

In Silico Screening of Ligand Databases: Methods and Applications

S. A. KHEDKAR, A. K. MALDE AND E. C. COUTINHO*

Department of Pharmaceutical Chemistry, Bombay College of Pharmacy, Kalina, Santacruz (East), Mumbai-400 098, India.

Pressure is mounting on the pharmaceutical industry to reduce both the cost of drugs and the time to market. The large number of targets made available in the last decade has created a new area for technologies that can rapidly identify quality lead candidates. Virtual screening is one such technology that is gaining increasing importance in the drug discovery process. Virtual screening is a reliable and inexpensive method currently being employed as a complementary approach to high-throughput screening. Virtual screening can be adopted irrespective of the structural information of the target receptor. In the absence of structural data, virtual screening using pharmacophore-based search is a major *in silico* tool. However, when the structure of the receptor is available, virtual screening using both pharmacophore-based and docking techniques can be employed. Both of these methods can be synergistically integrated to improve the drug design and development process. In this article, we provide an overview of methods for virtual screening – in particular, docking and pharmacophore-based – along with commercial algorithms implementing these methods, and a successful example in this arena. Further, we enumerate the potential for patenting such kind of studies.

The pharmaceutical industry is under ever increasing pressure to increase its success rate in bringing drugs to the market. Current efforts within the industry are directed at reducing the hit-to-drug timeline and increasing the number of quality candidate drugs that make the transition from discovery to the clinical phase. Enormous advances in genomics have resulted in large increase in the number of potential therapeutic targets. This growth in potential targets has increased the demand for reliable target validation, as well as technologies that can identify rapidly several quality lead candidates¹⁻⁴. Virtual screening (VS) technology is gaining increasing importance in the discovery process because it is a reliable and inexpensive method for identifying lead molecules⁵. It is seen as a complementary approach to high-throughput screening (HTS) and when coupled with structural biology, it enhances the success rate of identification of leads. Further, the advances in computational techniques have enabled VS to make a deep impact on the drug discovery process.

The tools for carrying out VS can be broadly categorized as receptor-based or ligand-based. The ligand-based methods use information provided by a

compound or a set of compounds that are known to bind to the desired target, and use this knowledge to identify other compounds with similar properties in databases. This is usually achieved by similarity and substructure searching⁶ or pharmacophore matching⁷ or 3D shape matching⁸. However, when the structure of the receptor is available, both pharmacophore-based and docking techniques can be employed. The latter method involves positioning each ligand into the binding site of the target, thus proposing a binding mode and affinity for each compound in the database. This information is then used to rank the compounds in order to select and experimentally test a small subset of hits for biological activity. In particular, both these methods can be synergistically integrated to improve the drug design and development process.

This article is intended to provide an overview of methods, in particular docking and pharmacophore searching, along with commercial and non-commercial software for *in silico* screening of ligand databases and some successful examples in this arena.

DATABASE STRUCTURE

The primary requirement for performing VS, by either

*For correspondence

E-mail: evans@bcpindia.org

docking or by pharmacophore search method, is the availability of a database (DB) of ligand molecules. A corporate database or one available from commercial vendors such as Available Chemicals Directory (ACD)⁹ or Cambridge Structural Database (CSD)¹⁰ can be employed for this purpose. Additionally, a database of known reagents and compounds which are readily synthesizable can be used for VS after primary filtration for 'drug-like' properties² such as variation of Lipinski's rule-of-five¹¹, number of rotatable bonds or the polar surface area. 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 should be assigned to 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.

For the generation of the 3D structure of a molecule during compilation of a database, one can use programs such as Corina¹² (Molecular Networks GmbH, <http://www.mol-net.com>), Concord and Confort (Tripos Inc., USA) or Converter (Accelrys Inc., USA). Sometimes it is necessary to assign partial charges to compounds in the database depending on the requirement of the 3D search method.

Another issue in VS is the handling of ligand flexibility. Conformational flexibility can be introduced in VS either through the database or the query, or during the search process. Handling conformational flexibility in the database involves storage of multiple conformations of each ligand in the DB. However, this is a practically intractable solution because the number of conformations of a ligand increases with an increase in the number of rotatable bonds. A small drug-like molecule with four rotatable bonds, if scanned at a resolution of 120° for each dihedral angle, yields $[360/120]^4 = 81$ conformations, and there is no firm basis for the ones to be included in the DB. A better solution is the hybrid approach. The combination of multi-conformer database and flexible searching provides an efficient and effective route. This hybrid approach is employed in the Catalyst BEST search method (Accelrys Inc., USA).

