Pharmacophore Modelling in Drug Discovery and Development
Module V
B. Pharm. VI Sem.

By
Gunjan Kalyani, M.Pharm.
Assistant Professor
Columbia Institute of Pharmacy
Raipur 493111, Chhattisgarh, India
Mobile: +91-8349204583
Mail ID: kalyani.gunjan@yahoo.in

Former Research Fellow
National Center for Natural Resources (NCNR)
Pt. Ravishankar Shukla University, Raipur 492010, 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 presentation 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, some very recent successes of the widely used pharmacophore generation methods in drug discovery has been mentioned.

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.

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.

Introduction

The ligand-based pharmacophore modelling 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, pharmacophore matching or 3D shape matching.

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 presentation 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 presentation concludes with some successful examples of drug discovery based on these approaches.

History and Evolution of 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".

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 to recognize pharmacophore patterns was developed by Günd, Wipke and Langridge at Princeton University in 1974.

History and Evolution of Pharmacophore concept

The demand for 3D structure searching software grew with the development of rapid 3D structure generation programs such as CONCORD, CORINA, AIMB and WIZARD.

3D search software like ALADDIN (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 (Chemical Abstract Services), DOCK (University of California at San Franscisco) and CAVEAT (University of California at Berkley).

The first commercial 3D searching system, MACCS-3D 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 – ChemDBS3D (Chemical Design Inc., USA), UNITY (Tripos Inc., USA) and Catalyst (Accelrys Inc., USA).

History and Evolution of Pharmacophore concept

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 by Barnum et al. (Accelrys Inc., USA), and GASP by Jones and Willett (Tripos Inc., USA).

Meanwhile, predictive models based on QSAR such as CoMFA (Tripos Inc., USA) by Cramer et al., Apex-3D (Accelrys Inc., USA) by Golander and Vorpagel and HypoGen by Teig et al. (Accelrys Inc., USA) also came into existence.

Pharmacophore model or query

A pharmacophore model consists of a few features organized in a specific 3D pattern. Each feature is typically represented as a sphere (although variants exist) with a radius determining the tolerance on the deviation from the exact position.

The features can be labeled as a single feature or any logic combination consisting of "AND," "OR," and "NOT" to combine different interaction patterns within one label. Additional features can describe forbidden volume interactions (typically to represent the receptor boundary).

A pharmacophore query is comprised of different features. The features represent molecular recognition motifs such as hydrogen bond acceptors or donors, anionic, cationic, hydrophobic, and aromatic groups. The radius of the sphere determines the strictness of the geometric constraint.

For features where the correct orientation of the interaction is important such as hydrogen bonds and the aromatic plane, a second feature can be used indicating the vector of the interaction (or the normal of the plane).

Pharmacophore model or query

A pharmacophore query can combine any of these features, with different radii and logic operations such as "AND," "OR," and "NOT." On the left a hypothetical pharmacophore query for BRAF kinase is given. Such pharmacophore features are typically used as queries to screen small molecule libraries of compounds.51 In these libraries all the compounds are present in their low-energy biorelevant conformations.

Each of these conformations is fitted to the pharmacophore query by aligning the pharmacophore features of the molecule and the query is composed. If a molecule can be fitted inside the spheres representing the query features it is considered a hit molecule. Often the pharmacophore query can be too complex to find hit molecules from a given library, and partial matching may be allowed. In such cases only certain features considered essential for activity are matched. Additional uses of such models are to align molecules or facilitate molecular docking simulations.
Depending on the situation and the type of experiment, multiple strategies are available to construct pharmacophore models, either manually or using automated algorithms.

Steps of Pharmacophore modelling

1. Input
2. Conformational Search
3. Feature extraction
4. Structure representation
5. Pattern Identification
6. Scoring

Pharmacophore modelling

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).

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).

Pharmacophore modelling

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 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 Distance Comparison (DISCO, HipHop (a part of CATALYST), Genetic Algorithm Superposition Program (GASP), Chem Diverse (3 and 4-point pharmacophore generation in Chem-X), SLATE, MOLMOD, MIMIC, Mapping Pharamcophores In Ligands (MPHIL), Dynamic Pharmacophore approach using molecular dynamics and receptor guided approaches. DISCO, Catalyst and GASP are widely used for pharmacophore identification.

