

# Pharmacophore Identification and Pseudo-Receptor Modeling

Gerhard Wolber<sup>1</sup> and Wolfgang Sippl<sup>2</sup>

<sup>1</sup>Computer-Aided Molecular Design, Institute of Pharmacy, Freie Universität Berlin, Königin-Luisenstr., Berlin, Germany;

<sup>2</sup>Medicinal Chemistry, Institute of Pharmacy, Martin-Luther-Universität Halle-Wittenberg, Halle (Saale), Germany

## OUTLINE

<b>I. Introduction</b>	<b>489</b>	<b>IV. Application Study: Novel Histamine H<sub>3</sub>-Receptor Antagonists</b>	<b>501</b>
A. Historical Background	490	A. Pharmacophore-Based Screening	501
B. Definitions	490	B. Pharmacophore Determination Process	502
C. Importance of the Pharmacophore Concept	491	C. Pharmacophore-Based Screening of Compound Libraries	504
D. Application of Pharmacophores	492	<b>V. Recent Developments and Outlook</b>	<b>505</b>
<b>II. Methodology</b>	<b>493</b>	<b>VI. Conclusions</b>	<b>507</b>
A. Pharmacophore Modeling	493	<b>References</b>	<b>507</b>
<b>III. Advanced Approaches</b>	<b>497</b>		
A. Structure-Based Pharmacophores	497		
B. Pseudo-Receptor Models	498		

*In theory, theory and practice are the same. In practice, they are not. Lawrence Peter Berra*

## I. INTRODUCTION

In the large majority of cases, the basis for a pharmacodynamic effect is the interaction of a certain substance with a biomacromolecule of physiological importance. Proteins—like enzymes, receptors, and ion channels—but also nucleic acids serve as physiological binding partners for small organic ligands. In all cases, a highly specific 3D binding epitope must exist, which serves as complementary binding site for a drug molecule. Compounds exerting similar activities on the same enzyme or receptor therefore possess—in most cases—closely related binding properties. That is, these molecules present structural elements of identical chemical features in sterically consistent locations to the macromolecule. The highest common denominator of a group of ligands exhibiting a similar biological effect recognized by the same binding site is named a “pharmacophore.” [1] In other words, a pharmacophore is an abstraction of the crucial molecular features responsible for the binding of a set of ligands to a macromolecular target.

As a practical matter, computer-aided molecular design is frequently split into disciplines that focus on either structure-based or ligand-based methods. When the 3D structure of a target protein and the binding site is available,

it is possible to invoke structure-based approaches. New candidate ligands may be docked into a particular binding site in order to study whether they can interact with the protein in an optimal way. If, however, knowledge about the structure of the macromolecular target is limited but a sufficient number of active analogues have already been discovered, then pharmacophore-based methods are applied to design novel active molecules. It may seem straightforward to develop new ligands for known proteins by applying structure-based approaches, but significant problems are involved. Induced fit mechanism, multiple binding modes, solvation, and entropic effects are some of the problems that must be overcome to end up with reliable models. Beside these problems, many target proteins of high pharmaceutical interest are membrane-bound receptors (e.g., G-protein coupled receptors [GPCRs]) and attempts to crystallize them have been only partially successful. Although twenty-four GPCR crystal structures from four different classes have been published at this point, structure-based 3D pharmacophore development with these structures remains a challenge. A larger amount of 3D structures for activated and inactivated structures would be necessary to be able to develop structure-based predictive models for agonists and inverse agonists, respectively.

In the absence of the 3D structure of a protein of interest or a biologically relevant conformation, ligand design may be performed by the use of a pharmacophore-based method. This is based on the assumption that several ligands bind to the same binding pocket of the protein. Thus, a flexible superposition can be identified, which represents the interaction pattern of the binding pocket from the view of the ligands.

## A. Historical Background

The idea that bioactive substances interact with receptors began in 1878 with Langley, who introduced the term “receptive substance.” [2] However, the term “receptor” was introduced several years later by Paul Ehrlich [3]. He also introduced the term “pharmacophore” to describe those parts of a molecule that are responsible for its activity. Together with the lock-and-key concept of Emil Fischer, it became clear that not all parts of a molecule—the “key”—are equally important for exerting its biological effect on the “lock.” [4] Thus, sometimes small variations of distinct parts of a molecule can dramatically influence the activity, whereas variations of other parts only cause minor changes in the biological activity. The concepts of Langley, Ehrlich, and Fischer constitute the cornerstones of modern drug discovery and development up to this day. Half a decade later, their concepts were confirmed in an impressive manner by the first solved crystal structures of protein–ligand complexes [5].

Even before the advent of computers and modeling software, simple pharmacophores were described in the literature and considered tools for the discovery of novel molecules. Based on initial structure–activity relationship considerations, simple 2D models were introduced in the 1940s. With the advent of computers and modeling programs, the idea of displaying and manipulating 3D structures became possible [6]. Kier and Marshall pioneered the development of the pharmacophore concept and its application in structure–activity relationships [7,8]. In the 1970s, Peter Gund implemented the first *in silico* screening method with a program to screen a substance library for pharmacophoric patterns [9]. The active analogue approach developed by Garland Marshall’s group was one of the first automated tools for pharmacophore generation. Marshall’s approach was the basis for many following pharmacophore modeling programs in that area. Since these early days, a variety of automated pharmacophore discovery programs have been developed in academia and by software developing companies (for review, see [10,11]).

## B. Definitions

The term pharmacophore is not always used by different groups of scientists in accordance with the official definition elaborated by the IUPAC working party, which states [1]: “A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interaction with a specific biological target structure and to trigger (or block) its biological response.” Many scientists use the term “pharmacophore” or “pharmacophoric group” to define distinct functional groups or substance classes possessing biological activity (e.g., sulfonamides or dihydropyridines). In this context, the term pharmacophore is mixed with another concept of structure and activity, namely “privileged structures.” The retrospective analysis of the chemical structures and scaffolds of drug molecules led to the detection of some structural motifs that are often associated with biological activity. Such motifs were called “privileged structures” by Evans et al to represent substructures that confer activity toward two or more different targets [12]. The idea behind this is that the privileged structure provides the scaffold and the substitution provides the reason for specificity. However, in terms of the IUPAC definition, the pharmacophore represents the common molecular interaction features of a set of molecules toward their receptor [13].

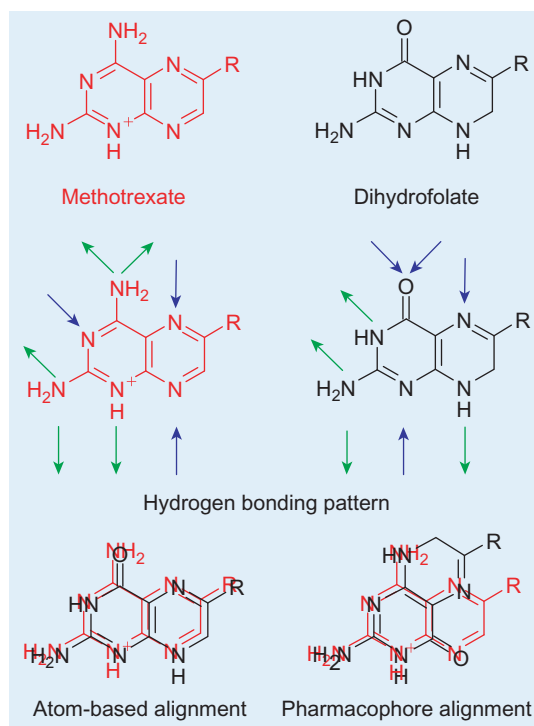
A pharmacophoric element (also called feature) is generally defined as an atom or a group of atoms (e.g., a hydrogen bond donor atom or an aromatic ring system) common to active compounds with respect to a target protein and essential for the activity. Thus, a pharmacophore model can also be regarded as the representation of a collection of pharmacophore features.

The above-described definition of a pharmacophore is based on a 3D point-of-view of molecules. It reflects the way medicinal chemists characterize the binding ability of molecules for a given target protein. However, depending on the different research areas, scientists have different views. Computational chemists often use the term pharmacophore in a more abstract way. Influenced by the structural representation of molecules, a set of topological connections is used to define the properties and dimension of a molecule in 2D. Here, the spatial and topological distribution of pharmacophoric features is converted to a lower dimensional representation (e.g., vectors). Such vectors, which represent pharmacophore descriptors, are called “fingerprints,” “keys,” “bitstrings,” or “correlation vectors,” depending on the type of information stored. The pharmacophore descriptors or fingerprints can be regarded as a transformed molecular representation instead of an explicit 3D structure. These fingerprints are often used to screen large compound libraries rapidly. In this chapter, we will only focus on 3D pharmacophore concepts.

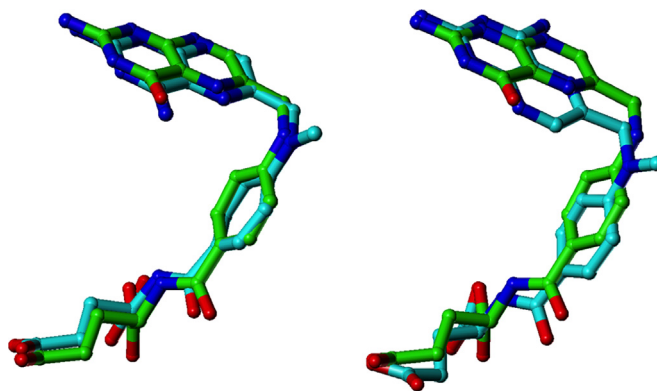
Starting from a preliminary pharmacophore model, a hypothetical receptor consisting of individual amino acid residues can be constructed surrounding a set of superimposed ligands. Guided by permanent correlation of biological data and model-derived calculated free energies of binding, a complex system is generated, mimicking the interaction pattern of a real binding site reasonably well. The resulting hypothetical receptor model is named “mini-receptor” or “pseudo-receptor,” and can be used to derive 3D quantitative structure–activity relationships (3D-QSAR). The concept was originally developed in the 1980s by several groups [14–17].

### C. Importance of the Pharmacophore Concept

A pharmacophore captures the concept of bio-isosterism by not only comparing topological similarities but structural groups at similar locations with the same chemical functionality. It is important to concentrate on the pharmacophoric features, since topological molecule characteristics are often misleading in the superpositioning of two molecules with respect to their binding mode. Figure 21.1 shows the well-known example of dihydrofolate reductase ligands [18]. For the two ligands shown, a topological overlay would result in an incorrect prediction



**FIGURE 21.1** Comparison of atom-based alignment and experimentally derived position of methotrexate and dihydrofolate in dihydrofolate reductase. Blue arrows indicate hydrogen bond acceptors, and green arrows show hydrogen bond donors.



**FIGURE 21.2** Comparison of the pharmacophore-based alignment of dihydrofolate reductase ligands (left) and the experimentally derived protein-based alignment (right).

of the binding mode. If the pharmacophoric features (the hydrogen bonding pattern in this example) are taken into account for the superimposition, the correct overlay mode can be deduced. The pharmacophore-based superposition is similar to the binding mode observed in the crystal structures of methotrexate and dihydrofolate with dihydrofolate reductase (Figure 21.2; PDB IDs 1RX2, and 1RB3).

The increasing number of accessible compounds that can currently be used as starting points for biological target screening makes it necessary to have fast and reliable *in silico* screening tools. Structure-based methods are often too slow to screen compound databases with millions of molecules virtually. Beside the speed, other problems in structure-based design and docking programs need to be addressed. For instance, most of the current docking programs do not take into account protein flexibility. Only recently were programs developed (e.g., AutoDock4 [19], GOLD Suite 5.2 [20], Glide [21] or FlexE [22]) that consider protein side-chain flexibility for docking. Other problems, which often occur in ligand docking, are the correct placement of water molecules within the binding site (which represents putative ligand binding partners), the treatment of solvation effects (on the ligand and protein site), and consideration of the internal strain of a docked ligand. Structure-based approaches are able to provide important information about the interaction between a ligand and a macromolecule, but the accurate prediction of the binding affinity is still an unsolved problem. A detailed discussion about the limitations of docking and scoring programs can be found in several reviews [23–28].