VS USING DOCKING

The docking process involves the simulation of molecular recognition events using the 3D structure of the receptor (obtained by X-ray crystallography or NMR or a

theoretical method such as comparative modelling). The 3D structure of the receptor is corrected and transformed into an appropriate format depending on the requirements of the docking program. The docking method then samples the conformational space of the binding site and scores each possible ligand pose, which is then taken as the predicted binding mode of that ligand. The ligand and receptor flexibility is an important aspect in VS because the conformations of the ligand, as well as the receptor binding site residues, might be different from their conformation in the bound state. The aspects of ligand and receptor flexibility are tackled to some extent in some of the docking programs (flexible search) or by pre-computing a DB of several conformers of each compound to be screened, as mentioned in the previous section. A majority of the docking programs explore ligand flexibility through a variety of algorithms. One way to explore the ligand's conformational space is to use Monte Carlo or simulated annealing methods^{13,14}, but these are often time consuming and not suitable for large-scale VS. The Genetic Optimization of Ligand Docking (GOLD¹⁵, Cambridge Crystallographic Data Center, UK; <http://www.ccdc.cam.ac.uk>) and the most recent version of AutoDock¹⁶ (Scripps Research Institute, USA) use genetic algorithm-based approaches to generate and select the best conformations of a ligand. Flex-X¹⁷ (Tripos Inc., USA) has a fragment-based approach and uses a set of low energy torsion angles derived from the CSD for each single bond. GOLD also enables the customization of torsional energy contributions *via* a statistical analysis of torsion angle distributions observed in CSD. The program Glide (Schrödinger Inc., USA; <http://www.schrodinger.com>) treats ligand flexibility by an exhaustive enumeration of the rotamer states for each rotatable bond coupled to a heuristic screen that rapidly eliminates conformations deemed unsuitable for binding to a receptor.

Majority of the docking methods treat the receptor as a rigid entity, which is an inaccurate but necessary approximation in order to limit the complexity and consequently the computational cost required to accurately sample the flexibility of the binding site. Researchers have used an ensemble of protein structures to incorporate protein flexibility. This ensemble of protein structures can be obtained by simulated annealing of the binding site or by use of several high resolution structures of the same target bound to different ligands or in *apo* form, or by use of an ensemble of NMR structures, or by use of multiple homology models. The programs AutoDock¹⁶ and Flex-E¹⁸ (Tripos Inc., USA)

incorporate protein mobility through rotamer libraries for flexible side chains and by use of an ensemble of protein structures to generate Boltzman-weighted grids with which the docking function is generated¹⁹.

The next important step that determines the success of the docking process is ranking the quality of different poses of the same molecule and then with respect to other molecules in the DB (Scoring). There are several scoring functions²⁰ available to achieve this task. Force field-based scoring functions such as AMBER²¹, CHARMM²² and CVFF²³ are based on atomic force fields in which single-point intermolecular interaction energy between the ligand and the receptor is used as a score for the given pose. Empirical scoring functions are based on an additive approximation of physicochemical properties such as H-bonding, hydrophobic interactions, entropic changes and metal-ion interactions, if present. Experimental binding energies of known ligand-receptor complexes are used to derive the coefficients of each term explaining a binding event. Therefore, the efficiency of these scoring functions depends on the variety and number of ligand-receptor complexes used to derive the scoring function and its validation on a large set of ligand-receptor complexes in the PDB. GOLD²⁴, FlexX¹⁷, SCORE²⁵, VALIDATE²⁶, ChemScore²⁷, Ludi²⁸ (Accelrys Inc., USA) and PLP²⁹ use an empirical scoring function to rank different poses. The GOLD scoring function has been validated on a large data set³⁰ and uses a genetic algorithm to rank the poses, and it is considered to be a reliable docking and scoring algorithm for VS. Another class of scoring functions, knowledge-based, is derived from the structures of ligand-receptor complexes using statistical mechanics. In contrast to empirical scoring functions, they do not require binding affinity data of ligand-receptor complexes and so are free to use information in ligand-receptor complexes deposited in the public or proprietary databases. Further, such scoring functions are expected to be less biased and more readily transferable to systems not included in the development of the scoring function. PMF³¹, Bleep³², SMOG³³ and DrugScore³⁴ (Cambridge Crystallographic Data Center, CCDC, UK) are some examples of knowledge-based scoring functions. A flow chart of the steps involved in VS by docking method is depicted in fig. 1.

