Doctoral Thesis

Development and application of new methods for the virtual screening of chemical databases

Author(s):
Bissantz, Caterina

Publication Date:
2002

Permanent Link:

https://doi.org/10.3929/ethz-a-004447026 →

Dissertation ETH No. 14771

Development and Application of new Methods for the Virtual Screening of Chemical Databases

A dissertation submitted to the
Swiss Federal Institute of Technology Zurich
for the degree of
Doctor of Natural Sciences

presented by

Caterina Bissantz

Pharmacist (approbierte Apothekerin)
University of Freiburg, Germany
Master of Science
Purdue University, USA

born June 15th, 1973
citizen of Germany

accepted on the recommendation of

Prof. Dr. Gerd Folkers, examiner
PD Dr.Didier Rognan, co-examiner
Prof. Dr. Leonardo Scapozza, co-examiner

2002

Zusammenfassung

Strategien zur Wirkstofforschung haben sich in den letzten Jahren drastisch verändert. Vor 20 Jahren bestand das Konzept vor allem daraus, Verbindungen, die ein Chemiker gemäss seiner Erfahrung und chemischen Intuition ausgewählt hat, zu synthesisiern und zu testen. Heutzutage sind die heutigen Strategien rationeller, indem sie Wissen über die molekulare Struktur des Zielproteines mit einbeziehen. Eine wichtige rationelle Methode zur Entdeckung neuer Leitstrukturen ist das "Virtuelle Screening". Virtuelles Screening ist ein Computermethode, die aus einer Datenbank all diejenigen Verbindungen heraussucht, die mit hoher Wahrscheinlichkeit an das Zielprotein binden. Virtuelles Screening wählt die Moleküle entweder gemäss ihrer Komplimentarität zu einer bestimmten Bindungstasche aus (Protein-basiertes Virtuelles Screening), oder gemäss ihrer Fähigkeiten, bestimmten Pharmakophoranforderungen zu genügen, die entweder über die Bindungstasche oder mit Hilfe bekannter Liganden definiert wurden (Pharmakophor-basiertes Virtuelles Screening).

Das Protein-basierte virtuelle Screening steht im Zentrum der hier vorgestellten Arbeit. Sie beinhaltet die Bewertung und Anwendung gegenwärtig erhältlicher Screeningprogramme, sowie die Entwicklung von Methoden, die die Anwendung von virtuellem Screening auf G-Protein-gekoppelte Rezeptoren ermöglichen.

In einer ersten Studie untersuchten wir, was man von einem virtuellen Screening erwarten kann. Dabei versuchten wir vor allem, folgende Fragen zu beantworten: Welches Dockingprogramm and welche Bewertungsfunktionen sollte man für das Screening eines neuen Proteins wählen? Hängt das beste Dockingprogramm/die beste Bewertungsfunktion

von der Art der Bindungstasche (Grösse, physiko-chemische Eigenschaften…) des Zielproteines ab? Und welche Trefferquote darf man von einem Screening erwarten? In dieser Studie vewendeten wir die Thymidinkinase des Herpes Simplex Virus Typ 1 (TK) und den Estrogenrezeptor α (ER) als Testproteine. Wir benützten hochauflösende Röntgenkristallstrukturen dieser Proteine für das Screening von Testdatenbanken, die bekannte TK- und ER-Liganden enthielten, wobei wir drei verschiedenen Dockingprogramme (Dock, Gold und FlexX) und sieben verschiedene Bewertungsfunktionen anwendeten. Wir konnten zeigen, dass man im Allgemeinen eine Trefferquote von 25 bis 70 % erwarten kann. Wir konnten weiterhin bestätigen, dass man mit "Consensus Scoring" bessere Resultate erhalten kann als mit der Anwendung von nur einer einzelnen Bewertungsfunktion. Es ist offensichtlich, dass es nicht eine grundsätzlich "beste" Kombination von Dockingprogramm und Bewertungsfunktion(en) gibt. Welche Kombination am besten ist, hängt vielmehr von der Art der Bindungstasche ab. Es erscheint uns jedoch relativ schwierig, allgemeine Regeln aufzustellen, mit denen man für ein bestimmtes Ziel die beste Kombination vorhersagen kann. Man sollte daher bei jedem Screening zuerst eine Testdatenbank, die bekannte Liganden und zufällig ausgewählte Verbindungen enthält, screenen, um die beste Kombination zu bestimmen.