Another reason pharmacophore-based approaches are often used in drug design is the missing 3D structure of many interesting macromolecules. Many current drug targets are membrane-bound and, despite recent progress in crystallizing GPCRs [29–31], only a small fraction of membrane proteins have been successfully crystallized. An additional challenge for some membrane-bound receptors is that ligands bind in the extracellular region, causing conformational changes in the protein that lead to a signaling response, such as the release of G-proteins in the case of GPCRs. It is still very difficult to capture the “right” conformation in a crystal, especially when it comes to modeling ligand function (e.g., agonists or antagonists). In the absence of an experimentally determined 3D protein structure, the use of indirect ligand-based approaches—including pharmacophores—is the only way to design novel bioactive molecules rationally [32].

#### D. Application of Pharmacophores

Pharmacophore modeling in computer-aided drug design is generally applied in three domains. The first is the definition of relevant pharmacophoric features in a drug molecule necessary to achieve a certain biological effect and to establish clear structure–activity relationships. A well-developed pharmacophore model, preferentially including information about the dimension of the receptor binding cavity, may be employed to design novel and more active molecules that fit the model. Often, such pharmacophore models are the starting point for 3D-QSAR analysis (e.g., CoMFA [33]), by which quantitative predictions may be made. The second is scaffold hopping, the practice of detecting molecules with different scaffolds (novel chemotypes) by virtually screening large compound libraries [34]. The third domain is the use of parallel pharmacophore-based screening in order to predict pharmacological profiles for lead structures *in silico*. The use of 3D pharmacophore models can hopefully predict unwanted side effect in very early stages of the drug-discovery process and therefore reduce the risk of late failure of drug candidates [35].

## II. METHODOLOGY

### A. Pharmacophore Modeling

To end up with a predictive pharmacophore model, it is necessary to start with reliable structural and biological data. First of all, it is important to have correct 3D structures of all compounds under study. Thus, atomic valences, bond orders, protonation state, and stereochemistry have to be checked carefully. Also, the consideration of different possible tautomers is necessary when the bioactive form is not known exactly. Another prerequisite is the existence of a similar binding mode of all ligands under study. Experimental data, from competition experiments or protein–ligand crystal structures, can clearly point out that the ligands interact with the same binding epitope in a similar way instead of on distinct binding sites.

The four steps in the development of a pharmacophore model are: (a) selection of a set of active ligands known to bind to the same target (same binding site); (b) conformational analysis for all ligands; (c) assignment of pharmacophoric features; and (d) molecular superimposition of the ligand conformations to develop a common 3D-pharmacophore. The majority of automated pharmacophore generation programs use qualitative pharmacophore models that do not consider the activity of the ligands. The ultimate goal of all these programs is to search for a unique conformation of all congeners, where most if not all assigned pharmacophoric features of the ligands are presented in a superimposed manner. Most of the programs are based on minimizing the root-mean-square (RMS) superposition error between conformations of the ligands under study while trying to increase the fit of the pharmacophoric features. To compare the different conformations for a data set of given active molecules, a superpositioning procedure is needed. The assignment of the pharmacophoric features and the generation of the ligand alignment is carried out in an automated way by most of the current pharmacophore modeling programs (e.g., Catalyst [36], DISCO [37], Galahad [38], LigandScout [39], Phase [40], MOE [41]). The scope of this chapter is not to describe all available software packages in detail, but to illustrate the different steps of the pharmacophore development process. For a recently published overview of current pharmacophore modeling programs, the reader is referred to the literature [10,42–44].

#### 1. Conformational Analysis of Ligand Molecules and Bioactive Conformation

Since molecules are flexible and not static, a conformational analysis has to be carried out first to generate an ensemble of low-energy conformations. This is probably one of the most critical steps in the pharmacophore discovery process, since the goal is not only to consider the global minima of a molecule but also to include the bioactive conformation as part of an ensemble of low-energy conformations.

In order to bind to a protein with high affinity, a ligand must match the binding pocket. The steric match will thereby depend primarily on the ligand conformation. Within a binding pocket, the ligand will not necessarily be present in its lowest energy conformation, as the gain in interaction energy with the receptor can compensate for a conformation with higher energy [45]. Still, it can be expected that for a high-affinity ligand, the bioactive conformation is at least energetically favorable, as otherwise the conformational energy cost would reduce binding affinity. The relation between a high energetic binding conformation and the loss of free energy of binding  $\Delta G$  is given by Equation 21.1:

$$\Delta G = -2.303RT \log K_i \quad (21.1)$$

Under physiological conditions ( $T = 310 \text{ K}$ ), the free energy (in  $\text{kcal mol}^{-1}$ ) and the binding affinity are related by

$$\Delta G = -1.42 \log K_i \quad (21.2)$$

Thus, if a compound binds in a conformation that deviates  $1.42 \text{ kcal mol}^{-1}$  from the global minimum structure, its affinity will be decreased by one order of magnitude. High-affinity compounds can thus be expected to bind in an energetically favourable conformation. To analyze the conformational space of molecules experimental and theoretical approaches are applied. Experimental techniques like NMR only provide information on one or a few conformations of a molecule. A complete overview about the conformational space of molecules can be gained only by theoretical techniques [46]. Correspondingly a variety of theoretical methods for conformational analysis has been developed. The most general conformational analysis methods are those that are able to identify all minima on the potential energy surface. However, as the number of minima dramatically increases with

the number of rotatable bonds, an exhaustive detection of all minima becomes a difficult and time-consuming task. Commonly used methods for this purpose are listed below (described in depth in [47]):

- Systematic search: Each bond is rotated incrementally and the resulting structures are minimized. Systematic search algorithms have the advantage of sampling the conformational space very well. In cases with a high number of rotatable bonds, this method may be computationally impracticable [48].
- Random search: In a random search, one can move from one region of the energy surface to a completely unconnected region in a single step. A commonly applied method is the Metropolis Monte Carlo scheme that starts with a minimized conformation A of a molecule. A random move on the energy-landscape is carried out (e.g., torsion angles are rotated by a random amount), and the structure is minimized. The potential energy of the output structure B is evaluated. If  $E_{pot}(B) < E_{pot}(A)$ , the new conformation is accepted. If  $E_{pot}(B) > E_{pot}(A)$ , the move may still be accepted depending on the transition probability, which in turn depends on the temperature. Monte Carlo methods efficiently sample the conformational space, but there is no guarantee—as with all random search tools—that the entire energetic landscape will be sampled. Another sampling technique applied to the problem of improved conformational searching is known as Poling [49]. Poling is implemented within Catalyst [36] allowing the generation of large multi-conformer virtual screening databases in a reasonable amount of time.
- Simulated annealing or molecular dynamics (MD) simulations: The aim of MD simulations is to reproduce the time-dependent motional behavior of a molecule. MD is based on molecular mechanics. It is assumed that the atoms in the molecule interact with each other according to the rules of an employed force field. MD simulations generate an ensemble of coordinates that does not only contain minimized structures, but rather provides a (limited) sampling of conformational space. In a simulated annealing MD protocol, the system temperature is periodically increased, resulting in a significant rise of kinetic energy, which makes it easier to overcome barriers of potential energy. Subsequently, the system is cooled down, thereby trapping the molecule in an energetically favorable conformation. MD simulation techniques for sampling the conformational space are quite time-consuming and are therefore used only for smaller ligand data sets. Again, there is no guarantee of sampling the entire potential energy surface [50].

There is an ongoing discussion in the literature about which ligand conformations (i.e., within which energy range) have to be considered in a pharmacophore generation process. Several recent studies on protein–ligand X-ray structures have shown that many conformational search tools yield ensembles, including the experimentally observed bioactive conformation [51]. The energy difference between the co-crystallized conformation of a ligand and its global minimum calculated with molecular mechanic programs is dependent on the force field employed. Therefore, a general energy range to be considered cannot be defined [52].

Which conformational analysis performs best? A clear-cut answer cannot be given, as it depends on the individual data set to be studied and the problems to be addressed. If only a limited number of ligands is considered, more computationally intensive methods such as the systematic search can be applied. If a compound library with hundreds of thousands of entries has to be converted into a multi-conformer database, faster simplified approaches have to be used (e.g., in Catalyst [36] or Omega [53]) [54].

## 2. Pharmacophore-Ligand Superposition Techniques

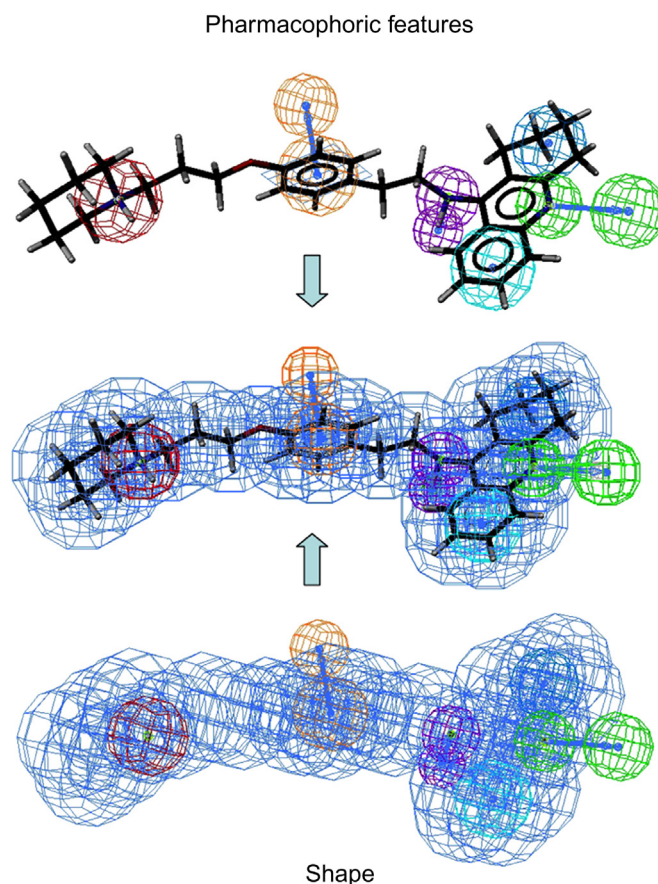
Three-dimensional pharmacophore–ligand superpositioning has to deal with the challenge of conformational flexibility. One possibility for addressing the problem is to perform the identification of common chemical features and the conformational search simultaneously (flexible alignment). Other approaches pre-generate conformations (rigid-body alignment), which makes the overlay algorithm faster, but the inclusion of relevant conformations must be ensured.

A well-known pharmacophore elucidation program using a flexible approach is GASP [55,56], which was initially developed by Jones and co-workers in the mid-1990s. The software is based on a genetic algorithm that simulates evolution by randomly combining and mutating chromosomes of an initial population. Each chromosome represents a potential flexible pharmacophore by encoding all torsion angles and by listing all feature mappings to a manually selected rigid reference compound. In each run, highest scoring chromosomes are selected according to a simple fitness function, and those are then mutated by applying random torsional rotations to cover conformational space during the alignment process. Today, most programs use rigid-body alignment techniques, because conformers only have to be calculated once and can be stored in a database, which saves computational time during the alignment.