At the end of the docking process, a set of best compounds is selected. It is not advisable to select the compounds merely on the basis of scoring or 'best ranking'. A recent study³⁵ that evaluated the performance

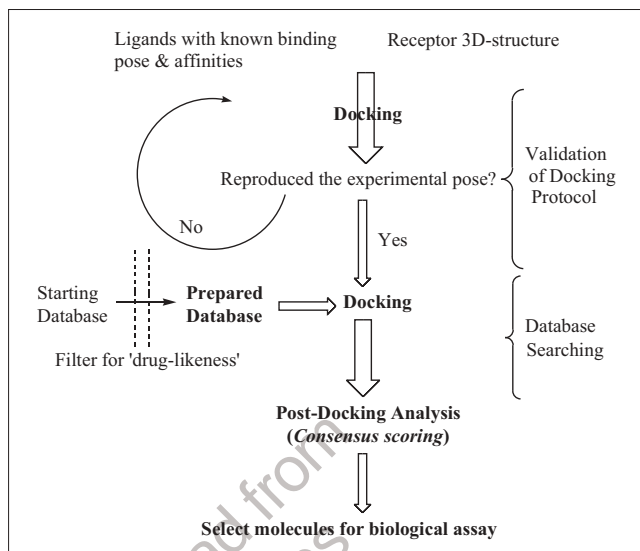


Fig. 1: Flow chart of the virtual screening process using docking method

of several scoring functions such as PLP, ChemScore and Dock found little or no correlation between predicted and experimental binding affinities. Therefore, it becomes crucial to use a post-analysis strategy to minimize the false positives in the selection list. In consensus scoring strategy³⁶⁻³⁸, a docking function is used to generate top-ranked poses and then multiple scoring functions are used to score these poses. Then, only the top-ranked compounds common to each scoring method are chosen for biological evaluation. It has been observed that almost all docking methods are able to identify the correct binding pose³⁹ but usually cannot rank it as the top-ranked binding pose for the ligand. Therefore, it is prudent to use consensus scoring of multiple poses rather than a top-ranked single pose.

Based on the above known limitations, it is strongly advisable to use a set of ligands with known binding modes and affinities, and multiple docking and scoring (consensus) algorithms, to calibrate the docking methodology for the target under study before undertaking VS of large databases.

VS USING PHARMACOPHORE METHODS

The docking methodologies discussed above help in screening databases only when the 3D structure of the target receptor is available. In the early phase of drug discovery, or in cases where the structure elucidation is beyond the capabilities of present day techniques, the docking technique cannot be used. A pharmacophore-based

search of databases can be carried out in absence of the receptor structure with the help of the information available from a set of ligands binding preferentially to the given receptor. The credit for first conceiving the concept of the pharmacophore goes to Paul Ehrlich, who devised a way (in early 1900s) to develop dyes through chromophores (i.e., the part of a molecule responsible for imparting colour)⁴⁰. A good review of the evolution and history of the pharmacophore concept has been published by Peter Gund⁴¹, who is also the author of the modern definition of a pharmacophore - "a set of structural features in a molecule that is recognized at a receptor site and is responsible for that molecule's biological activity"^{42,43}. The concept could not achieve its full utilization until the development of 3D database searching software in 1990. MOLPAT⁴⁴, the first computer program, was developed at Princeton University in 1974 by Gund, Wipke and Langridge to recognize pharmacophore patterns. The demand for 3D searching software reached a critical level with the development of rapid 3D structure generation programs such as CONCORD⁴⁵, CORINA^{12,46}, AIMB⁴⁷ and WIZARD⁴⁸. ALADDIN⁴⁹ (developed at Abbott and later commercialized by Daylight) and 3D-Search⁵⁰ (developed at Lederle) were developed by pharmaceutical companies, whereas academic and government institutions were responsible for CAST-3D⁵¹ (Chemical Abstract Services), DOCK⁵² (University of California at San Francisco) and CAVEAT⁵³ (University of California at Berkley). The first commercial 3D searching system, MACCS-3D⁵⁴, was developed by Güner *et al.* and was released in December of 1989. During the next 4 y, all of the technology that is available today was developed - ChemDBS-3D⁵⁵ (Accelrys Inc., USA), UNITY⁵⁶ (Tripos Inc., USA) and Catalyst⁵⁷ (Accelrys Inc., USA). The critical demand for the pharmacophore recognition software arose when the above-mentioned 3D searching technologies were made easily available. Though these 3D-search software had inbuilt query-generation tools, specialized pharmacophore generation software were also being developed. Most notable among them are DISCO⁵⁸ by Martin *et al.* (Tripos Inc., USA), HipHop⁵⁹ by Barnum *et al.* (Accelrys Inc., USA) and GASP⁶⁰ by Jones and Willett (Tripos Inc., USA). Meanwhile, predictive pharmacophore models such as CoMFA⁶¹ by Cramer *et al.* (Tripos Inc., USA), Apex-3D⁶² by Golander and Vorpagel (Accelrys Inc., USA) and HypoGen⁶³ by Teig *et al.* (Accelrys Inc., USA) were being developed. Detailed usage and validation of all of the pharmacophore development software are covered

elsewhere⁴¹. For timely reviews in the field, the reader may follow other references⁶⁴⁻⁶⁸.

