GALGOTIAS UNIVERSITY

PROJECT ON

Ligand based drug design and drug discovery

IN

BACHELOR OF PHARMACY

Submitted By

NARENDRA NISHAD
BACHELOR OF PHARMACY FINAL YEAR
Admission no 17SMAS102058

BRANCH OF PHARMACY

GALGOTIAS UNIVERSITY

SCHOOL OF MEDICAL ALLIED SCIENCE

Under the Supervision of

MS. SWATI VERMA
Assistant Professor

MAY 2021

TABLE OF CONTENT

DEPARTMENT OF PHARMACY

GALGOTIAS UNIVERSITY

CERTIFICATE

This is to certify that the work contained in this project on ligand based drug design and drug discovery submitted in partial fulfilment for the academic requirement in the degree of Bachelor of Pharmacy is the original work carries out by **NARENDRA NISHAD** during the academic year 2020–21, under the guidance of **Ms. SWATI VERMA (Assistant Professor) the work is completed and the ready for evaluation in partial fulfilment for the award of bachelor of pharmacy under galgotias university greater noida during the academic year 2020-21.**

Date :

Place:

Prof. PRAMOD KUMAR SHARMA

DEAN

SCHOOL OF MEDICAL

AND

ALLIED SCIENCE

GALGOTIAS UNIVERSITY

DEPARTMENT OF PHARMACY

DECLARATION

The project report entitled ligand-based drug design and drug dicovery is the compilation work of **Mr. Narendra nishad** under supervision of **MS. SWATI VERMA** Assistant Professor **Department of Pharmacy, GALGOTIAS UNIVERSITY Greater Noida U.P India**. All pictures, Figures and information used in project are taken from various sources are true and best of my knowledge.

Name and signature of candidate

(NARENDRA NISHAD)

ENROLLMENT No.1712102049

DEDICATION

I dedicate this thesis to my guider teacher MS. SWATI VERMA (Assistant professor) who taught me everything about this project and taught me the basics rules of life that are very useful and important for a person to live a healthy life. Mam taught that never too late to start a thing and achieve your goals. Mam you and your thoughts really motivate me in my life and my carrier so mam thank you for guiding me.

GALGOTIAS UNIVERSITY

ACKNOWLEDGEMENT

Praise to the almighty “GOD” who made me able to carry out the present study successfully, I feel highly privileged while starting my dissertation work with acknowledge of the genuine help and support received from others who made this research and project possible for me.

I deem it great pleasure record my heartfelt gratitude my supervisor for this valuable guidance.

I express my gratitude to my friend for their moral support, advise, affection, co-operation and functional freedom that made me to set goal and achieved this task

The episode of my acknowledgement would not be complete without mention thanks to **Dr.PK SHARMA DEAN OF OUR SMAS** for carving another milestone in my academic journey. His Pains taking efforts in suggesting guiding, designing, advising, improving the study of throughout the entire spanrch of investigation and for inexhaustible encouragement enabling me to the complete the research work

The most importantly I wish to express gratitude to my lovely parents **MR. MAHENDRA SINGH &MRS. SHREE BATI DEVI** for their patience and constant support they provide me with every opportunity to succeed.

Narendra Nishad

Ligand based drug design and drug discovery

Introduction

Computer-aided drug design is a very useful tool for drug designing as it minimizes the time utilized in identification, characterization, and structural optimization of novel drug candidate [1-5]. It can be utilized for rational design of drugs. Prodrugs are typically designed to increase the specificity or bioavailability of the original drug molecule [6-8].

Mainly two approaches are applicable in CADD. The first approach is the Ligand based drug design and structure-based drug design is the second approach we can apply for the design of new drug candidates. Ligand based drug design is an indirect approach to expedite the development of compounds which are pharmacologically active. The development is done by studying the molecules that interact with the biological target of interest [9]. While structure-based drug design utilized the knowledge of the 3D structure of the target molecule to identify and optimize the potential drug candidates [10-12].

Identification of the suitable target molecule is the first and the foremost step in the process of drug designing which is associated with a disease. Usually a key protein of a biochemical pathway which is associated with the target disease serve as a potential drug target [6, 13, 14].

Depending on the disease state, molecules referred to as lead compounds are identified or designed to inhibit or promote the concerned biochemical pathway [15-18]. The next step in the process of drug discovery is to optimize the lead molecules to maximize the interaction with the target molecules. In the process of lead optimization, CADD plays a very important role. Ligand-based drug design methods are used in the absence of information regarding the 3D structure of receptor. The information is confined only to the molecular structure and properties.

Due to the lack of an experimental structure, the known ligand molecules that bind to the drug target are studied to understand the structural and physicochemical properties of the ligands that correlate with the desired pharmacological activity [19-22].

Apart from the known ligand molecule, ligand-based methods may also include natural products or substrate analogues that interact with the target molecule having desired pharmacological effect [9]. Alternatively, in the presence of a 3D structure of the drug target, structure-based methods such as molecular docking or in-silico chemical alteration, are usually applied for lead optimization [23-27].

LITRATURE REVEIW

1. Correspondence to this author at the Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 Penn Street, Baltimore, MD 21201, USA; Tel: +1-410-706-7442; Fax: +1- 410-706-5017; E-mail: alex@outerbanks.umaryland.edu† Present Address: Institut Européen de Chimie et Biologie, Université Bordeaux 1, 2 rue Robert Escarpit, 33607 Pessac Cedex, France.
2. Hung-Jin Huang a , Kuei-Jen Lee b , Hsin Wei Yu a , Chien-Yu Chen a , Chih-Ho Hsu a , Hsin-Yi Chen a , Fuu-Jen Tsai b c & Calvin Yu-Chian Chen a b a Laboratory of Computational and Systems Biology , School of Chinese Medicine, China Medical University , Taichung , 40402 , Taiwan, ROC b Department of Bioinformatics , Asia University , Taichung , 41354 , Taiwan, ROC c Department of Medical Genetics, Pediatrics and Medical Research , China Medical University Hospital and College of Chinese Medicine, China Medical University , Taichung , 40402 , Taiwan, ROC Published online: 15 May 2012
3. A. Laboratory of Computational and Systems Biology, School of Chinese Medicine, China Medical University, Taichung, 40402, Taiwan, ROC. B. Department of Bioinformatics, Asia University, Taichung, 41354, Taiwan, ROC. C. Department of Medical Genetics, Pediatrics and Medical Research, China Medical University Hospital and College of Chinese Medicine, China Medical University, Taichung, 40402, Taiwan, ROC dCurrent address: Computational and Systems Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
4. Gregory L Wilson1 & Markus A Lill†1 1 Department of Medicinal Chemistry and Molecular Pharmacology, College of Pharmacy, Purdue University, 575 Stadium Mall Drive, West Lafayette, IN 47907, USA † Author for correspondence: Tel.: +1 765 496 9375 Fax: +1 765 494 1414 E-mail: mlill@purdue.edu
5. Department of Chemistry, Ganpat University, Mehsana, Gujarat, India 2 Department of Chemistry, MLS University, Udaipur, Rajasthan, India *Corresponding author: Prakash Prajapat prajapatprakash11@yahoo.in Department of Chemistry, Ganpat University, Mehsana-384 012, Gujarat, India. Tel: 09978996944 Citation: Prajapat P, Agarwal S, Talesara GL (2017) Significance of Computer Aided Drug Design and 3D QSAR in Modern Drug Discovery. J Med Org Chem Vol.1 No.1:1

