Pharmacophore Modeling in Drug Discovery and Development: An Overview

Santosh A. Khedkar, Alpeshkumar K. Malde, Evans C. Coutinho[*] and Sudha Srivastava[†]

Department of Pharmaceutical Chemistry, Bombay College of Pharmacy, Kalina, Santacruz (E), Mumbai 400 098, India; [†]Tata Institute of Fundamental Research , Homi Bhabha Road, Navy Nagar, Colaba, Mumbai 400 005, India

Abstract: Pharmacophore mapping is one of the major elements of drug design in the absence of structural data of the target receptor. The tool initially applied to discovery of lead molecules now extends to lead optimization. Pharmacophores can be used as queries for retrieving potential leads from structural databases (*lead discovery*), for designing molecules with specific desired attributes (*lead optimization*), and for assessing similarity and diversity of molecules using pharmacophore fingerprints. It can also be used to align molecules based on the 3D arrangement of chemical features or to develop predictive 3D QSAR models. This review begins with a brief historical overview of the pharmacophore evolution followed by a coverage of the developments in methodologies for pharmacophore identification over the period from inception of the pharmacophore concept to recent developments of the more sophisticated tools such as *Catalyst*, *GASP*, and *DISCO*. In addition, we present some very recent successes of the widely used pharmacophore generation methods in drug discovery.

Key Words: Pharmacophore modeling, fingerprints, virtual screening, database search.

1. INTRODUCTION

Virtual screening (VS) of databases is gaining increasing importance in drug discovery because it is a reliable and a low cost method for identifying lead molecules. In the pharmaceutical industry, which is under ever increasing pressure to increase its success rate to bring drugs to the market, VS is seen as a complementary approach to experimental high-throughput screening. VS coupled with structural biology has the capacity to enhance the success rate of lead identification. Further, the growth in the identification of potential targets has increased the demand for reliable target validation, as well as for technologies that can identify rapidly several quality lead candidates. The advances in computational techniques enable VS to make a significant impact on the drug discovery process [1-5].

There are several tools and methods available for performing VS of ligand databases depending on the availability of information on ligands and receptors. If the structure of the receptor is available, *molecular docking* is carried out to discover a lead molecule. However, several groups have shown that the protein structure itself is a good source of a pharmacophore and can be used as a first-screen before docking studies [6,7]. A pharmacophore-based search of 3D databases can be carried out even in the absence of information on the receptor structure. In many cases, the receptor structure is difficult to obtain, because the receptor is embedded in the transmembrane that poses an obstacle for crystallization, for example, the G-protein coupled receptors (GPCRs). A ligand or a set of ligands that bind to a particular receptor can be utilized efficiently to search a database for molecules with similar properties. The ligand-based pharmacophore modeling methods use information (*features*) provided by a compound or a set of compounds that are known to bind to the desired target, to identify other compounds in the corporate or commercial databases with similar properties. This is usually achieved by similarity and substructure searching [8], pharmacophore matching [9] or 3D shape matching [10]. The two methods – pharmacophore mapping and molecular docking complement each other and can be synergistically integrated to improve the drug design and development process.

This article is intended to provide an overview of pharmacophore identification and search methods along with commercial algorithms incorporating these methods, which are currently employed in *in silico* screening of ligand databases. The article concludes with some successful examples of drug discovery based on these approaches.

2. HISTORY AND EVOLUTION OF THE PHARMACOPHORE CONCEPT

The credit for the first use of the pharmacophore concept goes to Paul Ehrlich who devised a way to develop dyes through chromophores (the part of a molecule responsible for imparting color). He gave the first definition for a pharmacophore in 1890 as "a molecular framework that carries (*phoros*) the essential features responsible for a drug's (*pharmacon*) biological activity". The modern definition of pharmacophore as coined by Peter Günd is "a set of structural features in a molecule that is recognized at a receptor site and is responsible for that molecule's biological activity" [11]. The evolution and history of the pharmacophore concept has been reviewed by Peter Günd [12].

The pharmacophore concept could not achieve its full utility until the development of 3D database searching software in the 1990's. The first computer program, *MOLPAT*

*Address correspondence to this author at the Department of Pharmaceutical Chemistry, Bombay College of Pharmacy, Kalina, Santacruz (E), Mumbai 400 098, India; Tel: +91-22-2667 0871; Fax: +91-22-2667 0816;
E-mail: evans@bcpindia.org

1573-4064/07 $50.00+.00

[13], to recognize pharmacophore patterns was developed by Günd, Wipke and Langridge at Princeton University in 1974. The demand for 3D structure searching software grew with the development of rapid 3D structure generation programs such as *CONCORD* [14], *CORINA* [15,16], *AIMB* [17] and *WIZARD* [18]. 3D search software like *ALADDIN* [19] (Abbott Laboratories, later commercialized by Daylight Chemical Information Systems, Inc.) and *3D-Search* [20] (Lederle Laboratories) were developed by pharmaceutical companies, while academic and government institutions developed *CAST-3D* [21] (Chemical Abstract Services), *DOCK* [22] (University of California at San Franscisco) and *CAVEAT* [23] (University of California at Berkley). Later, the group of Marshall developed a pharmacophore method based on ligand structures called "active-analog" approach [24] and applied it to a set of ACE inhibitors [25]. Recently, they validated this pharmacophore model against available experimental information and found a good correlation [26].

The first commercial 3D searching system, *MACCS-3D* [27], was developed by Güner *et al.* and was released in December of 1989. During the next four years, all of the technology that is available today was developed – *ChemDBS-3D* [28] (Chemical Design Inc., USA), *UNITY* [29] (Tripos Inc., USA) and *Catalyst* [30] (Accelrys Inc., USA). The critical demand for the pharmacophore development software was reached when the above mentioned 3D searching technologies were widely available. Though most of these 3D searching software had inbuilt query generation tools, specialized pharmacophore generation software were also being developed. Most notable among them were *DISCO* [31] by Martin *et al.* (Tripos Inc., USA), *HipHop* [32] by Barnum *et al.* (Accelrys Inc., USA), and *GASP* [33] by Jones and Willett (Tripos Inc., USA). Meanwhile, predictive models based on QSAR such as *CoMFA* (Tripos Inc., USA) [34] by Cramer *et al.*, *Apex-3D* (Accelrys Inc., USA) [35] by Golander and Vorpagel and *HypoGen* [36] by Teig *et al.* (Accelrys Inc., USA) also came into existence. Detailed usage and validation of all the pharmacophore development software have been covered in the pharmacophore book edited by Güner [12]. For other timely reviews in the field, the reader should refer to references 37 to 41.

A very simple example of a pharmacophore is the one generated using *ALADDIN* for agonists of the dopamine D1 receptor. This pharmacophore contains a basic amine nitrogen (:N), a hydroxyl group (OH) and an aromatic ring (Ar) with the distances between N – O (6.8 – 8.3 Å), Ar – O (2.7 – 2.9 Å) and Ar – N (4.2 – 4.8 Å) as shown in Fig. (**1**). This model led to the discovery of a constrained analog of dopamine, A-68930, a ligand highly selective for the D1 receptor

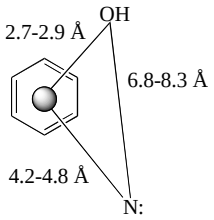

Fig. (1). An early pharmacophore model for dopamine D1 receptor agonists.

which had been synthesized earlier for a project targeting adrenergic receptors [42].

