

A combined ligand-based and target-based drug design approach for G-protein coupled receptors: application to salvinorin A, a selective kappa opioid receptor agonist

Nidhi Singh · Gwénaél Chevé · David M. Ferguson · Christopher R. McCurdy

Received: 21 June 2006 / Accepted: 15 August 2006 / Published online: 29 September 2006
© Springer Science+Business Media B.V. 2006

Abstract Combined ligand-based and target-based drug design approaches provide a synergistic advantage over either method individually. Therefore, we set out to develop a powerful virtual screening model to identify novel molecular scaffolds as potential leads for the human KOP (hKOP) receptor employing a combined approach. Utilizing a set of recently reported derivatives of salvinorin A, a structurally unique KOP receptor agonist, a pharmacophore model was developed that consisted of two hydrogen bond acceptor and three hydrophobic features. The model was cross-validated by randomizing the data using the CatScramble technique. Further validation was carried out using a test set that performed well in classifying active and inactive mole-

cules correctly. Simultaneously, a bovine rhodopsin based “agonist-bound” hKOP receptor model was also generated. The model provided more accurate information about the putative binding site of salvinorin A based ligands. Several protein structure-checking programs were used to validate the model. In addition, this model was in agreement with the mutation experiments carried out on KOP receptor. The predictive ability of the model was evaluated by docking a set of known KOP receptor agonists into the active site of this model. The docked scores correlated reasonably well with experimental pK_i values. It is hypothesized that the integration of these two independently generated models would enable a swift and reliable identification of new lead compounds that could reduce time and cost of hit finding within the drug discovery and development process, particularly in the case of GPCRs.

Electronic supplementary material Supplementary material is available in the online version of this article at <http://dx.doi.org/10.1007/s10822-006-9067-x> and is accessible for authorized users.

N. Singh · G. Chevé · C. R. McCurdy
Department of Medicinal Chemistry and Laboratory for Applied Drug Design and Synthesis, The University of Mississippi, University, MS 38677, USA

C. R. McCurdy
Department of Pharmacology, School of Pharmacy, The University of Mississippi, University, MS 38677, USA

C. R. McCurdy (✉)
Research Institute of Pharmaceutical Sciences, School of Pharmacy, The University of Mississippi, University, MS 38677, USA
e-mail: cmccurdy@olemiss.edu

D. M. Ferguson
Department of Medicinal Chemistry, The University of Minnesota, Minneapolis, MN 55455, USA

Keywords Salvinorin A · Pharmacophore · G-protein coupled receptor (GPCR) · Comparative modeling · Modeler · Molecular dynamics · Docking · Kappa opioid receptor · Homology · New molecular entities (NMEs)

Introduction

In recent years, virtual screening (VS) has emerged as a complementary method to high-throughput screening of large chemical databases in terms of time-efficiency and cost-effectiveness [1, 2]. VS methods are designed for searching large compound databases in silico and selecting a limited number of candidate molecules for testing to identify novel chemical entities that have the

desired biological activity [3]. There are two fundamental approaches to virtually screen databases for molecules fitting either a known pharmacophore [4, 5] or a 3D structure of the macromolecular target [6–12]. The power of pharmacophore-based methods for lead generation lies in their ability to suggest a diverse set of compounds potentially possessing a desired biological activity but which have totally different chemical scaffolds [13]. Structure-based VS searches a database for compounds that fit into the binding site of the target, based on their chemical feature matching and shape complementarity. Due to the exploitation of structural information, like shape and volume, taken directly from the active site, structure-based methods are more efficient than the pharmacophore-based methods [14]. However, structure-based studies require 3D knowledge of the target. Moreover, structure-based studies are affected by the inaccuracy of the molecular-docking methods [15, 16] that are, in turn, dependent on the computational algorithms for conformational sampling and scoring of different ligand-binding poses [8, 12, 17].

G-Protein coupled receptors (GPCRs) are major targets for drug discovery in the post-genomic era. They constitute a superfamily of membrane receptors of utmost importance in pharmaceutical research [18]. According to a recent analysis, GPCRs account for 40–50% of all current drug targets, with 26 out of the 100 top-selling drugs targeting this protein family [19–21]. GPCRs correspond to the largest gene family in the human genome, comprising about 5% of human genes [22, 23]. Thus, there is high interest in developing new lead compounds for most human GPCRs. Given the difficulty in elucidating the structure of these membrane proteins at the atomic level, traditional computer-aided drug design for GPCRs heavily rely on ligand-based modeling techniques [24, 25]. With the determination of X-ray crystal structure from two structurally related proteins, bacteriorhodopsin [26, 27] and bovine rhodopsin [28–31], homology modeling and structure-based approaches to GPCR drug discovery were made possible. Most of the modeled GPCR structures were only used to explain the binding of known agonists/antagonists retrospectively and their use in structure-based drug design was not common until recently. The probable reason for this may be the fact that the available rhodopsin structures are significantly different from each other, they have rather low sequence homology with other GPCRs, and interpreting data from mutagenesis experiments with respect to a GPCR has been problematic [32]. Therefore, it may be assumed that comparative modeling alone will not afford accurate structural information on

individual receptors required for ligand design and in silico screening. Despite this assumption, reports of such models being successfully applied in VS protocols provides some support towards the effectiveness of homology-based molecular modeling in identifying novel lead molecules [33–37].

Opioid receptors belong to the rhodopsin subclass within the GPCR superfamily [38]. The superfamily is characterized by a common structural motif of seven transmembrane-spanning (7TM) helices connected by intracellular (ILs) and extracellular loops (ELs) [39, 40]. The opioid receptors are ~60% identical to each other with greatest homology in the TM regions and the greatest diversity in their N- and C-termini and the ELs [41]. A major clinical application of opioid ligands has been as potent analgesics. However, undesired effects associated with opioids limit their wider use in the clinical management of pain.

Pharmacological studies suggest that multiple opioid receptor subtypes, namely, the mu opioid (MOP), the kappa opioid (KOP), the delta opioid (DOP) [42–50], and the more recently discovered NOP [51, 52] (also called LC132, ORL-1, or nociceptin/orphanin FQ receptor), might be responsible for the pharmacological effects of the opioid drugs, including the regulation of pain. Each receptor mediates unique pharmacological responses and is differentially distributed in the central and peripheral nervous systems. Particularly, activation of the KOP receptor produces many effects such as spinal antinociception, sedation, dysphoria, and diuresis [53–55]. Ligands targeting the KOP receptor have been of clinical interest due to their ability to provide effective pain control without the dependence and respiratory depression that is associated with MOP receptor activation by morphine (Fig. 1; 2) [53]. On one hand, selective KOP receptor antagonists are being explored for their effects in the treatment of a wide variety of areas including cocaine addiction, depression, and feeding behavior, and have been proposed as a treatment for psychosis and schizophrenia [54]. On the other hand, selective KOP agonists such as U50,488H (Fig. 1; 4) find potential applications as analgesics, diuretics, dysphorics, and as antipruritic agents [55].

Salvinorin A, a neoclerodane diterpene (Fig. 1; 1), has been identified as the most potent naturally occurring opioid agonist yet discovered with high selectivity and affinity for KOP receptor [56, 57]. It has no significant activity at other known opioid receptors, serotonergic, dopaminergic, adrenergic, muscarinic, cannabinoid and sigma receptors as well as various ionotropic and metabotropic receptors [58–60]. It is isolated from the leaves of widely available psychoactive plant, *Salvia divinorum* [61, 62]. This plant has been

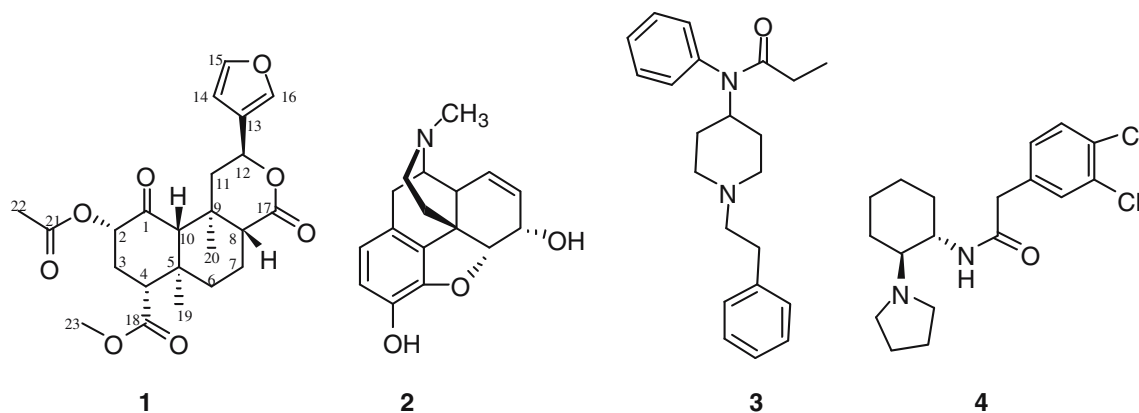


Fig. 1 Structures of salvinorin A (1), morphine (2), fentanyl (3), and U50,488H (4)

used in traditional, spiritual and ethnopharmacological practices by the Mazatec Indians of Oaxaca, Mexico to produce a hallucinogenic effect [56, 61, 63, 64]. Additionally, salvinorin A, being significantly more efficacious than U50,488H or U69,593 and similar in efficacy to dynorphin A (the endogenous peptide ligand for kappa opioid receptors), may be used to treat heroin, cocaine, alcohol and amphetamine dependency, depression, and even excessive marijuana use [56].

Historically the protonated amine, common to all opioid ligands, was considered essential to interact with the side chain of a conserved aspartic acid residue located in TM3 of the opioid receptors to elicit activity [65]. However, salvinorin A, lacking this nitrogen atom, common to all other classes of opioid receptor ligands, is unable to interact in the generally accepted manner and therefore, is the only known non-nitrogenous agonist identified at the KOP receptor [56]. Thus, it is not only structurally distinct from naturally occurring hallucinogens but other classes of KOP receptor agonists such as morphinans, benzomorphans, arylacetamides and 4-phenylpiperidines (Fig. 1). It can cross the blood brain barrier [66] and reportedly, has low toxicity in rodents [67, 68]. Given this, salvinorin A represents a promising lead compound for the development of more potent and selective KOP receptor ligands. However, due to the reported hallucinogenic activity and short duration of action in humans [69], the potential of salvinorin A, itself, as an analgesic is limited. Efforts towards the synthesis and evaluation of structurally related analogs that may have longer lasting analgesic properties and be devoid of psychotropic activity have been reported by several groups and may prove useful as potential therapeutic agents in the treatment of pain and other medical conditions.

The aim of this study was to develop a combined ligand-based and target-based drug design strategy in order to identify novel molecular scaffolds as potential lead molecules for the KOP receptor. Our computational approach is summarized in Fig. 2. The idea was to develop a robust VS model that would combine the speed of pharmacophore modeling and the accuracy of structure-based docking. Additionally, the information obtained from the putative binding site of the generated receptor model may facilitate the identification of key residues and physicochemical properties that may be the molecular determinants for ligand recognition. Simultaneously, structure–activity relationships (SAR) among available ligands for a given receptor may afford equivalent pharmacophores. It is understood that combining ligand-based and target-based models, developed independently from each other, may offer a

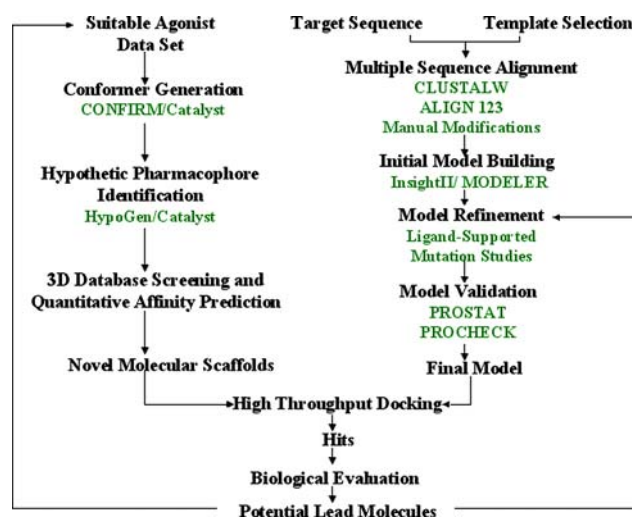


Fig. 2 Flowchart describing our combined target-based and ligand-based drug design approach

self-consistency check for spatial complementarities between them.

To achieve this, a series of molecular modeling studies including automated pharmacophore hypothesis generation and comparative protein modeling was carried out. A systematic study was initiated to develop a pharmacophore hypothesis that could serve the following two purposes: (a) to identify new molecular entities (NMEs) by using a 3D query to search chemical databases and (b) as a predictive tool for estimating biological activity of virtual compounds or compounds designed de novo on the basis of structure–activity analyses. Due to the absence of direct structural information of the KOP receptor–ligand complex, a comparative model of the hKOP receptor was also generated and validated to allow the computational exploration of potential ligand–receptor interactions at the atomic level. Mutational and ligand binding data reported in the literature was also considered to obtain models in order to generate more relevant geometry of the protein binding site [70]. This hKOP receptor model is expected to provide a structural basis for rational drug design methodologies and a tool for large-scale VS of chemical databases.

