

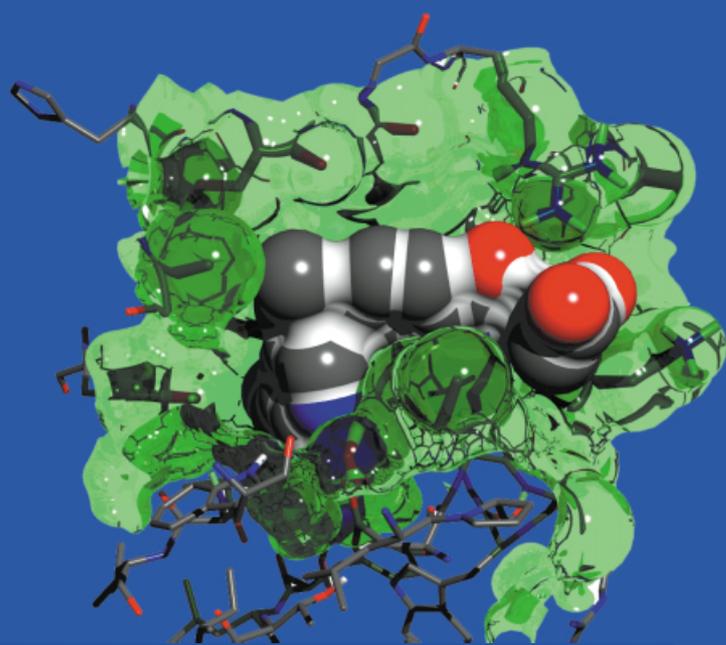
BIOforum

25. Jahrgang
Dezember 2002
Seiten 818–820

12

FORSCHUNG & ENTWICKLUNG

S O N D E R D R U C K



PROF. DR. MATTHIAS RAREY

Wirkstoffentwurf und Informatik

GIT VERLAG

A Wiley Company
www.gitverlag.com

Wirkstoffentwurf und Informatik



Matthias Rarey

Die Entwicklung neuer Medikamente ist eine komplexe Herausforderung von besonderer gesellschaftlicher und ökonomischer Bedeutung. Dass dieser Entwicklungsprozess vom Computer begleitet wird, ist aufgrund der enormen Datenmengen

selbstverständlich, hingegen ist die wissenschaftliche Rolle der Informatik nicht immer offensichtlich. Hinter den Kulissen von Computeranwendungen im Wirkstoffdesign zeigt sich aber, dass eine computergerechte Modellierung und bedarfsgerecht entwickelte Algorithmen die tragenden Säulen des computergestützten Wirkstoffdesigns sind.

Der Computer ist heute fester Bestandteil in vielen Bereichen der pharmazeutischen Forschung. Nur mit Hilfe von komplexen Datenbanken und Analyseverfahren ist man heute in der Lage, die Informationsflut aus modernen Hochdurchsatz-Experimenten der Biomedizin zu verwalten und diese für die weitere Forschung nutzbar zu machen. Die Informatik liefert das technische Handwerkszeug zur Bewältigung dieser Aufgaben. Vor allem bei der Entwicklung von computergestützten Analyse- und Vorhersageverfahren zeigt sich, dass neben der bedarfsgerechten Modellierung der Probleme auch maßgeschneiderte Algorithmen zur effizienten Lösung notwendig sind.

Algorithmische Probleme des computergestützten Wirkstoffentwurfs

Bei der Suche nach neuen Wirkstoffen kann man anhand der verfügbaren Information über das Zielprotein grob zwischen drei verschiedenen Szenarien unterscheiden. Im Idealfall ist die dreidimensionale Struktur des Proteins mittels Röntgenkristallographie oder NMR-Spek-