3. PHARMACOPHORE METHODS

Pharmacophore modeling provides a useful framework for a better understanding of the existing data, and can be used as a productive tool in the design of compounds with improved potency, selectivity and/or pharmacokinetic properties. Pharmacophore models are generated by analyzing structure-activity relationships and mapping common structural features of active analogs. The pharmacophore can be identified by direct method (using receptor–ligand complexes) or by indirect method (using only a collection of ligands that are known to interact with a given receptor, Fig. (**2**)). However direct methods are becoming extremely important because of the high rate at which protein structures are being determined. Depending on the level of automation of the process, these methods can be classified as manual or automatic (algorithm-based).

The manual method involves visual identification of structural and chemical features among the active molecules and those that are missing in the inactive ones. Then the spatial relationships (3D aspects) of the common features are measured in the development of a draft pharmacophore. This is then validated by logical and/or statistical methods. Finally the model is refined until desired results are obtained.

MOLPAT [13] was the first automated pharmacophore generation computer program. Since then many advances have taken place in automated methods which is reflected in the recent commercial programs like *Dis*tance *Co*mparison (*DISCO* [31]), *HipHop* [32] (a part of *CATALYST* [30]), *G*enetic *A*lgorithm *S*uperposition *P*rogram (*GASP* [32]), *Chem Diverse* (3 and 4-point pharmacophore generation in *Chem-X* [43]), *SLATE* [44], *MOLMOD* [45], *MIMIC* [46], *M*apping *Ph*aramcophores *I*n *L*igands (*MPHIL* [47]), Dynamic Pharmacophore approach using molecular dynamics [48] and receptor guided approaches. *DISCO*, *Catalyst* and *GASP* are widely used for pharmacophore identification.

The pharmacophore searching methods mainly differ in the manner of defining the query, handling conformational flexibility, and the approach of identifying the pharmacophore pattern.

3.1. Pharmacophore Generation

The ligand data set for construction of the pharmacophore model must be selected with great care. The type of ligand molecules, the size of the dataset and its chemical diversity affect the final pharmacophore model. The Carnell-Smith method, *RAPID* [49] and *HipHop* [32] do not take into consideration the activity data of molecules. *CLEW* [50] and the current version of *DISCO* [31] can consider information on inactive ligands that can be fruitfully utilized to indicate structural features that significantly decrease the activity. Models to predict the activity of unknown compounds can be derived using, for example, *HypoGen* [36] which utilizes a large enough set of diverse compounds (18 to 30) with different activity levels (4 to 5 orders of magnitude on the log scale). The pharmacophore generation methods such as *HipHop* [32], *HypoGen* [36], *MPHIL* [47] and *RAPID* [49]

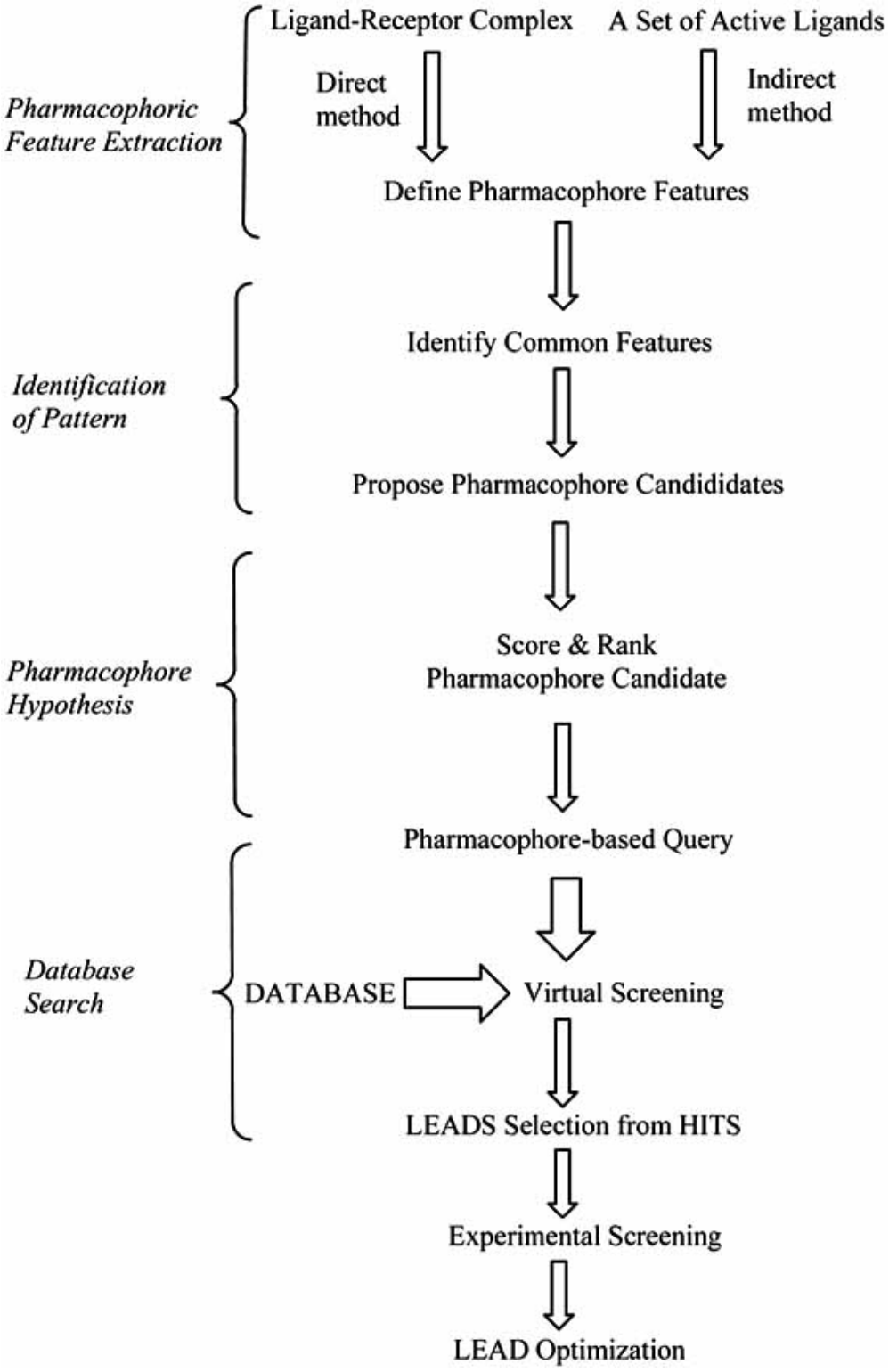

Fig. (2) Flow chart of the virtual screening process using the pharmacophore method.

are designed to handle small (less than 100 ligands) data sets. There are methods that use large data sets as input but then prune them into a smaller one by sorting the activities of ligands depending on the user specified cut off. Lastly, the data set, with molecules binding to the same pocket in the target, should be as diverse as possible, so as to get an accurate pharmacophore model. However, one should be aware of the fact that very different ligands may bind at different biding sites, resulting in a bad pharmacophore model.