Materials and methods

Computational resources

Molecular modeling was performed on a Silicon Graphics Octane2 R12000 dual processor workstation running on Irix 64 (SGI, 1600 Amphitheater Parkway, Mountain View, CA 94043). Pharmacophore Modeling was performed using *Catalyst 4.8* [71]. Sequence alignment was performed using *CLUSTALW 1.8* [72] and *Align123* in *InsightII 2005/HOMOLOGY* module [73]. Comparative Modeling was performed using *InsightII/MODELER*. Energy minimization and molecular dynamics simulations were performed using *InsightII/DISCOVER* module. Ramachandran plots and statistics were generated by *PROCHECK 3.4* [74]. For the purpose of docking, *GOLD 2.1* (Genetic Optimization for Ligand Docking) was used [75]. Visualization of docking results was performed using graphics facility of *Sybyl 7.2* [76].

Biological activity data and training set selection

The activity data for salvinorin A analogs, represented as K_i in nM, were obtained from the literature [57, 66, 77–80]. The chemical structures of the agonists are listed in Charts 1 and 2. The datasets were divided into a

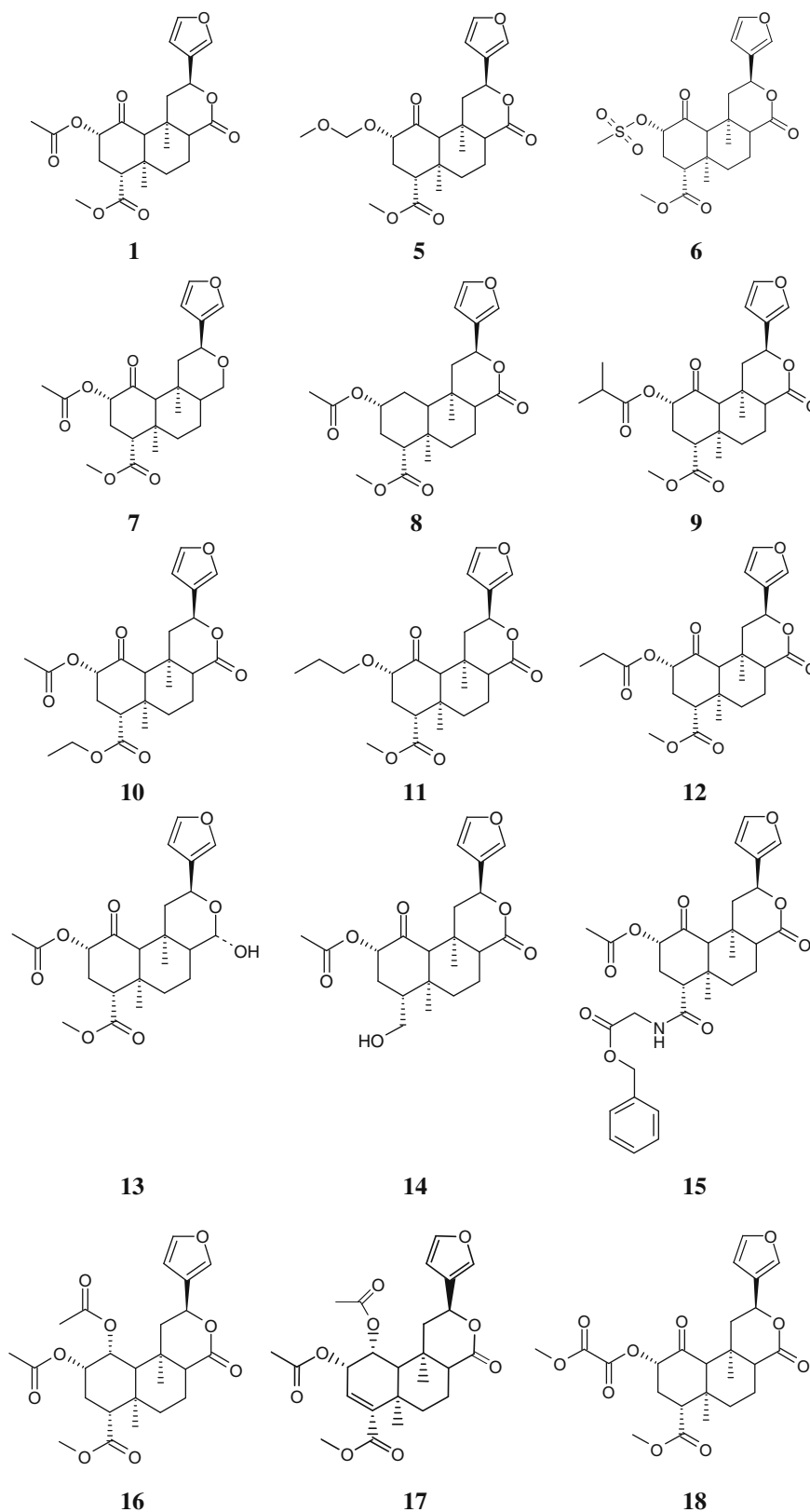
training set and a test set. The selection of a suitable training set is critical for the quality of automatically generated pharmacophore models [81]. To ensure the statistical relevance of the calculated model, the training set should contain a set of diverse compounds together with their activity values. These should originate from comparable binding assays and spread equally over at least 4–5 orders of magnitude. Each selected compound should add some new information to the model while avoiding redundancy and bias both in terms of structural features and activity range. The most active compounds should be included so that they provide information on most critical features required as a pharmacophore. Compounds that are inactive due to steric hindrance must be avoided since they cannot be handled by the currently available features in the *Catalyst*. On the basis of the above criteria, 15 compounds for the training set and 12 compounds for the test set were selected.

For estimation (prediction) purposes, the activity values were classified as follows: K_i (nM) ≤ 1 nM means the compounds are highly active (represented as +++); $1 \text{ nM} < K_i$ (nM) ≤ 100 nM means the compounds are moderately active (represented as ++); K_i (nM) > 100 nM means the compounds are inactive (represented as +). This classification scheme was created on the basis of the fact that salvinorin A, the lead compound had an activity (K_i) of ~ 1 nM, and subsequent optimization of this molecule has generated potent derivatives indicating that activity can reach subnanomolar levels. Therefore, a stricter cut-off value of less than 100 nM was chosen to consider any compound to be active. Because one of the primary objectives of pharmacophore generation is to utilize it for screening 3D drug-like chemical databases to identify lead compounds, this classification scheme is more consequential than actual prediction values and may help in the identification of novel, structurally diverse lead compounds.

Generation of pharmacophores

A pharmacophore model in *Catalyst* is generally referred to as a “hypothesis”, which consists of a collection of features necessary for the biological activity of the ligands oriented in 3D space [82]. In order to generate a pharmacophore, all molecules (both training and test sets) were built and minimized within the *Catalyst*. Conformation models for all molecules were generated using the quasi-exhaustive *Catalyst/ConFirm* module within the software, using the “best quality” conformational search option. A maximum of 250 conformations were generated using CHARMM force field parameters [83] and a constraint of 20 kcal/mol

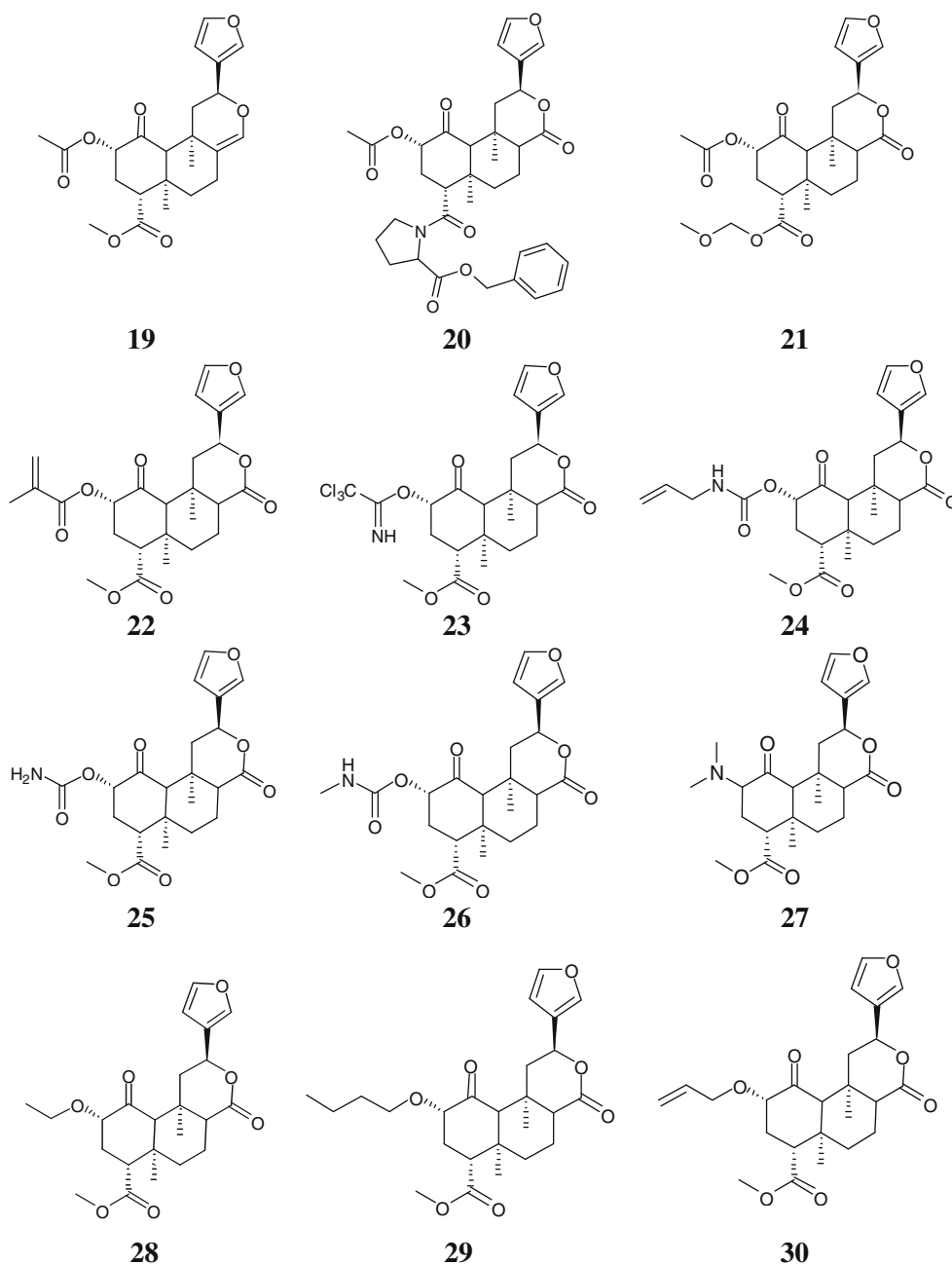
Chart 1 Structures of compounds in training set. All structures were drawn using *ChemDraw Ultra 8.0*



energy thresholds above the global energy minimum. *Catalyst* selects conformers using the Poling algorithm [84], that penalizes any newly generated conformer if it

is too close to an already formed conformer in the set. This method ensures maximum coverage in conformational space. All other parameters were set to the

Chart 2 Chemical structures of test set compounds. All structures were drawn using ChemDraw Ultra 8.0



default settings. An initial analysis revealed that hydrogen bond acceptor (HBA), hydrophobic (HYD), and ring aromatic (RA) features could effectively map all critical chemical/structural features of all the training set molecules. During the initial phase of the hypothesis generation, it was observed that only two features, i.e. HBA and HYD, out of those three mentioned above dominated most of the useful hypotheses generated by the *Catalyst*. Also, the maximum number of HBA features in those hypotheses was never more than two. Therefore, those two features were used to generate 10 pharmacophore hypotheses from the

training set, using a default uncertainty value of 3. The uncertainty value represents a ratio range of uncertainty in the activity value based on the expected statistical straggling of biological data collection. The *Catalyst/HypoGen* module can only generate a maximum of five features for a hypothesis. The best model was selected on the basis of a high correlation coefficient (r^2), lowest total cost, and rms values.

In general, HypoGen tries to construct a pharmacophore that correlates best the 3D arrangement of features in a given set of training compounds with the corresponding pharmacological activities. The result-

ing hypotheses are three-dimensional arrangements of several default feature types (e.g., HBA, hydrogen bond donor (HBD), HYD, RA, positive ionizable (PI)) located at defined positions (location constraints). These are surrounded by certain spatial tolerance spheres, assessing the area in space that should be assigned by the corresponding chemical functions of the matched molecule. Each of the features occupies a certain weight that is proportional to its relative contribution to biological activity. HBA, HBD, and RAs additionally include a vector, defining the direction of the interaction.

The HypoGen algorithm tries to find hypotheses that are common among the active compounds of the training set but do not reflect the inactive ones. The overall assumption is based on Occam's razor, which states that between otherwise equivalent alternatives, the simplest model is best [85]. Therefore, constructed models correlate best with measured activities and consists of as few features as possible. During a HypoGen run, three phases are passed through: (a) pharmacophores that are common among the most active compounds are elaborated in the constructive phase, (b) those pharmacophores that fit the inactive training set members are abolished in the subtractive phase, and (c) the remaining hypotheses are refined in the optimization phase. Thereby random translations of features, rotations of vectored features, and the removal or addition of features from the models are performed. Each perturbation is evaluated by consideration of three cost components, the Error, the Configuration, and the Weight Cost. By default, the 10 lowest cost hypotheses are written to the output file.