The wide field of rigid-body superpositioning algorithms (also referred to as “alignment techniques”) can roughly be divided into 3D geometry-based and linearized fingerprint-based (descriptor-based approaches). Fingerprint-based approaches create a linearized bit sequence representing the chemical feature properties of the alignment partners. They allow for fast computational similarity assessment. A very advanced example for fingerprint approaches is the Chemically Advanced Template Search (CATS) developed in Gisbert Scheider’s group [57]. However, no real 3D overlay is produced using such an implementation. With geometry-based approaches, chemical features represented as 3D points with optional geometry constraints (such as vectors or planes) are assigned to 3D conformations of the ligands. In a subsequent step, these algorithms attempt to minimize distances between those points while considering the assigned constraints like the parallel orientation of the planes or the overlap of the projection point of a vector. Computationally expensive solutions to this problem have been proposed relatively early and range from 3D maximum clique detection algorithms [58] as used in DISCO [59,60] to the sequential build-up of increasingly larger common feature configurations as employed in Catalyst [36], Phase [40], or MOE [41]. This approach becomes problematic if pharmacophore point tolerances shall be subsampled, and it results in geometric fuzziness of the resulting alignment. LigandScout [61] uses a novel and computationally more efficient pattern-matching technique [62] to identify an initial alignment. With this technique, it is possible to perform a geometrically more accurate alignment that also subsamples chemical feature point tolerances when performing high-throughput virtual screening.

### 3. Assignment of Pharmacophoric Elements

The assignment of pharmacophoric features shall be described using as an example the histamine H<sub>3</sub> receptor antagonist shown in Figure 21.3 [63]. Table 21.1 lists the pharmacophoric features assigned in the ligand structure by comparison with other known active antagonists. Thus, the protonated nitrogen atom of the piperidyl moiety can be translated into a center of a sphere with coordinates corresponding to the location of the nitrogen atom



**FIGURE 21.3** (top) Pharmacophoric features observed in the ligand (by comparison with other known active ligands). (bottom) The molecule’s shape can serve as an additional constraint in pharmacophore searches. The resulting pharmacophore is based on features and shape (middle).

**TABLE 21.1** Pharmacophoric features observed in the ligand shown in Figure 21.3

Feature	Color	Representation
positive charge	red	sphere
H-bond donor	magenta	sphere-vector-sphere
H-bond acceptor	green	sphere-vector-sphere
hydrophob. aliphatic	blue	sphere
aromatic ring	orange	plane, center of plane, vector
hydrophobic	light blue	sphere

and a radius defining a volume around this atom. If a molecule is compared to this pharmacophore model and its protonated nitrogen atom lies within the sphere, this pharmacophoric feature will be said to be matched. The bigger the sphere, the easier it will be for a ligand conformation to match the pharmacophoric features. Similarly, an aromatic or a hydrophobic aliphatic moiety can be defined by a center of sphere and radius. Hydrogen bond acceptors and donors are represented by vectors in order to account for the directionality of H-bonds, while aromatic rings can be either defined by spheres or the combination of center, plane, and vector. When defined this way, the orientation of the aromatic plane in respect to the rest of the molecule is considered, too.

Again, the shape of the molecule can be incorporated into a pharmacophore by translating the van der Waals volume into an additional feature. If the ligand is known to fill the binding pocket well, the available volume can be taken into account. The abstract definition of a molecule in form of a pharmacophore as defined in Figure 21.3 facilitates comparison with other molecules. In the given example, most features of the antagonist were considered for the generation of the pharmacophore model resulting in an fingerprint of the molecule. Depending on the number of features included in the model and the tolerances defined, it will be more or less difficult for other molecules to match the pharmacophore model. Matching a pharmacophore additionally depends strongly on the conformation adopted by the molecule that is compared to the pharmacophore model. Even a different conformation of the ligand might not match the pharmacophore model defined above. Thus, in order to increase the likelihood of a specific molecule fitting a pharmacophore model, each molecule of interest is associated with a conformational ensemble. When searching for similarities with the pharmacophore model, all conformations of a molecule are tested on the pharmacophore before the best fit is evaluated. The difficulty in defining a useful pharmacophore model lies in the restriction to essential pharmacophoric features observed in the active ligands.

#### 4. Model Quality and Pharmacophore-Based Virtual Database Screening

If a pharmacophore is used for database screening in order to retrieve new compounds based on the similarity of pharmacophoric features, a model is useful when it is able to identify known actives among a number of inactive molecules. In order to screen commercial compound databases with a pharmacophore model, a so-called multi-conformer database must first be generated. This means a set of conformations must be generated for all compounds deposited in the compound databases. Since such databases can include millions of compounds, fast algorithms are paramount. In addition, the conformation database should not lead to an explosion in storage requirements for the millions of conformers. Finally, the database program should be able to handle the pharmacophore search within a reasonable amount of time. The most widely available commercial programs for building large multi-conformation databases are Catalyst [36], UNITY [64], Omega [53], and MOE [41]. Whereas Catalyst, UNITY, and MOE are also used to carry out pharmacophore generation and pharmacophore searches, Omega can only be used to generate multi-conformer databases. A comparison of the performance of the different programs can be found in the literature [51].

Recently, freeware alternatives to the above-mentioned programs have become available, but they are yet to be evaluated in terms of their performance when used with pharmacophore development and virtual screening [65].

The pre-calculation of conformations bears the important advantage that the screening process is considerably faster and avoids a dramatic reduction of the conformational search space by falling into a local minimum [66,44]. With current computer hardware, the additionally required storage space no longer represents a limitation, and screening databases with pre-generated conformations are clearly preferred. These databases can be generated once and reused for subsequent virtual screening runs, which results in a considerable speed-up of the overall screening procedure.



Typically, a 3D pharmacophore is first created as a hypothesis and analyzed retrospectively to assess its predictive power before being used for a prospective virtual screening that should predict whether new molecules with a certain pharmacophore pattern actually bind to the protein under investigation. To assess retrospective screening performance, several virtual screening metrics have been established that can be used to evaluate model quality [67–69]. These metrics describe the ability of a pharmacophore model to match bioactive molecules and thus include them in the virtual hit list versus the ability to exclude biologically inactive molecules. The most frequently used metrics are described below. For a more extensive overview, the reader is referred to reviews on this topic [70,71].

*Sensitivity (Se)* describes the ratio of the retrieved true positive compounds (TP) in relation to all biologically active compounds in the database that is the sum of TP and the number of false negatives (FN). Sensitivity values range from 0 to 1, where a value of 0 means that the search did not yield any actives in the database and a value of 1 indicates that all active compounds could be retrieved.

$$Se = \frac{TP}{TP + FN}$$

*Specificity (Sp)* describes the portion of rejected biologically inactive compounds (true negatives, TN) divided by the sum of TN and the number of false positives (FP). Specificity ranges from 0 to 1, where a value of 0 means that none of the inactive compounds could be excluded and a value of 1 indicates that no inactive compounds fit the pharmacophore hypothesis.

$$Sp = \frac{TN}{TN + FP}$$

*Yield of Actives (Ya)* sets the amount of true positives (TP) in relation to the size of the hit list (n).

$$Ya = \frac{TP}{N}$$

*Enrichment Factor (EF)* measures the Ya proportionally to the fraction of actives in the database. In the formula below, *A* is the number of actives in the database and *N* is the total number of molecules in the database.

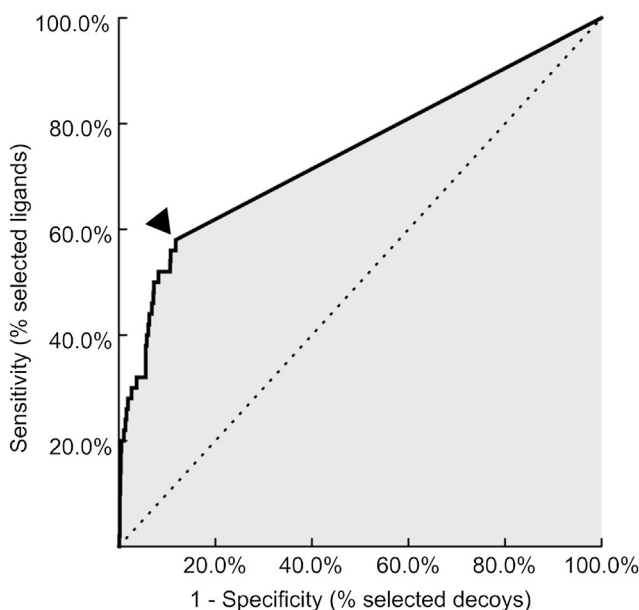
$$EF = \frac{Ya}{(A/N)}$$

Convenient and meaningful tools for the assessment of screening performance are Receiver Operating Characteristic (ROC) curves [72]. A ROC curve displays the sensitivity on the y-axis versus (1-Specificity) in the x-axis and thus gives a good impression of the quality of a model. An ideal curve would rise vertically along the y-axis until it reaches the maximum true positive rate (1), and then continue horizontally to the right, which means that the hit list contains all active compounds in the database and that none of the hits is a false positive. The diagonal line between the lower left and the upper right corner of the graph would represent the ROC curve of a random database search. Figure 21.4 shows an example for a ROC curve.

### III. ADVANCED APPROACHES

#### A. Structure-Based Pharmacophores

If the 3D structure of a protein–ligand complex is known from either X-ray crystallography or protein NMR, the most obvious way of deriving a picture of the relevant ligand interactions is to analyze the molecule’s complementarity within the corresponding protein binding site. A commonly used structure-based design approach is the previously mentioned molecular docking of ligands into a target binding pocket, assuming that the binding site is more or less rigid whereas the ligand is flexible. Molecular docking is still the most popular method for structure-based drug design. However, pharmacophore-based approaches have shown clear advantages regarding the computational demand and accuracy for virtual screening [61]. Especially with regard to the number of false positives—which are often observed in classical docking-based virtual screening—the idea of combining structural information derived from a protein–ligand complex and the use of a rapid pharmacophore-based screening technique is obvious. There is no competition between ligand-based and structure-based pharmacophore modeling, and both approaches can be used fruitfully in a complementary manner [73]. In fact, a variety of pharmacophore



**FIGURE 21.4** Example ROC plot for a pharmacophore screening. In this virtual screening run, approximately 60 percent of the biologically active molecules were identified by the model (i.e., part of the virtual hit list), while approximately 85 percent of the inactive compounds could be excluded (i.e., were not part of the virtual hit list). The black arrow illustrates the end of the virtual hit list.

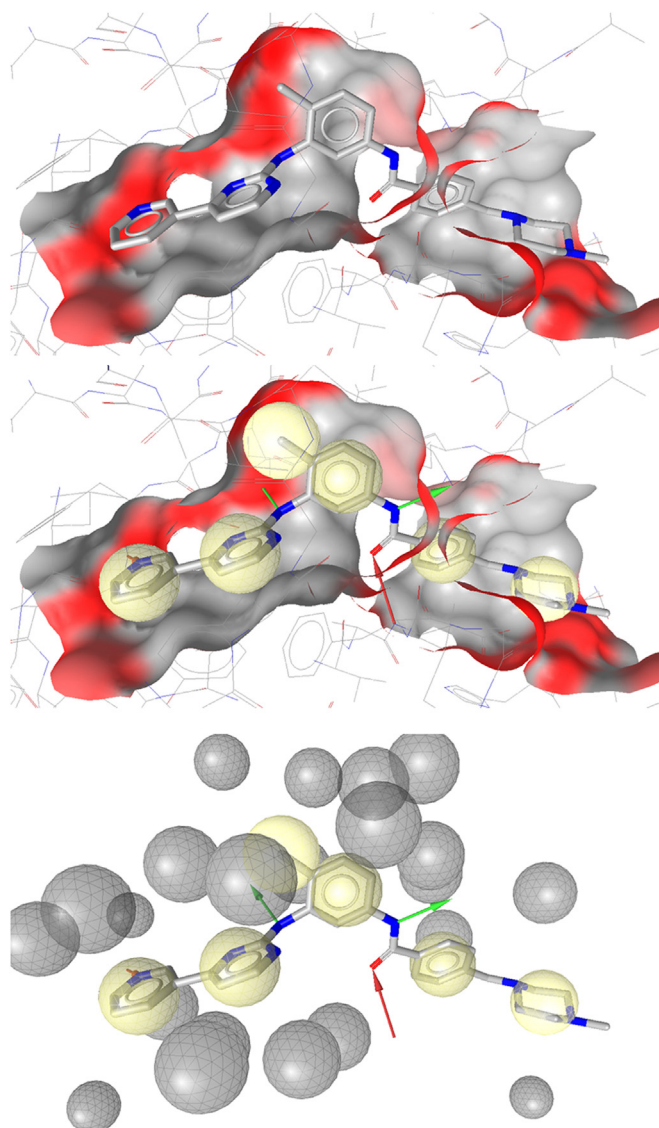
modeling programs allow one to take advantage of additional information provided by a protein or protein–ligand complex structure to help improve the reliability of the generated model. The development of several novel programs for deriving structure-based pharmacophores in the last few years has clearly shown that pharmacophore-based virtual screening is very successful in identifying novel bioactive molecules [74–78]. On the other side, it was also recognized that the consideration of pharmacophores in docking programs can increase reliability and accuracy. Several docking programs are now available which apply the pharmacophore concept to discriminate better between false and real binding modes (e.g., Glide [21], FlexX-Pharm [22], GOLD [20]).