In the next step, the features relevant to the pharmacophore discovery are extracted from the input ligands (*feature extraction*, Fig. (**2**)). Features can be defined depending on topology (phenyl ring and carbonyl group), function (H-bond donor/acceptor, acid, base, aromatic ring and hydrophobic group) and atom-based (3D position of atom and atom type). Topology-based and function-based features encounter some drawbacks; for example, the amine nitrogen can be classified as both a H-bond donor and acceptor. A functional group can be represented by its center. The center of an acid, base, H-bond donor/acceptor is usually defined as the position of the actual atom. A hydrophobic region or an aromatic ring can be described by its centroid. However, a vector representation is more accurate than a point representation since it imposes an additional constraint on bond directionality between the ligand feature and its complementary feature on the receptor. In addition, a hydrophobic group can be represented by a sphere and an aromatic ring by a plane and it's normal.

The selected features from each ligand are combined to form a representation of the whole structure. The program *RAPID* [49] represents ligand structures as a set of labeled points in 3D space, where each point is associated with an atom type (a feature). Another approach to represent a ligand structure is by a labeled graph with nodes representing the features and the edges representing the relations. A molecule can also be represented as a graph with atoms as vertices and the bonds as edges [51]. In another approach, a ligand structure is considered as a set of labeled points, together with the associated interpoint distances [52]. This type of representation is orientation-independent, in contrast to the 3D point set representation.

In the *pattern identification* phase (Fig. (**2**)), the features extracted from different ligand molecules are matched and pharmacophore candidates are proposed. A pattern or *configuration* is a set of features with their relative locations in 3D space. A ligand is said to match a pattern if it possesses a set of features and a conformation such that the features can be superimposed with the corresponding locations. The most popular approach to define a pattern is to find the *Maximal Common Substructure* (MCS) which has been implemented in *DISCO* [31], *RAPID* [49], *GAMMA* [53], and *GASP* [33]. This approach is based on the assumption that a common pharmacophore is responsible for the observed activity. The drawback of the MCS approach can be overcome by relaxing the requirement that all input ligands possess all the features (*Relax MCS approach*). This approach is used in the *MPHIL* [47] program.

Several algorithms have been developed for pattern identification [52]. The clique detection algorithm [52,55,56] is implemented in *DISCO* [31] and *MPHIL* [47] programs. *HipHop* [32] and *SCAMPI* [57] use an exhaustive search algorithm in which the search for a pattern starts with small sets of features, extending them until no larger common patterns exist. *HypoGen* [36] uses a similar approach but also incorporates information on activity in the pharmacophore derivation. This is done in three steps: the *constructive stage* identifies pharmacophore candidates that are common among the most active set of ligands, this is followed by the *subtractive stage* in which those candidates identified in step 1 that are also present in more than half of the least active ligands are removed, and the last step of *optimization* attempts to improve the score of the pharmacophore candidates that pass the subtractive stage, by simulated annealing. *GASP* [33] and *GAMMA* [53] are based on the genetic algorithm (GA) and also perform a conformational search as part of the GA run. In this way, molecular flexibility is simulated by applying the genetic operators to the first part of the chromosome. Changing the second part of the chromosome allows an exploration of ways to align the molecules and to identify the pharmacophore pattern. In the last step of pharmacophore generation, candidates are scored and ranked; a lower score indicates a greater possibility that the model has been obtained by chance correlation.

As mentioned in the introductory section, VS by pharmacophore searching can be more efficient in presence of a structural knowledge of the target receptor. Receptor-based approach for pharmacophore generation involves an analysis of the features of the active site and their spatial relationships; an active image of this is then used to construct the pharmacophore model. This gives rise to a large number of features and it is necessary to determine which of these are actually parts of the pharmacophore. The method takes as an input the 3D structure of the receptor (usually in PDB format) and a set of ligands with known activity. Using knowledge of the active site residues (from biochemical or structural studies), a program such as *LUDI* [58,59] generates an *interaction map* which is a complement of the receptor binding site (Fig. (**3**)). The functional features such as H-bond donors/acceptors and lipophilic groups are identified. Often it is necessary to prune the number of features, since queries with multiple features often fail to retrieve any hits from a database. Therefore, 3D queries composed of fewer features are generated by considering all possible combinations. *Catalyst* [30] uses these queries to search a database of ligands.

4. PHARMACOPHORE FINGERPRINTS

The *Chem-X* [43] software can be employed to define pharmacophore fingerprints in which a finite set of pharmacophores, also called as the *pharmacophore space*, is used to define the pharmacophore fingerprint or key. The pharmacophore fingerprints can be used to measure molecular similarity, to design libraries, to assess their diversity and to help search for novel active molecules. *Chem-X* uses a *n-point* pharmacophore model where *n* is the number of pharmacophore features (*centers*), usually 3 or 4. By default, there are seven *center* points important for ligand – receptor interactions. These are H-bond donor, H-bond acceptor, positively charged center, aromatic ring centroid, hydrophobic center, acidic center and basic center. A hydrophobic center is placed at the centroid of a lipophilic group. In this way, a 4-point pharmacophore is defined by four centers and six inter-center distances. Furthermore, a continuous range of inter-center distances is partitioned into a specified number of bins. Consider an example of a four-point pharmacophore having six sides defined by 15 distance ranges. This leads to a large space of 210×15^6 potential four-point pharmacophores, where 210 is the number of ways to chose 4 centers from the seven *center* types. This is a very high number and therefore seven to ten distance bins are usually used.

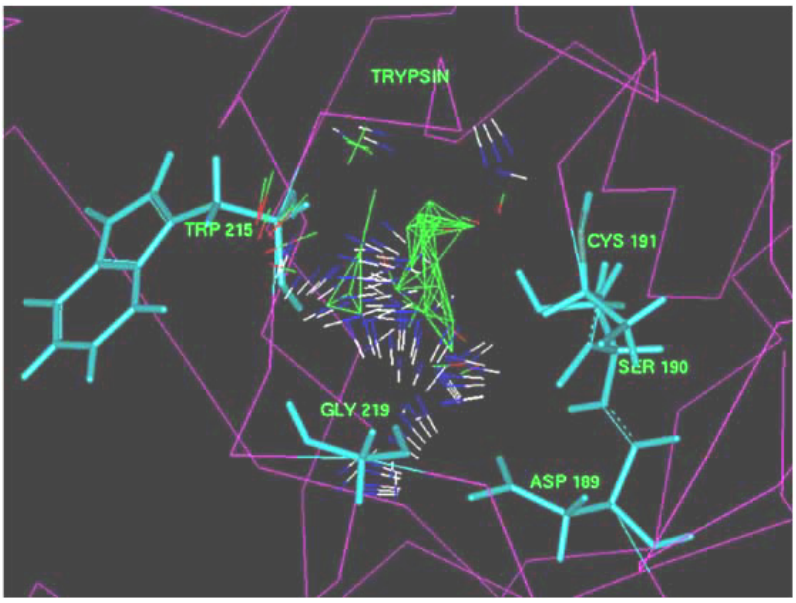

Fig. (3). *LUDI* interaction map of the active site of trypsin, a serine protease enzyme. Vectors that are half red and half green are hydrogen acceptor sites, with red for acceptor and green for acceptor antecedent. The blue and white vectors are hydrogen donor sites with white for the hydrogen and blue for the heavy atom. The hydrophobic sites are at the vertices of the lines that are entirely green. (Refer to online version for color figure).