Each feature of a hypothesis represents a certain order of magnitude of the compounds' activity. With the default setting of 0.302, the represented orders of magnitude are kept as close to 2 as possible. The weight component is a value that increases in a Gaussian form as these feature weight in a model deviate from the idealized value of 2. The configuration cost (also known as entropy cost) penalizes the complexity of the pharmacophore hypothesis space while the error cost penalizes the deviation between the estimated activities of the training set and experimentally determined values. The two additional theoretical cost calculations are the fixed cost and the null cost. The fixed cost represents the simplest model that fits all the data perfectly and is calculated by the minimum achievable error and weight cost in addition to the constant configuration cost. The null cost represents the maximum cost of a pharmacophore with no features and estimates activity data of every molecule in the training set as the average of all activities

in the set. Therefore, the contribution from the weight or configuration component does not apply. Its absolute value is equal to the maximum occurring error cost [86].

Assessment of the quality of pharmacophore hypotheses

Cost function analysis

The quality of the generated hypotheses was evaluated by considering the cost functions (represented in bits unit) relative to the null and fixed hypothesis calculated by the *Catalyst/HypoGen* module during hypothesis generation. The total cost of any hypothesis should be toward the value of the fixed cost to represent any useful model. It has been suggested in the *Catalyst* that the differences between the cost of the generated hypothesis and the null hypothesis cost should be as large as possible; a value of 40–60 bits difference may indicate that most probably it has a 75–90% chance of representing a true correlation in the data set used. Furthermore, the configuration cost for any generated hypothesis should be less than or equal to 17 (corresponds to 2^{17} pharmacophore models). Any value higher than 17 may indicate that the correlation from any generated pharmacophore is most likely due to chance since *Catalyst* cannot consider more than these models in the optimization phase and so the rest are left out of the process. The configuration cost value can be reduced by limiting the minimum and maximum features. The error cost should also be low; this value increases as the value of the rms increases. The rmsd represent the quality of the correlation between the estimated and the actual activity data.

Cross-validation test

To further assess the statistical significance of the pharmacophore hypotheses generated from the training set molecules a validation procedure, based on Fischer's randomization test [87] was applied. The point of this test is to establish the strong correlation between chemical structures and biological activity. The activity values of the training set molecules are scrambled randomly using the CatScramble technique, available in the *Catalyst/HypoGen* module, and new spreadsheets are created. The number of spreadsheets generated depends on the level of statistical significance one wants to achieve. For a 95% confidence level, 19 spreadsheets are created. For 98% and 99%

confidence levels, 49 and 99 spreadsheets, respectively, are created. In our validation test, we selected the 95% confidence level, and 19 spreadsheets were created by the CatScramble command.

Comparative modeling

Multiple sequence alignment

The peptide sequence of the hKOP receptor was extracted from the SWISSPROT protein sequence data bank (<http://us.expasy.org/sprot/>; accession no. P41145) [88]. The recently determined 2.2 Å X-ray structure of the bovine rhodopsin (PDB ID: 1U19) [31] available from the Protein Data Bank (PDB) [89] was used to build the hKOP receptor model. Three additional class A GPCR sequences, the human DOP (accession no.: P41143), MOP (accession no.: P35372), and NOP (accession no.: P41146) were used to obtain multiple sequence alignment. Sequence alignment was then derived with the *CLUSTALW 1.8* package [72] and align123 in *InsightII/HOMOLOGY* module using the BLOSUM matrices [90] for scoring the alignments. The obtained alignment was then manually refined and inspected for structural integrity in the structurally conserved regions present in TM helices across all the sequences.

Generation of comparative model

The resulting alignment was used as an input for the automated homology-modeling program MODELER *v.7.0* [91, 92] within the *InsightII* software suite. This program assigns atomic coordinates to regions structurally aligned with the template, builds intervening loops, optimizes the rotamers of amino acid side chains, and performs an initial energy optimization of the structure. MODELER generates protein 3D structures by satisfying spatial restraints imposed by the sequence alignment with the template structure. These spatial restraints consist of homology-derived restraints on the distances and dihedral angles in the target sequence based on its alignment with template structures, stereochemical restraints such as bond length and bond angle terms obtained from CHARMM22 forcefield [83], and statistical preferences for dihedral angles and non-bonded inter-atomic distances obtained from a representative set of protein structures. These restraints are expressed in terms of Probability Density Functions (PDF). A 3D protein model is obtained by optimizing the probability density

function with the variable target function procedure in cartesian space that employs methods of conjugate gradients and molecular dynamics with simulated annealing. To guarantee sufficient conformational sampling of each active site residue, several homology models are generated in this step.

Initial model refinement

Hydrogen atoms were added to the selected model. By explicitly specifying the pH as 7.0, the protonation states for all the amino acid side chains sets in accordance with their typical pK_a values. A consistent valence forcefield (CVFF) parameters, a distance dependent dielectric constant and a non-bond cut-off distance of 10.0 Å was employed. The initial model was energy-minimized, employing the *InsightII/DISCOVER* module, in a stepwise fashion following a standard procedure consisting of 500 steps of steepest descent and 1,000 steps of conjugate gradient minimization with an rms gradient of the potential energy of 0.001 kcal/mol·Å in each step. In the first step, all atoms other than hydrogens and terminal oxygen atoms were constrained. This was followed by restraining the backbone atoms allowing energy minimization of side chains of the protein. In the third step, the α -carbons of the entire model were held fixed. Finally, α -carbons of only the TM regions were held fixed while allowing free movement of loop atoms to optimize their geometry.

Generation and refinement of “agonist-bound” state model

After prepositioning the ligand in the receptor cavity according to the known SAR information [57, 66, 77–80] and mutation data [93, 114], the protein–ligand complex was minimized to relieve steric clashes between the non-bonded atoms. The minimization protocol for all calculations consisted of an initial minimization using 100 steps of steepest descent and 1000 steps of conjugate gradient algorithm. Omega (ω) torsions for all protein residues were forced to trans configuration using a force constant of 100 kcal/mol Å². Notably, the Phi–Psi (ϕ – ψ) torsions of residues of the TM domains were restrained using a force constant of 100 kcal/mol Å². An NVT (constant-volume/constant-temperature) molecular dynamics simulation using Andersen thermostat [94] was carried out where the resulting complex was gradually “heated” from a temperature of 100–600 K in steps of 100 K with 10 picoseconds (ps) simulation at each stage. At

600 K, the complex was simulated for another 50 ps followed by gradual “cooling” to 300 K in steps of 100 K with 10 ps at each stage. Finally, a 50 ps simulation at 300 K with a time-step of 1 femtosecond was carried out. The averaged structure obtained in the last 25 ps trajectory of MD was re-minimized until an rms of 0.05 kcal/mol Å was reached. During this minimization procedure, the restraints put on the TMs were reduced gradually. When the restraints were finally removed, the structure was submitted to the final minimization until an rms of 0.1 kcal/mol Å was reached.

Automated docking simulations

Docking runs were performed using the *GOLD 2.1* (Genetic Optimization for Ligand Docking) [76, 95] program that uses a genetic algorithm (GA) to explore the full range of ligand conformational flexibility and the rotational flexibility of selected receptor hydrogens. The mechanism for ligand placement is based on fitting points. The program adds fitting points to hydrogen-bonding groups on the protein and the ligand and maps acceptor points in the ligand on donor points in the protein and vice versa. Additionally, *GOLD* generates hydrophobic fitting points in the protein cavity onto which ligand CH groups are mapped. The GA optimizes flexible ligand dihedrals, ligand ring geometries, dihedrals of protein OH and NH₃⁺ groups, and the mappings of the fitting points. The docking poses are ranked based on a molecular mechanics like scoring function, which includes a hydrogen bond term, a 4–8 intermolecular van der Waals terms, and a 6–12 intramolecular van der Waals term for the internal energy of the ligand.

The protein input was prepared by removing salvinorin A from the final protein complex. Overall 15 ligands with known KOP receptor binding affinities [57, 66, 77–80] that spanned an activity range of 4 log orders were built and minimized for 1,000 steps each of steepest descent followed by conjugate gradient and finally by the BFGS method to a gradient of 0.001 kcal/mol Å or less using the *InsightII/BUILDER* module. All atom types and charges were assigned in *GOLD*. Standard default settings were used for all calculations. *GOLD* was allowed to terminate when the top three dockings were within 1.5 Å of each other. The binding pocket was defined as a spherical region of 12 Å radius centered from the side-chain hydrogen atom of amide of Gln115 (Q2.60). The selection of the most appropriate docking solutions was based on a suitable placement of the ligand with respect to key interactions with the binding site residues.

Numbering of residues and residue index

Residues are numbered according to their positions in the sequences retrieved from SWISS-PROT database [88]. A general index numbering system first described by Ballesteros and Weinstein [96] that applies to all rhodopsin-like GPCRs has been used. Each residue is identified by the TM number (1–7), the most conserved residue within each TM is assigned the number ‘50’ and the positions of all other residues are numbered relative to the most conserved amino acid in each TM segment. This indexing simplifies the identification of corresponding residues in different GPCRs.

Results and discussion

Pharmacophore generation

A set of 10 pharmacophore hypotheses were generated using 15 training set compounds listed in Chart 1. The results of the hypotheses, which include different cost values calculated during hypotheses generation along with rmsd, correlation (r^2), and pharmacophore features, are listed in Table 1. The value of total cost of each hypothesis was close to the fixed cost values, which is expected for good hypotheses. The entropy (configuration cost) values of the hypotheses were also within the allowed range. The difference between null hypothesis and the fixed cost and the total cost of the best hypothesis were 35.79 and 30.90 bits, respectively. These values are lower than the recommended values, which may be due to the high structural homology as well as rigidity of the molecules in the training set molecules. Hypotheses 1 and 2 had the best values in terms of total cost, error cost, rms differences, and the highest correlation and utilized identical features, namely, two HBAs and three HYDs. Hypothesis 2 estimated the activities almost similarly (data not shown) to hypothesis 1. Therefore, hypothesis 1 (Fig. 3) was selected as the best pharmacophore. Table 2 shows the actual and estimated K_i values of training set compounds calculated on the basis of hypothesis 1. It is notable that only one moderately active (++) compound was estimated to be inactive (+).

Pharmacophore assessment

The quality of the pharmacophore was assessed using the CatScramble technique in the *Catalyst*. The purpose of using this technique is to randomize the activity data among the training set compounds and to gener-

Table 1 Results obtained from pharmacophore hypothesis generation using the training set molecules^a

Hypo. no.	Total cost	Error cost	rms	Correlation (r^2)	Features ^b
1	73.01	55.34	0.81	0.93	HBA, HBA, HYD, HYD, HYD
2	73.85	55.74	0.84	0.92	HBA, HBA, HYD, HYD, HYD
3	77.09	59.41	0.86	0.86	HBA, HBA, HYD, HYD, HYD
4	78.25	60.58	1.16	0.84	HBA, HBA, HYD, HYD, HYD
5	78.37	60.44	1.15	0.84	HBA, HBA, HYD, HYD
6	79.22	59.97	1.12	0.86	HBA, HYD, HYD, HYD
7	79.41	61.27	1.20	0.83	HBA, HYD, HYD, HYD
8	79.55	61.36	1.21	0.83	HBA, HBA, HYD, HYD
9	79.87	61.97	1.24	0.81	HBA, HBA, HYD, HYD
10	79.92	62.22	1.25	0.81	HBA, HBA, HYD, HYD, HYD

^a Null cost = 103.91. Fixed cost = 68.12. Configuration cost = 16.54. All costs are in units of bits

^b HBA, Hydrogen Bond Acceptor; HYD, Hydrophobic

ate pharmacophore hypotheses using the same features and parameters used to develop the original pharmacophore hypothesis. If the randomized sets generate pharmacophores with similar or better cost values, rms, and correlation, then the original pharmacophore can be considered as generated by chance. The results of the randomization tests are listed in Table 3. The results obtained from the CatScramble runs indicated that hypotheses generated after randomization had no predictive value. This cross-validation technique provided confidence on the pharmacophore generated from the training set molecules.

Validation of pharmacophores

The selected pharmacophore was further validated by two techniques: (a) by assessing the predictive ability of the pharmacophore on test set molecules; (b) by incorporating an external set of negative controls consisting of five marketed drugs acting on the central nervous system, which target different G-protein coupled receptors. This validation should confirm that the pharmacophore does not predict those drugs as highly active (represented as +++).

Validation of pharmacophore using test set compounds

The validity of any pharmacophore model needs to be determined by applying that model to the test set to find out how correctly the model predicts the activity of the test set molecules and, most importantly, whether it can identify active and inactive molecules correctly. The selected pharmacophore was validated with a test set containing 12 compounds obtained from the same laboratories [57, 66, 77–80] as that of the training set compounds. This validation gives additional confidence in the usability of the selected pharmacophore. The estimated activities were scored using hypothesis 1 and reported in Table 4 (see supporting information, Fig. S1). Out of 12 compounds, 10 were accurately predicted. Only two moderately active compounds were classified as inactive. The selected pharmacophore clearly showed minimal failure in classifying compounds correctly. One of the most potent compounds reported (5 in Chart 1 and Table 2) was chosen to show mapping onto the selected hypothesis, as shown in Fig. 3. The “best fit” option was selected in all cases.