6. Chayan Acharya†, Andrew Coop, James E. Polli, and Alexander D. MacKerell Jr.* Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 Penn Street, Baltimore, MD 21201, USA.
7. Pure Appl. Chem. 84, 1837–1937 (2012). A collection of invited papers based on presentations on the Chemistry of Life theme at the 43rd IUPAC Congress, San Juan, Puerto Rico, 30 July–7 August 2011. ‡Corresponding author: Tel.: +55 16 3373 8095; Fax: +55 16 3373 9881; E-mail: aandrico@ifsc.usp.br

The Basics of QSAR

The well-known approaches for ligand-based drug design are the methods of pharmacophore modelling and QSAR. QSAR is a computational method used to quantify the correlation between the chemical structures of a series of compounds and its chemical and biological responses. The basic hypothesis behind QSAR method is that the similar structural and physiological properties yield similar activity [28,29]. Initially a group of lead molecules are identified which represents the desired biological activity of interest. A quantitative relationship is established between the physico-chemical features of the active molecule and the biological activity. The developed QSAR model is then used to optimize the active compounds to enhance the biological activity. The compounds predicted are then experimentally tested for the desired activity. The QSAR method can be thus used as a tool for identification of compound modifications with improved activity.

The general methodologies of QSAR is built upon a series of consecutive steps:

1) Identify the ligands with experimentally measured values of the desired biological activity. These ligands should have a large variation in biological activity.
2) Identify and determine the molecular descriptors associated with various structural and physicochemical properties of the molecules under study.
3) Discover the correlations between the molecular descriptors and the biological activity that can explain the variation in activity.
4) Test the statistical stability and predictive power of the QSAR model.

The biological activity of the series of compounds is experimentally measured and this will act as a dependent variable in QSAR modelling. The molecules are selected and are insilico modelled. To bring molecules to a stable configuration, the energy of the molecules are minimized by using various energy minimization procedures like molecular mechanics or quantum mechanical methods [21,30-33].

Next step is the generation of molecular descriptors for the set of molecules which affect the biological activity of the molecule. Molecular descriptors can be topological, structural, molecular weight variable, lipophilicity etc. depending upon the QSAR method, quantum or mechanical tools can be used to develop a mathematical relationship between the molecular descriptors and biological activity. The biological activity of the molecules

is dependent upon the molecular descriptors. They can increase as well as can lead to decrease in biological activity. The final step is the validation of model generated by using the set of molecules. The validation method adopted includes both internal as well as external validation, to test the statistical significance, predictive power and robustness of molecules.

The strength of the classical QSAR is that by using very simplistic mathematical relations involving various physicochemical properties and chemical substituents it is able to explain and predict biological activity of a series of similar molecules.
The molecular descriptors used for correlation with the activity were mostly representative of fragments of the parent molecule, including substituents on the parent. The advantage of using fragment-based descriptors is their ready availability for a wide-range of substituents computational ease and the ability to keep the mathematical implementation fairly understandable [34]

3D QSAR

Descriptors that describe 3D features of a molecule. As the name suggests, 3D QSAR method includes develop a QSAR model. Various geometric, physicalcharacteristics and quantum chemical descriptors may be used to describe the 3D features of the ligands in the 3D QSAR method. Such molecular descriptors are then combined to create a pharmacophore that can explain the biological activity of the ligands. A pharmacophore is defined as the 3D spatial orientation of various features, such as hydrogen bond donors or acceptors, which are essential for the desired biological activity [35-36].
The developed pharmacophore model is tested for the stability and statistical significance to obtain the final 3D QSAR model. There are several review articles available that elaborately discuss various techniques of 3D QSAR modeling [34, 35, 38-37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48]. To avoid redundancy, the following section will briefly describe the major 3D QSAR techniques currently in use for drug design. The concluding section will provide a detailed description of the CSP-SAR method developed in our laboratory along with applications of the method.

CoMFA

COMFA uses lennard-Jones and coulombic potential function to calculate steric and electrostatic interaction, respectively, which can cause unrealistically high values forthese energy terms due to the hyperbolic natures of the energy functions. An arbitrary cutoff value for thesepotential functions is assigned in CoMFA to avoid such behavior [49,50].

Comparative Molecular Field Analysis (CoMFA) [51] is one of the most widely used 3D QSAR methods. CoMFA was the first QSAR method to relate 3D shape-dependent steric and electrostatic properties of a molecule to its biological activity. In this method the molecules are aligned based on their 3D structures on a 3D grid and the values of steric and electrostatic potential energies are calculated at each grid point. Usually CoMFA assumes that the minimumenergy conformer is the bioactive conformer. For systems with known crystal structures, crystal coordinates may be used to define bioactive conformers. Field values corresponding to the potential energy terms are calculated at each grid point for every molecule and correlated with the biological activity. PCA or PLS methods are usually used for model development in CoMFA. The CoMFA model is then tested for statistical significance and robustness. The success and predictive ability of CoMFA models are highly sensitive on the alignment of the bioactive conformers [45, 52-53].

As the bioactive conformation is not necessarily the lowest conformation in the absence of the receptor [54-55], the assumption made by CoMFA in the selection of bioactive conformers and the corresponding alignment method may produce erroneous models. By neglecting the dynamical nature of the ligands CoMFA limits its applicability. Another limitation of CoMFA is the form of its energy function, as it does not explicitly account for hydrophobicity or hydrogen bond interactions [45, 53,56].

