E5.46 in an analogous way as in hH_3R agonists, no species influence should result for antagonists, as the same structural element as in agonistic compounds (i.e. the imidazole moiety) would be located at the same receptor point (i.e. E5.46), which is however not reflected in available experimental data. Species differences in the model proposed here could be explained via a hydrogen-bond cluster involving E5.46, T3.37 and Y4.57. While Y4.57 was anchored to T3.37 in the human H_3R, the mutation A3.37 would disrupt this interaction resulting in an increased conformational freedom of Y4.57, which could thus more easily interact with functionalities such as carbonyl-containing moieties present in ciproxifan [18] or A-304121 [24], which are most affected by species differences. Compounds which establish a salt-bridge interaction with E5.46 should be less affected by species differences as Y4.57 would not influence ligand binding.

A second assumption made during ligand docking was to consider all imidazole moieties in their protonated form. Although imidazole moieties are only slightly basic in solution (pK_a (imidazole) ~ 6.5), in proximity to an acidic residue (such as glutamic or aspartic acid) significant pK_a -shifts can result as shown for the imidazole moiety of the compound VUF5300 in proximity to E5.46. A corresponding pK_a -shift occurs in proximity to D3.32, thereby making a protonation of imidazole moieties very likely.

For validating the hH_3R binding site obtained from the simulation of the FUB836/hH_3R complex, additionally, a screening procedure was carried out against a focused library, as Verdonk and co-workers showed in a recent study [38], that virtual screening by protein-ligand docking can result in an artificial enrichment when screening against an unfocused library. As a more robust alternative they suggested a validation strategy in which docking scores of actives were compared to the scores obtained when docking a focused library comprising structures with one-dimensional properties, similar to the actives. When comparing the results of screening against an unfocused library (11.4% WDI compounds ranking among the 80% top ranked hH_3R antagonists) to the results obtained when screening against a focused library (23% WDI compounds ranking among the 80% top ranked hH_3R antagonists), a significantly better enrichment was obtained when screening against a non-focused library. This can be explained by the differences in chemical space spanned by the hH_3R antagonists and randomly chosen WDI compounds. Only 52 of the 470 randomly chosen compounds fulfilled the criterion required for a compound to be part of the focused library. Preselected WDI compounds with 1D properties, which resemble those of active hH_3R compounds, have per se an increased likelihood of representing a hit. Thus, top ranked structures from screening such a focused library represented interesting structures with potential affinity at the hH_3R.

Receptor-based virtual screening using the hH_3R binding site

The validated binding pocket was then applied as a filter in screening WDI and MDB. For this purpose, the GOLD genetic algorithm adapted for "library screening" was applied instead of the algorithm adopted for a "2-fold accelerated" screening used in the validation experiment. Although its computational performance regarding speed was significantly better, application of the "library screening" settings also resulted in a worse separation when applied in the validation experiment. In this regard, the hH_3R data set was—due to the high number of rotatable bonds—especially problematic. Although the application of the "2-fold accelerated" screening settings would be recommendable for highly flexible ligands, so far, these settings were computationally too demanding for a feasible screening.

In Fig. 10 the GoldScores obtained for docking WDI and MDB compounds were compared to the scores obtained for docking the hH_3R ligand data set. The distribution of hH_3R antagonists was shifted by a value of 20 to higher GoldScores indicating a satisfactory separation. With a GoldScore cut-off of 40, 66.5% of the validated hH_3R ligands were retrieved while reducing the number of WDI and MDB compounds to 1720 structures. Depending on the cut-off value chosen for visual inspection, a significant percentage of hH_3R compounds was however withheld by the applied filter, which could be problematic as no correlation existed between the docking score and the ligand affinity so that also some high-affinity compounds would be missed. Still, application of the docking procedure and a GoldScore cut-off of 40 increased the number of actives to 13.4% compared to 3.0% in the original database of 13,524 preselected WDI and MDB compounds mixed with 418 active hH_3R ligands.

Pharmacophore-based virtual screening

Due to their high flexibility and huge structural diversity, hH_3R antagonists also hampered the generation of pharmacophore models by standard means which normally include the identification of common features required for binding from a ligand set. This strategy has however the disadvantage that the entropic contribution to the free energy of binding is not sufficiently accounted for. If entropic binders (hydrophobic molecules comprising few functional moieties for which a high affinity results merely due to the fact that the desolvation is so favourable) and enthalpic binders (relatively polar compounds that fit the shape of the binding site in terms of steric and physico-chemical properties, yet have a high desolvation cost) are used for the generation of a common feature model (an automated strategy for the derivation of a pharmacophore model implemented in Catalyst), the presence of entropic binders will result in an underestimation of functional moieties present in the enthalpic binders.