As an illustration, the generation of a structure-based pharmacophore and its application for virtual screening of ABL tyrosine kinase inhibitors is given. STI-571 (Gleevec<sup>®</sup>) has been approved for the treatment of chronic myelogenous leukemia (CML) and was the first antitumor drug from the family of tyrosine kinase inhibitors [79]. Several crystal structures of STI-571 in complex with different tyrosine kinases (ABL, c-KIT, SYK) have been obtained in the last few years, showing that the compound can bind in varying conformations (open and closed conformation) to different forms of tyrosine kinases. In the case of ABL tyrosine kinase, STI-571 binds to the inactive enzyme form and prevents activation [80]. Several pharmacophores from the available X-ray structures of ABL in complex with STI-571 and analogues (PDB IDs 1IEP, 1FPU and 1OPJ) were generated [81]. In a straightforward approach, the different pharmacophore models were merged using the program LigandScout. The merged pharmacophore contained four lipophilic aromatic areas, two acceptor features, and eight excluded volume spheres. As an example, the structure-based pharmacophore extracted from the X-ray structure 1iep is shown in Figure 21.5. Subsequently a virtual screening was carried out using two different ligand databases. The first one was a collection of 2,765 drug-like ligands from the complexes in the Protein Databank (PDB): the second one was the Maybridge compound library (containing ~59,000 molecules). The pharmacophore model was able to identify all STI-571 entries from the PDB database and did not result in false positives. In addition, seven compounds from the Maybridge database were identified that might represent potential lead structures for the development of novel ABL tyrosine kinase inhibitors.

Several successful applications of the LigandScout program have been reported recently and have supported the feasibility of structure-based pharmacophores to identify novel active molecules [74–77].

## B. Pseudo-Receptor Models

Starting in the 1980s, a combination of pharmacophore modeling and structure-based design was introduced and referred to as “pseudo-receptor modeling” or “receptor mapping.” [82] Based on a preliminary pharmacophore



**FIGURE 21.5** 3D structure of Gleevec<sup>®</sup> (capped sticks) bound to ABL kinase (top). The structure-based pharmacophore generated with the program LigandScout is shown in the middle, and the extracted pharmacophore together with the excluded volumes (grey spheres) is shown at the bottom. Yellow spheres = hydrophobic features, green arrows = hydrogen bond donors, and red arrows = hydrogen bond acceptors.

model, a hypothetical receptor consisting of individual amino acid residues is constructed, surrounding a set of superimposed ligands. The placement of the individual amino acid residues is guided by experimental data (e.g., from site-directed mutagenesis data). Höltje successfully applied the receptor-mapping technique to several target proteins for which no 3D structure was available [83–85]. Using a data set of twenty 5-HT<sub>2A</sub> receptor antagonists from different chemical families, a pharmacophore was generated that was able to explain the SAR of the ligand [86]. The receptor mapping (i.e., the placement of the individual amino acids) was based on a homology model of the 5-HT<sub>2A</sub> receptor generated on the basis of the low-resolution 3D structure of bacteriorhodopsin (a related membrane protein). Using the derived pseudo-receptor, a predictive QSAR model could be obtained that was subsequently applied to design novel potent antagonists [86].

### 1. *Yak, PrGen, Flo*

Whereas the first pseudo-receptor models were generated more or less intuitively “by hand,” which sometimes resulted in irreproducible results, a broader distribution of this concept was achieved by the commercial software packages *Yak* and *PrGen* [15]. Both programs allow the generation of a pseudo-receptor in a more or

less automated way. In addition, guided by extensive correlation of experimental and model-derived free energies of binding, a host–guest system is created, mimicking the interaction at a real binding site reasonably well. The fundamental basis of a pseudo-receptor is the placement of the individual amino acid residues. In Yak and PrGen, ligand-specific interaction vectors (the pharmacophoric elements) are calculated and saturated with individual residues from a database of pre-calculated conformations of amino acids. Subsequently, a receptor minimization is carried out by relaxing all residues, keeping the position, orientation, and conformation of the ligands unchanged. To achieve a correlation between the experimentally derived binding affinities (or other biological data) and the calculated interaction energies, a coupling constant is introduced and the system is minimized (correlation-coupled minimization). In a next step, the ligand alignment is allowed to relax within the fixed pseudo-receptor (ligand relaxation). This process (i.e., correlation-coupled minimization followed by unconstrained ligand relaxation) is repeated several times until a highly correlated pseudo-receptor is obtained. To validate the generated pseudo-receptor, its ability to predict the binding affinities of novel ligands must be examined [85]. Therefore, classical QSAR methods such as cross-validation via leave-one-out and/or prediction of external test set compounds are applied. In case of test set or novel ligands, the molecules have to be placed equally to the training set molecules in the pseudo-receptor and have to be minimized applying the same protocol as for the training set ligands.

Another pseudo-receptor modeling approach has been developed by Bohacec et al [87]. Their program, Flo, generates an ensemble of low-energy conformers of each compound of a training set. The conformations are then optimized to minimize the internal energy and maximize the match of chemically similar moieties simultaneously. Then, a pseudo-receptor is composed of functional groups that will mimic the binding cavity. For example, a guanidinium group is selected to form hydrogen bonds with an acidic group of the ligands. The selected residues are positioned around the aligned training set ligands and anchored to the chemically complementary ligand atoms, applying a distance constraint. The remaining volume of the pseudo-receptor is filled with propane molecules to mimic a binding site's hydrophobic surface. In the last step, the pseudo-receptor is equilibrated—comparable to the PrGen approach—by applying several rounds of dynamics. While a pseudo-binding site is quite artificial, the method has the advantage of allowing the binding site to be visualized and used for ligand docking and structure-based design.

## 2. Quasar and Raptor

A further development of Vedani et al was the simplification of the atomistic pseudo-receptor concept (Yak and PrGen) to a quasi-atomistic receptor approach (named Quasar) [88]. Similar to the approach of Walters et al, who developed the program GERM [89], Quasar uses a 3D binding-site surrogate surrounding the ligands instead of a shell of amino acid residues. Each of the virtual particles bears relevant atomistic properties (e.g., H-bond donor, hydrophobic particle). Quasar not only takes into account one conformer per ligand but represents each ligand by an ensemble of low-energy conformations (called “fourth dimension”), thereby reducing the bias associated with the selection of a putative bioactive conformation. Binding of ligand molecules to a macromolecular binding pocket is often facilitated by an induced fit (i.e., the adaptation of a protein to the ligand topology). This effect, which is not considered in most of the pharmacophore and 3D-QSAR approaches, is considered by Quasar and Raptor [90] (the so-called “fifth dimension”). Quantitative models generated with these programs have therefore been named 4D- or 5D-QSARs [91].

## 3. Application of Pseudo-Receptor Models

The pseudo-receptor concept has been applied in recent years to analyze crucial ligand–receptor interaction sites and to establish 3D-QSARs for the prediction of biological activities of ligands [92]. A variety of application studies have shown that the pseudo-receptor concept is a versatile tool in establishing 3D-QSAR models that are often better in their predictive behavior compared to results obtained from classical 3D-QSAR approaches (e.g., CoMFA) [93]. Several application studies have been published that have shown the value—and the limitations—of this approach [82,94].

In a recently published study by Bohacec et al, the pseudo-receptor concept was successfully applied to identify novel small-molecule inducers of fetal hemoglobin [95]. Four available active compounds (Figure 21.6) were selected based on activity and diversity for the construction of an initial pharmacophore.

The initial pharmacophore was constructed using the Flo molecular modeling software [87]. The derived pharmacophore was then successfully tested on a larger ligand data set to see if it could distinguish between active and inactive compounds. Satisfied with the preliminary evaluation of the pharmacophore template, the authors used the model to design novel compounds. The model was sufficiently well defined to allow docking of 630

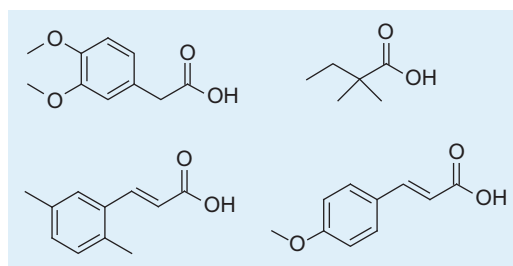


FIGURE 21.6 Molecules used to generate the pharmacophore for inducers of fetal hemoglobin.

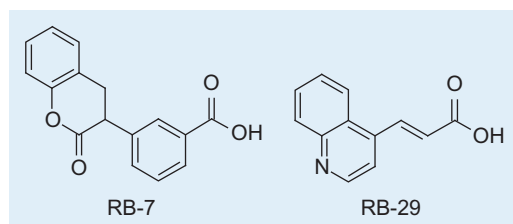


FIGURE 21.7 Molecular structures of the two most active inducers identified by the pseudo-receptor modeling.

compounds and the selection of thirty compounds for testing. Of the twenty-six compounds acquired and tested, four displayed significantly greater activity than previously identified ligands, showing the feasibility of using pseudo-receptor and docking to identify novel bioactive molecules. The structures of the two most potent molecules are shown in Figure 21.7.

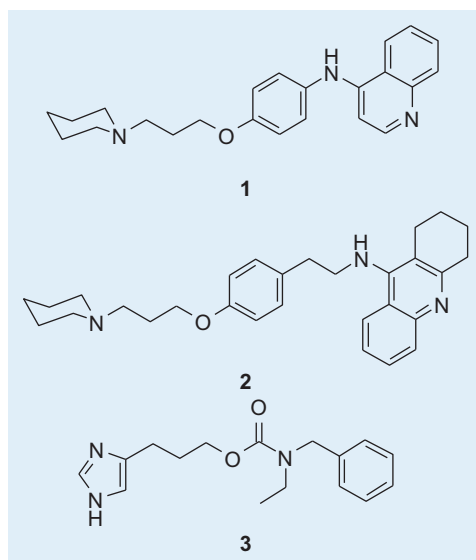
When working with pseudo-receptors and in general with quantitative-structure–activity relationships (QSAR) of any dimension, a word of caution is necessary with respect to the biological data that are used. These should preferably constitute binding affinities from a single laboratory, a prerequisite that is also true for all QSAR studies. Since the receptor models simulate interaction events ( $\Delta H$ ) in a highly simplified manner, the experimental data that are combined with them in a correlation analysis must be as close to the molecular level as possible. It is therefore nonsense to correlate the calculated interaction energies with biological *in vivo* data, because the receptor interaction can be blurred or even completely hidden by transport and other pharmacokinetic processes. Sometimes, even the use of *in vitro* data is dangerous if a reaction cascade separates the measured event from the receptor binding interaction. Also, the combination of biological data (e.g.,  $IC_{50}$  values) from different laboratories or assays is extremely dangerous. The reliability and meaning of any QSAR model (3D-QSAR, pseudo-receptor, 4D-QSAR, 5D-QSAR, 6D-QSAR [96]) should always be assessed by the ultimate test of usefulness, the prediction of new compounds [97]. Very often, QSAR models are internally validated but never tested on whether they are useful in designing novel, more potent compounds [98].