Fig. 3 (A) A representation of the top-ranked hypothesis. (B) Mapping of the most active compound (5) onto the selected pharmacophore hypothesis. The green and blue contours represent HBA and hydrophobic features (HYD), respectively

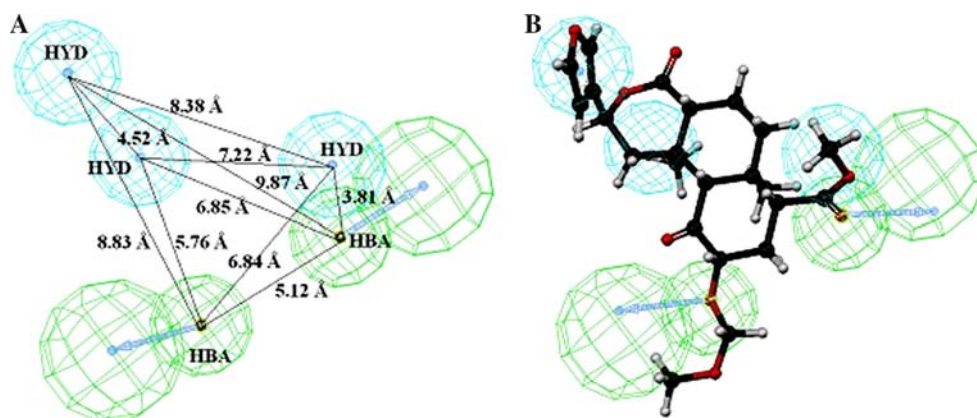


Table 2 Actual and estimated activities of training set molecules calculated on the basis of the top-ranked hypothesis

Name	Actual K_i (nM)	Estimated K_i (nM)	Activity scale ^a	Estimated activity scale	Reference
1	1.3	3.9	++	++	79
5	0.4	0.97	+++	+++	80
6	2.3	1.1	++	++	78
7	6	10	++	++	77
8	18	6.5	++	++	77
9	19	17	++	++	78
10	28.5	12	++	++	79
11	28.7	66	++	++	66
12	32.6	61	++	++	57
13	59	160	++	+	77
14	347	540	+	+	77
15	470	640	+	+	79
16	650	120	+	+	78
17	1,022	820	+	+	77
18	430	130	+	+	78

^a Activity scale: highly active ($1 \leq 100$ nM, +++), moderately active ($1 \leq 100$ nM, ++) and inactive (>100 nM, +)

Validation with negative controls

None of the five-marketed drugs mapped onto the selected pharmacophore (see supporting information, Table S1) and were estimated to be inactive (+) thus, confirming the validity of the model in a negative test.

Table 3 Results from cross-validation run using CatScramble

Validation	Total cost	Fixed cost	rms	Correlation (r^2)	Configuration cost
<i>Results for unscrambled</i>					
Hypo. no.	73.01	68.12	0.81	0.93	16.54
<i>Results for scrambled</i>					
Trial_1	86.96	65.16	1.21	0.70	16.07
Trial_2	78.90	67.23	1.86	0.53	14.38
Trial_3	81.24	64.14	1.80	0.66	15.54
Trial_4	91.97	65.00	1.66	0.34	16.87
Trial_5	81.84	67.12	1.89	0.49	15.60
Trial_6	83.89	66.67	2.11	0.53	16.69
Trial_7	87.74	65.92	1.92	0.48	15.98
Trial_8	95.46	67.81	1.43	0.68	13.17
Trial_9	88.87	65.45	1.76	0.64	15.03
Trial_10	83.98	67.81	1.91	0.71	16.46
Trial_11	86.51	64.32	1.79	0.50	15.92
Trial_12	87.79	64.96	2.05	0.46	16.77
Trial_13	89.45	66.78	2.25	0.73	16.82
Trial_14	87.57	67.42	1.87	0.62	15.96
Trial_15	89.90	65.89	1.85	0.51	15.61
Trial_16	84.75	65.76	1.12	0.59	16.95
Trial_17	86.66	67.32	1.96	0.56	16.76
Trial_18	87.44	67.67	1.41	0.71	15.88
Trial_19	83.59	66.76	1.93	0.61	15.74

Null cost = 103.91. All costs are in units of bits

Table 4 Actual and estimated activities of test set molecules calculated on the basis of the top-ranked hypothesis

Name	Actual K_i (nM)	Estimated K_i (nM)	Activity scale ^a	Estimated activity scale	Reference
19	6	8.6	++	++	77
20	210	170	+	+	79
21	99.6	140	+	+	79
22	42	17	++	++	78
23	64	36	++	++	78
24	120	100	+	+	78
25	3.2	6.3	++	++	66
26	83.0	130	++	+	66
27	90.9	160	++	+	66
28	7.9	6.4	++	++	66
29	35.8	38	++	++	66
30	60.1	38	++	++	66

^a Activity scale: highly active ($1 \leq 100$ nM, +++), moderately active ($1 \leq 100$ nM, ++) and inactive (>100 nM, +)

This validation corroborated the effectiveness of the derived pharmacophore model, which most likely will be useful in identifying NMEs that target hKOP receptor. Additionally, this pharmacophore may also be used in de novo design to develop totally novel drugs that satisfy the pharmacophoric requirements. Furthermore, it can guide chemists to synthesize new compounds during lead optimization.

Generation of KOP receptor model

Homology modeling has proved to be a very powerful technique, even when the template and the target are distantly related [97]. However, although a common general structure is likely to be shared by all GPCRs, including the KOP receptor, specific differences do exist between rhodopsin and other receptors. To date, bovine rhodopsin is the only GPCR for which an experimental structure has been elucidated and therefore, is the only suitable template for the construction of homology models of other members of the superfamily. This structure however, corresponds to an antagonist-bound conformation where retinal is covalently bound [28]. Homology models constructed from this structure are not directly suitable for docking purpose for a variety of reasons. For example, GPCRs are known to adopt different conformations based on the type of bound ligand (agonist/antagonist) [40, 98, 99]. In addition, the modeled binding site is too narrow which is mostly due to the rather flat nature of the retinal, and also because the E2 loop in bovine rhodopsin closes tightly upon the center of the receptor that restricts the access of the larger GPCR ligands into the active site. So far, a 3D structure of a GPCR in its activated state is not available, however,

several empirical models of the agonist-bound state have been generated using constraints from site-specific spin labeling and mutation studies [100–103], NMR [104] and cross-linking studies [105]. Although such structures need detailed verification, there are reports of successful VS of large compound databases for drug design projects in the literature [106, 107]. Several homology-modeling studies of the KOP receptor have been reported in the literature [108]. While the previously reported KOP receptor models offer insights into the molecular pharmacology of the receptor, they lack detail for purposes of drug design. This warranted the generation of a KOP receptor model that not only addresses the possible binding modes of agonists in the receptor binding site to afford a better understanding of the molecular recognition process at the KOP receptor but can serve as a tool for structure-based drug design. Furthermore, to generate a model closer to the agonist-bound state (activated state) the model was refined by minimizing the protein with a full agonist to achieve expansion of the active site.

Multiple sequence alignment

Sequence alignment is a central technique in homology modeling used to establish a one-to-one connection between the amino acids of the reference and the target protein [109]. It is particularly critical in the case of GPCRs where the only structural template available is that of bovine rhodopsin and the sequence identity with the model sequence falls below 19.5% in most cases [110], which is inappropriate for homology modeling. However, class A GPCRs are an exception owing to the fact that each TM contains one or two highly conserved residues that allows an explicit alignment [111], that may be combined with manual modifications to obtain a fairly accurate alignment. Also, the problem of low sequence identity may be addressed by performing a multiple sequence alignment of the retrieved protein sequences and that obtained from template bovine rhodopsin. This also helps to minimize the errors originating from the availability of a single template structure. To this end, sequences from four opioid receptors were used during the alignment process to improve the low confidence regions in the alignment and also point out potential areas of low reliability in the homology model. All four opioid receptors contain highly conserved fingerprint residues of family A receptors, such as Asn46 (N1.50) in TM1, Asp74 (D2.50) in TM2, Arg126 (R3.50) in TM3, Trp153 (W4.50) in TM4, Pro207 (P5.50) in TM5, Pro255 (P6.50) in TM6, and Pro299 (P7.50) in TM7;

Fig. 4 [112]. In addition, all four opioid receptors possess two conserved cysteine residues in the first and second ELs, that probably form a disulfide bridge which is implicated in constraining the conformation of the extracellular domains of the proteins and stabilizing the TM domains [113]. All of these highly conserved residues were carefully aligned with corresponding residues of the bovine rhodopsin, as well as the other sequences.

The global sequence identity between the KOP receptor and bovine rhodopsin is found to be approximately 19%. The sequence identity and similarity values for each TM domain are given in Table 5. Although the sequence identity between the TMs of the modeled receptor and the template is only between 12 and 29%, the model is believed to share the same fold as the template. This is supported by the fact that the modeled GPCRs include several conserved motifs with the structurally and functionally important amino acid residues either highly conserved or substituted by amino acids of high similarity (the LAXxD on TM2, the D(E)RY on TM3, and the NPxxY on TM7) [14]. The final multiple sequence alignment used for model building is shown in Fig. 4. Therefore, it may be assumed that the generated model is of higher accuracy in the TMs as compared to the loop regions, which are divergent and flexible to a greater extent. However, since the binding site is located in the TM domains, the uncertainty in the loop regions as well as the N- and C-terminal domains was not considered of much consequence to ligand screening.

Model building procedure and initial model refinement

The 3D structure of hKOP receptor was generated using the MODELER program interfaced with the *InsightII* software based on the recently published bovine rhodopsin crystal structure (PDB ID: 1U19; 2.2 Å resolution) as the template structure. The input parameters of MODELER were set to generate six models including three loop refinements, with the level of optimization set to high and the manual-definition disulfide bridges assumed to exist in the extracellular region by homology to hKOP receptor [113]. Amongst the six different models generated, the one showing the lowest energy, restraint violations, and the number of main chain and side chain bad conformations was selected for further refinement, see Table 6. Since it is known that N-terminal does not contribute significantly toward the binding of KOP receptor ligands, some of the overhanging residues in this region were truncated [93, 114].

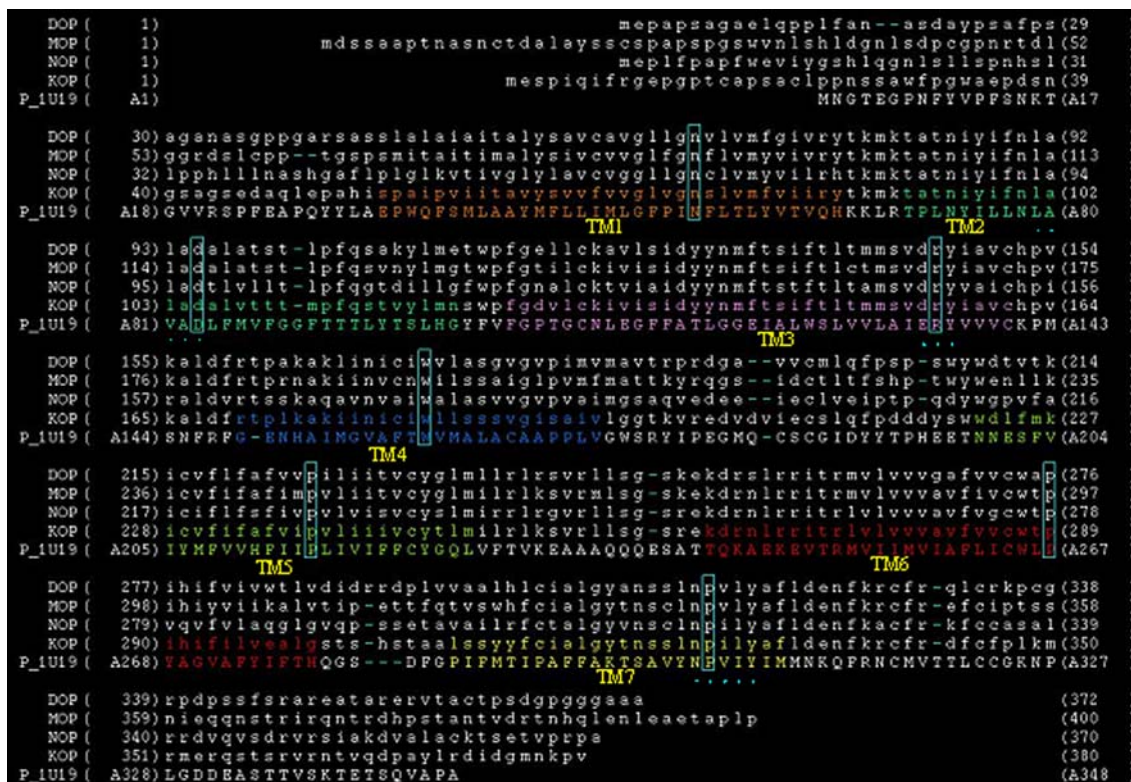


Fig. 4 Multiple sequence alignment of the hKOP receptor with the bovine rhodopsin (P_1U19) and the other class A GPCR sequences. Each TM is highlighted in a different color. The most

conserved residue for each TM helix across all the species is enclosed in a vertical box colored cyan. The conserved GPCR motifs have been indicated by cyan dots

The quality of the initial model was improved by subjecting it to a crude energy minimization protocol as detailed in the Materials and methods section. These minimizations helped relieve any steric clashes or improper geometries in the protein structure and produced a model with correct bond lengths and bond angles and where individual atoms are not too close together.