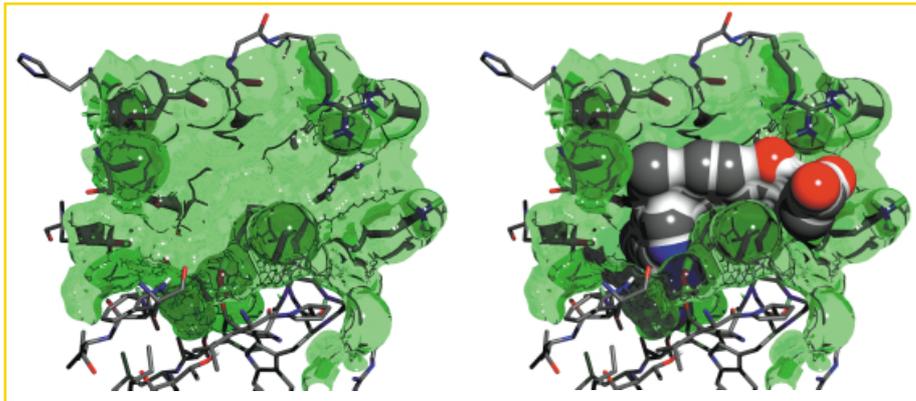


Abb. 1: Protein-Ligand-Komplex: Auf der linken Seite ist das aktive Zentrum des Proteins Dihydrofolat-Reduktase (DHFR) mit seiner molekularen Oberfläche dargestellt. Rechts sieht man den Komplex des Proteins mit dem Inhibitor Methotrexat.

trokopie aufgeklärt. Da sich eine Vielzahl von pharmazeutisch relevanten Zielproteinen der Strukturaufklärung entziehen, kennt man allerdings häufig nur niedermolekulare Verbindungen, die mit dem Zielprotein interagieren. Schließlich kann man den Fall, dass man keine Information über das Zielprotein auf molekularer Ebene hat, als drittes Szenario ansehen.

In allen drei Szenarien ergeben sich vielfältige Möglichkeiten zum computergestützten Wirkstoffentwurf und damit auch viele algorithmisch interessante Probleme [1]. Verschiedene Bereiche der Algorithmik wie zum Beispiel die Graphentheorie, die kombinatorische Optimierung oder die algorithmische Geometrie liefern Bausteine zur Lösung dieser Probleme.

Wirkstoffentwurf und Algorithmische Geometrie: Molekulares Docking

Bei Verfügbarkeit der dreidimensionalen Struktur des Proteins stellt die Vorhersage von molekularen Komplexen das zentrale Problem des computergestützten Wirkstoffentwurfs dar (molekulares Docking) [2]. Nur wenn man Interaktionen zwischen potentiellen Wirkstoffen und dem Zielprotein richtig vorhersagen kann, ist man in der Lage, die Strukturinformation über das Protein auch wirklich zu nutzen. Ob nun die Bindungsaffinität einer einzelnen Verbindung vorhergesagt, ein virtuelles Screening durchgeführt oder eine fokussierte kombinatorische Bibliothek entwickelt werden soll, in jedem Fall muss die Geometrie und Bindungsaffinität von Protein-Ligand-Komplexen so gut wie möglich prognostiziert werden. Gleichzeitig müs-

sen die notwendigen Computerressourcen, insbesondere die Rechenzeit, in Grenzen gehalten werden.

Die Grundlage zur Softwareentwicklung für komplexe Probleme wie molekulares Docking ist die Modellierung. Die Freiheitsgrade des Problems (relative Orientierung, Konformationsräume des Liganden, ggf. auch des Proteins) als auch die Zielfunktion, d. h. die energetischen Verhältnisse (geometrische Charakterisierung von molekularen Wechselwirkungen, Energiefunktionen), müssen in computergerechter Form beschreiben werden. Die im Institut SCAI entwickelte Docking-Software FlexX [3, 4] basiert auf einer diskreten Modellierung des Konformationsraums des Liganden. Jeder drehbaren Bindung wird auf der Basis von experimentellen Daten eine Menge von bis zu 12 möglichen Torsionswinkeln zugeordnet. Zur Beschreibung molekularer Wechselwirkungen wird ein geometrisches Modell auf der Basis von partiellen

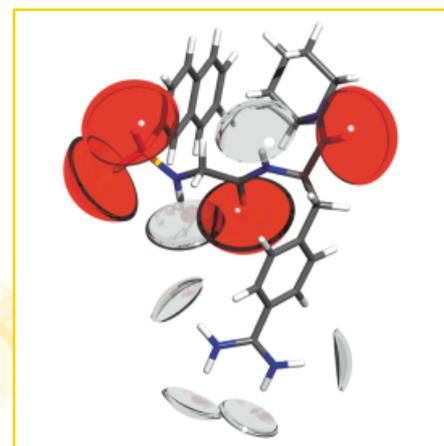


Abb. 2: Modellierung des Docking-Problems: Wasserstoffbrückenbindungsregionen an einem Liganden