Although in terms of ligand-specificity, a good fit between the ligand and the binding pocket is preferable, the goal in this work was to define a pharmacophore model, which was able to retrieve most of the ligands from the hH_3R subset, i.e. including also entropic binders. For this purpose, relatively loose pharmacophore models as shown in Fig. 4 were defined in the first instance. The choice of chemical moieties was thereby based on chemical functionalities observed in validated hH_3R antagonists and inspection of the binding pocket. The linker moiety abstracted by the orange sphere and the adjacent hydrophobic/π-electron rich system lay in a cleft between helices 3, 6 and 7 of the hH_3R model. In this region, the binding pocket was rather hydrophobic due to residues Y5.29, Y6.51, F7.39 and L7.42. In order to explain how polar groups could also be accommodated in this cleft, one could assume that potential hydrogen bond donor functions were present in this region, however involved in intramolecular hydrogen bond interactions. Thus, in order to establish an interaction with a polar ligand group, an intramolecular interaction would have to be broken up, resulting in a negligible netto-gain of enthalpic binding energy due to the introduced hydrogen bond acceptor. In case of polar groups such as carbamate, ester, urea or thiourea moieties which could be superimposed onto the aromatic/hydrophobic system described by the cyan sphere in Fig. 4, the presence of a π-electron system capable of establishing a π–π-interaction with Y5.29 and T-shape interaction with Y6.51 could represent the commonality. In order to augment the stringency of the pharmacophore model, the molecules' shape and forbidden volumes were included into the pharmacophore model of FUB836 (see Fig. 4, bottom).

Application of the pharmacophore model shown in Fig. 5 resulted in the retrieval of 316/428 ligands from the hH_3R database comprising 418 active hH_3R antagonists. In order to further increase these percentages, additional pharmacophore models were defined in a similar way based upon compounds FUB833 and FUB209. By combining a set of 3 pharmacophore models, 93% of the hH_3R ligand-dataset could be retrieved, while the number of WDI and MDB compounds was reduced to 2668 (=2.5% of the original database).

In a second screening, the application of a leave-one-out filter comprising more pharmacophoric features could then favour the retrieval of compounds that would better fit the physicochemical properties of the hH_3R binding site, which should ensure receptor selectivity. Further screening of the 2668 WDI and MDB compounds with the LOO filter reduced the number of hits to 320. When screening the

database of 418 inverse hH_3R antagonists, only 25% could be retrieved by the LOO-filter, however the filter was capable of filtering out 96% of the inactive ($pK_i < 6$) and 84 % of the moderately active hH_3R antagonists ($6 < pK_i < 7$). In order to ensure that compounds selected by the pharmacophore based screening could be accommodated into the hH_3R binding site, the 320 hits were docked into the hH_3R binding site and ranked according to their GoldScore. From the best scored complexes, seven compounds were chosen for experimental testing. All compounds showed affinity for the hH_3R with binding affinities ranging from 79 nM to 6.3 μM, thereby showing that the pharmacophore model and hH_3R binding site model used for ligand docking also had some predictive value.

Compared to the receptor-based virtual screening by docking, application of the pharmacophore-based search resulted in significantly improved results. While in the docking approach 66.6% of the hH_3R ligands were retrieved, though limiting the number of WDI and MDB compounds to approximately 1720 structures, application of a pharmacophore-based search allowed retrieval of 93% of active compounds, while reducing the number of WDI and MDB structures to 2668 compounds (2.5%). The ideal strategy for the flexible hH_3R ligand data set appeared to be however a combined approach comprising a pre-screening of commercial databases with relatively loose pharmacophore models that mainly reflect the available volume in the binding site (e.g. by considering shape queries of sterically demanding ligands and forbidden volumes derived from ligand superposition) and some essential requirements for binding such as the protonated head group. This way, the number of compounds for the subsequent docking was already significantly reduced which allowed the application of algorithm settings with better discriminatory properties.