4.1. Defining Pharmacophore Fingerprints

Every feature in the pharmacophore space is identified by a bit (0 or 1) that indicates the presence (1) or absence (0) of potential feature. Definition of features in the binary way is called fingerprints. Each bit in the fingerprint corresponds to a potential feature. Pharmacophore fingerprints are usually generated for a set of compounds instead of an individual one. In the direct mode, in which the structure and location of the binding site is known, pharmacophore fingerprints can be generated from complementary site-points in the binding pocket of the receptor. *Chem-X* [43] or *GRID* [60] can be used to automatically generate the site-points. From the site-points a set of complementary points can be generated to represent a hypothetical molecule that can bind to all possible positions in the binding pocket. This hypothetical molecule is then used to define pharmacophore fingerprints in the same manner as for a known ligand.

Pharmacophore fingerprints can be used in the analysis of molecular similarity by counting the number of bits common in a set of molecules or by calculating a similarity coefficient such as *Tanimoto coefficient* [61] (it ranges from 1 for identical molecules to 0 for molecules having no common bits). Pharmacophore fingerprints can also be used for assessing pharmacophore diversity. Furthermore, it can be used to design a diverse library with a smaller number of molecules without compromising the diversity.

Pharmacophore fingerprints, as defined above, assume only a single bit to represent the presence or absence of a pharmacophore. However, a set of ligands binding to the same receptor can have many pharmacophores in common and consequently these common pharmacophores are more likely the actual ones responsible for binding. Therefore, the frequency count for each pharmacophore is also included while defining the pharmacophore fingerprints.

5. DATABASES FOR USE WITH PHARMACOPHORES

5.1. Database Preparation

The availability of a database (DB) of ligand molecules is the primary requirements for performing VS either by the docking or pharmacophore-based search method. A corporate database or those available from chemical vendors such as Available Chemicals Directory (ACD) [62], Cambridge Structure Database (CSD [63]), World Drug Index (WDI [64]), or those in the public domain such as MayBridge [65], National Cancer Institute (NCI [66]) database, PubChem [67], ZINC [68], ChemDiv [69], *etc*., can be employed for this purpose. Additionally, a database of reagents and compounds of known chemistries, which are readily synthesizable, can be used for VS after a primary filtration of 'drug-like' properties such as variation of Lipinski's rule-of-5 [70], the polar surface area, *etc*. For lead-like properties one can filter the database based on rule-of-3 [71] (*i.e.*, molecular weight $\leq$ 300, number of H-bond donors/acceptors $\leq$ 3, ClogP $\leq$ 3, number of rotatable bonds $\leq$ 3, polar surface area $\leq$ 60 $Å^2$), so that after lead optimization the molecule fits into Lipinski's rule-of-5. A database could also be filtered to remove compounds with specific substructures associated with poor chemical stability or toxicity. Also, physically relevant ionization and tautomeric states must be assigned to the compounds in the DB. It is prudent to use all the relevant tautomers because there is no way of knowing *a priori* which tautomer is most likely to bind to the receptor.

3D coordinates of a molecule can be generated using programs such as *CORINA* [72] (Molecular Networks GmbH, Germany), *CONCORD* and *CONFORT* (Tripos Inc., USA [73]) or *CONVERTER* (Accelrys Inc., USA [74]). Sometimes it is necessary to assign partial charges to the com-

pounds in the database depending on the requirement of the 3D search method being used. Another issue in VS is the handling of conformational flexibility of ligands in a database. Conformational flexibility can be introduced in VS through incorporation of multiple conformations of every molecule in the database, by relaxing the query, or by letting the software generate multiple conformations during the search process. Storing multiple conformations of every ligand in the database is one way to handle conformational flexibility. However this is a practically intractable solution to the problem because the number of conformations of a ligand increases with the number of rotatable bonds. A small drug-like molecule with 4 rotatable bonds, and if scanned with a 120^{o} increment of the dihedrals, can generate 81 conformations, and there is no firm basis to which ones should be selected/included in the DB. A better option is the hybrid approach. A combination of a multi-conformational database along with a flexible search, provides an efficient and effective route to DB searches. This hybrid approach is employed in *Catalyst's BEST* search method (*BEST* is one of the two types of conformational analysis implemented in *Catalyst*, the other one is *FAST*).

6. APPLICATIONS OF PHARMACOPHORE

6.1. *De Novo* Design of Ligands

The pharmacophore can be used to design novel ligands that satisfy the constrains defined by the pharmacophore model. The *NEWLEAD* [75] method uses as input a set of disconnected molecule fragments that are consistent with the pharmacophore model. These fragments are then joined with connecting pieces that consist of small chemical groups. If the receptor structure is known, *LUDI* can be utilized to combine the identification of receptor-based pharmacophore with *de novo* design. Thus, the pharmacophore approach is an easy and fast method for searching established molecules, and in the absence of active ligands (usually at the start of new project), for designing novel molecules.

6.2. Database Searches Based on Pharmacophore

As explained above, a pharmacophore query is used to screen 3D database(s) of compounds, which on successful completion retrieves a set of compounds, called *hits* that match the pharmacophore query. Some of these hits might be known active compounds, but others might be entirely novel classes of compounds. Thus, pharmacophore searching can be used to discover novel lead compounds with unknown pharmacological properties. This diversity increases the chances that some of the compounds will pass all the stages of the drug development process.

6.3. Lead Optimization

The optimization of leads is a process of enhancing the binding affinity with simultaneous optimization of ADME characteristics. Both the above-mentioned methods, *pharmacophore searching* and *pharmacophore-based de novo design*, are capable of spawning totally new molecules containing the pharmacophore. Thus they have a good chance of being bioactive, but with a different pharmacokinetic/pharmacodynamic profile.

3D-QSAR models may be used to predict the biological activity of proposed molecules (*lead optimization*), using methods like *CoMFA* [34]. The crucial input for *CoMFA* and related 3D-QSAR analyses is the alignment of the molecules. Pharmacophore-based methods such as *DISCO, GASP*, *Apex-3D* and *HypoGen* may be used to define the rules for overlaying molecules. Current publications describe in detail how the alignments were performed, either explicitly using a pharmacophore discovery method, or implicitly using a methods such as fit atom, field fit, *etc*.

The geometric arrangement of features and the steric boundaries of the binding site can be inferred from the dataset using shape-enhanced pharmacophores generated by *DANTE* [76]. Unlike *CoMFA*, steric boundaries derived by *DANTE* have demonstrated to be surprisingly useful in prospective applications, by defining the '*limits of the playing field*', *i.e.*, constraining the space of possible molecules that a chemist should consider. *DANTE*'s shape-enhanced pharmacophore can easily be used as 3D database search queries, to screen databases composed of combinatorial libraries constructed around the lead. In *DANTE*, regions of the binding surface are marked either as '*sterically forbidden*' (active molecules in the dataset lie within that boundary, while inactive molecules in the dataset protrude beyond it), or '*terra incognita*', *i.e.* active molecules lie within that region, and define the extent of that surface, but no molecules in the dataset protrude beyond that region. This is a very important but underappreciated concept for molecular design in lead optimization, as medicinal chemists need to discover novel compounds, and need to explore regions of space hitherto unexplored. One can use a *DANTE* shape-enhanced pharmacophore to explicitly search for molecules that extend into these unexplored regions, to probe new regions of chemical space and to see if the properties improve during lead optimization.