Recently, a novel pseudo-receptor modeling method has been developed named Surfex-QMOD [99,100]. Two datasets, CDK2 inhibitors and muscarinic antagonists, have been chosen to test the performance of the method. Interestingly, the used dataset of congeneric CDK2 inhibitors showed that induced binding pockets can be quite congruent with the enzyme's active site but that model predictivity within a chemical series does not necessarily depend on congruence.

## IV. APPLICATION STUDY: NOVEL HISTAMINE H<sub>3</sub>-RECEPTOR ANTAGONISTS

### A. Pharmacophore-Based Screening

An example from the author's laboratory shall give the reader an informative picture of the pharmacophore generation process and its application to develop novel bioactive compounds [63]. The example deals with antagonists of the human histamine H<sub>3</sub> receptor (hH<sub>3</sub>R). hH<sub>3</sub>R is a GPCR for which no exact 3D structures is available, as is the case for many other GPCRs. Although we have a crystal structure for the H<sub>1</sub> receptor in the meantime (PDB ID 3RZE), it is still challenging to derive information about the right physiological receptor



**FIGURE 21.8** Molecular structures of h<sub>3</sub>R antagonists used for the pharmacophore development.

conformation by structure-based modeling techniques. The H<sub>3</sub> receptor 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 [101]. Although ligands addressing this receptor are already known, the discovery of alternative lead structures represents a challenging goal in drug design [102]. Experimental structure–activity data for the h<sub>3</sub>R antagonists can be summarized as follows. The pharmacological results suggest that a protonatable nitrogen atom (either in an aromatic imidazole or in a saturated ring system) and an aromatic system separated by a certain distance seem to constitute a potent h<sub>3</sub>R antagonist. Additional polar moieties in the spacer can enhance the antagonistic activity (Figure 21.8).

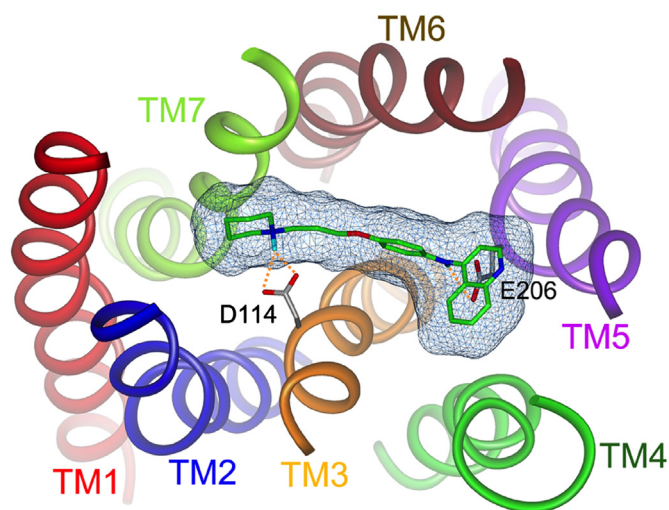
## B. Pharmacophore Determination Process

Due to their high flexibility and huge structural diversity, h<sub>3</sub>R antagonists also provide difficulties in the generation of pharmacophore models by standard means, which normally include the identification of common features required for binding from a ligand set. A dataset of 418 ligands for which h<sub>3</sub>R binding affinities were determined in a [<sup>3</sup>H]N<sup>α</sup>-methylhistamine assay is available (pK<sub>i</sub> from 5 to 10). A pharmacophore able to discriminate between active and inactive antagonists should be developed on the basis of the known antagonists and be used virtually to screen compound libraries for novel structurally diverse h<sub>3</sub>R antagonists.

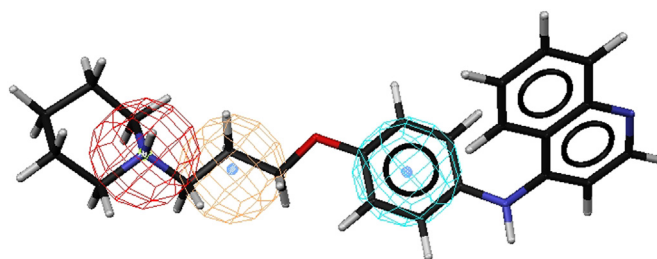
For the available ligand dataset, a multi-conformer database was generated using the Catalyst software. An energy cut-off of 20 kcal mol<sup>-1</sup> from each energetic minimum structure was set in order to avoid high-energy structures. In a first step, three individual pharmacophore models were generated based on the potent antagonists **1**, **2**, and **3**. The bioactive conformation of the ligands was deduced from a conformational analysis of semi-rigid h<sub>3</sub>R antagonists and an extensive docking study carried out on a homology model of the h<sub>3</sub>R (Figure 21.9; for details see [103]). The docking study showed that the homology model is able to explain the interaction of the ligands, which is in accordance with known biochemical data (e.g., site-directed mutagenesis data). However, a receptor-based virtual screening was not very successful in discriminating active from inactive antagonists. Therefore, the idea was to carry out a pharmacophore-based virtual screening.

Defining a pharmacophore model upon a ligand has the advantage that the individual features are already correctly aligned in space. In order to account for the great structural variability of h<sub>3</sub>R antagonists, the pharmacophores were defined as loosely as possible in order to retrieve most of the validated h<sub>3</sub>R ligands as hits.

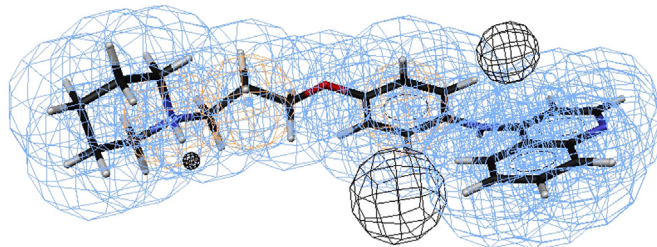
Once a pharmacophore capable of retrieving known h<sub>3</sub>R antagonists had been defined, it can be used in subsequent screening procedures of commercial compound libraries. As an example, the pharmacophore generated on the basis of compound **1** is shown in Figure 21.10.



**FIGURE 21.9** Interaction of compound **1** with the hH<sub>3</sub>R binding site as obtained from the docking study. Only the two important acidic amino acid residues of the binding site are shown for clarity. H-bonds are shown as orange dotted lines.



**FIGURE 21.10** Pharmacophoric features defined based upon compound **1**. Red sphere = any positively charged element; orange sphere = aromatic or hydrophobic group; cyan sphere = aromatic ring.



**FIGURE 21.11** Pharmacophore model based on compound **1** including the shape feature (van der Waals volume) and the two forbidden volumes (black spheres).

The choice of chemical features was based on functionalities observed in validated hH<sub>3</sub>R antagonists and inspection of the binding pocket of the homology model. The linker moiety and the adjacent hydrophobic/ $\pi$ -electron-rich system of the ligands lie in a cleft between trans-membrane region (TM) 3, 6, and 7 of the hH<sub>3</sub>R. In this region, several aromatic residues border the binding site that are able to interact with the electron-rich system in the hH<sub>3</sub>R antagonists. No pharmacophoric features were defined upon the 4-aminoquinoline moiety, as a high degree of chemical diversity is observed in active ligands within this region. Any restriction of chemical features was thus avoided.

Apparently, the derived pharmacophore model is too loose fitting for screening a compound database. Thus, the van der Waals volume of ligand **1** was included as an additional constraint into the pharmacophore model. Default parameters were used for the definition of the shape query. Finally, forbidden volumes (black spheres) were defined in order to account for the fact that some ligands extending into these areas were inactive, although they resembled other active compounds. [Figure 21.11](#) shows the ligand **1** fitted into the complete pharmacophore model.

Using this model, 316 compounds from the 418 ligand dataset were found as hits in a pharmacophore search using the Catalyst [36] program. Ninety-three percent of the ligands with highest activity were retrieved by the pharmacophore model; less satisfactorily, 54 percent of a set of inactive compounds could also pass the pharmacophoric filter (Figure 21.11, top). Application of the pharmacophore filter for screening the Maybridge Database (MDB) and the World Drug Index (WDI) resulted in 249 and 929 hits, respectively. Thus, 70 percent of the active and moderate active hH<sub>3</sub>R ligands (with a pK<sub>i</sub> > 7) were retrieved by the pharmacophore. Meanwhile, from the pool of MDB and WDI ligands (MDB: 59,000 compounds; WDI: 48,000 compounds), 98.9 percent could be excluded.

The filter was still quite loose, however, so that a subsequent definition of further pharmacophoric features could result in a better separation of in/actives. In order to further increase the percentage of active hH<sub>3</sub>R ligands found during the virtual screening, further pharmacophore models were defined in a similar way based upon compounds 2 (K<sub>i</sub> = 0.33 nM) and 3 (K<sub>i</sub> = 69 nM). For the definition of the pharmacophore derived from compound 2, the three features described above were again used in combination with a shape query and forbidden volumes. The third individual pharmacophore model was defined based on ligand 3 (see Figure 21.12), capable of retrieving 68 percent of ligands deposited in the hH<sub>3</sub>R database. By combining the three pharmacophore models, 369 of 398 (93 percent) hH<sub>3</sub>R ligands with a pK<sub>i</sub> > 7 could be obtained, while only 2668 (2.5 percent) compounds were obtained as hits when screening the MDB and WDI database with 107,599 total structures deposited. The small percentage of structures from commercial databases matching the pharmacophores showed that the generated models were stringent enough for a reasonable screening.

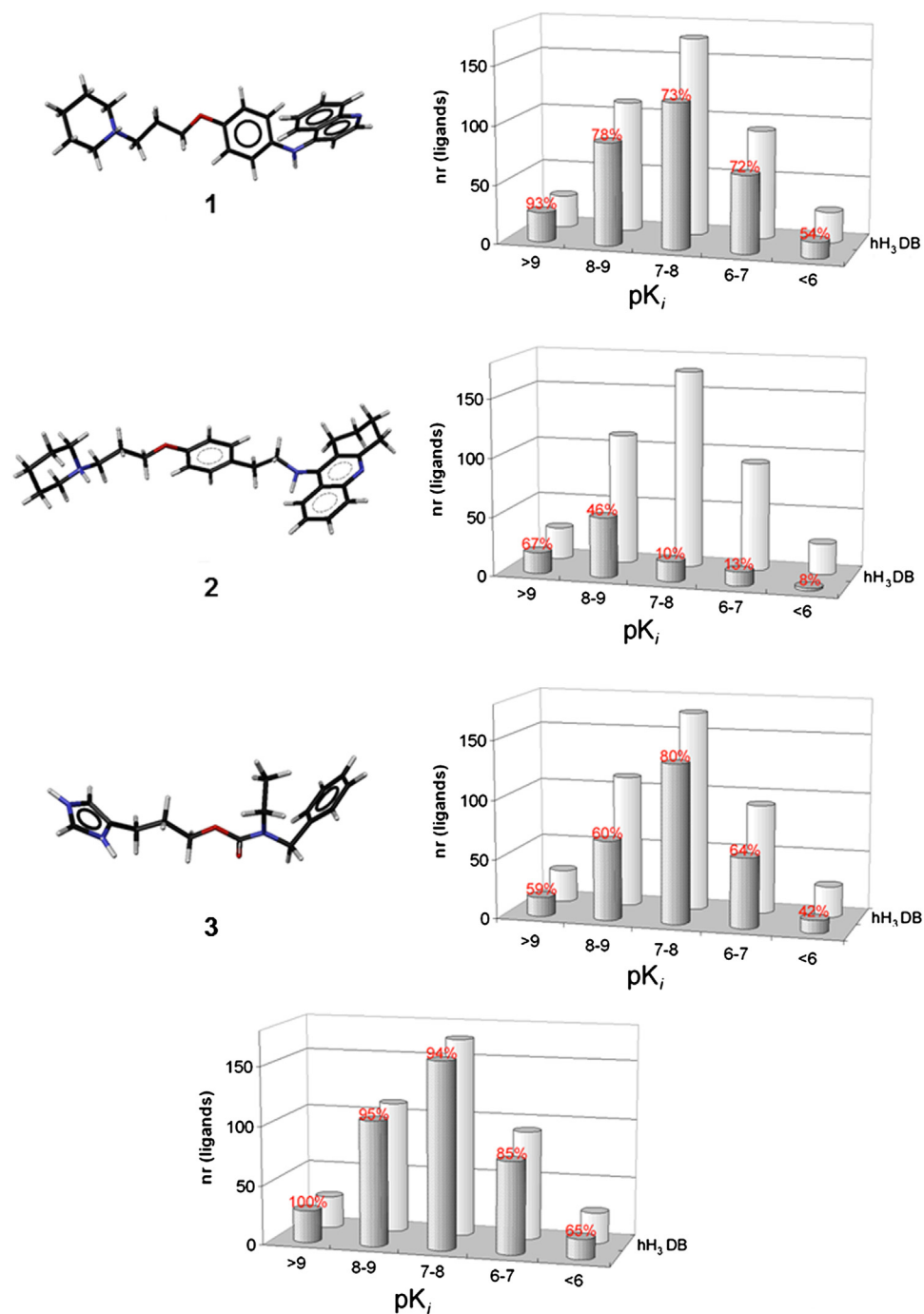
### C. Pharmacophore-Based Screening of Compound Libraries