Wir haben diese Screeningstrategie angewendet, um Liganden für humanes Serumalbumin (HSA) zu finden. Wir haben die für ein virtuelles HSA-Screening beste Kombination von Dockingprogramm und Bewertungsfunktion(en) über ein 2-stufiges Validierungsverfahren bestimmt. Zuerst bewerteten wir die Dockingprogramme Dock, Gold und FlexX in Kombination mit sieben Bewertungsfunktionen nach ihrer

Fähigkeit, Farbstoffe, die an HSA binden, von Farbstoffen, die zwar eine ähnliche Struktur haben, aber dennoch nicht an HSA binden, zu unterscheiden. In der zweiten Validierungsstufe testeten wir die Programme dann über das Screening einer Testdatenbank, die bekannte HSA Liganden und zufällig ausgewählte Verbindungen enthielt. Aufgrund der Resultate der Validierung entschieden wir, FlexX als Dockingprogramm und FlexX und Dock als Bewertungsfunktionen zu verwenden. Da HSA zwei Bindungsstellen für kleine organische Moleküle besitzt, führten wir auch zwei getrennte Screenings durch. Von den 33 experimentell getesteten virtuellen Hits zeigten 12 wirklich Affinität zu HSA (Trefferquote: 36.4 %). Diese Hits binden mit mikromolarer Affinität. Von drei dieser Hits testeten wir noch einige Analoga. Diese zeigten teilweise eine noch höhere Affinität.

Die weiteren hier vorgestellten Studien beschäftigen sich mit G Protein-gekoppelten Rezeptoren (GPCR). GPCRs stellen eine Familie von Membranproteinen mit grosser pharmazeutischer Bedeutung dar. Bisher ist nur von einem einzigen Mitglied dieser Familie (bovines Rhodopsin) eine hochauflösende Röntgenstruktur gelöst worden. Die Anwendung von virtuellem Screening auf GPCRs stellt daher eine besondere Herausforderung dar, da man hier mit Homologiemodellen anstelle von hochauflösenden Röntgenstrukturen arbeiten muss. Bevor wir virtuelles Screening wirklich auf einen bestimmtes GPCR anwendeten, führten wir eine Referenzstudie durch, um zu untersuchen, ob GPCR Homologiemodelle sich wirklich für virtuelles Screening eignen. Wir haben dabei zwischen einem Screening für neue Antagonisten, bei dem man ein Model eines "Antagonisten-gebundenen" Rezeptorzustandes benutzt, und einem Screening für neue Agonisten, bei dem man ein Model

eines "Agonisten-gebundenen" Rezeptorzustandes benutzt, zu unterscheiden. Beide Modelarten wurden mit bovinem Rhodopsin als Vorlage konstruiert. In unserer Studie benutzten wir fünf humane Testrezeptoren (Dopamin D3 Rezeptor, Muscarin M1 Rezeptor, Vasopressin V1a Rezeptor, β2-Adrenozeptor, δ-Opioidrezeptor). Wir konnten zeigen, dass sowohl das Screening für GPCR Antagonisten als auch für GPCR Agonisten möglich ist. Die erhaltenen Trefferquoten liegen im gleichen Bereich wie diejenigen, die wir in unserer ersten Studie erzielt haben, in der wir hochauflösende Röntgenstrukturen als Zielstrukturen benutzten. Um geeignete "Antagonisten-gebundenen" Modelle zu erhalten, reicht es aus, das ursprüngliche, von Rhodopsin abgeleitete Model mit einem bekannten Antagonisten in der Bindungstasche zu minimieren. Ein geeignetes "Agonisten-gebundenes" Model zu erhalten, ist schwieriger, da Rhodopsin im inaktivierten Zustand kristalliert worden ist. Man muss daher bei der Modelkonstruktion die Konformationsänderungen, die bei der Rezeptoraktivierung stattfinden, nachahmen. Wir haben dafür eine Prozedur entwickelt, die eine manuelle Rotation von Helix 6 um die eigene Achse und eine anschliessenden Minimierung mit mehreren bekannten Agonisten in der Bindungstasche beinhaltet.

Nachdem wir die Eignung von GPCR Homologie Modellen für virtuelles Screening bestätigt hatten, haben wir diese Strategie auf den lysophosphatischen Rezeptor LPA_1 (EDG2) angewendet. Es wird vermutet, dass dieser Rezeptor eine Rolle bei der Myelinbildung und damit bei Erbkrankheiten, bei denen eine Myelindysfunktion vorliegt, spielt. Neue EDG2 Liganden könnten daher nützliche Tools sein, um die Funktion dieses Rezeptors weiter zu untersuchen. Wir generierten eine Liste von 50 virtuellen Hits, die jetzt noch experimentel überprüft werden müssen.