Generation of “agonist-bound” hKOP receptor model

Binding of a ligand to the active site of a protein is typically associated with local and, perhaps, also glo-

bal structural rearrangement of the receptor (Induced-fit effect). As a result, structure-based drug design preferentially relies on the structures of protein–ligand complexes containing bound ligands. Keeping this in mind, the next step was to develop a protein–ligand complex that would offer a more detailed and accurate picture of the interactions and structural complementarities between the ligands and the active site. Such a protein–ligand complex may be used more appropriately to perform docking simulation studies and may provide more meaningful structural models.

Therefore, to obtain the 3D coordinates of the agonist-bound state of the KOP receptor, salvinorin A was manually placed within the putative binding pocket guided by SAR, supported by experimental data [93, 114]. It should be noted that due to the lack of a protonable nitrogen atom common to all GPCR ligands, from the ligand structure, the generally accepted key anchoring interaction for agonist association (basic nitrogen forming a salt bridge with the highly conserved Asp138 (D3.11)) in TM3 for GPCRs cannot be established which makes the docking of salvinorin A challenging [115]. However, site directed mutagenesis studies did provide sufficient information

Table 5 Sequence identity and similarity between conserved regions of the hKOP receptor and the bovine rhodopsin

	No. of residues	% Identity	% Similarity
TM1	33	18.18	66.67
TM2	32	25.00	56.25
TM3	36	22.22	69.44
TM4	26	11.54	50.00
TM5	28	28.57	71.43
TM6	36	22.22	63.89
TM7	24	16.67	50.00

Table 6 Analysis of the hKOP receptor models using MODELER and PROSTAT

	Energy	PDF violations	Bond lengths ^a	Bond angles ^a	Helix Phi–Psi ^a	Omega ^a	C α -Chirality ^a
KOP1	759.64	1372.26	0	3	1	0	0
KOP2	771.67	1393.38	0	2	4	0	0
KOP3	802.88	1554.58	1	7	6	0	0
KOP1_I1	2790.34	3305.33	0	13	1	0	0
KOP2_I1	2790.39	3305.18	1	14	1	0	0
KOP3_I1	2810.18	3286.29	1	11	3	0	0
1U19	–	–	0	5	2	0	0

^a The values represent the number of violations found (with standard deviation >5.0 from reference values) for respective geometrical parameters. The first three are receptor models for hKOP while the last three are corresponding loop refinements

of the key residues involved in binding of the ligand to the receptor that directed suitable placement of the ligand into the putative binding pocket [93, 114]. This was followed by a series of energy minimization and restrained MD simulations performed iteratively on the protein–ligand complex, the general procedure of which has been detailed in the Materials and methods section. The preferred binding mode was chosen on the basis of compatibility with mutation and binding data, as well as energetic considerations. Care was taken to maintain the conserved α -helical geometry of the TM domains during the refinement. The refined 3D structure of hKOP receptor generated by the homology modeling method is displayed in Fig. 5. The backbone rmsd between the final minimized homology model and the template crystal structure is 3.55 Å, while that between the starting homology-modeled structure and the template was only 1.58 Å.

Protein structure validation

One essential requisite for a model is to have a good stereochemistry. The refined structure was evaluated for overall quality using available analysis procedures. These analyses compare specific properties of the model with those for known high quality protein structures. For this purpose two protein analysis programs were used: PROCHECK and PROSTAT.

PROCHECK is a well-known protein structure checking program that provides an assessment of the overall quality of the structure and highlight regions that may need further investigation. The phi–psi plots are shown in Fig. 6, while the more detailed results are listed in Table 7. The plot clearly shows the vast majority of the membrane-spanning amino acids are in a phi–psi distribution consistent with right-handed alpha helices. The remaining residues that fall into the random or beta configuration geometries are very short segments and are primarily in the loop regions of the

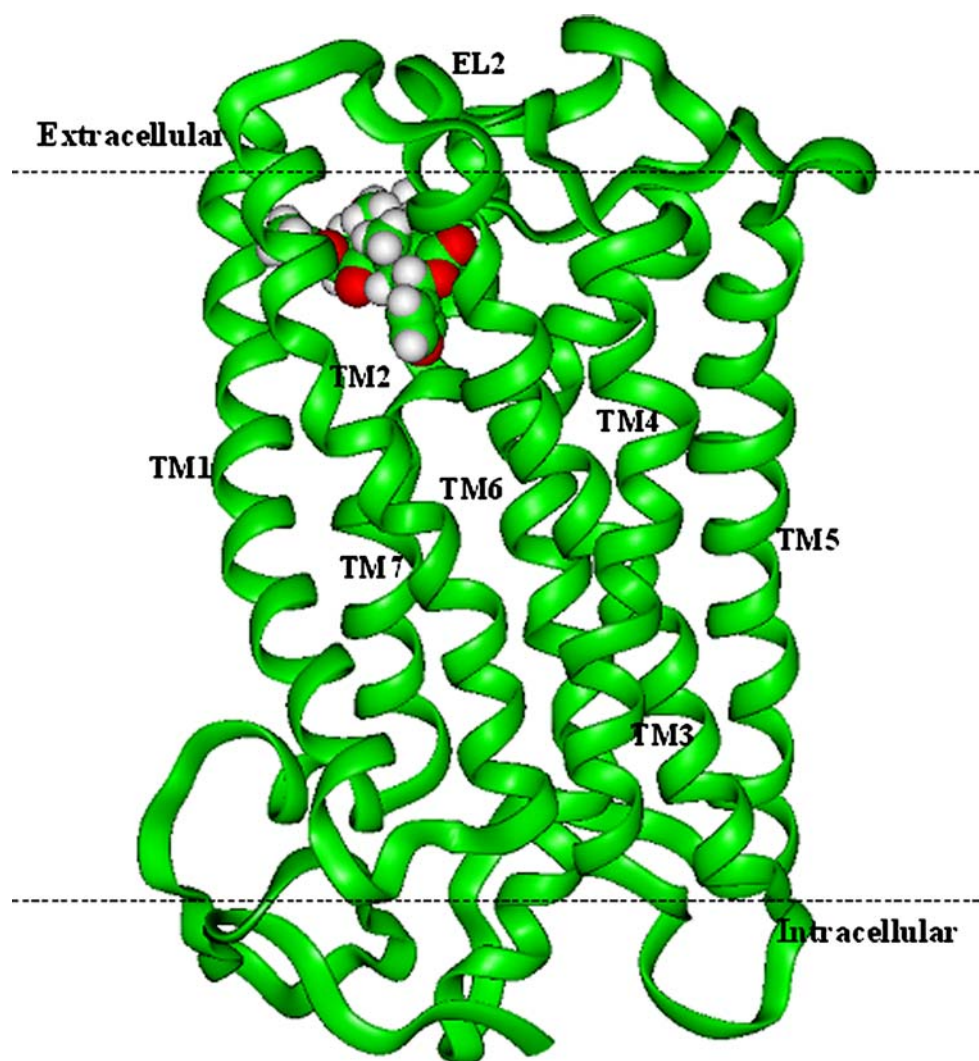
protein. A quick check of amino acid side chain integrity did not reveal any amino acid side chain clashes with one another or with the backbone. Only five non-glycyl residues are outside the allowed regions. Analysis of the model reveals that these outlying residues are located far from the putative active site of the protein. Figure 6 shows that the structure is reasonable overall because the spot distribution for the homology modeled structure was similar to the standard X-ray structure of bovine rhodopsin. The results showed that our modeled structure was comparable with the standard structures.

Another protein structure analysis program, PROSTAT, was used for structure validation. The program measures all bond distances, bond angles, and dihedral angles for the given protein structure and compare them with its own database. It then lists the number of instances where structural features differ significantly from average values in known proteins. The cut-off used for significantly different was 10 standard deviations from the reference value. No spurious angle or bond length was detected in our model. The results are listed at the bottom of Table 7. Using these geometric criteria, the hKOP receptor model compares well with the X-ray structure of bovine rhodopsin.

Salient features of the model

There is a plethora of structural information available for GPCRs from a variety of experimental studies on different receptors. Since modeling of GPCRs is based on the assumption that they share considerable structural commonalities, the existence of such features may serve the purpose of validating the model. Accordingly, the final modeled structure of hKOP receptor was scrutinized for its similarity to the low-resolution structure of the bovine rhodopsin and for presence of common structural features inferred from several studies on GPCRs.

Fig. 5 General topology of the hKOP receptor model. The backbone of the KOP receptor is represented as oval ribbon in green. Salvinorin A, represented as a CPK model colored by atom type, is docked inside the trans-membrane recognition site (see Materials and methods section for details)

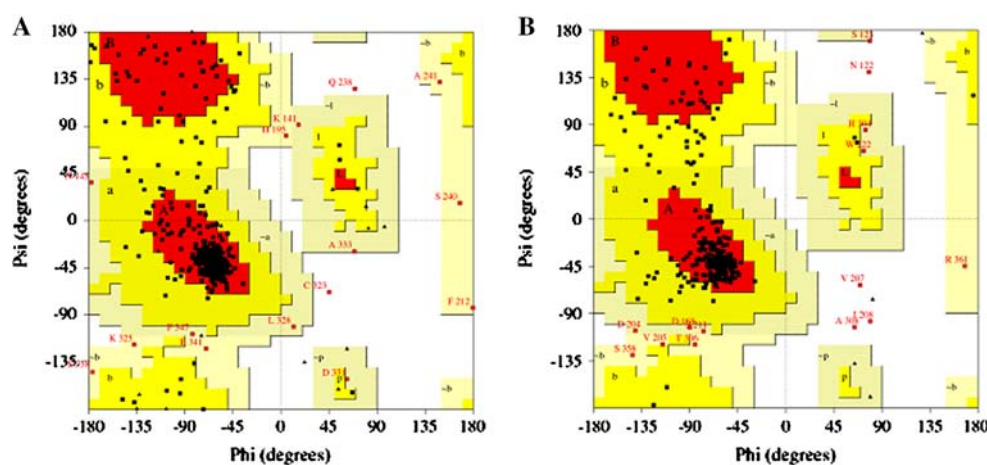


Conserved residues and motifs

A highly conserved residue in the GPCR family, Asn77 (N1.50) being a hydrophilic polar residue, is sited in the intracellular side in helix 1 and faces towards the core of the modeled structure. The L_{Axx}D motif of GPCRs is also observed in the TM2 of the with the side chain of highly conserved Leu101 (L2.46) and Ala102 (A2.47) pointing towards the central core of the helix bundles. It is known that conserved cysteine residue in GPCRs may form an intramolecular disulfide bridge that links ELs of the receptors. In view of that a disulfide bond was observed between Cys131 (EL2) and Cys210 (C3.25) in the modeled structure. Another important motif, D(E)RY that has been connected to the regulation of interaction between the G-protein with the receptor [29], exists in most GPCRs. This microdomain is also conserved in the hKOP receptor.

The side chains of both Asp155 (D3.49) and Arg156 (R3.50) are placed in hydrophobic surroundings (Thr94 (T2.39), Met152 (M3.46), Lys174 (L4.40), Ile272 (I6.33), Leu275 (L6.36) and Val276 (V6.37)). Interestingly, the charged interaction between Asp155 (D3.49) and Arg156 (D3.50) resulting in the formation of a salt bridge that serves to constrain the rhodopsin-like receptors in its inactive state [28] was not observed. This interaction was however, present in the initial model of hKOP receptor. It has been hypothesized that one of the key steps in receptor activation is breaking of this salt bridge [116, 117]. It is likely that refinement of the model with a full agonist may have induced changes allowing the receptor to assume a conformation closer to the activated state. Yet, another conserved microdomain in most GPCRs is the NPxxY motif [118]. The side chain nitrogen in Asn326 (N7.49) of this microdomain has been found to H-bond with

Fig. 6 Ramachandran plot of (A) the X-ray structure of bovine rhodopsin (B) and the homology-modeled structure of final hKOP receptor. The different colored areas indicate “disallowed” (white), “generously allowed” (light yellow), “additional allowed” (yellow), and “most favored” (red) regions (refer to Table 6)



the carboxylate of Asp105 (D2.50). Such an interaction has been demonstrated for Gonadotropin releasing hormone and Serotonin-2A receptors as well as in other GPCRs [119, 120]. Additionally, the helix–turn–helix fold of the NPxxY(x)5.6F region appears to be stabilized by a hydrophobic interaction between aromatic side chains of Tyr330 (Y7.53) and Phe337 (F7.60).

The aromatic cluster

Within the opioid receptor family, the conserved residues present in the aromatic cluster are Tyr139 (Y3.33), Phe231 (F5.43), and His291 (H6.52). The aromatic residues highly conserved in the larger family of GPCRs include Trp183 (W4.50), Phe235 (F5.47), Phe283 (F6.44), and Trp287 (W6.48), Fig. 7.