CoMSIA

Comparative Molecular Similarity Indices (CoMSIA) [57] is a 3D QSAR technique similar to CoMFA. However, unlike CoMFA, the molecular field expression of CoMSIA includes hydrophobic, hydrogen-bond donor and acceptor terms in addition to steric and coulombic contributions. CoMSIA also calculates the similarity indices instead of interaction energies by comparing each ligand molecule with a common probe with a radius of 1Å, and charge, hydrophobicity and hydrogen bond properties equal to 1 [58]. CoMSIA uses bell-shaped Gaussian function to describe steric,

electrostatic and hydrophobic components of the energy function. Unlike CoMFA, this allows CoMSIA to avoid the use of an arbitrary cutoff value for the energy calculations. Similarity indices corresponding to CoMSIA molecular fields define the ligand-protein binding interaction [59].

Catalyst

Efforts have been made to include the conformational flexibility in 3D QSAR methodology. CATALYST is one of the most popular 3D QSAR software packages that uses conformational variation during model development. CATALYST uses the poling algorithm [60] to sample conformational space for the ligand molecules. Typically, 250 conformers are generated in this process with a default cutoff value of 20 kcal/mol above the energy of global minimum conformation. Spatial orientations of the functional groups are used to develop the pharmacophore hypothesis and the estimated and observed activity values are compared to evaluate the QSAR models. The most common properties or functional groups used to define the pharmacophoric features are:

1. Hydrogen-bond acceptor
2. Hydrogen-bond donor
3. Positively charges group (basic)
4. Negatively charged group (acidic)
5. Aromatic ring
6. Aliphatic hydrophobic moieties
7. Aromatic hydrophobic moieties

The pharmacophore generation process is divided into constructive and subtractive phases. During the constructive phase compounds having activity greater than a cutoff value are used to build a pharmacophore hypothesis. In the subtractive phase any pharmacophore that fits more than half of the inactive compounds is rejected. A cost value is assigned to each selected pharmacophore based on its prediction error, feature weight and complexity.

CATALYST is equipped to overcome most of the drawbacks of the previous 3D QSAR methods. However, there are a few limitations in CATALYST. The conformation generator of CATALYST creates a maximum of 250 conformers which

may not include all accessible conformations for flexible ligand molecules. Hence, CATALYST may fail to include the bioactive conformer of the active compounds, which in turn may lead to incorrect pharmacophore models. CATALYST also does not generate models that include both the physico-chemical properties and pharmacophoric features

CSP-SAR

***Principle*.** CSP-SAR is a novel method for developing 3D QSAR models based on the Conformationally Sampled Pharmacophore (CSP) method developed in our laboratory [44, 61, 62]. This method is designed for ligands with conformational flexibility and avoids problems associated with ligand alignment. As discussed before, the active or bound conformation of a ligand molecule does not need to be the lowest-energy conformation in the absence of the target molecule; even the active conformer may not belong to the ensemble of low energy conformers [63]. In order to maximize the potential inclusion of the bioactive conformers of the ligand molecules in the model, a rigorous sampling of the conformational space for each ligand is essential. Unlike other pharmacophore development methods, CSP considers all accessible conformations of each ligand molecule for pharmacophore development. Thus, CSP maximizes the probability of including the bioactive conformer in the model. Once all accessibly conformations of the ligands have been generated, typically *via* molecular dynamics simulations (see below), it is necessary to extract descriptors of the conformational properties of the ligands for use in model development. This requires the selection of pharmacophore features for the set of ligand molecules, and represents a critical step in the CSP approach.

Typically, these features include hydrogen bond donor and acceptors,hydrophobic groups or any other structural feature that may be important for the biological activity. Available SAR data on the system of interest may guide but not limit this selection procedure. For example, the CSP approach was successfully applied to opioids using previously defined functional groups, such as the basic nitrogen known to be essential for opioid activity, as well as identifying novel functional groups during model development [44, 64, 65].

Studies involving relatively large ligands that have not previously been subjected to significant SAR studies pose more difficulty in the functional group selection

process. In the absence of any existing model functional group selection involves the user considering all functional groups that might have any effect on the biological activity. After the identification of all possible chemical features that may serve as pharmacophoric points, all possible distances, angles and dihedral angles between the feature points need to be considered. Once these descriptors have all been identified they are regressed against each other to eliminate redundant descriptors from further analysis (e.g. if descriptors have a correlation coefficient (r^2) greater than, for example, 0.8 one of those may be removed from further consideration). The remaining descriptors are then systematically regressed against the biological data, with those having correlation coefficient (r^2) less than a cutoff value (typically 0.01) with respect to the biological data discarded from further analysis. Initially an extensive calculation of the structural descriptors is highly recommended. For relatively large ligand molecules the number of possible structural descriptors can be quite high, being on the order of 100,000 or more. However, automation of this procedure readily allows the selection of descriptors for additional analysis to be performed.

The nature the descriptors, which include selected pharmacophore features in combination with all accessible conformations of each ligand, is the key feature of the CSP approach. This combination requires that the descriptors be treated as probability distributions that include, for example, all possible distances between two pharmacophore features or all possible angles between three pharmacophore features, and so on. To better elucidate this concept, we will expand on published results of a CSP study of bile acid conjugates and their transporter (Apical Sodium-dependent Bile acid Transporter or ASBT) [43]. Presented in Fig. (**2**) are three conjugates of the bile acid **9**, **2** and **21**, on which three pharmacophore points are shown (note that in the original study, a total of 30 pharmacophore points were initially considered on a total of 13 compounds). For the present example, three conjugates shown in Fig. (**2**) will be considered. Each of these conjugates was subjected to MD simulations to obtain all possible conformations from which probability distributions of descriptors based on the pharmacophore features were determined. Onedimensional descriptors associated with the NG-OA distance and OA-NG-CG angles are compounds **9** (red), **2** (blue) and **21** (turquoise) [43]. As is evident, each conjugate samples a range of conformations as represented by the probability distributions. It is these distributions that represent the individual descriptors and the degree of

overlap between the descriptors (see following paragraph) may be used as independent variables for model development. In addition, the descriptors may be developed in two or more dimensions. An example of 2D probability distributions for the two structural descriptors . From the distributions it is evident that **9** and **2** share high degree of structural similarity with respect to the given descriptors, while **21** did not sample conformational space similar to either **9** or **2**. Accordingly, based on this qualitative analysis, **9** and **2** would be predicted to have similar activity *vs* **21**. Notably, this analysis did not require any alignment of the ligands, simply a comparison of the probability distributions of the selected pharmacophore features. The lack of a requirement for structural alignment represents another strength of the CSP approach.

While use of the CSP approach in a qualitative manner is of utility, as described below, quantitative analysis is required to predict inhibition constants, potencies and so on. This requires that the degree of overlap of the probability distributions of the individual ligands be determined, yielding overlap coefficients that may be used directly in regression analysis. 1D overlap coefficient of a single structure descriptor between two ligands can be calculated using the following relation for discrete probability density
functions [43, 66],

Computational Method.