Durch die Sequenzierung des humanen Genoms sind die Aminosäuresequenzen mehrerer Hundert GPCRs bekannt. Damit wir dreidimensionelle Modelle von all diesen GPCRs erstellen können, ist es notwendig, den Modellierungsprozess zu automatisieren. Wir haben daher ein Programm (GPCRalign) entwickelt, das automatisch die Aminosäuresequenzen der sieben transmembranen Domainen (TM) überlagert. Von einer gegebenen Aminosäuresequenz ausgehend, berechnet das Programm zuerst die ungefähre Lage der sieben transmembranen Domainen voraus. Die Sequenz wird dann einer GPCR Familie zugeordnet (Rhodopsin-ähnliche Rezeptoren, Calcitonin-ähnliche Rezeptoren, metabotropic Glutamate-ähnliche Rezeptoren,…). Dafür werden konservierte, familienspezifische Aminosäuren verwendet. Das Programm sucht entweder direkt nach dem Vorhandensein dieser charakteristischen Aminosäuren in Form von ‘pattern’, oder aber sie werden zuerst in sogenannte ‘position-specific scoring matrices’ (PSSM) umgewandelt. Dazu werden multiple Überlagerungen von kurzen Sequenzregionen, die diese charakteristischen Aminosäuren enthalten (motifs), verwendet. Sobald die Familie bestimmt ist, können die TMs überlagert werden. Wenn eine TM charakteristische Aminosäuren enthält, kann dies über die Lage dieser Aminosäuren erreicht werden. Bei allen anderen TMs, die keine charakteristische Aminosäuren enthalten, wird stattdessen ein Algorithmus verwendet, der Blosum Matrices zur Bewertung der möglichen Überlagerungen benutzt. Wir haben bisher insgesamt 236 humane GPCRs überlagert (205 rhodopsin-ähnliche Rezeptoren, 20 secretin- ähnliche Rezeptoren and 11 metabotropic glutamate- ähnliche Rezeptoren).

Zusammenfassend ist die vorliegende Arbeit der Untersuchung des Leistungsvermögens und der Anwendung von zur Zeit erhältlichen

Screeningprogrammen gewidmet sowie der Ausweitung ihrer Anwendung von hochauflösenden Röntgenstrukturen zu GPCR Homologiemodellen, was zum ersten Mal rationellen GPCR Ligandendesign ermöglicht. Zudem können die hier gesammelten Informationen auch die Anwendung von virtuellem Screening auf andere Proteinfamilien, für die nur Homologiemodelle zur Verfügung stehen, unterstützen.

Summary

Drug discovery strategies have undergone a radical change in the last years. 20 years ago, a common approach was to synthesize and test compounds that have been selected based on the chemical intuition and knowledge of the medicinal chemist. The strategies are today more rationl', using knowledge of the molecular target. One important approach to rational lead finding is virtual screening (VS). Virtual screening is a computational method to select from a chemical database those compounds that are the most likely to bind to a target protein. Virtual screening selects compounds either by their complementary to a defined active site (protein-based virtual screening) or by their fulfilling specific pharmacophoric requirements that can be defined using the structure of the receptor or known ligands (pharmacophore-based virtual screening).

The herein presented work is focused on protein-based virtual screening, involving the evaluation and application of currently available virtual screening tools, and the development of methods that enable us to carry out virtual screening against G protein-coupled receptors.

In a first study, we investigated what can be expected from a virtual screening, trying to answer especially the following questions: Which docking tool and scoring function(s) should be chosen when setting up a virtual screening for a new target protein? Does the best docking tool/scoring function depend on the type of active site (size, physicochemical properties etc) to be screened? And which hit rates can we expect from a virtual screening? In this study, we used the herpes simplex virus type 1 thymidine kinase (TK) and the estrogen receptor α (ER) as test cases. Screening test databases containing known TK and ER ligands and

random compounds against high-resolution X-ray structures of the proteins using three docking programs (Dock, Gold and FlexX) and seven scoring functions, we could show that we can generally expect to achieve hit rates between 25 and 70%. We could furthermore confirm that consensus scoring gives better results than single scoring. It is obvious that there is not one 'best' combination of docking tool and scoring functions, but the best performing combination depends on the type of binding site to screen. Since it seems rather difficult to derive general rules to predict the best combination for a target, the docking/scoring combination used for a screening should generally be validated by first screening a test database including known ligands against the protein.

