

Modellierung von Protein-Ligand Komplexen als Basis einer rationalen Arzneistoffentwicklung

Gerd Folkers

Departement Pharmazie, ETH, CH-Zürich

Zusammenfassung

Die Entdeckung fast unseres gesamten Arzneimittelschatzes ist die Folge von glücklichen Verkettungen von Zufällen auf den Wegen moderner, aber auch antiker Forschung. Die englische Sprache verwendet dafür den Ausdruck Serendipity. Nichtsdestoweniger haben sich Arzneimittel im Laufe der Zeit drastisch gewandelt. Vom Kräutertee der Vorfahren sind wir zu sehr stark wirksamen Monopräparaten als Therapeutika gekommen. Von einer beträchtlichen Zahl dieser Monopräparate haben wir durch moderne analytische und molekularbiologische Verfahren gelernt, an welchen Orten sie wirken und was sie auf molekularer Ebene tun. In der Regel löst die Wechselwirkung mit einem körpereigenen Eiweiß spezieller Bauart, dem Rezeptor, die biologische Antwort aus. Dies führt zu der Vorstellung, daß man Arzneistoffe mit vorberechneten Eigenschaften, wie auf dem Reißbrett, entwerfen können sollte, wenn man Struktur und Eigenschaften des Arzneistoff-Rezeptor-Komplexes besäße. Die Serendipity bei der Suche und Entwicklung von Arzneistoffen ließe sich sehr reduzieren. Damit würden auch weniger Substanzen einer tierexperimentellen Testung unterworfen.

Die Praxis zeigt, daß tatsächlich eine Kombination von biophysikalischen, theoretischen und biologischen in vitro Methoden uns dem Entwurf von Arzneistoffen sehr nahe bringen. Es gibt aber Probleme grundlegender Art, die dem reduktionistischen Ansatz in der Arzneimittelforschung entgegenstehen. Sie verhindern, daß auf dem Computerbildschirm durch eine einfache Schlüssel-Schloß Einpassung eines Wirkstoffmoleküls in eine Rezeptorstruktur ein Arzneimittel entsteht.

Ist jedoch einmal ein Wirkstoff gefunden, sei es durch reinen Zufall, durch Serendipity, Design oder ethnologische Heiltradition, können die theoretischen Modellierungsverfahren wesentlich zu einer Optimierung und zum Ausbau der ganzen therapeutischen Klasse beitragen.

Summary: Modeling of protein-ligand interaction complexes as a basis of the rational design of drugs.

Most of our today's drugs have been discovered by the lucky coincidences called serendipity. The drugs itself however, have tremendously changed since the ancient times. From the herbal teas as the remedies of our ancestors, we have switched to modern therapeutics, mostly containing only one, very high active compound. Modern analytical methods and the approaches of recent molecular biology have revealed the molecular mechanisms of a lot of this high active drugs. In most of the cases, interaction of the drug molecule with a special protein in the body, located on the surface of the cell releases the biological response.

This may lead to the idea, that knowledge of the structure of the protein-ligand complex enables us to design a drug with certain properties on the computer screen. If successful, serendipity could be considerably reduced as well as animal testing.

As a matter of fact, it can be shown, that the combination of biophysics, modeling, and biological in vitro methods brings us very close to that goal. However, basic problems, for instance in the handling of complex systems are opposed to the reductionistic approach in the drug research. They prevent the simple application of the lock and key principle in making a drug by fitting a putative drug molecule into the binding site of a receptor.

Nevertheless, whenever a biologically active compound has been discovered, be it by pure chance, serendipity, design or ethnological pharmacy, molecular modeling or drug design will be of great help in optimizing the active compounds and enrichment of the whole new class of therapeutics.