For a more stringent screening, a leave-one-out (LOO) filter was defined on the pharmacophoric features of 1. The Catalyst LOO model consisted of a combination of five individual pharmacophore models, each lacking one pharmacophoric feature found in compound 1 at a time, with the exception of the positive ionizable group and the spacer moiety that were required in all models. The screening of the 2,668 WDI and MDB compounds with the LOO filter reduced the number of hits to 320. In order to ensure that compounds selected by the pharmacophore-based screening could be accommodated into the hH<sub>3</sub>R binding site, the 320 hits were docked into the hH<sub>3</sub>R binding site using the GOLD [20] program and ranked according to their docking scores. From the top-ranked complexes, seven MDB compounds were selected for experimental testing. The selection of the seven compounds was guided by a cluster analysis in order to select the most structurally diverse compounds among the top-ranked molecules. All compounds showed affinity for the hH<sub>3</sub>R with binding affinities ranging from 79 nM to 6.3 μM, thereby showing that the pharmacophore and hH<sub>3</sub>R binding site model can be used to identify novel active antagonists. Two compounds, BTB-08079 and RJC-03033, were found to be active in the nanomolar range (Figure 21.13) [63].

In order to determine the structural similarity between the seven retrieved MDB compounds and the 418 hH<sub>3</sub>R ligands, we calculated similarity indices on the basis of different fingerprint systems (MACCS keys and graph-3-point pharmacophore fingerprints in MOE [41]). Using the different fingerprint systems, low similarities were observed between the seven MDB compounds and the original hH<sub>3</sub>R antagonists. For the most potent hit (BTB-08079; 79 nM), the lowest similarity to the original hH<sub>3</sub>R ligand structures was observed. The dimethylaminofuran fragment, which was already known from the potent histamine H<sub>2</sub> receptor antagonist Ranitidine, was not reported before as a structural element of potent H<sub>3</sub>R antagonists.

Compared to the receptor-based virtual screening, application of the pharmacophore-based search resulted in significantly improved results. In the docking approach, 66.6 percent of the hH<sub>3</sub>R ligands were retrieved, limiting the number of WDI and MDB compounds to approximately 1,720 structures, but application of a pharmacophore-based search allowed retrieval of 93 percent of active compounds, while reducing the number of WDI and MDB structures to 2,668 compounds (2.5 percent). The ideal strategy for the flexible hH<sub>3</sub>R ligand data set, however, appeared to be 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. In order to ensure that compounds selected by the pharmacophore-based screening fit into the binding site, docking of this subset of ligands resulted in a selection of candidates for biological testing.





**FIGURE 21.12** Enrichment of hH<sub>3</sub>R ligands by pharmacophore search based on compound 1, 2, and 3. The percentage of hH<sub>3</sub>R ligands retrieved by the individual pharmacophore model within each pK<sub>i</sub>-cluster is depicted. The percentage of ligands found in each cluster (dark columns) is written in red numbers and compared to the population of pK<sub>i</sub> clusters of all hH<sub>3</sub>R compounds in the 418 ligand data set (light grey columns).

## V. RECENT DEVELOPMENTS AND OUTLOOK

3D pharmacophores have evolved as important tools for describing protein–ligand interactions, and the number of examples that successfully predict biological activity using pharmacophore methods is constantly growing. A multitude of therapeutic areas is covered, such as anti-viral drug discovery [104,105], the discovery of novel anti-bacterial agents [106,107], and modeling of GPCR ligand interaction (Figure 21.14) [108,109], among others.

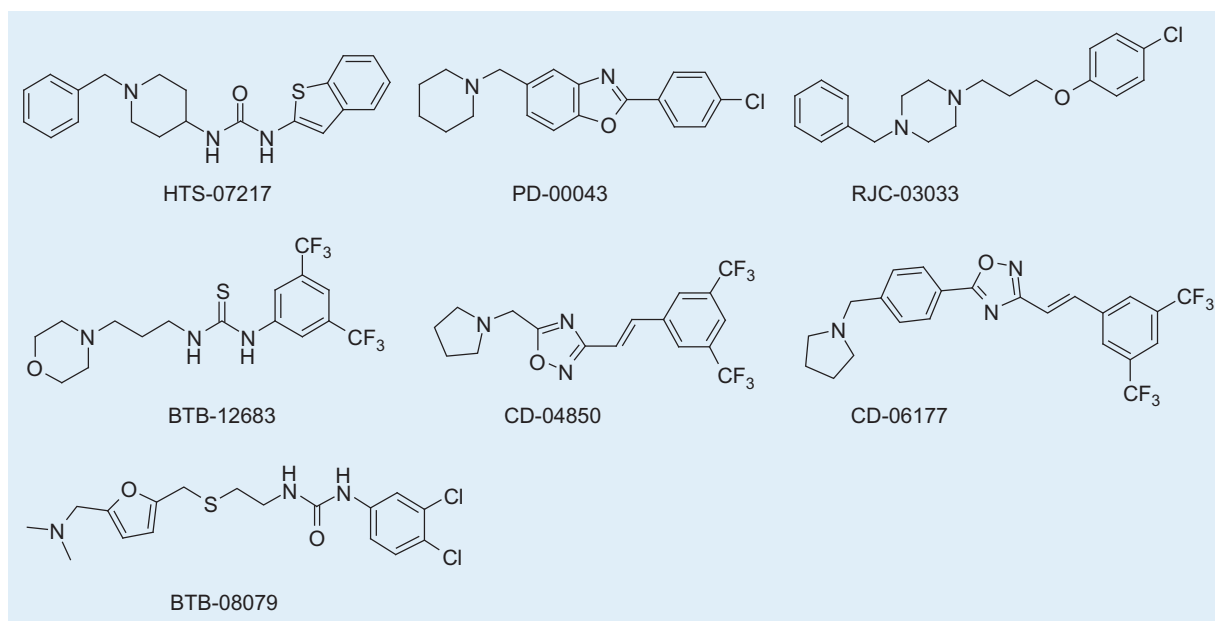


FIGURE 21.13 Selected hits from the pharmacophore-based virtual screening.

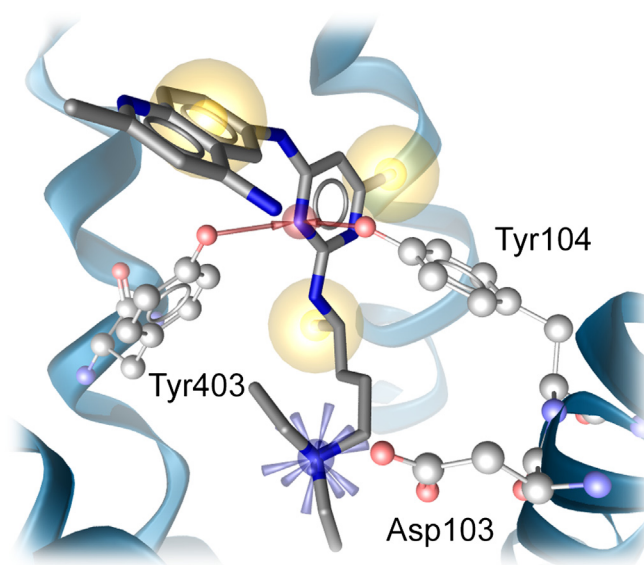


FIGURE 21.14 LigandScout 3D pharmacophore model example for the newly discovered muscarinic acetylcholine receptor antagonist NSC23766. Red arrows = H-bond donors; yellow spheres = lipophilic contact areas; blue rays = positively charged groups.

Technically, all these approaches follow the sequence of careful model design, prediction, and selection of existing compounds from libraries with previously unknown biological activity or *de novo* design to fulfil pharmacophoric requirements. The crucial step remains model design and careful retrospective validation before performing a prospective virtual screening, which is followed by purchasing compounds and biological tests. If model building is done carefully, it can lead to successful predictions and rationalization of a possible ligand-binding mode. Mainly due to the better usability of recent programs, 3D pharmacophores are now also used aside of the typical workflow of virtual screening against a single protein-binding site. The program Ligandscout, for example, now includes a module for fragment-based screening that aims at supporting fragment-based *de novo* design [110].

Another important new application is parallel screening against several pharmacophores to predict multi-target effects [111,112]. Although this can also be done using various other molecular modeling techniques, such as classical QSAR or descriptor approaches [113], multi-target predictions using 3D pharmacophores have the advantage that results remain easy to interpret and provide intuitive starting points for further optimization.

As previously mentioned pharmacophore screening comparisons [43] show, algorithm development for compound/pharmacophore mapping remains a challenge. This leaves room for further algorithmic improvements, such as the implementation of advanced filter methods like bloom filtering as implemented in recent tools like Pharmer [114] or LigandScout [44].

Another recent approach developed by Rognan's group uses 3D fingerprints based on pharmacophoric features to compare protein-binding sites in terms of ligand binding properties (including druggability). In a comparable fashion, pharmacophore fingerprints can be used to describe subpocket similarity [115].

## VI. CONCLUSIONS

In spite of the recent success and popularity of pharmacophore-based drug design, one should not forget the limitations of pharmacophore modeling. As with any other model, we should be aware of the abstraction that is applied to generate these models. All pharmacophore approaches are based on molecular mechanical abstractions. Thus, properties associated with the interaction of electrons (e.g., polarization effects) are not considered. Another limitation in many pharmacophore-based approaches is the neglect of the dynamic nature of protein–ligand interaction. Although novel pharmacophore generation programs allow the parallel consideration of multiple/alternative pharmacophores (e.g., Catalyst [38], LigandScout [39]), modeling different binding modes is still a challenge. It is becoming increasingly clear that for some protein binding sites, one has to be prepared to consider different binding modes and therefore different pharmacophores [116–118].

Whereas in the past, pharmacophore models have been mainly generated using ligand-based strategies, novel programs have been developed and applied successfully in the last few years by combining structure-based and pharmacophore-based approaches. This is mainly influenced by the rapidly growing number of protein–ligand 3D structures that are the basis for such combined approaches. Closely related to this, one can observe a general merging of different techniques in molecular modeling studies—pharmacophore modeling, 3D-QSAR, *de novo* design, and docking [119,120]—that might be helpful for future drug design studies.