Interhelix hydrogen bonding network

In addition to hydrophobic interactions, an extensive interhelix hydrogen bond interaction exists between the helices that help stabilize the TM domains of the hKOP receptor. This cluster consists of several highly conserved residues across opioid receptors; Asn141 (N3.35), Ser145 (S3.39) and Thr148 (T3.42). Several residues highly conserved across all known GPCRs; Asn77 (N1.50), Asp105 (D2.50), Trp287 (W6.48), Asn322 (N7.45), Asn326 (N7.49), and Tyr330 (Y7.53) were also found within the cluster.

Binding site analysis

The binding cavity for salvinorin A type agonists appears to be formed essentially by TM2 and TM7. Mutational analysis has also suggested the binding site for the KOP receptor primarily between the amino acid residues of TM2 and TM7, implicates involvement of EL2 as well with no significant contribution from the TM1, TM3, TM5, TM6 and extracellular domains [114]. Consistent with the mutational and biochemical data [93, 114], the active site appears to be essentially hydrophobic with side chains of several hydrophobic residues (Tyr66 (Y1.39), Tyr119 (Y2.64), Ile137 (I3.33), Tyr312 (Y7.35), Tyr313 (Y7.36), Ile316 (I7.39) and Tyr320 (Y7.43)) extending into the deduced

Table 7 Results of protein structure check by PROCHECK and PROSTAT

	1U19	hKOP
Residues in most favored regions [A,B,L]	242 79.9%	221 76.2%
Residues in additional allowed regions [a,b,l,p]	47 15.5%	55 19.0%
Residues in generously allowed regions [-a,-b,-l,-p]	12 4.0%	9 3.1%
Residues in disallowed regions	2 0.7%	5 1.7%
Number of non-glycyl and non-proline residues	303 100%	290 100%
Number of end residues	2	2
Number of glycine residues (shown in triangle)	23	9
Number of proline residues	20	11
Total number of residues	348	312
Overall PROCHECK score ^a	0.18	-0.47
Number of bond distances with significant deviations ^b (PROSTAT)	0	0
Number of bond angles with significant deviations ^b (PROSTAT)	0	0
Number of dihedral angles with significant deviations ^b (PROSTAT)	0	1

^a Recommended value >-0.50 and investigation is needed for <-1.0

^b Number of instances for which the property differs more than 10 standard deviations from the reference value

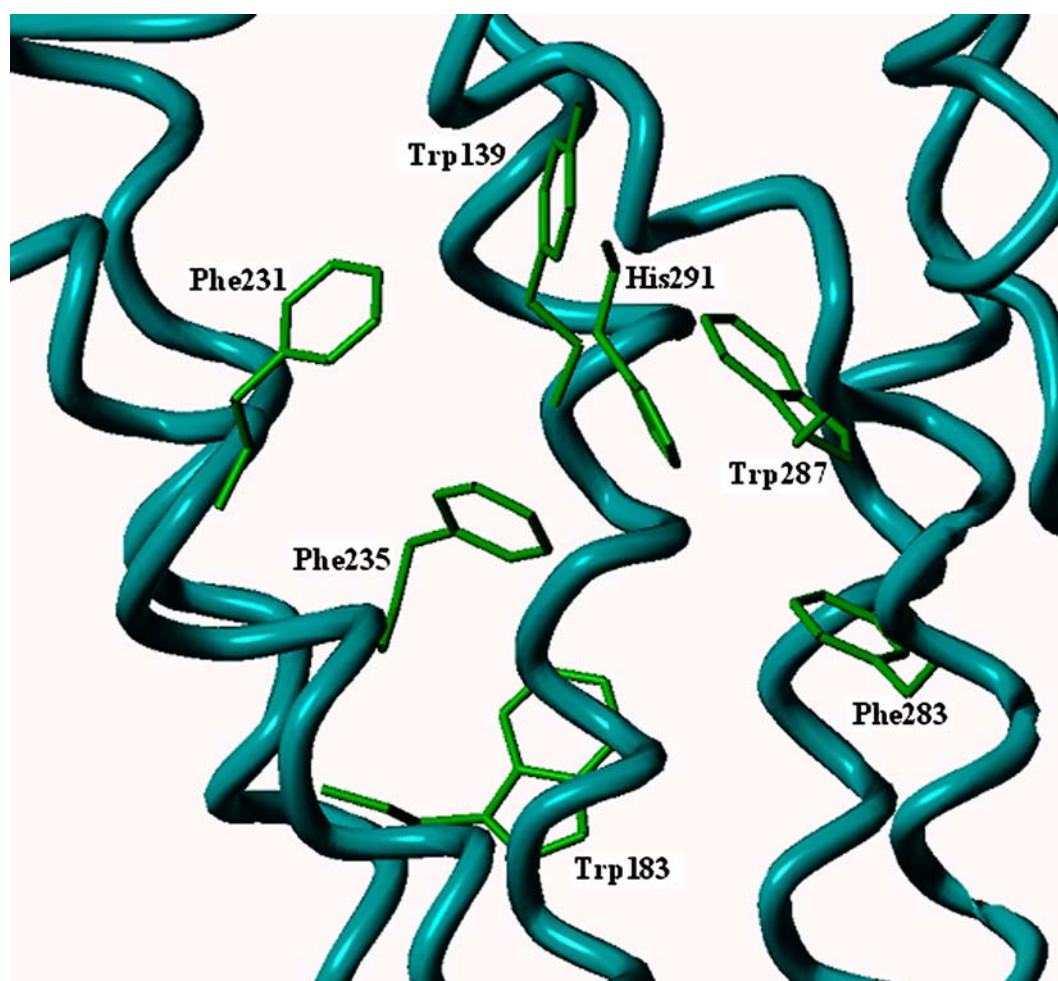


Fig. 7 Interhelical network of aromatic-aromatic interactions in the hKOP receptor model

binding pocket for salvinorin A, shown in Fig. 8. As is obvious from the Fig. 5, salvinorin A appears to be positioned between the TM helices, closer to the extracellular domain of the receptor.

Detailed analysis of the final hKOP receptor model reveals several critical ligand–receptor contact points, which are in agreement with the known SAR and mutational data for the corresponding amino acid residues [93,114]. The ester oxygen at position 2 (between C-2 and C-21) is observed to be involved in H-bonding with the amide side-chain of Gln115 (Q2.60), shown in Fig. 9a. This residue has been shown to be one of the key amino acid residues involved in binding of salvinorin A to the hKOP receptor [114]. This is also in compliance with our ligand-based pharmacophore model where the ester oxygen at the position 2 of the ligand is implicated as HBA, refer Fig. 9b. Another hydrogen bonding interaction between the carbonyl oxygen (C-18) at position 4 of salvinorin A and Tyr313

(Y7.36) was also observed. Additionally, the methyl group (C-23) at position 4 ester shows a hydrophobic interaction with the phenol ring of Tyr313 (Y7.36) residue. Indeed, a similar interaction has been proposed by the developed pharmacophore model as well. Site directed mutation studies [93, 114] indicate that this residue (Tyr313) may be involved in hydrophobic interaction. However, the Tyr313Ala mutation results in an approximate 22-fold decreased affinity, which may be explained on the basis of the loss of both interactions established by this residue [93]. The furan ring of salvinorin A shows π -stacking interactions with Tyr320 (Y7.43) that lies at the base of the putative binding pocket. Similar inferences may be derived from the pharmacophore model. It is interesting to note that mutation studies by Kane et al. [114] suggest involvement of this residue in π - π interactions or other hydrophobic interactions, however, Yan et al. [93] in their study have implicated this residue in H-bonding

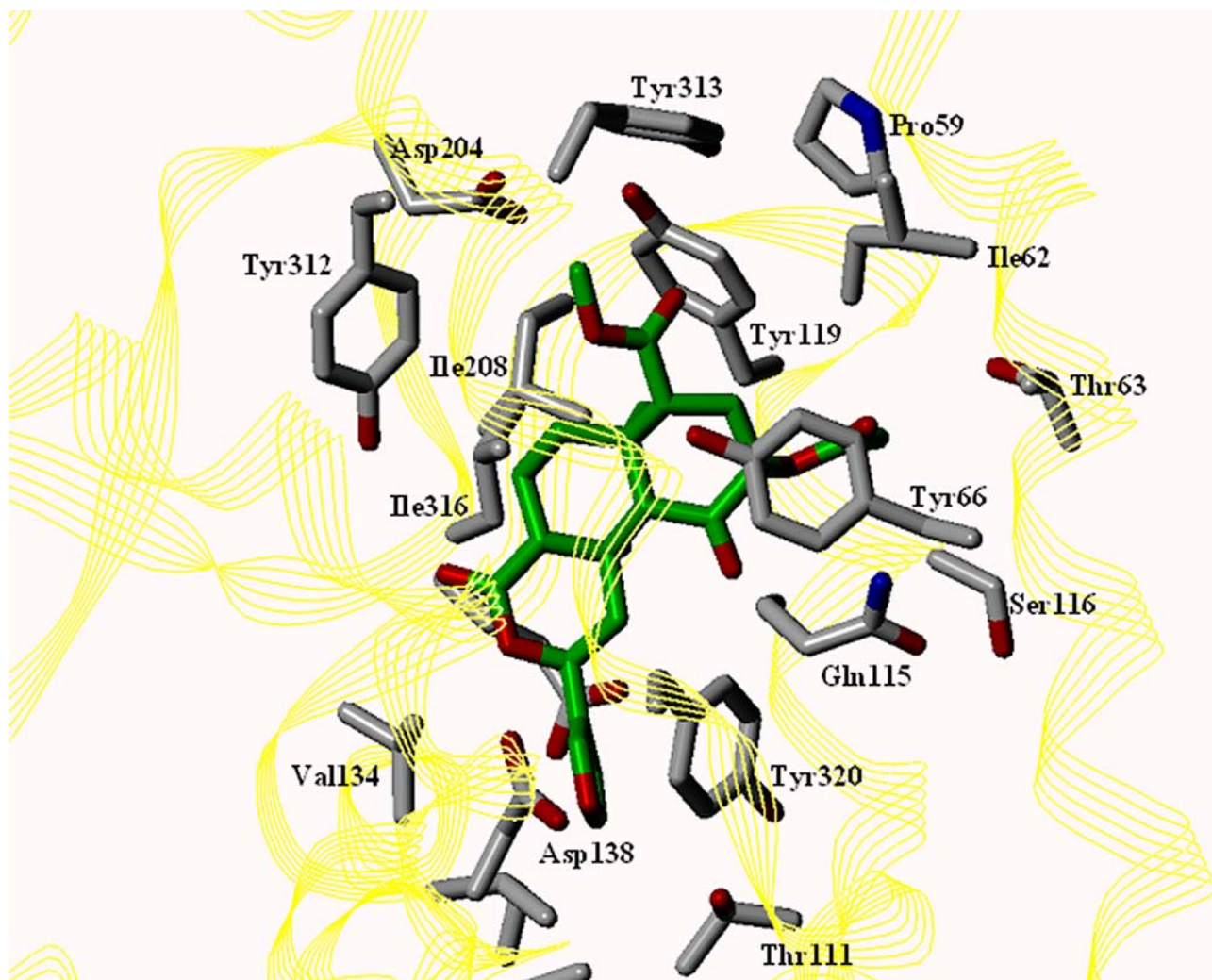


Fig. 8 Salvinorin A in the putative binding pocket of agonist-bound conformation of hKOP receptor, represented as yellow ribbon. Side chains of selected residues within 5.0 Å of docked