We applied this screening strategy to the discovery of high affinity ligands for human serum albumin (HSA). The best docking/scoring combination was first evaluated using a two-step validation protocol. The docking programs Dock, Gold and FlexX in combination with seven different scoring functions were analyzed in terms of their ability to discriminate between dye compounds that bind to HSA and dye compounds that have a similar structure but nevertheless do not bind HSA. In the second validation step, we then screened a test database containing known ligands and random compounds. Based on this validation, we decided to use FlexX as docking tool in combination with the FlexX and Dock scoring function functions for rescoring. Since HSA has two binding sites for small molecules, separated screenings were carried out. Experimental testing of 33 virtual hits confirmed 12 of the compounds as true hits, giving a hit rate of 36.4 %. These hits bind in the low micromolar range. Testing of analogs of three of the hits allowed us to discover further compounds with even slightly higher affinity.

The other studies presented here involve G protein-coupled receptors (GPCR), a family of membrane proteins with high pharmaceutical importance. There is only a high-resolution X-ray structure of one GPCR (bovine rhodopsin) solved. The application of virtual screening to GPCRs is therefore especially challenging since it has to be carried out using homology models instead of high-resolution X-ray structures. Before applying virtual screening to a specific GPCR, we thus carried out a reference study to investigate if it is really possible to use homology models as targets for VS. We have hereby to distinguish between screening for antagonists using a model of an 'antagonist-bound' receptor state and screening for new agonists using a model of an 'agonist-bound' receptor state. Both type of models were constructed using bovine rhodopsin as template. Using five human test receptors (dopaminergic D3 receptor, muscarinic M1 receptor, vasopressin V1a receptor, β2-adrenergic receptor, δ-opioid receptor), we could show that both screening for GPCR antagonist and GPCR agonists is indeed feasible. We could obtain hit rates in the same range as in our first study using high-resolution X-ray structures as targets. To obtain suitable 'antagonist-bound' models, it was sufficient to refine the initial rhodopsin-based model by minimization with a known antagonist docked into the binding site. To obtain suitable 'agonist-bound' models was more difficult since the rhodopsin template has been crystallized in its inactivated state. We therefore have to somehow simulate the conformational changes taking place during receptor activation. This was achieved by a new modeling procedure that involves a manual rotation of helix 6 around its helical axis followed by a minimization with several different agonists docked into the binding pocket.

Having shown that GPCR homology models are indeed suitable for virtual screening, we applied this strategy to the lysophosphatidic receptor LPA_1 (EDG2). This receptor might play a role in myelination and thus in inherited diseases in which myelin dysfunctions are present. New EDG2 ligands could be useful tools to further investigate the role of this receptor. From our screening a list of 50 virtual hits was obtained. It remains to experimentally verify their affinity.

Due to the sequencing of the human genome, there are the amino acid sequences of several hundreds GPCRs known. To enable us to generate three-dimensional models of all these GPCRs, it is necessary to develop tools that automatize the modeling process. We consequently developed a new program (GPCRalign) to automatically align the amino acid sequences of the seven transmembrane domains (TM). Starting from a given amino acid sequence, the program first predicts the rough locations of the seven transmembrane helices. The sequence is then assigned to one of the GPCR superfamilies (rhodopsin-like receptors, calcitonin-like receptors, metabotropic glutamate-like receptors, …) using highly conserved, family-specific amino acids. These highly characteristic amino acids are either directly searched for in form of patterns, or they are first transformed into position-specific scoring matrices (PSSM) using a multiple sequence alignment of short sequence regions that include these conserved amino acids (motifs). Once the superfamily is determined, the transmembrane domains can be aligned. All TMs bearing characteristic amino acids can be aligned using the relative location of the conserved amino acids. For all other TMs, for which no such characteristic was found, an algorithm that uses Blosum matrices for scoring putative alignments is used. Until now,

we have aligned 236 human GPCRs (205 rhodopsin-like receptors, 20 secretin-like receptors and 11 metabotropic glutamate-like receptors).

In conclusion, the presented work was dedicated to investigate the capabilities of the currently available virtual screening tools and to expand its application from high-resolution X-ray structures to homology models of GPCRs what opens the door to rational GPCR ligand design. This might also help to use virtual screening for other protein classes for which only homology models are available.