7. APPLICATIONS OF SOME WIDELY USED PHARMACOPHORE GENERATION METHODS

7.1. *DISCO*

DISCO [31] (DIStance COmparison) is a fast, automated, systematic analysis introduced by Yvonne Martin and colleagues at Abbott Laboratories to discover (a) how many pharamacophores can explain the data and which conformations and superposition rules are to be used to explain the data; (b) the trade-off between a low RMS for superposition and inclusion of more points in the model; and (c) the trade-off between a low RMS and inclusion of higher-energy conformations in the model. *DISCO* searches over the input conformations of a set of structures to find the pharmacophore; by default identifying positive, negative, hydrogen-bond donor, hydrogen-bond acceptor, and hydrophobic ligand points and hypothetical complementary receptor site points which are common to the set. It proposes a superposition rule and the bioactive conformation of each molecule. Instead of considering a conformation to be a three-dimensional object in space, *DISCO* considers a conformation to be a set of interpoint distances. A point can have more than one label; for example, the hydroxyl oxygen is labeled as both a hydrogen-bond acceptor and donor. The distances are calculated not only between the locations of atoms, but also between points located at the hypothetical position of the complementary atoms of a macromolecule. *DISCO* uses the

Bron-Kerbosh clique-detection algorithm [54-56] for the distance comparisons. Clique detection comes from graph theory. A clique is a subgraph in which every node is connected to every other node. The clique detection algorithm finds the largest clique in a reference graph which is contained in every other graph in the set. For the pharmacophore purpose, nodes are the feature elements and connections are the feature-to-feature distances. The graph of the molecule is said to match with the reference if, for any of the molecule's conformers, there exists a set of elements of the same class as that of the reference and the inter-element distances match those of the reference within a specified tolerance.

The molecule with the least number of conformations is used as a reference molecule. *DISCO* takes each conformation of the reference molecule in turn and compares it to all conformations of the other molecules. The cliques identified are examined in an attempt to find one that is common to at least one conformation of every molecule. This process is repeated for every conformation of the reference molecule. If no solution is found, the tolerances on the clique detection process are increased until a solution is found, or the maximum tolerance is reached. The output from a *DISCO* run is a ranked list of all possible pharmacophore mappings where each feature of a pharmacophore should be present in all molecules. This requirement may result in good pharmacophores being missed; hence *DISCO* also has the option of finding solutions where some molecules are excluded from the model. The pharmacophore model, after a suitable validation, can be used subsequently for database searches to discover new hits. The *DISCO* based pharmacophoric alignments have also been used as an input alignment for subsequent CoMFA analysis [77].

Flower and colleagues have reported successful applications of *DISCO* in combination with database searching for discovery of muscarinic M3 receptor antagonists [78]. These molecules find use in the treatment of irritable bowel syndrome, chronic obstructive airway disease and urinary incontinence. Three molecules, a flexible selective muscarinic M3 antagonist and two rigid non-selective muscarinic M3 antagonists, were used to derive the pharmacophore models. Five models were generated, out of which two 4-point models were selected based on visual inspection of the structural superposition. These two models were subsequently subjected to a *UNITY* search of a database comprising of molecules from the Astra Charnwood, UK 'in-house' compound bank, Aldrich, Maybridge, Specs & Biospecs and Bionet. The first model yielded 176 hits, while a search using the second model resulted in 172 hits. There were 172 hits common to the two sets, which were screened for muscarinic M3 receptor antagonist activity. Three compounds with pA_2 value of 4.83, 5.54 and 6.67 were found to be good hits. The best hit, the most potent of the above three was found to be a competitive antagonist with a simple chemical structure and limited similarity to existing M3 receptor antagonists (Fig. (**4**)).

Fossa *et al.* [79] have reported a pharmacophore model of phosphodiesterase (PDE) type III inhibitors, using *DISCO* methodology. These inhibitors, implicated for the treatment of congestive heart failure, bind to the cAMP site of the enzyme. Thirteen structurally diverse molecules, along with the

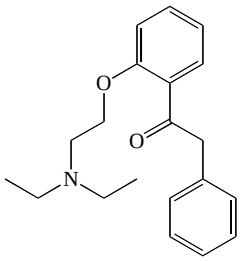

Fig. (4). A competitive muscarinic M3 receptor antagonist discovered using pharmacophore based database search.

anti conformation of cAMP, with varying biological activity were used to derive the pharmacophore models. Three models (Model A, a six point model; Model B, an eight point model and Model C, an eleven point model) with various combinations of molecules in the dataset were generated. The models were validated using *GRID* based calculations of interaction energies and subsequent 3D-QSAR using principal component analysis (PCA). The pharmacophore models could explain the observed selectivity and potency of various inhibitors and provided the basis for the design and synthesis of more selective inhibitors.

Fossa *et al.* [80] have also reported a topographical model for the PDE IV catalytic site based on the pharmacophore model derived using *DISCO* for a set of PDE IV inhibitors. The PDE IV inhibitors are useful for treatment of asthma and chronic obstructive pulmonary disease. A set of eighteen structurally diverse PDE IV selective inhibitors with a homogenous distribution of pharmacological activities (six compounds sets, each with moderate, good and optimal potency) were used to derive the pharmacophore models. The models were validated by calculating the molecular electrostatic potential and hydropathic fields with *MOPAC* and *HINT* respectively, for the molecular conformations selected by *DISCO*. Three models were constructed using different combination of inhibitors. Model A, with all eighteen molecules, is a four-point model which gives the requirement for basic PDE IV inhibition. Model B, derived using compounds with good and optimal potency, is a six point model that gives the requirements for good PDE IV inhibition. Model C, derived using compounds with optimal potency, is an eight point model that gives insights into the requirements for a potent PDE IV inhibitor. Comparison of the three models provided important structural insights for the design of novel and selective PDE IV inhibitors.

7.2. *GASP*

GASP [33] (Genetic Algorithm Similarity Program) is a program based on genetic algorithm for superposition of flexible ligands to derive pharmacophore models. *GASP* was developed as part of a research project to evaluate the utility of Genetic Algorithms (GAs) for tackling combinatorial problems in molecular recognition. *GASP* exploits some of the methods that were developed in the docking program *GOLD* [81]. The *GASP* program does not require *a priori* knowledge of either the constraints or the nature of the pharmacophoric pattern, to run successfully. The only input required is a set of molecules. *GASP* uses two unique features to find pharmacophore alignments: a genetic algorithm

to drive the evolution of better models and a unique fitness function which takes into account the protein-ligand interactions of different acceptors and donors.

A GA is a computer program that mimics the progress of evolution by manipulating a population of data-structures known as chromosomes. Staring from an initial randomly generated population of chromosomes, the GA repeatedly applies two genetic operators, crossover and mutation. Crossover combines chromosomes, while mutation introduces random perturbations. Both operations work on parent chromosomes that are randomly selected from the existing population with a bias towards the fittest, thus introducing an evolutionary pressure into the algorithm. The fitness measures how good a solution to the problem under consideration is encoded in that chromosome. The emphasis on the survival of the fittest ensures that, over time, the population moves towards the optimum solution. In the present context, this corresponds to the best possible structural overlay of a series of active molecules that are presumed to bind to a biological receptor in a similar fashion. Given a set of active molecules, *GASP* selects one of them as a 'base molecule', to which the other molecules are fitted. The chromosome in *GASP* encodes a range of information that is necessary to ensure an appropriate overlay of a molecule onto a base molecule. Each chromosome contains binary strings that encode angles of rotation about the rotatable bonds in the molecules, and integer strings that map hydrogen-bond donor protons, acceptor lone pairs and ring centers in the base molecule to corresponding sites in each of the other molecules. A least-squares fitting process is used to overlay molecules onto the base molecule in such a way that as many as possible of the structural equivalences suggested by the mapping are formed. The fitness of a decoded chromosome is then a combination of the number and similarity of overlaid features, the volume integral of the overlay, and the van der Waals energy of the molecular conformations.