Pharmacophore modelling provides a useful framework for better understanding of the existing data and can be used as a predictive 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 analogues. The pharmacophore can be identified by direct method (using a receptor-ligand complex) or by indirect method (using only a collection of ligands that have experimentally been observed to interact with a given receptor). However, direct methods are becoming extremely important because of the increase in the rate at which protein structures are being determined. A flow chart of steps involved in VS by the pharmacophore method is depicted in fig. 2.

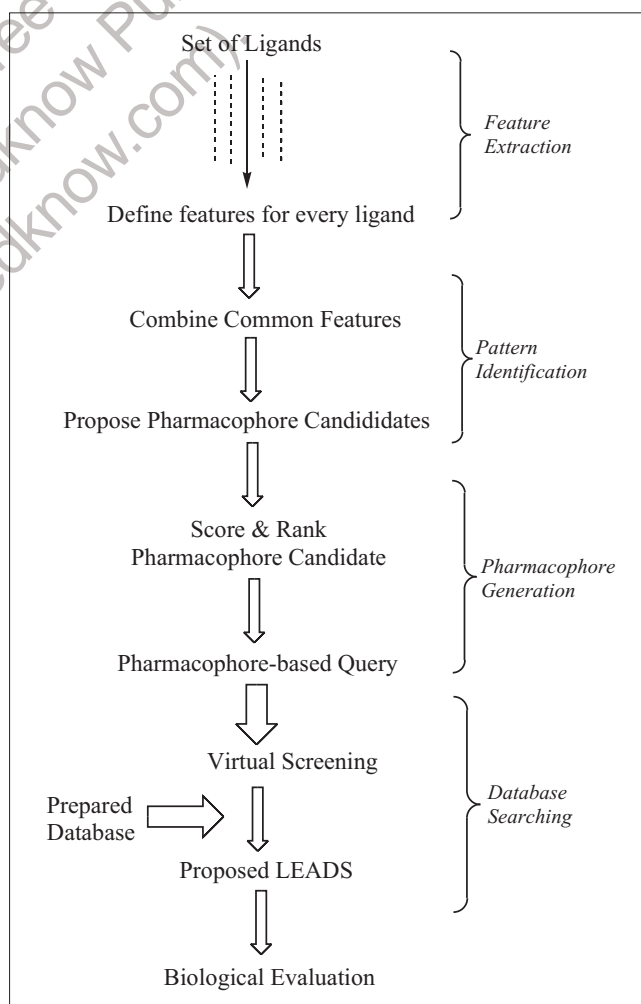


Fig. 2: Flow chart of the virtual screening process using pharmacophore method

Pharmacophore generation:

Pharmacophore generation starts with the selection of a set of ligands for which the pharmacophore model is to be constructed. The type of ligand molecules, the size of the dataset and its diversity have a great impact on the resulting pharmacophore model⁶⁷. The Carnell-Smith method⁶⁹, RAPID⁷⁰ and HipHop^{59,64,71} assume that ligands have the same activity and thus do not consider activity data. The information of inactive ligands can also be used to indicate structural features that significantly decrease the activity. The programs that use information of inactive ligands are CLEW⁷² and the current version of DISCO⁵⁸. Pharmacophore models can also be used to predict the activity of unknown compounds. For deriving such models, HypoGen⁶³ utilizes a large enough set of diverse compounds (about 18-30) with different activity levels (4-5 orders of magnitude). Most of the currently available methods such as HipHop, HypoGen, MPHIL⁷³ and RAPID are designed to handle small (less than 100 ligands) data sets. Other methods use larger data sets as input but convert them into a smaller one by sorting the activities of the ligands, depending on a user-specified cut-off. Lastly, one should remember that though the dataset should be as diverse as possible to get an accurate pharmacophore model, very different ligands may bind at different binding sites, resulting in a misleading pharmacophore model⁷⁴.