Binding mode analysis

The redocking of the identified seven hits (using default docking settings and no docking constraints) showed that they are interacting in a similar way with the H_3 receptor as observed for the other hH_3R ligands under study. The basic nitrogen atom, which is incorporated in a piperidine, piperazine, morpholine, pyrrolidine or dimethylamino group, makes a hydrogen bond to D3.32, whereas the lipophilic parts are interacting with several aromatic residues. The amino acids Y2.61, F7.39 and W7.40 form an aromatic cage nearby the negatively charged glutamate residue (Fig. 12). The docked antagonists are stabilised by a hydrogen bond to D3.32, and show in addition π-cation interactions between their lipophilic and protonated head groups and the residues of the aromatic cage. The lipophilic dichlorophenyl group of BTB-08079 is bound in the hydrophobic cavity nearby helix 5 (Y3.33, L4.56, A4.60, L5.39, and A5.42), where also the aromatic systems of FUB181 or UCL2190 are bound (Figs. 7 and 8). Comparable orientations were derived for the other six compounds. The calculated GoldScores for the seven compounds provide only a qualitative explanation of their biological activities, an observation known from a variety of docking studies [51]. However, the most active compound among the detected hits showed the highest docking score (Table 3).

The analysis of the structural similarity between the VS hits and the original dataset showed that the applied VS strategy was able to identify novel lead structures. Only one compound (RJC-03033) shows higher structural similarity to one of the 418 original antagonists, as indicated by a Tanimoto coefficient above 0.8 using the MACCS keys. Interestingly, the most potent hit BTB-08079 shows no similarity towards the original hH3R ligand data set (calculated either with the MACCS keys or the graph-3-point pharmacophore fingerprint). The dimethylaminofuran fragment, which is already known from the potent histamine H_2 receptor antagonist Ranitidine, has not been reported so far as structural element for a potent H_3R antagonist. Therefore, BTB-08079 represents an interesting lead structure for the further development of novel hH_3R antagonists.

References

1. Arrang JM, Garbarg M, Schwartz JC (1983) Nature 302:832
2. Leurs R, Bakker RA, Timmerman H, De Esch IJ (2005) Nature 4:107
3. Bakker RA (2004) Inflamm Res 53:509
4. Hancock AA, Esbenshade TA, Krueger KM, Yao BB (2003) Life Sci 79:3043
5. Cowart M, Altenbach R, Black L, Faghih R, Zhao C, Hancock AA (2004) Mini Rev Med Chem 4:979
6. Stark H, Kathmann M, Schlicker E, Schunack W, Schlegel B, Sippl W (2004) Mini Rev Med Chem 4:965
7. De Esch IJ, Belzar KJ (2004) Mini Rev Med Chem 4:955
8. McLeod RL, Rizzo CA, West Jr RE, Aslanian R, McCormick K, Bryant M, Hsieh Y, Korfmacher W, Mingo GG, Varty L, Williams SM, Shih NY, Egan RW, Hey JA (2003) J Pharmacol Exp Ther 305:1037
9. Krause M, Stark H, Schunack W (2001) Curr Med Chem 8:1329
10. Meier G, Apelt J, Reichert U, Graßmann S, Ligneau X, Elz S, Leurquin F, Ganellin CR, Schwartz JC, Schunack W, Stark H (2001) Eur J Pharm Sci 13:249
11. Evers A, Klebe G (2004) J Med Chem 47:5381
12. Evers A, Hessler G, Matter H, Klabunde T (2005) J Med Chem 48:5448
13. Bissantz C, Bernard P, Hibert M, Rognan D (2003) Proteins 50:5
14. Byvatov E, Sasse BC, Stark H, Schneider G (2005) ChemBioChem 6:997
15. Becker OM, Marantz Y, Shacham S, Inbal B, Heifetz A, Kalid O, Bar-Haim S, Warshaviak D, Fichman M, Noiman S (2004) Proc Natl Acad Sci USA 101:11304
16. De Esch IJ, Timmerman H, Menge WM, Nederkoorn PH (2000) Arch Pharm Pharm Med Chem 333:254