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 type of ligand molecules, the size of the dataset and its chemical diversity affect the final pharmacophore model. The Carnell Smith method, RAPID and HipHop do not take into consideration the activity data of molecules.

CLEW and the current version of DISCO 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 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, HypoGen, MPHIL and RAPID are designed to handle small (less than 100 ligands) data sets.

Pharmacophore Generation

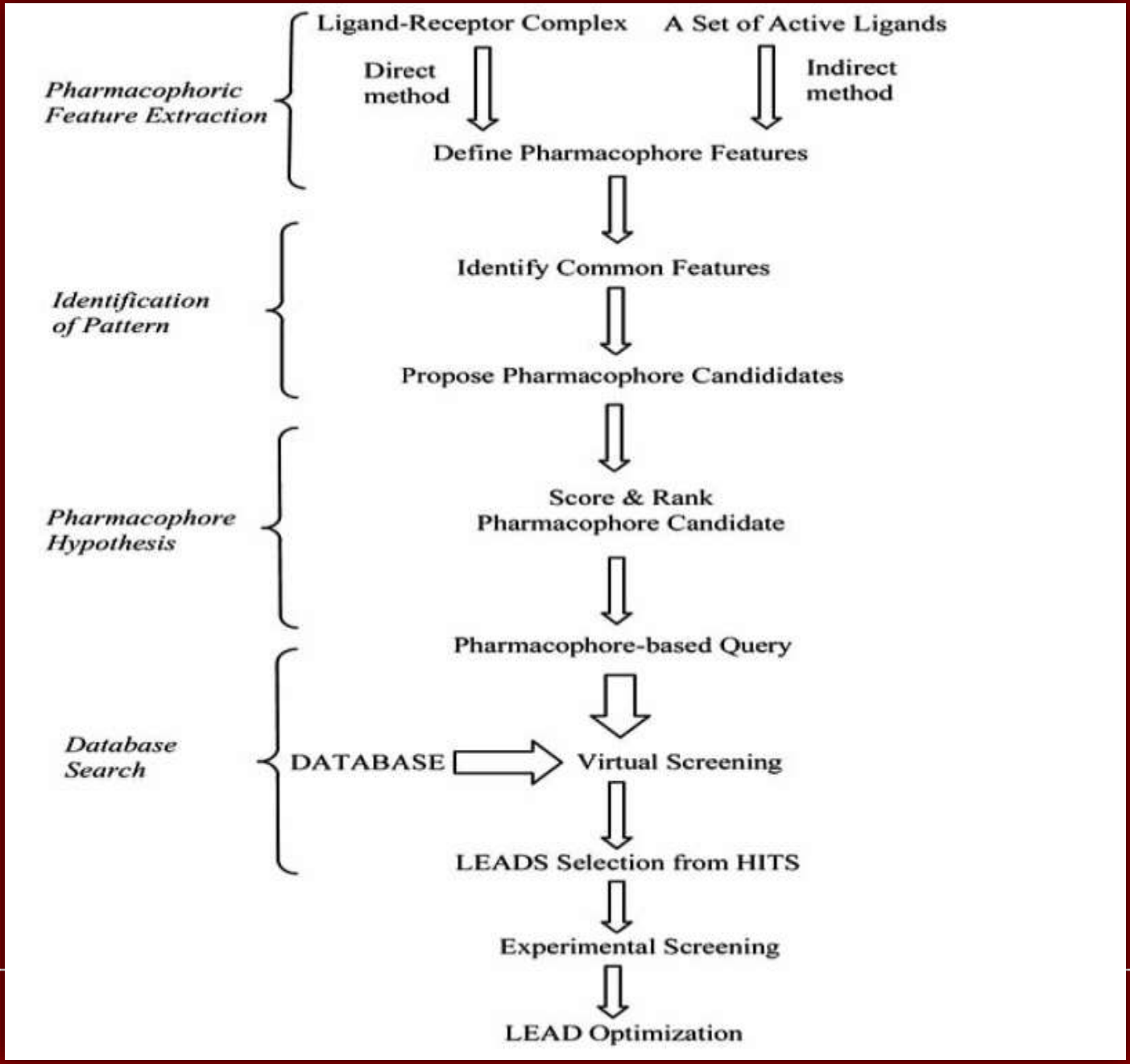
There are methods that use large data sets as input but then prune them into a smaller one by sorting the activities of ligands depending on the user specified cut off.