In the next step, the features relevant to the pharmacophore model are extracted from the input ligands (feature extraction). 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 (3D position of an atom and its type)⁶⁷. Both topology-based and function-based features have some drawbacks. For example, hydroxyl oxygen can be classified as both an H-bond donor and acceptor. A simple way to represent a functional group is by its centre. The centre of an acid, base, H-bond donor/acceptor is usually defined as the position of an actual atom. For a hydrophobic region or an aromatic ring, the centre is defined as the centroid of the group. Furthermore, 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^{71,75}. In addition, a hydrophobic group can be represented by a sphere and an aromatic ring by a plane and its normal.

From each ligand structure, the selected features are

combined to form a representation of the whole structure. The RAPID method represents ligand structures as a set of labelled points in three-dimensional space, where each point is associated with a feature⁷⁰. Another approach by Takahashi *et al.*⁷⁶ represents a ligand structure through a labelled graph, with nodes representing the features and the edges representing the relations. Different methods that use this representation are given in the literature⁷⁷. In another approach, a ligand structure is considered a set of labelled points, together with the associated interpoint distances⁷⁷. This type of representation is orientation independent, in contrast to the 3D point-set representation.

The features extracted from different ligand molecules are matched and pharmacophore candidates are proposed. This is called pattern identification. A pattern or configuration is a set of relative locations in 3D space, each associated with a feature. 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). The methods that have adopted this approach are DISCO⁵⁸, RAPID⁷⁰, GAMMA⁷⁸ and GASP⁶⁰. This approach assumes that a common pharmacophore is responsible for the observed activity. This is an inaccurate assumption in the MCS approach and can be overcome by relaxing the requirements ("Relax MCS" approach) that all input ligands possess all the features. This philosophy is used in the MPHIL⁷⁴ method.

Algorithms for pattern identification have also been developed. (For a detailed description, see reference 67). The Clique detection algorithm⁷⁹⁻⁸¹ has been implemented in DISCO and MPHIL methods. Clique detection has its origin in 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 also present in every other graph in the set. HipHop and SCAMPI use an exhaustive search algorithm in which the search for a pattern starts with small sets of features and extends until no larger common pattern exists. HypoGen uses a similar approach but also incorporates activity information into the pharmacophore derivation process. This is done in three steps: the constructive stage identifies pharmacophore candidates that are common among the set of most active ligands, followed by the subtractive stage in which those pharmacophore candidates identified in the constructive stage and 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 annealing⁶⁷. GASP⁶⁰ and GAMMA⁷⁸ are based on the genetic algorithm⁸² and also perform a conformational search as part of the GA run. Thus, molecular flexibility is simulated by generating multiple conformations of the given ligand followed by an exploration of different 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 is a result of chance correlation. A detailed description of scoring methods implemented in techniques that relax the MCS requirements, and scoring in genetic algorithms like GASP and GAMMA is given in the literature^{59,67}.

As mentioned in the introductory section, VS by pharmacophore searching is more efficient when structural knowledge of the target receptor is available. The receptor-based approach for pharmacophore generation involves analyzing features of the active site and the spatial relationship among them, and then an active image of this is used to construct the pharmacophore model. Such an operation 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 begins with a 3D structure of the receptor (usually in PDB format) and a set of ligands with known activity. Using the information of the active site residues (from biochemical or structural studies), a program such as Ludi^{25,28} (Accelrys Inc., USA) generates an interaction map which is a complement of the receptor-binding site. The ligands with known activity are used to identify the functional features such as H-bond donors/acceptors and lipophilic groups from the interaction map in the active site. Often it is necessary to filter/limit the number of features since queries with multiple features may fail to retrieve any hits from the database. Therefore, 3D queries composed of fewer features are generated by considering all possible combinations. Catalyst⁵⁷ uses these queries to search the ligand database.

SUCCESSFUL APPLICATION OF VS METHODS

DNA gyrase inhibitors:

A 3D database search using Ludi and Catalyst was

performed to find novel scaffolds of DNA gyrase inhibitors⁸³. The ACD database of 350,000 compounds and a portion of the Roche compound inventory (RCI) were screened to obtain 150 weak inhibitors. These novel DNA gyrase inhibitors binding to the ATP-binding site have seven structurally different scaffolds. The optimization of the indazole scaffold provided a DNA gyrase inhibitor which was 10 times more potent than novobiocin.