The primary requirement of the CSP method is adequate conformational sampling for the ligand molecules. In order to achieve a complete sampling of conformational space rigorous molecular dynamics (MD) simulations [67] are an essential part of CSP. However, other sampling methods such as systematic grid search [68-69], fragment-based search [47, 70], random search or Monte Carlo (MC) simulations [71], distance geometry [72], genetic algorithm [73-74], simulated annealing [75, 76], taboo search [77] etc. can also be applied for conformational sampling purpose as long as exhaustive sampling of conformational space can be assured. A detail description of these searching algorithms can be found in several review articles and book chapters [66, 78-79].

Empirical force fields [20] are an integral part of *in silico* modeling. Any molecular force field such as CHARMM [71], AMBER [80], MMFF [80-81] or OPLS [82]

which is suitable for small molecules can be used for CSPSAR modeling. However, it is important that the force field used accurately model the structural properties of the molecules of interest. Test of this accuracy may be performed by quantifying the ability of the force field to reproduce minimum energy geometries with those obtained from quantum mechanical (QM) calculations or high resolution crystals structures such as those obtained from the Cambridge Structural Database [83]. In addition, the use of QM methods allows the ability of a force field to reproduce the change in energy as a function of ligand conformation to be validated and optimized, as required. Proper treatment of the conformational energies is particular important for the CSP approach as it is based on conformational distributions. Methods for force field validation and optimized have been described elsewhere [47, 84].

CSP-SAR models developed by our laboratory used MD simulations as a tool for conformational search. MD generates consecutive conformations of a molecular system

using Newton's second law of motion in which the force acting on a system along with velocities of the atoms in the system are used to predict new conformations by integrating over time[85]*. The time evolution of the position and velocity of the molecular system is estimated from the analytical solution of the differential equation of motion. For flexible ligands replica-exchange MD simulations [86-87] are preferably employed for sampling the conformational space.Replica-exchange MD simulation methods reduce the probability of a molecular system getting trapped in local minimum energy region during a simulation facilitating complete sampling of the accessible conformational space.

In this method a number of replicas of the same system are simultaneously simulated at different temperatures and with coordinates or other properties swapped between the replicas performed at regular interval. The probability of the exchange of two replicas is subjected to the Metropolis Criterion [88] thereby assuring that the system maintains a proper Boltzmann distribution. MD simulations of each replica are typically performed using 20ns Langevin dynamics [89] with an integration time step of 0.002ps in the presence of an implicit solvent model [90-91], such as Generalized Born Continuum Solvent Model (GBMV) [92],

[93]. Usually 20ns simulations yield conformational convergence for flexible ligands with moderate size (~ 650 Daltons); testing that additional simulation time does not

lead to additional sampling is often adequate to verify that the full range of accessible conformations of the molecule has been sampled. Coordinate frames are saved from the MD trajectories and used to determine the conformational distribution of the structural descriptors from which the overlap coefficients are calculated. During *in silico* modeling the protonation state of any ionizable chemical group should be properly assigned based on the pH of the experimental condition used to measure the biological activity.

1D and 2D probability distributions of various pharmacophoric feature points are obtained by analyzing the trajectories from the MD simulations. Overlap coefficients of the conformational distributions are combined with the physico-chemical properties of the ligands to obtain a set of molecular descriptors. The molecular descriptors are subjected to single-variable as well as multivariable linear regression (MLR) analysis against the biological activity of interest. All possible combinations of the molecular descriptors are subjected to MLR analysis to identify the combination of descriptors (candidate models) that can explain the variability of the biological activity of the ligands. To avoid overfitting, any combination of independent variables having correlation between each other greater than 0.8 are not included for multivariable regression.Akaike information criteria [94, 95] is applied to rank the candidate models for systems with more than one statistically significant quantitative models. Simple SHELL scripts may be used to automate the process of capturing snapshots from the MD trajectories and calculating the overap coefficients of the structural features. MLR analysis for all possible combinations of molecular descriptors and calculation of AIC values of the selected candidate models can also be automated using statistical software like R in conjunction with a SHELL script. The combination of CSP approach with 3D QSAR method, named CSP-SAR, thus potentially can capture information on the bioactive conformation in model development which facilitates an understanding of the biological interactions dictating activity without any available ligand-target 3D structure.

Applications.

The CSP method was developed and first successfully applied on opioid ligands [44,62,63,]. CSP was used to study both peptidic as well as non peptidic opioids and the derived pharmacophore model distinguished opioid agonists from the

antagonists. Using qualitative CSP models for opioid ligands Bernard and coworkers [44] discovered that DPI2505, a compound previously suggested to be a antagonist, may act as an agonist. The qualitative model was also used to design novel opioid ligands.

Subsequent application of quantitative CSP for opioid ligands [63] yielded efficacy and affinity models that were able to distinguish between ligands that differed by a single substitution on an aromatic ring. These efforts also discovered a novel hydrophobic moiety imporant for efficacy and affinity that had not been identified in previous studies. This represented a significant advance in our understanding of opioid SAR, as previous thinking assumed that the hydrophobic moiety was limited to aromatic groups, whereas as the CSP approach showed that aliphatic moieties could also serve as the hydrophobic groups in certain ligands. Notably, the models of opioid developed by the CSP methods encompassed low molecular weight, nonpeptidic opioids as well as peptidic ligands. Previous opioid models were not able to bridge this gap.

The ability of the CSP method to overcome this is based on the inclusion of all conformations in model development, the lack of the need to align molecules, a particular problem when both nonpeptidic and peptidic ligands are being studied and the inclusion of a large number of possible pharmacophore features in model development. Indeed, that later consideration led to the identification of the novel hydrophobic moieties in the selected opioids. Recently, the inhibition requirement of hASBT using amino-piperidine conjugates of bile acids was studied using the CSP-SAR method [69]. CSP-SAR models developed for hASBT inhibition successfully identified structural and physico-chemical descriptors that explained the variance of the biological activity. Despite the fact that the inhibitors used in this study had a narrow range of activity, theconformational sampling feature of the CSP-SAR method was able to facilitate identification of the information from the molecular descriptors necessary to explain the activity. The quantitative CSP-SAR models developed in this study was able to distinguish between very-potent inhibitors (<16□M) from moderately-potent (>16 □M) inhibitors with some exceptions. However, further qualitative analysis was able to overcome the limitation of the quantitative models. Qualitative CSP-SAR demonstrated that very subtle chemical modifications in some inhibitors led to the formation of salt-bridge interaction resulting in conformational restriction associated to poorer binding affinity. This study established the strength of

CSP-SAR method to capture the effect of such small chemical modification on biological activity and it emphasizes the utility of both quantitative and qualitative CSP approaches.