Keywords

Wirkstoffdesign, Bioinformatik, Chemoinformatik, Software, molekulares Docking, molekulare Ähnlichkeit

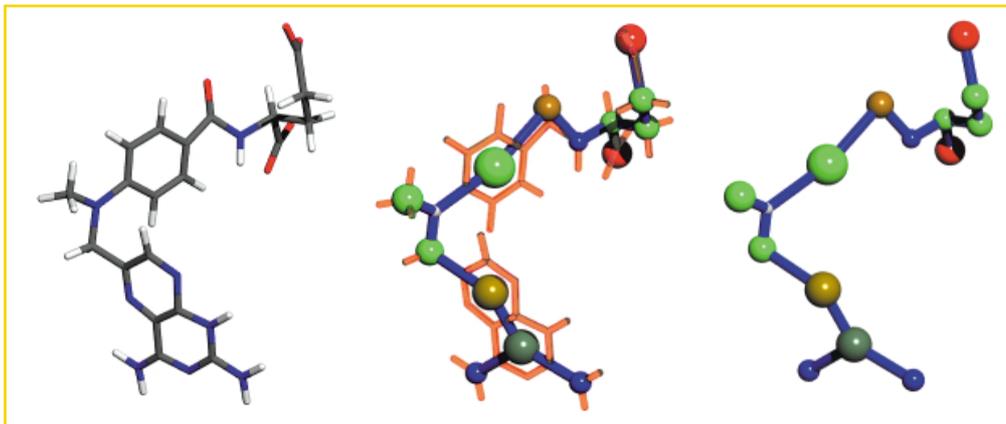


Abb. 3: Molekül (links) und die Umsetzung in einen Feature Tree (rechts). Jeder Baumknoten stellt eine (funktionale) Gruppe des Moleküls dar. Physiko-chemische Eigenschaften der Gruppen werden durch entsprechende Knotenmarkierungen repräsentiert.

Kugelflächen verwendet (Abb. 2). Die energetische Bewertung erfolgt durch eine einfache, schnell auswertbare Scoring-Funktion, die die wesentlichen Energiebeiträge (Salz- und Wasserstoffbrücken, hydrophobe Kontakte, Entropie) numerisch erfasst [5].

Auf der Basis dieser Modellierung lässt sich molekulares Docking als ein geometrisches Optimierungsproblem beschreiben. Zur Lösung verwendet FlexX eine inkrementelle Aufbaustrategie: Der Ligand wird zunächst in kleinere Fragmente zerlegt und einzelne Fragmente werden in das aktive Zentrum des Proteins platziert. Aufbauend auf diese initialen Platzierungen werden weitere Fragmente sukzessive in energetisch günstiger Weise angebaut bis das Molekül vollständig im aktiven Zentrum des Proteins eingepasst ist. Methoden aus den klassischen Teilgebieten der Algorithmik spielen in allen Phasen des Dockingverfahrens eine wichtige Rolle. So werden Graphalgorithmen bei der Fragmentierung der Moleküle eingesetzt, ein Mustererkennungsalgorithmus bei der Fragmentplatzierung und geometrische Datenstrukturen zur Identifikation auftretender molekularer Interaktionen und Überlappungen beim inkrementellen Komplexaufbau.

In der Summe ergibt sich so eine Software für molekulares Docking, die im Minutenbereich Protein-Ligand-Komplexe und Bindungsaffinitäten vorhersagen kann. In ca. 2/3 der Fälle entsprechen die vorgeschlagenen Komplexstrukturen experimentellen Daten, die Genauigkeit der Affinitätsvorhersage ist hoch genug, um signifikante Anreicherungen (Faktor 5–10 über eine randomisierte Auswahl) beim virtuellen Suchen nach neuen Wirkstoffen zu erzielen.