## References

- [1] Wermuth CG, Ganellin CR, Lindberg P, Mitschler LA. Glossary of terms used in medicinal chemistry (IUPAC recommendations 1997). *Annu Rep Med Chem* 1998;33:385–95.
- [2] Langley JN. On the reaction of cells and nerve-endings to certain poisons, chiefly as regards the reaction of striated muscle to nicotine and to curari. *J Physiol* 1905;33:374–413.
- [3] Ehrlich P, Morgenroth J. Über Haemolysine. Dritte Mitteilung [About haemolysins. Third report]. *Berl Klin Wochenschr* 1900;37:453–7. [in German].
- [4] Fischer E. Einfluss der Konfiguration auf die Wirkung der Enzyme [Influence of configuration on the action of enzymes]. *Ber Dtsch Chem Ges* 1894;27:2985–93. [in German].
- [5] Perutz MF. X-ray analysis of haemoglobin. *Science* 1963;140:863–9.
- [6] Marshall GR. Binding-site modelling of unknown receptors. In: Kubinyi H, editor. *3D QSAR in drug design: theory methods and applications*. Leiden: ESCOM; 1993. p. 80–116.
- [7] Barry CD, Ellis RA, Graesser S, Marshall GR. Display and manipulation in three dimensions. In: Faiman M, Nievergelt J, editors. *Pertinent concepts in computer graphics*. Chicago: University of Illinois Press; 1969. p. 104–53.
- [8] Kier LB, Aldrich HS. A theoretical study of receptor site models for trimethylammonium group interactions. *J Theor Biol* 1974;46:529–41.
- [9] Gund P, Wipke WT, Langridge R. Computer searching of a molecular structure file for pharmacophoric patterns. *Comput Chem Res Educ Technol* 1974;3:5–21.
- [10] Poptodorov K, Liu T, Hoffmann RD. Pharmacophore model generation software tools. In: Langer T, Hoffmann RD, editors. *Pharmacophores and pharmacophore searches*. Weinheim, Germany: Wiley-VCH; 2006. p. 17–47.
- [11] Sanders MP, Barbosa AJ, Zarzycka B, Nicolaes GA, Klomp JP, de Vlieg J, et al. Comparative analysis of pharmacophore screening tools. *J Chem Inf Model* 2012;52:1607–20.
- [12] Evans BE, Rittle KE, Bock MG, DiPardo RM, Freidinger RM, Whitter WL, et al. Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J Med Chem* 1988;31:2235–46.
- [13] Wermuth CG. Pharmacophores: historical perspective and viewpoint from a medicinal chemist. In: Langer T, Hoffmann RD, editors. *Pharmacophores and pharmacophore searches*. Weinheim, Germany: Wiley-VCH; 2006. p. 1–13.
- [14] Snyder JP, Rao SN, Koehler A, Vedani A. Minireceptors and pseudoreceptors. In: Kubinyi H, editor. *3D QSAR in drug design: theory methods and applications*. Leiden: ESCOM; 1993. p. 336–54.

- [15] Vedani A, Zbinden P, Snyder JP. Pseudo-receptor modelling: a new concept for the three-dimensional construction of receptor binding sites. *J Recept Res* 1993;13:163–77.
- [16] Höltje HD, Anzali A. Molecular modelling studies on the digitalis binding site of the Na<sup>+</sup>/K<sup>+</sup>-ATPase. *Pharmazie* 1992;47:691–8.
- [17] Frühbeis H, Klein R, Wallmeier H. Computergestütztes Moleküldesign (CAMD) – ein Überblick [Computer aided molecular design (CAMD): An overview]. *Angew Chem Int Ed Engl* 1987;26:403–18. [in German].
- [18] Böhm HJ, Klebe G, Kubinyi H. Wirkstoffdesign: Der Weg zum Arzneimittel [Drug design: The path to drugs]. Heidelberg, Germany: Spektrum Akademischer Verlag; 1996. [in German].
- [19] Osterberg F, Morris GM, Sanner MF, Olson AJ, Goodsell DS. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins* 2002;46:34–40.
- [20] Gold, Cambridge Crystallographic Data Centre, Cambridge, UK. Available from: <<http://www.ccdc.cam.ac.uk/>>.
- [21] Glide, Schrödinger, Portland, OR. Available from: <<http://www.schroedinger.com/>>.
- [22] FlexE, BiosolveIT, St. Augustin, Germany. Available from: <<http://www.biolsolveit.de/>>.
- [23] Klebe G. Virtual ligand screening. *Drug Discov Today* 2006;11:580–94.
- [24] Tame JR. Scoring functions: the first 100 years. *J Comput Aided Mol Des* 2005;19:445–51.
- [25] Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, et al. A critical assessment of docking programs and scoring functions. *J Med Chem* 2005;49:5912–31.
- [26] Moitessier N, Englebienne P, Lee D, Lawandi J, Corbeil CR. Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *Br J Pharmacol* 2008;153:S7–26.
- [27] Yuriev E, Ramsland PA. Latest developments in molecular docking: 2010–2011 in review. *J Mol Recognit* 2013;26:215–39.
- [28] Sousa SF, Ribeiro AJ, Coimbra JT, Neves RP, Martins SA, Moorthy NS, et al. Protein-ligand docking in the new millennium: a retrospective of 10 years in the field. *Curr Med Chem* 2013;20:2296–314.
- [29] Venkatakrisnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular signatures of G-protein-coupled receptors. *Nature* 2013;494:185–94.
- [30] Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, et al. Structure of the human glucagon class B G-protein-coupled receptor. *Nature* 2013;499:444–9.
- [31] Hollenstein K, Kean J, Bortolato A, Cheng RKY, Doré AS, Jazayeri A, et al. Structure of class B GPCR corticotropinreleasing factor receptor 1. *Nature* 2013;499:438–43.
- [32] Hackling A, Ghosh R, Perachon S, Mann A, Höltje H-D, Wermuth CG, et al. N-(4-(2-Methoxyphenyl)piperazin-1-yl)alkylcarboxamides as dopamine D<sub>2</sub>- and D<sub>3</sub>-receptor ligands. *J Med Chem* 2003;46:3883–99.
- [33] Cramer III RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc* 1988;110:5959–67.
- [34] Ahlstrom MM, Ridderstrom M, Luthman K, Zamora I. Virtual screening and scaffold hopping based on GRID molecular interaction fields. *J Chem Inf Model* 2005;45:1313–23.
- [35] Steindl TM, Schuster D, Wolber G, Laggner C, Langer T. High-throughput structure-based pharmacophore modelling as a basis for successful parallel virtual screening. *J Comput Aided Mol Des* 2006;20:703–15.
- [36] Catalyst, Accelrys Software, San Diego, CA. Available from: <<http://www.accelrys.com/>>.
- [37] DISCOtech, St. Louis, MO. Available from: <<http://www.tripos.com/>>.
- [38] Galahad, Tripos, St. Louis, MO. Available from: <<http://www.tripos.com/>>.
- [39] LigandScout, Inteligand, Wien, Austria. Available from: <<http://www.inteligand.com/>>.
- [40] Phase, Schrödinger, Portland, OR. Available from: <<http://www.schroedinger.com/>>.
- [41] MOE, Chemical Computing Group, Montreal, Canada. Available from: <<http://www.chemcomp.com/>>.
- [42] Patel Y, Gillet VJ, Bravi G, Leach A. A comparison of the pharmacophore identification programs: catalyst, DISCO, and GASP. *J Comput Aided Mol Des* 2002;16:653–81.
- [43] Spitzer G, Heiss M, Mangold M, Kirchmair J, Wolber G, Liedl K. One concept, three implementations of 3D pharmacophore-based virtual screening: distinct coverage of chemical search space. *J Chem Inf Model* 2010;50:1241–7.
- [44] Seidel T, Ibis G, Bendix F, Wolber G. Strategies for 3D pharmacophore-based virtual screening. *Drug Discov Today Technol* 2010;7:e221–8.
- [45] Boström J, Norrby P-O, Liljefors T. Conformational energy penalties of protein-bound ligands. *J Comput Aided Mol Design* 1998;12:383–96.
- [46] Agrafiotis DK, Gibbs AC, Zhu F, Izrailev S, Martin E. Conformational sampling of bioactive molecules: a comparative study. *J Chem Inf Model* 2007;47:1067–86.
- [47] Höltje H-D, Sippl W, Rognan D, Folkers G. Molecular modelling: basic principles and applications. 3rd ed. Weinheim, Germany: Wiley-VCH; 2007.
- [48] Marshall GR, Barry CD, Bosshard HE, Dammkoehler RA, Dunn DA. The conformational parameter in drug design: the active analog approach. In: Olsen EC, Christoffersen RE, editors. *Computer-assisted drug design*. ACS Monograph 112. Washington DC: American Chemical Society; 1979. p. 205–26.
- [49] Smellie A, Teig S, Towbin P. Poling: promoting conformational variation. *J Comp Chem* 1994;16:171–87.
- [50] Kearsley SK, Smith GM. An alternative method for the alignment of molecular structures: maximizing electrostatic and steric overlap. *Tetrahedron Comput Methodol* 1990;3:615–33.
- [51] Kirchmair J, Wolber G, Laggner C, Langer T. Comparative performance assessment of the conformational model generators omega and catalyst: a large-scale survey on the retrieval of protein-bound ligand conformations. *J Chem Inf Model* 2006;46:1848–61.
- [52] Boström J. Reproducing the conformations of protein-bound ligands: a critical evaluation of several popular conformational searching tools. *J Comput Aided Mol Design* 2001;15:1137–52.
- [53] Omega, OpenEye Scientific Software, Santa Fee, NM. Available from: <<http://www.eyesopen.com/>>.