The CSP method has also been applied and discussed by other researchers in the context of 3D QSAR [54-60]. CSP models developed by Bernard and coworkers, demonstrated the importance of including extensive conformational sampling in model development. This motivated other researchers to include conformational sampling during the development of 3D QSAR models for other flexible systems. Gilbert and coworkers applied the concept of CSP method in their work by considering a set of representative conformations of the flexible ligands to develop selective inhibitors of DAT/SERT using CoMFA and CoMSIA methods [88]. Mallik and coworkers [95] used the CSP method for developing 3D pharmacophore for the 13-residue cyclic peptide, compstatin, an anti-complement peptide and other related peptidic analogues. Using the CSP methodology the researchers were able to distinguish between active and inactive analogues. The researchers also extended the original CSP work by Bernard and coworkers, by including dihedral angles as a pharmacophoric descriptor to capture 3D structural features of the peptidic ligands. The inclusion of multiple conformers instead of using only the lowest-energy conformer yielded a stable and predictive model. Kalaszi and coworkers [90] developed a novel 3D QSAR method based on thermodynamic properties to predict bioactive conformation of flexible ligands using conformational analysis of the ligand molecules. In two recent studies Lexa and coworkers [91] and Kirschner and coworkers [92] used replica exchange molecular dynamics to explore the conformational space accessible by peptidic ligands with breast cancer inhibiting properties. They used ligand-based 3D QSAR method to identify the bioactive conformations of the active ligands. Conformational analysis of the larger active peptides allowed them to explain the activity of the existing ligands as well as discover novel smaller peptidic ligands with full biological activity. The successful works of these researchers confirm the validity and importance of CSP approach in ligand-based 3D QSAR modeling for flexible molecules. In addition, CSP has also been mentioned in several review articles [93-94] as a novel method to utilize the dynamical behavior of flexible biomolecules to explain ligand-protein binding.

Validation of the CSP Method.

To assess theperformance of CSP method as compared to more traditional 3D QSAR approaches, additional calculations were performed as part of the present study. These involved a comparative study of the inhibition pharmacophore model for hASBT based on the thirteen ligands in **G1** and **G3** groups as described by Gonzalez and coworkers [55] with a model developed presently using the Catalyst approach. Similar to the observation of Gonzalez and coworkers inclusion of the compounds in **G2** group did not yield statistically significant model (r^2 = 0.55).

Catalyst model development was performed using Discovery Studio 2.1 CatalystTM (Accelrys, San Diego, CA). The best conformation generation method as implemented in

CatalystTM was used to generate up to 250 conformers of each ligand based on a 20 kcal/mol energy cutoff. Ten hypotheses were generated using the conformers of the ligands and their Ki values using five molecular features, such as hydrogen bond donor, hydrogen bond acceptor, hydrophobic, positively ionizable group and negatively ionizable group. Out of the ten hypotheses, the hypothesis yielding the lowest total cost was selected for further analysis. The best inhibition model generated by Catalyst consisted of five features including one hydrogen-bond acceptor, one hydrogen-bond donor, two hydrophobic moieties and one positively ionizable group. The most potent inhibition in the set, compound **9**, mapped all the five features of the pharmacophore ; 3-OH represented the hydrogen-bond acceptor, 7-OH represented the hydrogen bond donor, C-19 and D-ring represented the two hydrophobes and the basic piperidine nitrogen depicted the positively ionizable group feature. Top three CSP-SAR inhibition models also consisted of structural descriptors representing similar features e.g. 7-OH, basic piperidine nitrogen (positively ionizable group) and hydrophobic moieties close to C-19 and D-ring such as centroid of B and C rings and C-20. In addition, CSP-SAR models included structural descriptors involving the relative orientation of □-substituent with respect to the steroidal nucleus. However,

CSP-SAR models did not contain any descriptor that explicitly considered 3-OH. CSP-SAR models also included physico-chemical descriptors such as GB energy (electrostatic component of solvation free energy) and logP (octanol/water partition coefficient). This is a clear advantage of CSP-SAR method over Catalyst as there is no simple tool in Catalyst that can combine structural features with physico-chemical descriptors. The Catalyst model yielded r^2 of 0.849 while the r^2 of the best model

reported by Gonzalez and coworkers was 0.813. However, the CSP-SAR model yielded better RMSD value than the Catalyst model (Table **1**).

Table **1** represents the observed and estimated Ki values of the ligands based on CSP and Catalyst methods.Moreover, the best CSP-SAR model included only two descriptors to explain the activity while five descriptors were used by Catalyst model for the same set of compounds. From the comparison of the inhibition models developed by the two methods it is evident that CSP and Catalyst yielded very similar fitting quality. Nevertheless, the Catalyst method did not provide any tool to explanation of the variation in activity due to subtle chemical modifications; while CSPSAR qualitative model was able to explain such variation in activity *via* salt-bridge interaction.

Limitations.

One limitation of the CSP-SAR method is that the selection of pharmacophoric features of the ligands is user dependent. The selection of functional groups is often facilitated by previous studies though all the chemical groups present on the ligands must be considered. This limitation may be overcome by considering probability distributions between all possible distances, angles and dihedral angles involving all chemical groups that may impact on biological activity. A second limitation is the computational requirement. As extensive sampling of conformational space is required, extended MD simulations must be performed on each ligand. While this step is computationally demanding, the accessibility of commodity computing minimizes this limitation. In addition, once the conformational sampling of a ligand is completed and the generated conformations stored, further analysis may be performed to identify additional structural or physio-chemical properties that correlate with biological activity without redoing the MD simulation or other sampling procedure.

CONCLUSIONS

Ligand-based drug design is inherently a complicated problem as this approach is restricted to considering only one side of the actual biochemical process. It has been shown in many cases that receptor molecules and/or ligands undergo significant conformational changes to facilitate their interaction [95-96]. While traditional pharmacophore approaches often did not take into account ligand conformational flexibility by only using minimum energy conformations of the ligands, more recent methods include a large number of conformations during model development. Though such methods offer significant improvements, they are still limited by including a finite range of conformations as well as requiring alignment of the ligands under study. The CSP method largely overcomes these limitations including all accessible conformations of the ligands and using the overlap of probability distributions of pharmacophore features during model development. In addition, the CSP-SAR method may readily be combined with physicochemical properties. The utility of this approach has been demonstrated in a number of studies in our laboratories as well as by other workers. Clearly, ligand-based drug design is an effective method to understand the features of ligands important for their biological activity in the absence of the receptor structure.