Pajeva and Wiese [82] have reported a successful application of *GASP* to deduce the pharmacophore for a set of structurally diverse molecules that bind to the P-glycoprotein (P-gp) veramapil binding site. The P-gp modulators find application as multi-drug reversal (MDR) agents. Nineteen structurally diverse substrates and modulators were used in the study. A pharmacophore model with two hydrophobic points, three hydrogen-bond acceptor points and one hydrogen-bond donor point was obtained in the study. The pharmacophore model revealed that the binding affinity of the drugs depends on the number of pharmacophore points simultaneously involved in the interaction with P-gp. The authors proposed the following hypothesis to explain the broad structural variety of the P-gp substrates and modulators: (a) the veramapil binding site of the P-gp has several points that can participate in hydrophobic and hydrogen-bonding interactions; (b) different drugs can interact with different receptor points in different binding modes. The study provides a good model for prediction of binding of molecules to P-gp.

7.3. *Catalyst*

Catalyst [30] is one of the widely used software for pharmacophore generation and database searching. The software contains four components: *ConFirm*, *HypoGen*, *HipHop* and *CatSearch*. In *Catalyst*, the pharmacophores are generally referred to as hypothesis. The primary objective of *ConFirm* is to generate a moderate number of conformations for a given molecule while adequately covering its conformational space within a defined energy threshold. The approach is based on the 'poling' algorithm. If the biological activity data is included during the pharmacophore generation, then *HypoGen* is employed. Alternatively, when no activity data is considered during the hypothesis generation and only common chemical features are considered, then *HipHop* is used. *CatSearch* is the database searching tool in *Catalyst* which involves a rapid screening process followed by a rigorous atom-by-atom mapping in which a fairly comprehensive set of features including customizable chemical features (e.g. donors, acceptors, and hydrophobes), exclusion spheres and inclusion volumes (shape) are considered. In addition to the traditional pharmacophore based searching, shape similarity and partial match searching is also available. A pharmacophore model, or hypothesis, consists of a three-dimensional configuration of chemical functions surrounded by tolerance spheres. A tolerance sphere defines an area in space that should be occupied by a specific type of chemical functionality. Each chemical function is assigned a weight, which describes its relative importance within the hypothesis. A larger weight indicates that the feature is more important in conferring the activity than the other composite parts of the hypothesis. *HypoGen* and *HipHop* have been used as alignment tools. *HypoGen* can include the biological activity as a dependent property in the alignment phase.

Böhm and colleagues [83] have applied *Catalyst* along with a database search using *LUDI*, to discover novel bacterial DNA gyrase inhibitors. The pharmacophore obtained consists of a hydrogen bond donor, a hydrogen bond acceptor, and a lipophilic region. Compounds from the Available Chemical Dictionary (ACD) numbering about 350,000 and a part of the Roche compound inventory (RCI) were reduced to 3000 compounds using *in silico* screening. After screening 3000 molecules for *in vitro* DNA gyrase inhibition, 150 inhibitors (14 scaffolds) were found to exhibit weak activity. Among the 14 scaffolds, 7 scaffolds were determined as novel DNA gyrase inhibitors, which bind to the ATP binding site. The optimization of the indazole scaffold provided a structurally novel inhibitor (Fig. (**5**)), 10 times more potent as a DNA gyrase inhibitor than novobiocin.

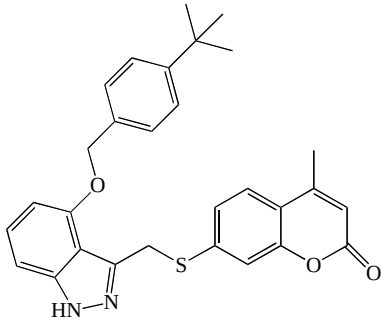

Fig. (5). A novel DNA gyrase inhibitor, discovered using pahrmacophore-based database search, 10 times more potent than novobiocin.

Laggner and colleagues [84] have reported pharmacophore models for three protein targets involved in sterol metabolism. A dataset of 23 structurally diverse molecules with binding affinity data for emopamil binding protein (EBP), ERG2 (fungal counter-part of EBP) and sigma-1 receptor were used in the study to derive a pharmacophore model with the *HypoGen* module of *Catalyst*. These three enzymes of sterol metabolism share a high affinity for various structurally diverse compounds. Three pharmacophore models with one positive ionizable group and four hydrophobic features in common, but with different spatial arrangements were derived and validated. The study showed that hydrogen bond interactions are not required for high-affinity inhibitor binding. The models were subsequently used in a database search. The databases used were the World Drug Index (WDI, 48,405 molecules) and a 3,525 metabolite subset of the KEGG (Kyoto Encyclopedia of Genes and Genomes) COMPOUND database. In the virtual screen, the drugs that were reported previously to bind to one or several of these proteins were retrieved along with 11 new hits which were then tested experimentally. Inhibitors with nanomolar binding affinity were discovered. The ERG2 pharmacophore model, when searched against 3525 metabolites, successfully retrieved 10 substrates among the top 28 hits. The models can be used as screens for chemically diverse and putative endogenous ligands.

There are studies reported in the literature where several pharmacophore generation methods have been applied to the same datasets and evaluated for their ability to generate known pharmacophores deduced from protein-ligand complexes extracted from the Protein Data Bank. Andrew Leach and colleagues [85] have reported one such study wherein *DISCO*, *GASP* and *Catalyst/HipHop* were applied to five datasets comprising of thrombin (seven ligands), cyclin dependent kinase 2 (CDK2, six ligands), dihyrofolate reductase (DHFR, six ligands), HIV reverse transcriptase (HIV-RT, ten ligands) and thermolysin (six ligands). The datasets were selected on the basis of an abundance of crystallographic information and diversity of the ligands. A set of ligands for a known protein-ligand complex were identified using the *Relibase+* program. The pharmacophoric features for each ligand in a protein family were deduced by examining the protein-ligand complex in *Relibase+* and by reference to the literature. A target pharmacophore was then defined as the set of pharmacophoric features that is common to all ligands. Each of the above-mentioned programs was then tested by its ability to generate the target pharmacophore. The program evaluation consisted of two phases: a 'rigid search' using the bound conformations and later a 'flex search' in which the conformational space available to the ligand is explored. The evaluation criteria involved calculating the RMSD between the hypothesis generated by the program and the target pharmacophore (generated by *Relibase+* based on the crystal structure of protein-ligand complex) and the number of misses in the hypothesis which finally gives the extent to which the hypothesis is representative of the ligands. The study showed that *GASP* and *Catalyst* outperformed *DISCO* at reproducing the five target pharmacophores. For the CDK2, DHFR and thermolysin datasets, *GASP* ranked first while *Catalyst* delivered best results for thrombin and HIV-RT datasets. In each of the five datasets, *DISCO* consistently performed the worst, giving satisfactory results with just two of the datasets (CDK2 and HIV-RT). The main difference between *DISCO* and *Catalyst* or *GASP* is that the former only uses distances between features to superimpose ligands. This might be the reason for the poor performance of *DISCO*, especially when dealing with complex target pharmacophores. It is difficult to differentiate between *GASP* and *Catalyst*; both programs have their own strengths and weaknesses.