17. De Esch IJ, Mills JE, Perkins TD, Romeo G, Hoffmann M, Wieland K, Leurs R, Menge WM, Nederkoorn PH, Dean P, Timmerman H (2001) J Med Chem 44:1666
18. Stark H, Sippl W, Ligneau X, Arrang JM, Ganellin CR, Schwartz JC, Schunack W (2001) Bioorg Med Chem Lett 11:951
19. Ligneau X, Morisset S, Tardivel-Lacombe J, Gbahou F, Ganellin CR, Stark H, Schunack W, Schwartz JC, Arrang JM (2000) Br J Pharmacol 131:1247
20. Lovenberg TW, Pyati J, Chang H, Wilson SJ, Erlander MG (2000) J Pharmacol Exp Ther 293:771
21. Ballesteros JA, Weinstein H (1995) Methods Neurosci 25:366
22. Uveges AJ, Kowal D, Zhang Y, Spangler TB, Dunlop J, Semus S, Jones PG (2002) J Pharmacol Exp Ther 301:451
23. Jacobsen ML, Rimvall K, Hastrup S, Pettersson IV, Wulff BS (2003). Poster presentation: XXXII Annual Meeting of the European Histamine Research Society, Noordwijkerhout, Netherlands
24. Yao BB, Hutchins CW, Carr TL, Cassar S, Masters JN, Bennani YL, Esbenshade TA, Hancock AA (2003) Neuropharmacology 44:773
25. Fahmy K, Jäger F, Beck M, Zvyaga TA, Sakmar TP, Siebert F (1993) PNAS 90:10206
26. Okada T, Fujiyoshi Y, Silow M, Navarro J, Landau EM, Shichida Y (2002) PNAS 99:5982
27. Lorenzi S, Mor M, Bordi F, Rivara S, Rivara M, Moroni G, Bertoni S, Ballabeni V, Barocelli E, Placi PV (2005) Bioorg Med Chem 13:5647
28. Axe FU, Bembenek SD, Szalma S (2006) J Mol Graph Mod 24:456
29. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) Science 289:739
30. Canutescu AA, Shelenkov AA, Dunbrack RL (2003) Protein Sci 12:2001
31. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) J Mol Biol 267:727
32. Briggs JM, Madura JD, Davis ME, Gilson MK, Antosiewicz J, Luty BA, Wade RC, Bagheri B, Ilin A, Tan RC, McCammon JA. UHBD (University of Houston Brownian Dynamics), 5th edn. University of Houston
33. van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ (2005) J Comput Chem 26:1701
34. Schuler LD, Daura X, van Gunsteren WF (2001) J Comp Chem 22:1205
35. Schlegel B, Sippl W, Höltje HD (2005) J Mol Model 12:49
36. Pan Y, Huang N, Cho S, MacKerell AD (2003) J Chem Inf Comput Sci 43:267
37. Triballeau N, Acher F, Brabet I, Pin JP, Bertrand HO (2005) J Med Chem 48:2534
38. Verdonk ML, Berdini V, Hartshorn MJ, Mooij WT, Murray CW, Taylor RD, Watson P (2001) J Chem Inf Comput Sci 44:793
39. Cheng Y, Prusoff WH (1973) Biochem Pharmacol 22:3099
40. Lu ZL, Saldanha JW, Hulme EC (2001) J Biol Chem 276:34098
41. Heitz F, Holzwarth JA, Gies JP, Pruss RM, Trumpp-Kallmeyer S, Hibert MF, Guenet C (1999) Eur J Pharmacol 380:183
42. Shin N, Coates E, Murgolo NJ, Morse KL, Bayne M, Strader CD, Monsma FJ (2002) Mol Pharmacol 62:38
43. Wieland K, Bongers G, Yamamoto Y, Hashimoto T, Yamatodani A, Menge WM, Timmerman H, Lovenberg TW, Leurs R (2001) J Pharm Exp Ther 299:908
44. Warhurst DC, Steele JC, Adagu IS, Craig JC, and Cullander C (2003) J Antimicrob Chemother 52:188
45. Apelt J, Ligneau X, Pertz HH, Arrang JM, Ganellin CR, Schwartz JC, Schunack W, Stark H (2002) JMed Chem 45:1128
46. Stark H, Purand K, Ligneau X, Rouleau A, Arrang JM, Garbarg M, Schwartz JC, Schunack W (1996) J Med Chem 39:1157
47. Flohil JA, Vriend G, Berendsen HJ (2002) Proteins Structure Funct Gen 48:593
48. Konvicka K, Guarnieri F, Ballesteros JA, Weinstein H (1998) Biophys J 75:601
49. Nickel T, Bauer U, Schlicker E, Kathmann M, Göthert M, Sasse A, Stark H, Schunack W (2001) Br J Pharmacol 132:1665
50. Faghih R, Dwight W, Vasudevan A, Dinges J, Conner SE, Esbenshade TA, Bennani YL, Hancock AA (2002) Bioorg Med Chem Lett 12:3077
51. Tame JR (2005) J Comput Aided Mol Des. 19:445