Wirkstoffentwurf und Kombinatorische Optimierung: Molekulare Ähnlichkeit

Steht die 3D-Struktur des Zielproteins nicht zur Verfügung, so stellt die molekulare Ähnlichkeit zwischen niedermoleku-

laren Verbindungen das zentrale Entwurfswerkzeug für Wirkstoffe dar. Computerverfahren, die Moleküle auf Ähnlichkeit analysieren, gibt es sehr viele, das Spektrum reicht vom einfachen Vergleich physiko-chemischer Moleküleigenschaften über Beschreibungen der Molekülstruktur durch Bitstrings bis zur räumlichen Überlagerung von Molekülen [6, 7].

Wie schon beim molekularem Docking spielt die Modellierung bei der Ähnlichkeitsanalyse eine entscheidende Rolle. Die Software Feature Trees [8] geht hier einen unkonventionellen Weg. Moleküle werden durch eine Baumstruktur beschrieben, die gegenüber den herkömmlichen linearen Deskriptoren den Vorteil besitzt, dass sie die relative Anordnung physiko-chemischer Gruppen viel besser konserviert. Beim Vergleich zweier Moleküle, die durch ihre Feature Trees repräsentiert werden, ergibt sich aber ein nicht triviales, algorithmisches Problem: Welcher Teil des Baumes soll mit welchem Teil des anderen verglichen werden? Dieses Matching von Teilbäumen unter Maximierung der molekularen Ähnlichkeit lässt sich als kombinatorisches Optimierungsproblem beschreiben und mit Hilfe eines klassischen Paradigmas aus der Algorithmik, der dynamischen Programmierung, effizient lösen. In nur wenigen Millisekunden können so Moleküle mit höherer Genauigkeit verglichen werden als mit den klassischen linearen Deskriptoren.

Als weiterer Vorteil des Feature Tree Deskriptors ist seine Kompatibilität zur kombinatorischen Chemie zu nennen. Weiterentwicklungen der Software ermöglichen die vollständige Suche in virtuellen kombinatorischen Bibliotheken mit mehr als 10^{18} Molekülen in wenigen Minuten auf einer einfachen Workstation [9].

Modellierung und Algorithmenentwurf aus einer Hand

Molekulares Docking und die Beschreibung molekularer Ähnlichkeit sind zwei

typische Beispiele für algorithmische Probleme aus dem computergestützten Wirkstoffentwurf. Beiden Problemen ist gemein, dass zu ihrer Bearbeitung eine Mischung aus komplexen Modellen und effizienten Algorithmen notwendig ist. Die Modellierung spielt dabei eine besondere Rolle. Sie muss zum einen nahe genug an den realen Sachverhalten liegen, um überhaupt korrekte Vorhersagen zu ermöglichen. Auf der anderen Seite muss sie aber auch die Entwicklung effizienter Algorithmen ermöglichen. Eine weitere Gemeinsamkeit der Probleme ist, dass sich die Softwareentwicklung zwar auf bekannte Algorithmen stützt, diese aber in sehr spezifischer Weise adaptiert und weiterentwickelt werden müssen, um in der praktischen Anwendung erfolgreich zu sein. Die Modellierung der Anwendung und der Algorithmenentwurf sind somit eng miteinander verwoben und begründen damit die Notwendigkeit einer interdisziplinären Bio- und Chemoinformatik.

Literatur

- [1] Rarey, M.: Algorithmen für den computergestützten Wirkstoffentwurf, 2002, Habilitationsschrift

Weitere Literatur beim Autor

Prof. Dr. Matthias Rarey

forscht und lehrt am Zentrum für Bioinformatik der Universität Hamburg (www.zbh.uni-hamburg.de); Mitgründer der Firma BioSolveIT GmbH (Sankt Augustin; www.biosolveit.de). Studium der Informatik an der UGH Paderborn, Promotion (1996) und Habilitation (2001) in Informatik an der Universität Bonn während seiner Forschungstätigkeit am Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen (SCAI). Seine Habilitationsschrift „Algorithmen für den computergestützten Wirkstoffentwurf“ wurde mit dem NRW Wissenschaftspreis 2002 ausgezeichnet.

Zentrum für Bioinformatik
Universität Hamburg
Institut für Algorithmisches Molekulares Design
Bundesstraße 43
20146 Hamburg
rarey@zbh.uni-hamburg.de