Investigation of the structural and physico-chemical features of the ligands of a drug target can indicate the nature of interactions that are essential for the desired pharmacological response. The method can also predict novel molecular structures with features facilitating the interaction with the target molecule. As stated above, there are several different methodologies to perform ligand-based modeling. However, proper understanding of the underlying principle of the chosen method is highly recommended for successful application of these methods to complex biological systems.

REFERENCE

[1] Chang, C.; Ekins, S.; Bahadduri, P.; Swaan, P.W. Pharmacophorebased discovery of ligands for drug transporters. Adv. Drug Deliv. Rev., 2006, 58, 1431-1450.

[2] Ekins, S.; Mirny, L.; Schuetz, E.G. A ligand-based approach to understanding selectivity of nuclear hormone receptors PXR, CAR, FXR, LXRa, and LXRb. Pharm. Res., 2002, 19, 1788-1800.

[3] Ekins, S.; Waller, C.L.; Swaan, P.W.; Cruciani, G.; Wrighton, S.A.; Wikel, J.H. Progress in predicting human ADME parameters in silico. J. Pharmacol. Toxicol. Methods, 2000, 44, 251-272.

[4] van de Waterbeemd, H.; Gifford, E. ADMET in silico modelling: towards prediction paradise? Nat. Rev. Drug Discov., 2003, 2, 192- 204.

[5] Ekins, S. Computer applications in pharmaceutical research and development. John Wiley & Sons; Hoboken, NJ, 2006.

[6] Balakrishnan, A.; Polli, J.E. Apical sodium dependent bile acid transporter (ASBT, SLC10A2): a potential prodrug target. Mol. Pharm., 2006, 3, 223-230.

[7] Takakura, Y.; Hashida, M. Macromolecular drug carrier systems in cancer chemotherapy: macromolecular prodrugs. Crit. Rev. Oncol. Hematol., 1995, 18, 207-231.

[8] Tolle-Sander, S.; Lentz, K.A.; Maeda, D.Y.; Coop, A.; Polli, J.E. Increased acyclovir oral bioavailability via a bile acid conjugate. Mol. Pharm., 2004, 1, 40-48.

[9]Kurogi, Y.; Guner, O.F. Phamacophore modeling and threedimensional database searching for drug design using catalyst. Curr. Med. Chem., 2001, 8, 1035-1055.

[10] Marrone, T.J.; Briggs, J.M.; McCammon, J.A. Structure-based drug design: computational advances. Annu. Rev. Pharmacol. Toxicol., 1997, 37, 71-90.

[11] Gane, P.J.; Dean, P.M. Recent advances in structure-based rational drug design. Curr. Opin. Struct. Biol., 2000, 10, 401-404.

[12] Jhoti, H.; Leach, A.R. Structure-based Drug Discovery. Springer: 2007.

[13] Zhong, S.; Macias, A.T.; MacKerell, A.D., Jr. Computational identification of inhibitors of protein-protein interactions. Curr. Top. Med. Chem., 2007, 7, 63-82.

[14] Chen, F.; Hancock, C.N.; Macias, A.T.; Joh, J.; Still, K.; Zhong, S.; MacKerell, A.D. Jr.; Shapiro, P. Characterization of ATPindependent ERK inhibitors identified through in silico analysis of the active ERK2 structure. Bioorg. Med. Chem. Lett., 2006, 16, 6281-6287.

[15] Hafner, C.; Schmitz, G.; Meyer, S.; Bataille, F.; Hau, P.; Langmann, T.; Dietmaier, W.; Landthaler, M.; Vogt, T. Differential gene expression of Eph receptors and ephrins in benign human tissues and cancers. Clin. Chem., 2004, 50, 490-499.

[16] Dobrzanski, P.; Hunter, K.; Jones-Bolin, S.; Chang, H.; Robinson, C.; Pritchard, S.; Zhao, H.; Ruggeri, B. Antiangiogenic and antitumor efficacy of EphA2 receptor antagonist. Cancer Res., 2004, 64, 910-919.

[17] Cheng, N.; Brantley, D.; Fang, W.B.; Liu, H.; Fanslow, W.; Cerretti, D.P.; Bussell, K.N.; Reith, A.; Jackson, D.; Chen, J. Inhibition of VEGF-dependent multistage carcinogenesis by soluble EphA receptors. Neoplasia, 2003, 5, 445-456.

[18] Torres, G.E.; Gainetdinov, R.R.; Caron, M.G. Plasma membrane monoamine transporters: structure, regulation and function. Nat. Rev. Neurosci., 2003, 4, 13-25.

[19] Mason, J.S.; Good, A.C.; Martin, E.J. 3-D pharmacophores in drug discovery. Curr. Pharm. Des., 2001, 7, 567-597.

[20] González, P. M.; Acharya, C.; MacKerell, A.D. Jr.; Polli, J.E. Inhibition requirements of the human apical sodium-dependent bile acid transporter (hASBT) using aminopiperidine conjugates of glutamyl-bile acids. Pharm. Res., 2009, 26, 1665-1678.

[21] Bernard, D.; Coop, A.; MacKerell, A.D. Jr. Conformationally sampled pharmacophore for peptidic delta opioid ligands. J. Med. Chem., 2005, 48, 7773-7780.

[22] Loew, G.H.; Villar, H.O.; Alkorta, I. Strategies for indirect computer-aided drug design. Pharm. Res., 1993, 10, 475-486.

[23] Koehn, F.E.; Carter, G.T. The evolving role of natural products in drug discovery. Nat. Rev. Drug Discov., 2005, 4, 206-220.

[24] Lee, K.H. Anticancer drug design based on plant-derived natural products. J. Biomed. Sci., 1999, 6, 236-250.

[25] Lee, K.H.; Huang, B.R.; Tzeng, C.C. Synthesis and anticancer evaluation of certain alpha-methylene-gamma-(4-substituted phenyl)-gamma-butyrolactone bearing thymine, uracil, and 5- bromouracil. Bioorg. Med. Chem. Lett., 1999, 9, 241-244.

[26] Kuntz, I.D. Structure-based strategies for drug design and discovery. Science, 1992, 257, 1078-1082.

[27] Guner, O.F. Pharmacophore perception, development, and use in drug design. 1999.

[28] Akamatsu, M. Current state and perspectives of 3D-QSAR. Curr. Top. Med. Chem., 2002, 2, 1381-1394.