- [54] Boström J, Greenwood JR, Gottfries J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *J Mol Graph Mod* 2003;21:449–62.
- [55] Jones G, Willet P. GASP: genetic algorithm superimposition program. In: Güner OF, editor. *Pharmacophore perception, development, and use in drug design*. La Jolla, CA: International University Line; 2000. p. 85–106.
- [56] Jones G, Willett P, Glen RC. A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. *J Comput Aided Mol Des* 1995;9:532–49.
- [57] Reutlinger M, Koch CP, Reker D, Todoroff N, Schneider P, Rodrigues T, et al. Chemically advanced template search (CATS) for scaffold-hopping and prospective target prediction for “orphan” molecules. *Mol Inf* 2013;32:133–8.
- [58] Brint AT, Willet P. Algorithms for the identification of three-dimensional maximal common substructures. *J Chem Inf Comput Sci* 1987;27:152–8.
- [59] Martin YC. DISCO: what we did right and what we missed. In: Güner OF, editor. *Pharmacophore perception, development, and use in drug design*. La Jolla, CA: International University Line; 2000. p. 51–66.
- [60] Martin YC, Bures MG, Danaher EA, DeLazzer J, Lico I, Pavlik PA. A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. *J Comput Aided Mol Des* 1993;7:83–102.
- [61] Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J Chem Inf Model* 2005;45:160–9.
- [62] Wolber G, Dornhofer AA, Langer T. Efficient overlay of small organic molecules using 3D pharmacophores. *J Comput Aided Mol Des* 2006;20:773–88.
- [63] Schlegel B, Meier R, Laggner C, Schnell D, Langer T, Seifert R, et al. Generation of a homology model of the human histamine H3 receptor for ligand docking and pharmacophore based screening. *J Comput Aided Mol Des* 2007;21:437–53.
- [64] UNITY, Tripos, St. Louis, MO. Available from: <<http://www.tripos.com/>>.
- [65] Ebejer JP, Morris GM, Deane CM. Freely available conformer generation methods: how good are they? *J Chem Inf Model* 2012;52:1146–58.
- [66] Leach AR, Gillet VJ, Lewis RA, Taylor R. Three-dimensional pharmacophore methods in drug discovery. *J Med Chem* 2010;53:539–58.
- [67] Kirchmair J, Ristic S, Eder K, Markt P, Wolber G, Laggner C, et al. Fast and efficient in silico 3D screening: toward maximum computational efficiency of pharmacophore-based and shape-based approaches. *J Chem Inf Model* 2007;47:2182–96.
- [68] Laggner C, Wolber G, Kirchmair J, Schuster D, Langer T. Pharmacophore-based virtual screening in drug discovery. In: Varnek A, Tropsha A, editors. *Chemoinformatics: an approach to virtual screening*. Royal Society of Chemistry; 2008. p. 76–119.
- [69] Langer T, Wolber G. Pharmacophore definition and 3D searches. *Drug Discov Today Technol* 2004;1:203–7.
- [70] Truchon J-F, Bayly CI. Evaluating virtual screening methods: good and bad metrics for the “early recognition” problem. *J Chem Inf Model* 2007;47:488.
- [71] Kirchmair J, Markt P, Distinto S, Wolber G, Langer T. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection—What can we learn from earlier mistakes? *J Comput-Aided Mol Des* 2007;22:213–28.
- [72] Triballeau N, Acher F, Brabet I, Pin JP, Bertrand HO. Virtual screening workflow development guided by the “receiver operating characteristic” curve approach: application to high-throughput docking on metabotropic glutamate receptor subtype 4. *J Med Chem* 2005;48:2534–47.
- [73] Hindle SA, Rarey M, Buning C, Lengauer T. Flexible docking under pharmacophore type constraints. *J Comput Aided Mol Des* 2002;16:129–49.
- [74] Barreca ML, De Luca L, Iraci N, Rao A, Ferro S, Maga G, et al. Structure-based pharmacophore identification of new chemical scaffolds as non-nucleoside reverse transcriptase inhibitors. *J Chem Inf Model* 2007;47:557–62.
- [75] Krovat EM, Fruhwirth KH, Langer T. Pharmacophore identification, in silico screening, and virtual library design for inhibitors of the human factor Xa. *J Chem Inf Model* 2005;45:146–59.
- [76] Maccari R, Ottana R, Ciarleo R, Vigorita MG, Rakowitz D, Steindl T, et al. Evaluation of in vitro aldose reductase inhibitory activity of 5-arylidene-2,4-thiazolidinediones. *Bioorg Med Chem Lett* 2007;17:3886–93.
- [77] Laggner C, Schieferer C, Poles G, Hoffmann RD, Glossmann H, et al. Discovery of high-affinity ligands of sigma1 receptor, ERG2, and emopamil binding protein by pharmacophore modelling and virtual screening. *J Med Chem* 2005;48:4754–64.
- [78] Spannhoff A, Heinke R, Bauer I, Trojer P, Metzger E, Gust R, et al. Target-based approach to inhibitors of histone arginine methyltransferases. *J Med Chem* 2007;50:2319–25.
- [79] Capdeville R, Buchdunger E, Zimmermann J, Matter A. Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. *Nat Rev Drug Discov* 2002;7:493–502.
- [80] Mol CD, Fabbro D, Hosfield DJ. Structural insights into the conformational selectivity of STI-571 and related kinase inhibitors. *Curr Opin Drug Discov Devel* 2004;5:639–48.
- [81] Wolber G, Kosara R. Pharmacophores from macromolecular complexes with ligandscout. In: Langer T, Hoffmann RD, editors. *Pharmacophores and pharmacophore searches*. Weinheim, Germany: Wiley-VCH; 2006. p. 131–50.
- [82] Schleifer KJ. Concept and applications of pseudoreceptors. In: Langer T, Hoffmann RD, editors. *Pharmacophores and pharmacophore searches*, 2006. Weinheim, Germany: Wiley-VCH; 2006. p. 117–30.
- [83] Höltje H-D. Theoretical studies on drug-receptor model complexes. *Pharm Acta Helv* 1979;54:125–34.
- [84] Marrer S, Höltje H-D. Molecular modelling study of calcium channel activating and blocking 1,4-dihydropyridine derivatives: a hypothetical molecular receptor model. *J Protein Chem* 1989;8:420–1.
- [85] Sippl W, Stark H, Höltje H-D. Development of a binding site model for histamine H<sub>3</sub>-receptor agonists. *Pharmazie* 1998;53:433–7.
- [86] Höltje H-D, Jendretzki UK. Construction of a detailed serotonergic 5-HT<sub>2a</sub> receptor model. *Arch Pharm (Weinheim)* 1995;328:577–84.
- [87] McMartin C, Bohacek RS. QXP: powerful, rapid computer algorithms for structure-based drug design. *J Comput Aided Mol Des* 1997;11:333–44.

- [88] Vedani A, Zbinden P. Quasi-atomistic receptor modelling: a bridge between 3D QSAR and receptor fitting. *Pharm Acta Helv* 1998;73:11–18.
- [89] Walters DE, Hinds RM. Genetically evolved receptor models: a computational approach to construction of receptor models. *J Med Chem* 1994;37:2527–36.
- [90] Lill MA, Vedani A, Dobler M. Raptor: combining dual-shell representation, induced-fit simulation, and hydrophobicity scoring in receptor modelling: application toward the simulation of structurally diverse ligand sets. *J Med Chem* 2004;47:6174–86.
- [91] Lill MA, Vedani A. Combining 4D pharmacophore generation and multidimensional QSAR: modelling ligand binding to the bradykinin B2 receptor. *J Chem Inf Model* 2006;46:2135–45.
- [92] Hahn M, Rogers D. Receptor surface models. 2. Application to quantitative structure-activity studies. *J Med Chem* 1995;38:2091–102.
- [93] Schleifer KJ. Pseudoreceptor model for ryanodine derivatives at calcium release channels. *J Comput Aided Mol Des* 2000;14:467–75.
- [94] Vedani A, Dobler M, Dollinger H, Hasselbach KM, Birke F, Lill MA. Novel ligands for the chemokine receptor-3 (CCR3): a receptor-modeling study based on 5D-QSAR. *J Med Chem* 2005;48:1515–27.
- [95] Bohacek R, Boosalis MS, McMartin C, Faller DV, Perrine SP. Identification of novel small-molecule inducers of fetal hemoglobin using pharmacophore and “PSEUDO” receptor models. *Chem Biol Drug Des* 2006;67:318–28.
- [96] Vedani A, Dobler M, Lill MA. Combining protein modeling and 6D-QSAR: simulating the binding of structurally diverse ligands to the estrogen receptor. *J Med Chem* 2005;48:3700–3.
- [97] Sippl W, Contreras JM, Parrot I, Rival YM, Wermuth CG. Receptor-based QSAR and design of novel acetylcholinesterase inhibitors. *J Comput Aided Mol Des* 2001;15:395–410.
- [98] Doweiko AM. 3D-QSAR illusions. *J Comput Aided Mol Des* 2004;18:587–96.
- [99] Varela R, Cleves AE, Spitzer R, Jain AN. A structure-guided approach for protein pocket modeling and affinity prediction. *J Comput Aided Mol Des* 2013;27:917–34.
- [100] Jain AN. QMOD: physically meaningful QSAR. *J Comput Aided Mol Des* 24:865–878.
- [101] Ligneau X, Perrin D, Landais L, Camelin JC, Calmels TP, Berrebi-Bertrand I, et al. BF2.649 [1-{3-[3-(4-Chlorophenyl) propoxy]propyl} piperidine, hydrochloride], a nonimidazole inverse agonist/antagonist at the human histamine H<sub>3</sub> receptor: preclinical pharmacology. *J Pharmacol Exp Ther* 2007;320:365–75.
- [102] Stark H, Kathmann M, Schlicker E, Schunack W, Schlegel B, Sippl W. Medicinal chemical and pharmacological aspects of imidazole-containing histamine H<sub>3</sub> receptor antagonists. *Mini Rev Med Chem* 2004;4:965–77.
- [103] Schlegel B, Sippl W, Hölftje H-D. Molecular dynamics simulations of bovine rhodopsin: influence of protonation states and different membrane-mimicking environments. *J Mol Model* 2005;12:49–64.
- [104] Murgueitio MS, Bermudez M, Mortier J, Wolber G. In silico virtual screening approaches for anti-viral drug discovery. *Drug Discov Today Technol* 2012;9:e219–25.
- [105] Kirchmair J, Distinto S, Liedl K, Markt P, Rollinger JM, Schuster D, et al. Development of anti-viral agents using molecular modeling and virtual screening techniques. *Infect Disord Drug Targets* 2011;11:64–93.
- [106] Perdih A, Kovač A, Wolber G, Gobec S, Solmajer T. Discovery of novel benzene 1,3-dicarboxylic acid inhibitors of bacterial MurD and MurE ligases by structure-based virtual screening approach. *Bioorg Med Chem Letters* 2009;19:2668–73.
- [107] Perdih A, Wolber G, Solmajer T. Molecular dynamics simulation and linear interaction energy study of d-Glu-based inhibitors of the MurD ligase. *J Comput Aided Mol Des* 2013;27:723–38.
- [108] Levay M, Krobot KA, Wittig K, Voigt N, Bermudez M, Wolber G, et al. NSC23766, a widely used inhibitor of Rac1 activation, additionally acts as a competitive antagonist at muscarinic acetylcholine receptors. *J Pharmacol Exp Ther* 2013;347:69–79.
- [109] Leschner J, Wennerberg G, Feierler J, Bermudez M, Wolber G, Welte B, et al. Interruption of the ionic lock in the bradykinin B2 receptor results in constitutive internalization and turns several antagonists into strong agonists. *J Pharmacol Exp Ther* 2013;344:85–95.
- [110] Mortier J, Rakers C, Frédéric R, Wolber G. Computational tools for in silico fragment-based drug design. *Curr Top Med Chem* 2012;12:1935–43.
- [111] Schuster D. 3D pharmacophores as tools for activity profiling. *Drug Discov Today Technol* 2010;7:e205–11.
- [112] Meslamani J, Li J, Sutter J, Stevens A, Bertrand H-O, Rognan D. Protein–ligand-based pharmacophores: generation and utility assessment in computational ligand profiling. *J Chem Inf Model* 2012;52:943–55.
- [113] Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein pharmacology by ligand chemistry. *Nat Biotech* 2007;25:197–206.
- [114] Koes DR, Camacho CJ. Pharmer: efficient and exact pharmacophore search. *J Chem Inf Model* 2011;51:1307–14.
- [115] Wood DJ, de Vlieg J, Wagener M, Ritschel T. Pharmacophore fingerprint-based approach to binding site subpocket similarity and its application to biosostere replacement. *J Chem Inf Model* 2012;52:2031–43.
- [116] Steuber H, Zentgraf M, La Motta C, Sartini S, Heine A, Klebe G. Evidence for a novel binding site conformer of aldose reductase in ligand-bound state. *J Mol Biol* 2007;369:186–97.
- [117] Steuber H, Zentgraf M, Gerlach C, Sotriffer CA, Heine A, Klebe G. Expect the unexpected or caveat for drug designers: multiple structure determinations using aldose reductase crystals treated under varying soaking and co-crystallisation conditions. *J Mol Biol* 2006;363:174–87.
- [118] Specker E, Bottcher J, Brass S, Heine A, Lilie H, Schoop A, et al. Unexpected novel binding mode of pyrrolidine-based aspartyl protease inhibitors: design, synthesis, and crystal structure in complex with HIV protease. *Chem Med Chem* 2006;1:106–17.
- [119] Sippl W. Receptor-based 3D QSAR analysis of estrogen receptor ligands: merging the accuracy of receptor-based alignments with the computational efficiency of ligand-based methods. *J Comput Aided Mol Des* 2000;14:559–72.
- [120] Evers A, Gohlke H, Klebe G. Ligand-supported homology modelling of protein binding-sites using knowledge-based potentials. *J Mol Biol* 2003;334:327–45.