8. PATENTING THE PHARMACOPHORES?

The answer is now YES. Though there are no reports of patents for QSAR studies, the pharmacophores are being protected under Intellectual Property Rights. The credit for the first application of a patent using such a knowledge-based concept goes to Biogen. In 1998, Biogen applied for a world patent of pharmacophore (WO 98/04913) in which all compounds derived from a 3D database search of the described pharmacophore were included. Peptor Ltd. filed a patent (US 6,343,257) that involves the process of developing a pharmacophore, its use in VS and the use of the hits to design new compounds. Another patent of a pharmacophore covers Hepatitis C NS3 protease inhibitors. This patent (WO 98/46630) claims all compounds that fit the pharmacophore model that in turn represent the structure for inhibitors of Hepatitis C NS3 protease. Another patent filed for a pharmacophore is US 2002/0013372 for the identification of CYP2D6 inhibitors.

SUMMARY AND OUTLOOK

A substantial increase in the number of target proteins is anticipated as a result of the completion of several genome projects. This opens more avenues for the application of pharmacophores in 3D searches to find new lead molecules with higher affinity. Currently, the indirect methods are being used to a great extent but an increasing number of protein structures being determined will shift the focus on the direct methods to identify (receptor-based) pharmacophores.

Pharmacophores play a key role in computer-aided drug design, especially in the absence of a receptor structure. The supremacy of pharmacophore methods for drug design and development lies in their ability to suggest a diverse set of compounds with the potential to possess a desired biological activity, but which have totally different chemical scaffolds. It must also be recognized that not all the SAR datasets have a pharmacophore, and it is essential to discover if a pharmacophore exists. Also, a major caveat associated with pharmacophore approach is that several pharmacophores may be possible within a single binding site and one pharmacophore may not describe all the possible ligands. Furthermore, it should be remembered that a pharmacophore is a necessary but insufficient condition for the ligand to interact at the receptor site and other factors like transport properties and size must also be considered.

ACKNOWLEDGEMENTS

The Department of Science and Technology (DST), New Delhi is acknowledged for their support through the FIST program (SR/FST/LS1–083/2003). Santosh Khedkar and Alpeshkumar Malde thank the Council of Scientific and Industrial Research (CSIR), New Delhi for financial support.