Sterol metabolism:

Laggner and colleagues⁸⁴ have reported pharmacophore models for three protein targets involved in sterol metabolism. Twenty-three structurally diverse molecules with binding affinity data for EBP (emopamil binding protein), ERG2 (fungal counterpart of EBP) and the 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 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 the hydrogen-bonding interactions are not required for high-affinity inhibitor binding. The models were subsequently used to search the database [such as World Drug Index (WDI), Kyoto Encyclopedia of Genes and Genomes (KEGG) and COMPOUND database]. In virtual screening, 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.

CAN ONE PROTECT KNOWLEDGE-BASED CONCEPTS?

The answer is now YES. Though there are no patents yet of QSAR studies, the pharmacophores are being protected under Intellectual Property Rights. The credit for the first application of patenting such a knowledge-based concept goes to Biogen. In 1998, Biogen applied for a world patent of pharmacophores (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 development of a pharmacophore, its use in VS and use of the hits to design new compounds. Another patent of pharmacophores covers Hepatitis C NS3 protease inhibitors. This patent (WO 98/46630) claims all compounds that fit the pharmacophore model that in turn

represents the structure for inhibitors of Hepatitis C NS3 protease. Another patent filed for pharmacophores is US 2002/0013372 for the identification of CYP2D6 inhibitors.

ACKNOWLEDGEMENTS

The support provided by the Department of Science and Technology (DST), New Delhi, through the FIST program (SR/FST/LS1-083/2003) is acknowledged with thanks; and the authors SAK and AKM thank the Council of Scientific and Industrial Research (CSIR), New Delhi, for Research Fellowships.