[29] Verma, R. P.; Hansch, C., Camptothecins: A SAR/QSAR Study. Chem. Rev. 2009, 109, 213-235. [30] Duchowicz, P.R.; Castro, E.A.; Fernandez, F.M.; Gonzalez, M.P. A new search [30] Algorithm for QSPR/QSAR theories: normal boiling points of some organic molecules. Chem. Phys. Lett., 2005, 412, 376-380.

[31] Wade, R.C.; Henrich, S.; Wang, T. Using 3D protein structures to derive 3D-QSARs. Drug Discov. Today: Technol., 2004, 1, 241- 246.

[32] Halloway, M.K. A priori prediction of ligand affinity by energy minimization. Persp. Drug Discov. Des., 1998, 9, 63-84.

[33] Bohl, C.E.; Chang, C.; Mohler, M.L.; Chen, J.; Miller, D.D.; Swaan, P.W.; Dalton, J.T. A ligand-based approach to identify quantitative structure-activity relationships for the androgen receptor. J. Med. Chem., 2004, 47, 3765-3776.

[34] Winkler, D.A. The role of quantitative structure-activity relationships (QSAR) in biomolecular discovery. *Brief. Bioinform.,* **2002**, *3,* .

[35] Chang, C.; Swaan, P.W. Computational approaches to modeling drug transporters. *Eur. J. Pharm. Sci.,* **2006**, *27,* 411-424.

[36] IUPAC Compendium of Chemical Terminology, 2nd Edition. **1997**.

[37] Wermuth, C.G.; Langer, T. *Pharmacophore identification. In 3DQSAR in Drug Design. Theory, Methods, and Applications*. ESCOM Science Publishers: **1993**.

[38] Hopfinger, A.J.; Wang, S.; Tokarski, J.S.; Jin, B.; Albuquerque, M.; Madhav, P.J.; Duraiswami, C. Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. *J. Am. Chem. Soc.,* **1997**, *119,* 10509-10524.

[39] Ekins, S.; Mirny, L.; Schuetz, E.G. A ligand-based approach to understanding selectivity of nuclear hormone receptors PXR, CAR, FXR, LXRa, and LXRb. *Pharm. Res.*, **2002**, *19*, 1788-1800.

[40] Ekins, S.; Waller, C.L.; Swaan, P.W.; Cruciani, G.; Wrighton, S.A.; Wikel, J.H. Progress in predicting human ADME parameters *in silico*. *J. Pharmacol. Toxicol. Methods*, **2000**, *44*, 251-272.

[41] Kurogi, Y.; Guner, O.F. Phamacophore modeling and three dimensional database searching for drug design using catalyst. *Curr. Med. Chem.*, **2001**, *8*, 1035-1055.

[42] Mason, J.S.; Good, A.C.; Martin, E.J. 3-D pharmacophores in drug discovery. *Curr. Pharm. Des.*, **2001**, *7*, 567-597.

[43] González, P. M.; Acharya, C.; MacKerell, A.D. Jr.; Polli, J.E. Inhibition requirements of the human apical sodium-dependent bile acid transporter (hASBT) using aminopiperidine conjugates of glutamyl-bile acids. *Pharm. Res.*, **2009**, *26*, 1665-1678.

[44] Bernard, D.; Coop, A.; MacKerell, A.D. Jr. Conformationally sampled pharmacophore for peptidic delta opioid ligands. *J. Med. Chem.*, **2005**, *48*, 7773-7780

[45] Akamatsu, M. Current state and perspectives of 3D-QSAR. *Curr. Top. Med. Chem.*, **2002**, *2*, 1381-1394.

[46] Winkler, D.A. Overview of Quantitative Structure-Activity Relationships (QSAR). In *Molecular Analysis and Genome Discovery*, Rapley, R.; Harbron, S., Eds. 2004.

[47] *Molecular Operating Environment, Chemical Computing Group*, Montreal, Quebec, Canada, http://www.chemcomp.com..

[48] Van Drie, J.H. Pharmacophore discovery: lessons learned. *Curr. Pharm. Des.*, **2003**, *9*, 1649-1664.

[49] Folkers, G.; Merz, A.; Rognan, D. *CoMFA: Scope and Limitations in 3D QSAR in Drug Design*. **1993**.

[50] Flower, D.R. *Predicting chemical toxicity and fate*. **2002**.

[51] Flower, D.R. *Drug design: cutting edge approaches*. **2002**.

[52] Gohda, K.; Mori, I.; Ohta, D.; Kikuchi, T. A CoMFA analysis with conformational propensity: an attempt to analyze the SAR of a set of molecules with different conformational flexibility using a 3DQSAR method. *J. Comput. Aided Mol. Des.*, **2000**, *14*, 265-275.

[53] Ghose, A.K.; Viswanadhan, V.N.; Wendoloski, J.J. A knowledgebased approach in designing combinatorial or medicinal chemistry libraries for drug discovery. *J. Comb. Chem.,* **1999**, *1*, 55-68.

[54] Bostrom, J.; Norrby, P.O.; Liljefors, T. Conformational energy penalties of protein-bound ligands. *J. Comput. Aided Mol. Des.,* **1998**, *12*, 383-396.

[55] MacKerell, A.D. Jr. Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.,* **2004**, *25*, 1584-1604.

[56] Hasegawa, K.; Arakawab, M.; Funatsu, K. Rational choice of bioactive conformations through use of conformation analysis and 3-way partial least squares modeling. *Chemom. Intell. Lab. Syst.,* **2000**, *50*, 253-261.

[57] Klebe, G.; Abraham, U.; Mietzner, T. Molecular similarity indices in a c

[58] Reiser, B.; Faraggi, D. Confidence intervals for the overlapping coefficient: the normal equal variance case. *Statistician,* **1999**, *48*, 413-418.

[59] Leach, A.R. *Molecular modelling: principles and applications.* Addison-Wesley Longman Ltd: **2001**.

[60**]** Ghose, A.K.; Jaeger, E.P.; Kowalczyk, P.J.; Peterson, M.L.; Treasurywala, A.M. Conformational searching methods for small molecules. I. Study of the sybyl search method. *J. Comp. Chem.,* **1993**, *14*, 1050-1065.

[61] Wiberg, K.B.; Boyd, R.H. Application of strain energy minimization to the dynamics of conformational changes. *J. Am.*m *Chem. Soc.,* **1972**, *94*, 8426-8430.

[62] Chen, I.J.; Foloppe, N. Conformational sampling of drug-like molecules with MOE and catalyst: implications for pharmacophore modeling and virtual screening. *J. Am. Chem. Soc.,* **2008**, *48*, 1773–1791.