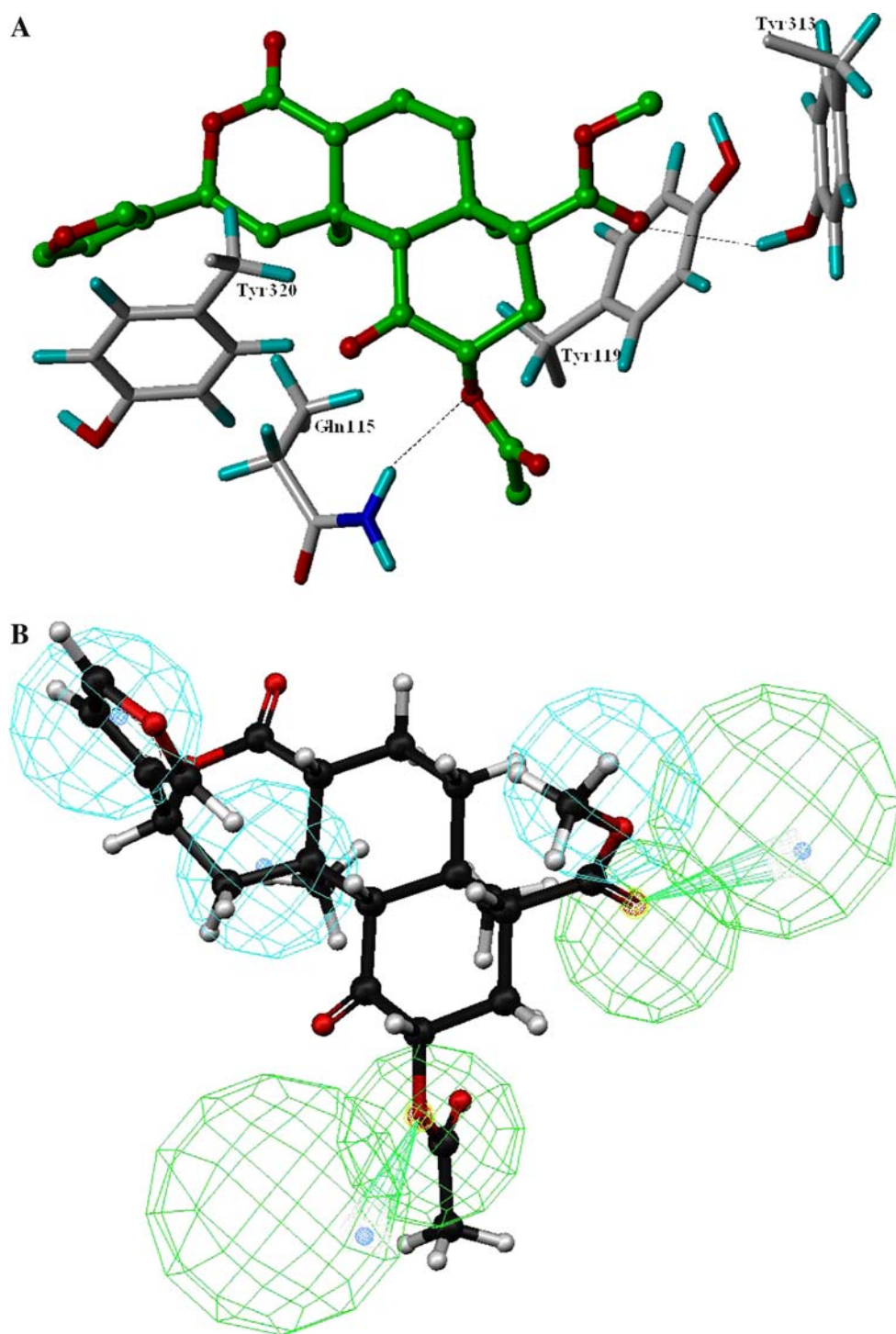
ligand are displayed. Salvinorin A is displayed as green carbon atoms. Receptor side chains are displayed as white carbon atoms

with the furan oxygen of salvinorin A. A closer inspection of the putative binding pocket of our model reveals that this furan ring oxygen is 3.55 Å away from the hydroxyl functionality in Tyr320 (Y7.43) thus, it is possible that this residue may possibly be involved in H-bond formation as well. Hydrophobic interactions are also observed between Tyr119 (Y2.64) and methyl group (C19) at position 5 in salvinorin A. This interaction is consistent with the predicted pharmacophore model. It was also observed that EL2 moved away (towards extracellular region) during the refinement resulting in expanded binding pocket in order to accommodate ligands of varying sizes. Consequently, the strong correlation between the observed ligand–receptor interactions and the reported experimental data provide additional validation to the developed hKOP receptor model.

Docking simulations

The task of docking 15 known derivatives of salvinorin A (see supporting information; Fig. S2), having agonistic activity at hKOP receptor, was undertaken to authenticate the model by calculating the correlation between the pK_i values of the ligands and their respective docking scores. The GA program *GOLD* was used to perform automated molecular docking. The reliability of the applied docking protocol was assessed by docking salvinorin A into the active site of modeled hKOP receptor. The key characteristic of a good docking program is its ability to reproduce the experimental binding modes of ligands [121]. To test this, a ligand is taken out of the 3D structure of its protein–ligand complex and docked back into its binding site. The docked binding mode is then compared with the

Fig. 9 (A) Proposed binding mode and important interactions of salvinorin A in the final hKOP model. Salvinorin A is displayed as green carbon atoms. Receptor side chains are displayed as white carbon atoms. Hydrogen bonds are shown as dotted lines. (B) Mapping of salvinorin A onto the selected top-ranked pharmacophore hypothesis. The green and blue contours represent HBA and HYD features, respectively



experimental binding mode, and an rmsd between the two is calculated; a prediction of a binding mode is considered successful if the rmsd is below a certain value (usually 2.0 Å). The rmsd between the docked binding mode and the binding mode for salvinorin A in the generated protein–ligand complex was within this cut-off limit (1.61 Å).

This protocol was then similarly applied to all ligands. For each ligand, 10 solutions were generated. Generated ligand docking solutions mutually differ by 1.5 Å rmsd. An analysis of the best docking modes of various docked ligands exhibited common binding characteristics, which may be due to the close structural similarity existing among them. For the majority

of the ligands, top-scoring poses showed relevant binding interactions including formation of the key H-bond with Gln115 (Q2.60). The plot of *GOLD* score versus the binding affinities (pK_i) is shown in Fig. 10. As logically expected, a robust correlation ($r^2 = 0.80$) between the experimental binding affinity and the *GOLD* score was achieved. Thus, the model can explain variations in the pK_i values, which, in turn, is related to binding affinities of the ligands, implying that it had been constructed with reasonable accuracy. The results from the docking simulations are self-consistent in that the best docking modes of various ligands exhibited common binding characteristics among them, supporting the validity of the model and its suitability for structure based drug design.

Furthermore, the structure-based and ligand-based methods were quantitatively compared by plotting the predicted activities of the test set compounds utilized to validate the pharmacophore model against their respective *GOLD* scores obtained from docking them into the receptor model. The plot of *GOLD* scores versus predicted activities (pK_i) for these compounds is shown in Fig. 11 (see supporting information, Table S2). Initially, a correlation of 0.67 was obtained however, after removing one outlier (compound 27), the correlation improved to 0.85, which is an excellent correlation taking into consideration that a computationally driven GPCR model was used for the docking analysis and the only scoring function used was *GOLD* score. This result further reinforces the idea that our ligand-based approach and the receptor-based study are indeed converging.

Conclusions

A composite VS model is proposed where ligand-based and target-based information has been synergistically

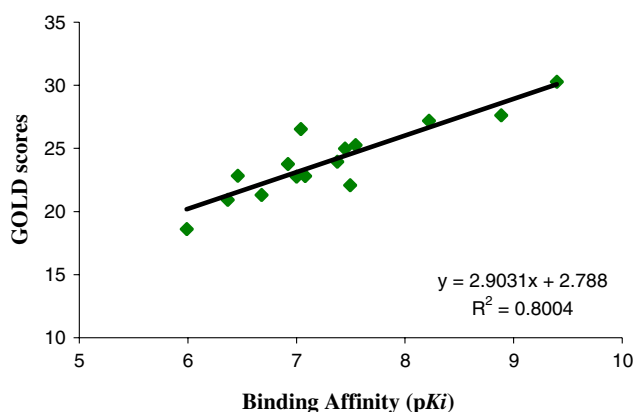


Fig. 10 Correlation plot of binding affinity versus *GOLD* docking score for salvinorin A derivatives

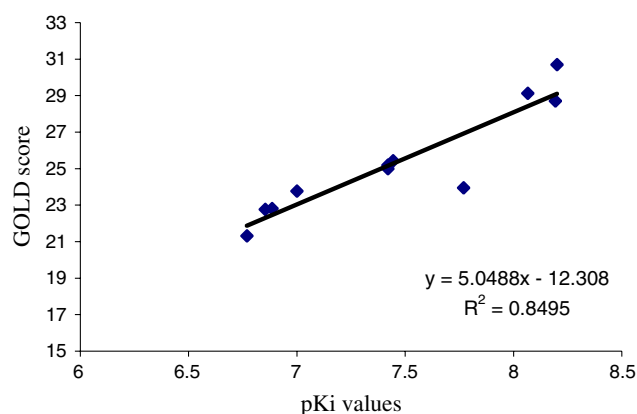


Fig. 11 Correlation plot of the predicted activities of the test set compounds obtained from pharmacophore model versus their *GOLD* docking scores obtained from hKOP model. Thus, we may conclude that our validated model indeed presents an appropriate platform for structure-based drug design and may be used for GPCR lead finding and compound optimization

integrated to improve the effectiveness and performance of database searching of the drug discovery process via VS. Such a combination should lower the imprecision embedded into either systems and allow the weakness of one approach to be compensated by the strength of the other, leading to a more reliable solution than obtained by each method separately.

In this regard, a ligand-based quantitatively predictive pharmacophore was generated and validated for salvinorin A-based KOP receptor agonists. The resulting best hypothesis consisted of two HBA and three hydrophobic features. The reliability of the pharmacophore in quantitative terms was verified in several validation procedures. The pharmacophore model elaborated within this study may therefore, prove useful as a virtual high-throughput screen to identify novel molecular scaffolds as potential lead molecules.

In addition, a computational approach was utilized to determine the complete 3D structure of hKOP receptor utilizing the recently published 2.2 Å resolution X-ray crystal structure of bovine rhodopsin. The model was refined through molecular dynamics and energy minimization. The quality of the model was assessed using several validation tools like PROCHECK and PROS-TAT. A model structure of the protein–ligand complex was also obtained which may be used as a tool to enhance our structure-based understanding of salvinorin A recognition at the hKOP. Our proposed receptor–ligand binding mode rationalized the observed single-point mutation data and SAR data. The model also helped to identify several critical ligand–receptor contact points in the binding site that might serve as the key

molecular determinants of agonist recognition by the hKOP receptor. Binding site residues such as Gln115 (Q2.60) and Tyr313 (Y7.43) were observed to interact via hydrogen bond while some other residues like Tyr119 (Y2.64), Tyr320 (Y7.43) and Tyr313 (Y7.43) were found to interact via hydrophobic and/or aromatic interactions.

The model was further validated by docking a series of salvinorin A derivatives into the binding pocket of the receptor model. A significant correlation was obtained between the binding affinities and the docking scores. This underscores the predictive ability of the model in identifying salvinorin A derivatives with high binding affinity.

Remarkably, the independently generated pharmacophore and comparative model of hKOP receptor complemented each other and provided similar conclusions. Whereas the ligand-based pharmacophore model revealed binding features, which are already captured by ligands, the steric factors were less well accounted for or completely neglected. In an attempt to overcome this, inclusion of complementary information from the 3D receptor model was made that allowed for a comprehensive molecular recognition process. In the future, we propose to employ a hybrid computational screening approach for the discovery of several classes of novel and potent KOP receptor ligands. It is anticipated that the employed hybrid approach may be more effective for lead discovery from a large chemical database than either pharmacophore-based or structure-based screening methods alone. Thus, in consideration of the enormous efforts of in vitro screening of large compound libraries, the in silico database mining method presented in our study aims at enabling a fast and reliable identification of new lead compounds that could decisively reduce time and cost of hit finding within the drug discovery and development process, particularly in the case of GPCRs.

Supporting information available

A representative graph for the activity prediction of the test set compounds (Fig. S1) and chemical structures, binding affinities and the corresponding *GOLD* scores of the compounds used in automated docking simulation studies (Fig. S2). Pharmacophore validation results with negative controls are listed in Table S1. The estimated activities of the test set molecules (derived from the pharmacophore model) and their respective *GOLD* scores (obtained from homology derived hKOP receptor model) were used to generate Fig. 11 (Table S2).

Acknowledgements Support for this work was provided by NIDA grant DA017360.

References

- Bleicher KH, Bohm HJ, Muller K, Alanine AI (2004) *Curr Opin Chem Biol* 8:287
- Oprea TI, Matter H (2004) *Curr Opin Chem Biol* 8:349
- Bajorath J (2001) *J Chem Inf Comput Sci* 41:233
- Good AC, Krystek SR, Mason JS (2000) *Drug Discov Today* 5:61
- Patel Y, Gillet VJ, Bravi G, Leach AR (2002) *J Comput Aided Mol Des* 16:653
- Schneider G, Böhm HJ (2002) *Drug Discov Today* 7:64
- Abagyan R, Totrov M (2001) *Curr Opin Chem Biol* 5:375
- Bissantz C, Folkers G, Rognan D (2000) *J Med Chem* 43:4759
- Taylor RD, Jewsbury PJ, Essex JW (2002) *J Comput Aided Mol Des* 16:151
- Brooijmans N, Kuntz ID (2003) *Annu Rev Biophys Biomol Struct* 32:335
- Lyne PD (2002) *Drug Discov Today* 7:047
- Halperin I, Ma B, Wolfson H, Nussinov R (2002) *Proteins* 47:409
- Guner O (ed) (2000) *Pharmacophore perception, development and use in drug design*. International University Line, La Jolla, CA
- Marrone TJ, Briggs JM, McCammon JA (1997) *Ann Rev Pharmacol Toxicol* 37:71
- Dixon JS (1997) *Proteins Suppl* 1:198
- Tame JRH (1999) *J Comput Aided Mol Des* 13:99
- Gohlke H, Klebe G (2001) *Curr Opin Struct Biol* 11:231
- Schwalbe H, Wess G (2002) *Chembiochem* 3:915
- Flower DR (1999) *Biochim Biophys Acta* 1422:207
- Wise A, Gearing K, Rees S (2002) *Drug Discov Today* 7:235
- Drews J (2000) *Science* 287:1960
- Takeda S, Kadowaki S, Haga T, Takaesu H, Mitaku S (2002) *FEBS Lett* 520:97
- Collins FS (2004) *Nature* 431:931
- Marriott DP, Dougall IG, Meghani P, Liu YJ, Flower DR (1999) *J Med Chem* 42:3210
- Flohr S, Kurz M, Kostenis E, Brokovich A, Fournier A, Klabunde T (2002) *J Med Chem* 45:799
- Pebay-Peyroula E, Rummel G, Rosenbusch JP, Landau EM (1997) *Science* 277:1676
- Belrhali H, Nollert P, Royant A, Menzel C, Rosenbusch JP, Landau EM, Pebay-Peyroula E (1999) *Struct Fold Des* 7:909
- Palczewski K, Kumasaka T, Hori T, Behnke CA, Moto-shima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) *Science* 289:739
- Teller DC, Okada T, Behnke CA, Palczewski K, Stenkamp RE (2001) *Biochemistry* 40:7761
- Ballesteros J, Palczewski K (2001) *Curr Opin Drug Discov Dev* 4:561
- Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V (2004) *J Mol Biol* 342:571
- Beeley NRA, Sage C (2003) *Targets* 2:19
- Evers A, Klabunde T (2005) *J Med Chem* 48:1088
- Varady J, Wu X, Fang X, Min J, Hu Z, Levant B, Wang S (2003) *J Med Chem* 46:4377