REFERENCES

- Dean, N.M., *Curr. Opin. Biotechnol.*, 2001, 12, 622.
- Walters, W.P. and Murcko, M.A., *Adv. Drug Deliv. Rev.*, 2002, 54, 255.
- Carr, R. and Jhoti, H., *Drug Discov. Today*, 2002, 7, 522.
- Hertzberg, R.P. and Pope, A.J., *Curr. Opin. Chem. Biol.*, 2000, 4, 445.
- Walters, W.P., Stahl, M.T. and Murcko, M.A., *Drug Discov. Today*, 1998, 3, 160.
- Mestres, J. and Knegtel, R.M.A., *Perspect. Drug Des. Discov.*, 2000, 20, 191.
- Mason, J.S., Good, A.C. and Martin, E.J., *Curr. Pharm. Des.*, 2001, 7, 567.
- Srinivasan, J., Castellino, A., Bradley, E.K., Eksterowicz, J.E., Grootenhuys, P.D.J., Putta, S. and Stanton, R.V., *J. Med. Chem.*, 2002, 45, 2494.
- MDL Information Systems, <http://www.mdli.com>.
- Cambridge Structural Database, <http://www.ccdc.cam.ac.uk>.
- Lipinski, C.A., *Adv. Drug Discov. Res.*, 1997, 23, 3.
- Sadowski, J., *J. Comput.-Aided Mol. Des.*, 1997, 11, 53.
- McMartin, C. and Bohacek, R.S., *J. Comput.-Aided Mol. Des.*, 1997, 11, 333.
- Liu, M. and Wang, S., *J. Comput.-Aided Mol. Des.*, 1999, 13, 435.
- Jones, G., Willett, P. and Glen, R.C., *J. Mol. Biol.*, 1995, 245, 43.
- Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J., *J. Comput. Chem.*, 1998, 19, 1639.
- Rarey, M., Kramer, B., Lengauer, T. and Klebe, G., *J. Mol. Biol.*, 1996, 261, 470.
- ClauBen, H., Buning, C., Rarey, M. and Lengauer, T., *J. Mol. Biol.*, 2001, 308, 377.
- Osterberg, F., *Proteins Struct. Funct. Genet.*, 2002, 46, 34.
- Ajay, A. and Murcko, M.A., *J. Med. Chem.*, 1995, 38, 4953.
- Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.A., *J. Am. Chem. Soc.*, 1995, 117, 5179.
- MacKerell, A.D., Bashford, D., Bellott, R.L., Dunbrack, R.L., Jr., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., III, Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D. and Karplus, M., *J. Phys. Chem. B*, 1998, 102, 3586.
- Dauber-Osguthorpe, P., Roberts, V. A., Osguthorpe, D.J., Wolff, J., Genest, M. and Hagler, A.T., *Proteins*, 1988, 4, 31.
- Jones, G., Willett, P., Glen, R.C., Leach, A.R. and Taylor, R., *J. Mol. Biol.*, 1997, 267, 727.
- Bohm, H.J., *J. Comput.-Aided Mol. Des.*, 1994, 8, 243.
- Head, R.D., Smythe, M.L., Oprea, T.I., Waller, C.L., Green, S.M. and Marshall, G.R., *J. Am. Chem. Soc.*, 1996, 118, 3959.
- Eldridge, M.D., Murray, C. W., Auton, T.R., Paolini, G.V. and Mee, R.P., *J. Comput.-Aided Mol. Des.*, 1997, 11, 425.
- Bohm, H.J., *J. Comput.-Aided Mol. Des.*, 1992, 6, 61.
- Gehlhaar, D.K., Verkhivker, G.M., Rejto, P.A., Sherman, C.J., Fogel, D.R., Fogel, L.J. and Freer, S.T., *Chem. Biol.*, 1995, 2, 317.
- Nissink, J.W.M., Murray, C., Hartshorn, M., Verdonk, M.L., Cole, J.C. and Taylor, R., *Proteins*, 2002, 49, 457.
- Muegge, I., *Perspect. Drug Des. Discov.*, 2000, 20, 99.
- Mitchell, J.B.O., Laskowski, R.A., Alex, A., Forster, M.J. and Thornton, J.M., *J. Comp. Chem.*, 1999, 20, 1177.
- Ishchenko, A.V. and Shakhnovich, E.I., *J. Med. Chem.*, 2002, 45, 2770.
- Gohlke, H., Hendlich, M. and Klebe, G., *J. Mol. Biol.*, 2000, 295, 337.
- Pearlman, D.A. and Charifson, P.S., *J. Med. Chem.*, 2001, 44, 3417.
- Charifson, P.S., Corkery, J.J., Murcko, M.A. and Walters, W.P., *J. Med. Chem.*, 1999, 42, 5100.
- Stahl, M. and Rarey, M., *J. Med. Chem.*, 2001, 44, 1035.
- Wang, R.X. and Wang, S.M., *J. Chem. Inf. Comput. Sci.*, 2001, 41, 1422.
- Dixon, J.S., *Proteins Struct. Funct. Genet.*, 1997, Suppl. 1, 198.
- Ehrlich, P., *Chem. Ber.*, 1909, 42, 17. Quoted by Ariens, E.J., *Progr. Drug Res.*, 1966, 10, 429.
- Gund, P., In: *Pharmacophore Perception, Development and Use in Drug Design*, Güner, O.F. Eds., International University Line (UIL) Biotechnology Series: La Jolla, 2000, 3.
- Gund, P., *Ann. Reports Med. Chem.*, 1979, 14, 299.
- Gund, P., In: *Progress in Molecular and Subcellular Biology*, vol. 5, Hahn, F.E., Eds., Springer-Verlag, Berlin, 1979, 117.
- Gund, P., Wipke, W.T. and Langridge, R., In: *Proceedings of International Conference on Computers in Chemistry Research and Education*, Elsevier, Amsterdam, 1974, 33.
- Pearlman, R.S., *Chem. Des. Aut. News*, 1987, 2, 1.
- Hiller, C. and Gasteiger, J., In: *Software-Entwicklung in der Chemie*, vol. 1, Gasteiger, J., Ed., Springer, Berlin, 1987, 53.
- Wipke, W.T. and Hahn, M.A., *Tetrahedron Comput. Meth.*, 1988, 1, 141.
- Dolata, P.D., Leach, A.R. and Prout, K., *J. Comput.-Aided Mol. Des.*, 1987, 1, 73.
- Van Drie, J.H., Weininger, D. and Martin, Y.C., *J. Comput.-Aided Mol. Des.*, 1989, 3, 225.
- Sheridan, R. P., Nilakantan, R., Rusinko, A. III., Bauman, N., Haraki, K.S., Venkataraghavan, R., *J. Chem. Inf. Comput. Sci.*, 1989, 29, 255.
- Fisanick, W., Cross, K.P., Rusinko, A. III., *J. Chem. Inf. Comput. Sci.*, 1992, 32, 664.
- Kuntz, I.D., *Science*, 1992, 257, 1078.
- Bartlett, P., Shea, G.T., Telfer, S.J. and Waterman, S., In: Roberts, S.M., Ley, S.V., and Campbell, M.M. Eds., *Molecular Recognition: Chemical and Biological Problems*, Royal Society in Chemistry, Cambridge, 1989, 182.
- Christie, B.D., Henry, D.R., Güner, O.F. and Mook, T.E., In: Raitt, D.I., Ed., *Online Information '90, 14th International Online Information Meeting Proceedings*, Learned Information, Oxford, 1990, 137.
- Murrall, N.W. and Davies, E.K., *J. Chem. Inf. Comput. Sci.*, 1990, 30, 312.
- Hurst, T., *J. Chem. Inf. Comput. Sci.*, 1994, 34, 190.
- Sprague, P.W., In: *De Novo Design*, Müller, K., Eds., ESCOM Science Publishers, Leiden, 1995, 1.
- Martin, Y.C., In: *Classical and 3D QSAR in Agrochemistry*, Hansch C. and Fujita T., Eds., American Chemical Society, Washington D.C., 1995, 318.
- Barnum, D., Green, J., Smellie, A. and Sprague, P., *J. Chem. Inf. Comput. Sci.*, 1996, 36, 563.
- Jones, G., Willett, P. and Glen, R.C., *J. Comput.-Aided Mol.*