REFERENCES

[1] Dean, N. M. *Curr. Opin. Biotechnol.*, **2001**, *12*, 622-25.
[2] Walters, W. P.; Murcko, M. A. *Adv. Drug Deliv. Rev.*, **2002**, *54*, 255-71.
[3] Carr, R.; Jhoti, H. *Drug Discov. Today*, **2002**, *7*, 522-27.
[4] Hertzberg, R. P.; Pope, A. J. *Curr. Opin. Chem. Biol.*, **2000**, *4*, 445-51.
[5] Walters, W. P.; Stahl, M. T.; Murcko, M. A. *Drug Discov. Today*, **1998**, *3*, 160-78.
[6] Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. *ACS Symp Ser.*, **1999**, *719*, 226-238.
[7] Joseph-McCarthy, D.; Alvarez, J. C. *Proteins*, **2003**, *51*, 189-202.
[8] Mestres, J.; Knegtel, R. M. A. *Perspect. Drug Des. Discov.*, **2000**, *20*, 191-207.
[9] Mason, J. S., Good, A. C.; Martin, E. J. *Curr. Pharm. Des.*, **2001**, *7*, 567-97.
[10] Srinivasan, J.; Castellino, A.; Bradley, E. K.; Eksterowicz, J. E.; Grootenhuis, P. D. J.; Putta, S.; Stanton, R. V. *J. Med. Chem.*, **2002**, *45*, 2494-2500.
[11] Ehrlich, P. *Chem. Ber.*, **1909**, *42*, 17-47.
[12] Günd, P. In *Pharmacophore Perception, Development and Use in Drug Design*, O. F. Güner, ed.; UIL Biotechnology Series, La Jolla, **2000**, pp. 3-12.
[13] Günd, P.; Wipke, W. T.; Langridge, R. In *Proc. Intl. Conf. on Comp. in Chem. Res. and Edu.* Elsevier, 1974, Amsterdam, pp. 33–39.
[14] Pearlman, R. S. *Chem. Des. Aut. News*, **1987**, *2*, 1-6.
[15] Sadowski, J. *J. Comput.-Aided Mol. Des.*, **1997**, *11*, 53-60.
[16] Hiller, C.; Gasteiger, J. In *Software-Entwicklung in der Chemie*, J. Gasteiger, Ed., Springer, **1987**, Berlin, pp. 53-66.
[17] Wipke, W. T.; Hahn, M. A. *Tetrahedron Comput. Meth.*, **1988**, *1*, 141-67.
[18] Dolata, P. D.; Leach, A. R.; Prout, K. *J. Comput.-Aided Mol. Des.*, **1987**, *1*, 73-86.
[19] van Drie, J. H.; Weininger, D.; Martin, Y. C. *J. Comput.-Aided Mol. Des.*, **1989**, *3*, 225-51.
[20] Sheridan, R. P.; Nilakantan, R.; Rusinko, A. III.; Bauman, N.; Haraki, K.S.; Venkataraghavan, R. *J. Chem. Inf. Comput. Sci.*, **1989**, *29*, 255-60.
[21] Fisanick, W.; Cross, K. P.; Rusinko, A. III *J. Chem. Inf. Comput. Sci.*, **1992**, *32*, 664-74.
[22] Kuntz, I. D. *Science*, **1992**, *257*, 1078-82.
[23] Bartlett, P.; Shea, G. T.; Telfer, S. J.; Waterman, S. In *Molecular Recognition: Chemical and Biological Problems*, S. M. Roberts; S. V. Ley; M. M. Campbell, Eds., Cambridge, Royal Society in Chemistry, **1989**, pp. 182-96.
[24] Beusen, D.D., G.R. Marshall In *Pharmacophore Perception, Development, and Use in Drug Design*, Güner, O. F. Eds., **2000**, International University Line, La Jolla, CA., pp. 21-45.
[25] Mayer, D.; Naylor, C. B.; Motoc, I.; Marshall, G. R. *J. Comput.-Aided Mol. Des.*, **1987**, *1*, 3-16.
[26] Kuster, D. J.; Marshall, G. R. *J. Comput.-Aided Mol. Des.*, **2005**, *19*, 609-615.
[27] Christie, B. D.; Henry, D. R.; Güner, O. F.; Moock, T. E. In *Online Information '90, 14th International Online Information Meeting Proceedings*, D. I. Raitt, Ed., Learned Information, Oxford, **1990**, pp. 137-61.
[28] Murrall, N. W.; Davies, E. K. *J. Chem. Inf. Comput. Sci.*, **1990**, *30*, 312-16.
[29] Hurst, T. *J. Chem. Inf. Comput. Sci.*, **1994**, *34*, 190-96.
[30] Sprague, P. W. In *De Novo Design*, K. Müller, Ed., ESCOM Science Publishers, **1995**, Leiden, pp. 1-20.
[31] Martin, Y. C. In *Classical and 3D QSAR in Agrochemistry*, C. Hansch, T. Fujita, Eds., American Chemical Society, Washington D.C., **1995**, pp. 318-29.
[32] Barnum, D.; Green, J.; Smellie, A.; Sprague, P. *J. Chem. Inf. Comput. Sci.*, **1996**, *36*, 563-71.
[33] Jones, G.; Willett, P.; Glen, R. C. *J. Comput.-Aided Mol. Des.*, **1995**, *9*, 532-49.
[34] Cramer, R. D. III.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.*, **1988**, *110*, 5959-67.
[35] Golender, V. E.; Vorpagel, E. R. In *Drug Design: Theory, Methods, and Applications*, H. Kubinyi, Ed., ESCOM, Leiden, **1993**.
[36] Li, H.; Sutter, J.; Hoffmann, R. In *Pharmacophore perception, development, and use in drug design*, O. F. Güner, Ed., International University Line (IUL), Biotechnology Series, La Jolla, **2000**, pp 171-89.
[37] Kurogi, Y.; Güner, O. F. *Curr. Med. Chem.*, **2001**, *8*, 1035-55.
[38] Güner, O. F. *Curr. Topics Med. Chem.*, **2002**, *2*, 1321-32.
[39] van Drie, J. H. *Curr. Pharm. Des.*, **2003**, *9*, 1649-64.
[40] Dror, O.; Shulman-Peleg, A.; Nussinov, R.; Wolfson, H. J. *Curr. Med. Chem.*, **2004**, *11*, 71-90.
[41] Glennon, R. A.; Dukat, M.; Liao, L. *Curr. Top Med. Chem.*, **2004**, *4*, 631-44.
[42] Martin, Y. C. *J. Med. Chem.*, **1992**, *35*, 2145-54.
[43] van Drie, J.; Weininger, D.; Martin, Y. *J. Comput.-Aided Mol. Des.*, **1989**, *3*, 225-51.
[44] Mills, J. E. J.; de Esch, I. J. P.; Perkins, T. D. J.; Dean, P. M. *J. Comput.-Aided Mol. Des.*, **2001**, *15*, 81-96.
[45] Harris, D. L.; Loew, G. H. *Bioorg. Med. Chem.*, **2000**, *8*, 2527-38.
[46] Mestres, J.; Rohrer, D. C.; Maggiora, G. M. *J. Comput. Chem.*, **1997**, *18*, 934-54.
[47] Holliday, J. D.; Willett, P. *J. Mol. Graph. Model.*, **1997**, *15*, 221-32.
[48] Kamat, S. A.; Coutinho, E. C. *J. Biosci.*, **1996**, *21*, 599 -611.
[49] Finn, P. W.; Kavraki, L. E.; Latombe, J. C.; Motwani, R.; Shelton, C.; Venkatasubramanian, S.; Yao, A. *Computational geometry: theory and applications*, **1998**, *10*, 263-70.
[50] Dolata, D.; Parrill, A.; Walters, W. *SAR QSAR Environ. Res.*, **1998**, *9*, 53-81.
[51] Takahashi, Y.; Satoh, Y.; Suzuki, H.; Sasaki, S. *Anal. Sci.*, **1987**, *3*, 23-28.
[52] Brint, A. T.; Willett, P. *J. Chem. Inf. Comput. Sci.*, **1987**, *27*, 152-58.
[53] Handschuh, S.; Wagener, M.; Gasteiger, J. *J. Chem. Inf. Comput. Sci*. **1997**, *38*, 220-32.
[54] Dror, O.; Shulman-Peleg, A.; Nussinov, R.; Wolfson, H. *Curr. Med. Chem.*, **2004**, *11*, 71-90.
[55] Handschuh, S.; Wagener, M.; Gasteiger, J. *J. Mol. Model.*, **2000**, *6*, 358-78.
[56] Bron, C.; Kerbosch, J. *Commun. ACM*, **1973**, *16*, 575-77.
[57] Chen, X.; Rusinko, A III; Trophsa, A.; Young, S. *J. Chem. Inf. Comput. Sci.*, **1999**, *39*, 887-96.
[58] Bohm, H. J. *J. Comput.-Aided Mol. Des.*, **1994**, *8*, 243-56.
[59] Bohm, H. J. *J. Comput.-Aided Mol. Des.*, **1992**, *6*, 61-78.
[60] Goodford, P. J. *J.Med. Chem.*, **1985**, *28*, 849-57.
[61] Tanimoto, T. T. *An Elementary Mathematical Theory of Classification and Prediction*, IBM Internal Report, **1958**.
[62] MDL Information Systems, http://www.mdli.com.
[63] Cambridge Structural Database, http://www.ccdc.cam.ac.uk.
[64] World Drug Index, http://www.derwent.com
[65] Maybridge database, http://www.maybridge.com
[66] NCI database, http://cactus.nci.nih.gov
[67] PubChem database, http://cactus.nci.nih.gov
[68] Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Model.*, **2005**, *45*, 177-82; http://blaster.docking.org/zinc
[69] http://www.chemdiv.com
[70] Lipinski, C. A. *Adv. Drug Discov. Res.*, **1997**, *23*, 3-25.
[71] Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A. *Drug Discov. Today*, **2003**, *8*, 876-77.
[72] Sadowski, J. *J. Comput.-Aided Mol. Des.*, **1997**, *11*, 53-60.
[73] Tripos Associates Inc., 1699 S. Hanley Rd., St. Louis, MO 631444, USA.; http://www.tripos.com
[74] Accelrys Inc., San Diego, CA, USA; http://www.accelrys.com
[75] Tschinke, V.; Cohen, N. C. *J. Med.Chem.* **1993**, *36*, 3863-70.
[76] VanDrie, J. H., *J. Chem. Info. Comput. Sci.*, **1997**, *37*, 38-42.
[77] Kamano, Y.; Kotake, A.; Hashima, H.; Inoue, M.; Morita, H.; Takeya, K.; Itokawa, H.; Nandachi, N.; Segawa, T.; Yukita A.; Saitou, K.; Katsuyama, M.; Pettit, G. R. *Bioorg. Med. Chem.*, **1998**, *6*, 1103-15.
[78] Marriot, D. P.; Dougall, I. G.; Meghani, P.; Liu, Y.; Flower, D. *J. Med. Chem.*, **1999**, *42*, 3210-16.
[79] Fossa, P.; Boggia, R.; Mosti, L. *J. Compu.t-Aided Mol. Des.*, **1998**, *12*, 361-72.
[80] Fossa, P.; Boggia, R.; Mosti, L. *Quant. Struct.-Act. Relat.*, **2001**, *20*, 17-22.
[81] Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.*, **1997**, *267*, 727-48.

[82] Pajeva, I. K.; Wiese, M. *J. Med. Chem.*, **2002**, *45*, 5671-86.
[83] Boehm, H. J.; Boehringer, M.; Bur, D.; Gmuender, H.; Huber, W.; Klaus, W.; Kostrewa, D.; Kuehne, H.; Luebbers, T.; Meunier-Keller, N.; Mueller, F. *J. Med. Chem.*, **2000**, *43*, 2664-74.
[84] Laggner, C.; Schieferer, C.; Fiechtner, B.; Poles, G.; Hoffmann, R. D.; Glossmann, H.; Langer, T.; Moebius, F. F. *J. Med. Chem.*, **2005**, *48*, 4754-64.
[85] Patel, Y.; Gillet, V. J.; Bravi, G.; Leach, A. R. *J. Comput.-Aided Mol. Des.*, **2002**, *16*, 653-81.

Received: 16 June, 2006 **Revised: 22 December, 2006** **Accepted: 08 January, 2007**