35. Bissantz C, Bernard P, Hibert M, Rognan D (2003) *Proteins* 50:5
36. Evers A, Hessler G, Matter H, Klabunde T (2005) *J Med Chem* 48:5448
37. Evers A, Klebe G (2004) *J Med Chem* 47:5381
38. Kieffer BL (1995) *Cell Mol Neurobiol* 15:615
39. Ulloa-Aguirre A, Stanislaus D, Janovick JA, Conn PM (1999) *Arch Med Res* 30:420
40. Gether U (2000) *Endocr Rev* 21:90
41. Chen Y, Mestek A, Liu J, Yu L (1993) *Biochem J* 295:625
42. Evans CJ, Keith DE Jr, Morrison H, Magendzo K, Edwards RH (1992) *Science* 258:1952
43. Kieffer BL, Befort K, Gaveriaux-Ruff C, Hirth CG (1992) *Proc Natl Acad Sci USA* 89:12048
44. Yasuda K, Raynor K, Kong H, Breder CD, Takeda J, Reisine T, Bell GI (1993) *Proc Natl Acad Sci USA* 90:6736
45. Meng F, Xie GX, Thompson RC, Mansour A, Goldstein A, Watson SJ, Akil H (1993) *Proc Natl Acad Sci USA* 90:9954
46. Chen Y, Mestek A, Liu J, Hurley JA, Yu L (1993) *Mol Pharmacol* 44:8
47. Fukuda K, Kato S, Mori K, Nishi M, Takeshima H (1993) *FEBS Lett* 327:311
48. Knapp RJ, Malatynska E, Fang L, Li X, Babin E, Nguyen M, Santoro G, Varga EV, Hruby VJ, Roeske WR et al (1994) *Life Sci* 54:PL463
49. Wang JB, Johnson PS, Persico AM, Hawkins AL, Griffin CA, Uhl GR (1994) *FEBS Lett* 338:17
50. Mansson E, Bare L, Yang D (1994) *Biochem Biophys Res Commun* 202:1431
51. Mollereau C, Parmentier M, Maillieux P, Butour JL, Moisan C, Chalon P, Caput D, Vassart G, Meunier JC (1994) *FEBS Lett* 341:33
52. Meunier JC (1997) *Eur J Pharmacol* 340:1
53. Pogozheva ID, Przydzial MJ, Mosberg HI (2005) *AAPS J* 7:E434
54. Metcalf MD, Coop A (2005) *AAPS J* 7:E704
55. Rees DC (1992) *Prog Med Chem* 29:109
56. Roth BL, Baner K, Westkaemper R, Siebert D, Rice KC, Steinberg S, Ernsberger P, Rothman RB (2002) *Proc Natl Acad Sci USA* 99:11934
57. Chavkin C, Sud S, Jin W, Stewart J, Zjawiony JK, Siebert DJ, Toth BA, Hufeisen SJ, Roth BL (2004) *J Pharmacol Exp Ther* 308:1197
58. Siebert DJ (1994) *J Ethnopharmacol* 43:53
59. Wang Y, Tang K, Inan S, Siebert D, Holzgrabe U, Lee DY, Huang P, Li JG, Cowan A, Liu-Chen LY (2005) *J Pharmacol Exp Ther* 312:220
60. Yan F, Roth BL (2004) *Life Sci* 75:2615
61. Ortega A, Blount JF, Manchand PS (1982) *J Chem Soc Perkins Trans* 1:505
62. Wasson RG (1962) *Bot Mus Leaflet Harv Univ* 20:77
63. Valdes LJ, Butler WM, Hatfield GM, Paul AG, Koreeda M (1984) *J Org Chem* 49:4716
64. Valdes LJ III, Diaz JL, Paul AG (1983) *J Ethnopharmacol* 7:287
65. Rees DC, Hunter JC (1990) *Comprehensive medicinal chemistry*. Pergamon, New York, p 805
66. Beguin C, Richards MR, Wang Y, Chen Y, Liu-Chen LY, Ma Z, Lee DY, Carlezon WA Jr, Cohen BM (2005) *Bioorg Med Chem Lett* 15:2761
67. Mowry M, Mosher M, Briner W (2003) *J Psychoactive Drugs* 35:379
68. Valdes LJ III (1994) *J Psychoactive Drugs* 26:277
69. McCurdy CR, Sufka KJ, Smith GH, Warnick JE, Nieto MJ (2006) *Pharmacol Biochem Behav* 83:109
70. Evers A, Gohlke H, Klebe GJ (2003) *Mol Biol* 334:327
71. *Catalyst*, version 48, Accelrys: Burlington, MA
72. Thompson JD, Higgins DG, Gibson TJ (1994) *Nucleic Acid Res* 22:4673
73. *InsightII*, Molecular Modeling Program Package, Accelrys: 9685 Sranton Road, San Diego, CA 92121–3752, USA
74. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) *J Appl Cryst* 26:283
75. *GOLD version 21*, Cambridge Crystallographic Data Centre: Cambridge, UK
76. *Sybyl version 72*, Tripos Inc, St Louis, MO, USA
77. Munro TA, Rizzacasa MA, Roth BL, Toth BA, Yan F (2005) *J Med Chem* 48:345
78. Harding WW, Tidgewell K, Byrd N, Cobb H, Dersch CM, Butelman ER, Rothman RB, Prisinzano TE (2005) *J Med Chem* 48:4765
79. Lee DY, He M, Kondaveti L, Liu-Chen LY, Ma Z, Wang Y, Chen Y, Li JG, Beguin C, Carlezon WA Jr, Cohen B (2005) *Bioorg Med Chem Lett* 15:4169
80. Lee DY, Karnati VV, He M, Liu-Chen LY, Kondaveti L, Ma Z, Wang Y, Chen Y, Beguin C, Carlezon WA Jr, Cohen B (2005) *Bioorg Med Chem Lett* 15:3744
81. Li H, Sutter J, Hoffman R (1999) In: Guner OF (ed) *Pharmacophore: perception, development and use in drug design*. International University Line, La Jolla CA, p 173
82. Gund P (1977) In: Hahn FE (ed) *Progress in molecular and subcellular biology*, vol 5. Springer-Verlag, New York, p 117
83. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) *J Comp Chem* 4:187
84. Smellie A, Teig SL, Towbin P (1999) *J Comput Chem* 16:171
85. English philosopher and Franciscan Monk, William of Ockham, 1285
86. Sutter J, Guner OF, Hoffman R, Li H, Waldman M (1999) In: Guner OF (ed) *Pharmacophore: perception, development and use in drug design*. International University Line, La Jolla CA, p 501
87. Fischer R (1966) *The design of experiments*, 8th edn, Chapter II. Hafner Publishing Co, New York
88. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilboud S, Schneider M (2003) *Nucleic Acids Res* 31:365
89. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) *Nucleic Acids Res* 28:235
90. Henikoff S, Henikoff JG (1992) *Proc Natl Acad Sci USA* 89:10915
91. Sali A, Blundell TL (1993) *J Mol Biol* 234:779
92. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F et al (2000) *Annu Rev Biophys Biomol Struct* 29:291
93. Yan F, Mosier PD, Westkaemper RB, Stewart J, Zjawiony JK, Vortherms TA, Sheffler DJ, Roth BL (2005) *Biochemistry* 44:8643
94. Andersen HC (1980) *J Chem Phys* 72:2384
95. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) *J Mol Biol* 267:727
96. Ballesteros JA, Weinstein H (1995) *Methods Neurosci* 25:366
97. Costanzi S, Lambertucci C, Vittori S, Volpini R, Cristalli G (2003) *J Mol Graph Model* 21:253
98. Hulme EC, Lu ZL, Ward SDC, Allman K, Curtis CA (1999) *Eur J Pharmacol* 375:247
99. Ghanouni P, Steenhuis JJ, Farrens DL, Kobilka BK (2001) *Proc Natl Acad Sci USA* 98:5997
100. Farrens DL, Altenbach C, Yang K, Hubbell WL, Khorana HG (1996) *Science* 274:768

101. Kim JM, Altenbach C, Kono M, Oprian DD, Hubbell WL, Khorana HG (2004) *Proc Natl Acad Sci USA* 101:12508
102. Hubbell WL, Altenbach C, Hubbell CM, Khorana HG (2003) *Adv Protein Chem* 63:243
103. Hanson SM, Francis DJ, Vishnivetskiy SA, Kolobova EA, Hubbell WL, Klug CS, Gurevich VV (2006) *Proc Natl Acad Sci USA* 103:4900
104. Patel AB, Crocker E, Reeves PJ, Getmanova EV, Eilers M, Khorana HG, Smith SO (2005) *J Mol Biol* 347:803
105. Han SJ, Hamdan FF, Kim SK, Jacobson KA, Bloodworth LM, Li B, Wess J (2005) *J Biol Chem* 280:34849
106. 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
107. Hiramoto T, Nonaka Y, Inoue K, Yamamoto T, Omatsu-Kambe M, Matsuura H, Gohda K, Fujita N (2004) *J Pharmacol Sci* 95:81
108. (a) Subramanian G, Paterlini MG, Larson D, Portoghese PS, Ferguson DM (1998) *J Med Chem* 41:4777; (b) Pogozheva ID, Lomize AL, Mosberg HI (1998) *Biophys J* 75:612; (c) Cappelli A, Anzini M, Vomero S, Menziani MC, De Benedetti PG, Sbacchi M, Clarke GD, Mennuni L (1996) *J Med Chem* 39:860; (d) Metzger TG, Paterlini MG, Portoghese PS, Ferguson DM (1996) *Neurochem Res* 21:1287; (e) Strahs D, Weinstein H (1997) *Protein Eng* 10:1019
109. Lee KW, Briggs JM (2004) *Proteins* 54:693
110. Archer E, Maigret B, Escrieut C, Pradayrol L, Fourmy D (2003) *Trends Pharmacol Sci* 24:36
111. Oliveira L, Hulsen T, Lutje Hulsik D, Paiva AC, Vriend G (2004) *FEBS Lett* 564:269
112. Waldhoer M, Bartlett SE, Whistler JL (2004) *Annu Rev Biochem* 73:953
113. (a) Karnik SS, Khorana HG (1990) *J Biol Chem* 265:17520; (b) Kurtenbach E, Curtis CA, Pedder EK, Aitken A, Harris AC, Hulme EC (1990) *J Biol Chem* 265:13702; (c) Dohlman HG, Caron MG, DeBlasi A, Frielle T, Lefkowitz RJ (1990) *Biochemistry* 29:2335
114. Kane BE, Nieto MJ, McCurdy CR, Ferguson DM (2006) *FEBS J* 273:1966
115. (a) Thirstrup K, Elling CE, Hjorth SA, Schwartz TW (1996) *J Biol Chem* 271:7875; (b) Hjorth SA, Thirstrup K, Schwartz TW (1996) *Mol Pharmacol* 50:977 (c) Xue JC, Chen C, Zhu J, Kunapuli S, DeRiel JK, Yu L, Liu-Chen LY (1994) *J Biol Chem*, 269:30195
116. Ballesteros J, Kitanovic S, Guarnieri F, Davies P, Fromme BJ, Konvicka K, Chi L, Millar RP, Davidson JS, Weinstein H, Sealfon SC (1998) *J Biol Chem* 273:10445
117. Scheer A, Fanelli F, Costa T, De Benedetti PG, Cotecchia S (1996) *EMBO J* 15:3566
118. Fritze O, Filipek S, Kuksa V, Palczewski K, Hofmann KP, Ernst OP (2003) *Proc Natl Acad Sci USA* 100:290
119. Zhou W, Flanagan C, Ballesteros JA, Konvicka K, Davidson JS, Weinstein H, Millar RP, Sealfon SC (1994) *Mol Pharmacol* 45:165
120. Sealfon SC, Chi L, Ebersole BJ, Rodic V, Zhang D, Ballesteros JA, Weinstein H (1995) *J Biol Chem* 270:16683
121. Folkers G (1998) In: Coddling PW (ed) *Structure-based drug design experimental and computational approaches*. Kluwer Academic Publishers, Norwell, MA USA, p 271