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).

Features can be defined depending on topology (phenyl ring and carbonyl group), function (Hbond 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.

Flow chart of the virtual screening process using the pharmacophore method.



Pharmacophore Generation

The selected features from each ligand are combined to form a representation of the whole structure.

In the pattern identification phase, 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, RAPID GAMMA, and GASP.

Pharmacophore Generation

This is done in three steps:

1. The constructive stage identifies pharmacophore candidates that are common among the most active set of ligands.

2. 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.

3. The last step of optimization attempts to improve the score of the pharmacophore candidates that pass the subtractive stage, by simulated annealing.

In this way, molecular flexibility is simulated by applying the genetic operators.

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.

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.

Applications of Pharmacophore

De Novo Design of Ligands

The pharmacophore can be used to design novel ligands that satisfy the constrains defined by the pharmacophore model. 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.

Database Searches Based on Pharmacophore

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.

Applications of Pharmacophore

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.

Limitations of pharmacophore methods

1. Despite the abundance of successful cases of drug design relying on pharmacophore modeling, as with any method, it is not failsafe and one should be cautious about the limitations of this technique.

2. The major limitation in virtual screening by pharmacophore is the absence of good scoring metrics. Whereas docking simulations are based on scoring functions trying to predict the affinity, and similarity searches utilize similarity metrics such as the Tanimoto score, pharmacophore queries do not have a reliable, general scoring metric which are not complementary with the receptor binding site, rendering them inactive despite being a perfect match.

3. A second limitation is the dependency of a pharmacophore-based virtual screen on a pre-computed conformation database. These databases only contain a limited number of low-energy conformations per molecule.[95,96] It may be possible that an active molecule cannot be identified as the conformation is missing.

Limitations of pharmacophore methods

4. Finally, a major limitation is that there is no one clear way to construct a pharmacophore query. In many cases, pharmacophore models are able to retrieve molecules, but different models may have worked.

5. In conclusion, plenty of experience and a certain dose of serendipity may be required for successful results. The influence of expert knowledge for in silico screening, also known as the in cerebro step, has been demonstrated during the virtual screening challenge.

6. While target identification, prediction of side effects, and ADME-tox profiling appear to be promising applications for pharmacophore modeling, success is limited for new molecule classes as information is lacking for such compounds or targets.

Summary and Outlook

A. 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.

B. 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.

C. 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.

Summary and Outlook

D. It must also be recognized that not all the SAR datasets have a pharmacophore, and it is essential to discover if a pharmacophore exists.

E. 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.

F. 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

A. D.A.V. Public School, ACC Colony, Jamul – 490024, Durg, Chhattigarh, India (1990 – 2004)

B. University Institute of Pharmacy, Pt. Ravishankar Shukla University, Raipur, Chhattisgarh, India (2005 – 2009)

C. Shri Rawatpura Sarkar Institute of Pharmacy, Kumhari, Durg, Chhattisgarh, India (2010 – 2012)

D. Royal College of Pharmacy, Raipur, Chhattisgarh, India (March – October 2013)

E. National Center for Natural Resources (NCNR), Pt. Ravishankar Shukla University, Raipur, Chhattisgarh, India (2013 – 2017)

F. Columbia Institute of Pharmacy, Raipur 493111, Chhattisgarh, India (2017 – till date)

G. Gratitude to Resp. Prof. Shiv Shankar Shukla Sir, Columbia Institute of Pharmacy, Raipur 493111, Chhattisgarh, India (Mentor)

H. Gratitude to Resp. Prof. Atanu Kumar Pati, Vice Chancellor, Gangadhar Meher University, Sambalpur, Odisha

I. Gratitude to Resp. Prof. Mitashree Mitra, Dept. of Anthropology, PRSU, Raipur, Chhattisgarh.

J. Gratitude to Resp. Vishal S. Deshmukh Sir (M.Pharm. Supervisor)

Thanks for studying !!!!