[63] Saunders, M. Stochastic exploration of molecular mechanics energy surfaces. Hunting for the global minimum. *J. Am. Chem. Soc.,* **1987**, *109*, 3150–3152..

[64] Blaney, J.M.; Dixon, J.S. *Distance geometry in molecular modeling, in. Reviews in Computational Chemistry*. VCH Publishers: New York **1994**.

[65] Nair, N.; Goodman, J.M. Genetic algorithms in conformational analysis. *J. Chem. Inf. Comput. Sci.,* **1998**, *38*, 317–320.

[66] Parrill, A.L. Evolutionary and genetic methods in drug design. *Drug Discov. Today,* **1996**, *1*, 514-521.

[67] Corcho, F.J.; Filizola, M.; Perez , J.J. Evaluation of the iterative simulated annealing technique in conformational search of peptides. *Chem. Phys. Lett.,* **2000**, *319,* 65-70.

[68] Guarnieri, F.; Weinstein, H. Conformational memories and the exploration of biologically relevant peptide conformations: an Illustration for the gonadotropin-releasing hormone. *J. Am. Chem. Soc.*, **1996**, *118*, 5580-5589.

[69] Cvijovicacute, D.; Klinowski, J. Taboo search: an approach to the multiple minima problem. *Science*, **1995**, *267*, 664-666.

[70] Foloppe, N.; Chen, I.J. Conformational sampling and energetics of drug-like molecules. *Curr. Med. Chem.*, [Epub ahead of print] **2009**.

[71] Becker, O.M.; MacKerell, A.D. Jr.; Roux, B.; Watanabe, M. *Computational biochemistry and biophysics*. **2001**.

[72] Brooks, B.R.; Brooks, C.L. III; MacKerell, A.D. Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C. Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A.R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R.W.; Post, C.B.; Pu, J.Z.; Schaefer, M.; Tidor, B.; Venable, R.V.; Woodcock, H.L.; Wu, X.; Yang, W.; York, D.M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comp. Chem.*, **2009**, *30*, 1545- 1614.

[73] Case, D.A.; Darden, T.A. e. a. *AMBER 9*, University of California, San Francisco **2006**.

[74] Halgren, T.A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comp. Chem.*, **1996**, *17*, 490-519.

[75] Halgren, T.A. Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *J. Comp. Chem.*, **1996**, *17*, 616-641.

[76] Jorgensen, W.L.; Tirado-Rives, J. The OPLS potential functions for proteins: energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, **1988**, *110*, 1657-1666.

[77] The Cambridge Structural Database. Available from: http://www ccdc.cam.ac.uk/products/csd/

[78] Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell, A.D. Jr. CHARMM general force field (CGenFF): a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comp. Chem.*, **2010**, *31*.

[79] Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.*, **1999**, *314*, 141-151.

[80] Rhee, Y.M.; Pande, V.S. Multiplexed-replica exchange molecular dynamics method for protein folding simulation. *Biophys. J.*, **2003**, *84* (2 Pt 1), 775-786.

[81] Metropolis, N.; Rosembluth, A.; Rosembluth, M.; Teller, A. Equation of state calculations by fast computing machins. *J. Chem. Phys.*, **1953**, *21*, 1087-1092.

[82] Allen, M.P.; Tildesley, D.J. *Computer simulation of liquids*. Clarendon Press: Oxford, **7**; p 385 p.

[83] Im, W.P.; Lee, M.S.; Brooks, C.L. Generalized born model with a simple smoothing function. *J. Comput. Chem.*, **2003**, *24*, 1691- 1702.

[84] Lee, M.S.; Feig, M.; Salsbury, F.R.; Brooks, C.L. New analytic approximation to the standard molecular volume definition and its application to generalized born calculations. *J. Comput. Chem.*, **2003**, *24*, 1348-1356.

[85] Lee, M.S.; Salsbury, F.R.; Brooks, C.L. Novel generalized born methods. *J. Chem. Phys.*, **2002**, *116*, 10606-10614.

[86] Akaike, H. Likeliood of a model and information Criteria. *J. Economet.*, **1981**, *16*, 3-14.

[87] McQuarrie, A.D.R.; Tsai, C.L. *Regression and Time Series Model Selection*. World Scientific: Singapore, **1998**.

[88] Gilbert, K.M.; Boos, T.L.; Dersch, C.M.; Greiner, E.; Jacobson, A.E.; Lewis, D.; Matecka, D.; Prisinzano, T.E.; Zhang, Y.; Rothman, R.B.; Rice, K.C.; Venanzi, C.A. DAT/SERT selectivity of flexible GBR 12909 analogs modeled using 3D-QSAR methods.
Bioorg. Med. Chem., **2007**, *15*, 1146-1159..

[89] Mallik, B.; Morikis, D. Development of a quasi-dynamic pharmacophore model for anti-complement peptide analogues. *J. Am. Chem. Soc.*, **2005**, *127*, 10967-10976.

[90] Kalaszi, A.; Imre, G.; Jakli, I.; Farkas, O. Identification of the bioactive conformation for mucin epitope peptides. *J. Mol. Struct., Theochem*, **2007**, *823*, 16-27.

[91] Lexa, K.W.; Alser, K.A.; Salisburg, A.M.; Ellens, D.J.; Hernandez, L.; Bono, S.J.; Michael, H.C.; Derby, J.R.; Skiba, J.G.; Feldgus, S.; Kirschner, K.N.; Shields, G.C. The search for low energy conformational families of small peptides: Searching for active conformations of small peptides in the absence of a known receptor. *Int. J. Quantum Chem.*, **2007**, *107*, 3001-3012.

[92] Kirschner, K.N.; Lexa, K.W.; Salisburg, A.M.; Alser, K.A.; Joseph, L.; Andersen, T.T.; Bennett, J.A.; Jacobson, H.I.; Shields, G.C. Computational design and experimental discovery of an antiestrogenic peptide derived from alpha-fetoprotein. *J. Am.*
Chem. Soc., **2007**, *129*, 6263-6268.
[93] Bodnar, R.J. Endogenous opiates and behavior: *Peptides*, **2007**, *28*, 2435-2513.
[94] Remy, I.; Wilson, I.A.; Michnick, S.W. Erythropoietin receptor activation by a ligand-induced conformation change. *Science*, **1999**, *283*, 990-99
[95] Kunishima, N.; Shimada, Y.; Tsuji, Y.; Sato, T.; Yamamoto, M.; Kumasaka, T.; Nakanishi, S.; Jingami, H.; Morikawa, K. Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor
[96] Humphries, M.J. Integrin activation: the link between ligand binding and signal transduction. *Curr. Opin. Cell Biol.*, **1996**, *8*, 632-640.