- Des.**, 1995, 9, 532.
61. Cramer, R.D. III., Patterson, D.E. and Bunce, J.D., **J. Am. Chem. Soc.**, 1988, 110, 5959.
 62. Golender, V.E. and Vorpapel, E.R., In; Kubinyi, H., Eds., **Drug Design: Theory, Methods and Applications**, ESCOM, Leiden, 1993.
 63. Li, H., Sutter, J. and Hoffmann, R., In; Pharmacophore perception, development, and use in drug design, Güner, O.F., Ed., International University Line (IUL), Biotechnology Series, La Jolla, 2000, 171.
 64. Kurogi, Y. and Güner, O.F., **Curr. Med. Chem.**, 2001, 8, 1035.
 65. Güner, O.F., **Curr. Topics Med. Chem.**, 2002, 2, 1321.
 66. van Drie, J.H., **Curr. Pharm. Des.**, 2003, 9, 1649.
 67. Dror, O., Shulman-Peleg, A., Nussinov, R. and Wolfson, H.J., **Curr. Med. Chem.**, 2004, 11, 71.
 68. Glennon, R.A., Dukat, M. and Liao, L., **Curr. Topics Med. Chem.**, 2004, 4, 631.
 69. Crandell, C. and Smith, D., **J. Chem. Inf. Comput. Sci.**, 1983, 23, 186.
 70. Finn, P.W., Kavradi, L.E., Latombe, J.C., Motwani, R., Shelton, C., Venkatasubramanian, S. and Yao, A., **Computational Geometry: Theory and Applications**, 1998, 10, 263.
 71. Clement, O.A. and Mehl, A.T., In; Güner, O.F., Ed., Pharmacophore perception, development, and use in drug design, International University Line (IUL): La Jolla, CA, USA, 2000, 69.
 72. Dolata, D., Parrill, A. and Walters, W., **SAR and QSAR in Environmental Research**, 1998, 9, 53.
 73. Holliday, J. and Willet, P., **J. Mol. Graph. Model.**, 1997, 15, 221.
 74. Ghose, A.K., Viswanadhan, V.N. and Wendoloski, J.J.J., **Receptors and Signal Transduction**, 2001, 21, 357.
 75. Mayer, D., Naylor, C., Motoc, I. and Marshall, G.J., **J. Comput.-Aided Mol. Des.**, 1987, 1, 3.
 76. Takahashi, Y., Satoh, Y., Suzuki, H. and Sasaki, S., **Anal. Sci.**, 1987, 3, 23.
 77. Brint, A.T. and Willett, P., **J. Chem. Inf. Comput. Sci.**, 1987, 27, 152.
 78. Handschuh, S., Wagener, M. and Gasteiger, J., **J. Mol. Model.**, 2000, 6, 358.
 79. Bron, C. and Kerbosch, J., **Communications of the ACM**, 1973, 16, 575.
 80. Kuhl, F.S., Crippen, G.M. and Friesen D.K., **J. Comput. Chem.**, 1984, 5, 24.
 81. Gardiner, E.J., Artymiuk, P.J. and Willett, P., **J. Mol. Graph. Model.**, 1997, 15, 245.
 82. Jones, G., In; Schleyer, P.V.R., Allinger, N.L., Clark, T., Gasteiger, J., Kollman, P.A., Schaefer, III, H.F., Schreiner, P.R., Eds., **Encyclopedia of computational chemistry**, John Wiley & Sons, Chichester, UK, 1998, 342.
 83. Boehm, H.J., Boehringer, M., Bur, D., Gmuender, H., Huber, W., Klaus, W., Kostrewa, D., Kühne, H., Luebbbers, T., Meunier-Keller, N., Mueller, F., **J. Med. Chem.**, 2000, 43, 2664.
 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.

Accepted 2 November 2006

Revised 9 June 2006

Received 15 September 2005

Indian J. Pharm. Sci., 2006, 68 (6): 689-696