Keywords: Computer Aided Drug Design (CADD), ligand-protein interaction, molecular modeling, drug research and development, site-directed mutagenesis, nuclear magnetic resonance spectroscopy, substance libraries, receptor, homology modeling, multiple binding mode

1 Einleitung

Arzneimittel wirken dadurch, daß sie im Körper, nachdem sie an eine bestimmte Stelle gelangt sind, dort mit den biologischen Strukturen Informationen austauschen oder diese blockieren (Abb. 1). Der Austausch der Informationen geschieht über Komplexierung des Arzneistoffmoleküls, des Liganden und des komplementären biologischen Systems, des Rezeptorproteins, der Membran, eines Enzyms oder auch der DNS. Wenn man ein neues Arzneimittel entwickeln möchte, versucht man diese Vorgänge im molekularen Detail zu verstehen. Dazu muß ein interdisziplinäres System von Theorie und Experiment existieren, das um den eigentlichen Auslöser der Arzneimittelwirkung, die Ausbildung des Protein-Ligand-Komplexes, gruppiert ist (Blundell et al., 1992) (Abb. 2).

Die Dynamik der Arzneistoff-Rezeptor-Wechselwirkung selbst ist nicht direkt beobachtbar, sie muß modelliert werden. Die Zeitspanne, in der ein Enzym einen Arzneistoff umsetzt, kann sehr unterschiedlich

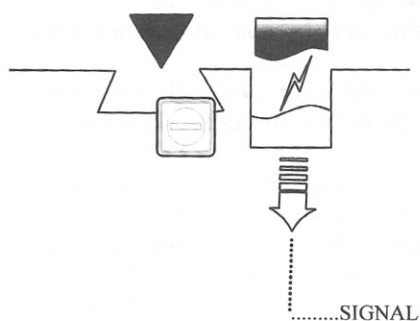


Abbildung 1: Stark schematisierte Wechselwirkung eines aktivierenden und eines blockierenden (schwarzes Dreieck) Arzneistoffs bei der Wechselwirkung an seiner biologischen Erkennungsstruktur, beispielsweise dem Rezeptor. Beide Elemente fixieren die biologische Struktur in einem bestimmten Zustand. Der aktive Zustand (Erkennung des rechten Elements) bewirkt die Weitergabe einer Information und löst über einen Signalweg eine biologische Antwort aus. Die Blockade des Signalwegs bewirkt die Unterdrückung einer biologischen Antwort und kann eine den Organismus schädigende Reaktion unterbinden.

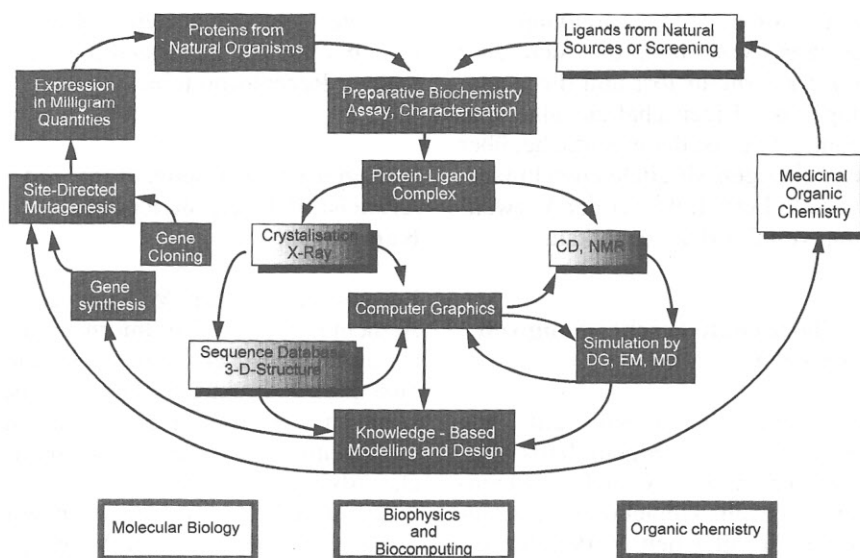


Abbildung 2: Der „Drug Design Bicycle“ beschreibt die optimale Vernetzung experimenteller und theoretischer Verfahren im rationalen Entwicklungsprozeß für eine Wirksubstanz.

sein, von einigen tausendstel Sekunden bis hinauf zu einer halben Sekunde. Darin ist nicht eingerechnet, daß selbst dann, wenn der Ligand (Arzneistoff) sich nahe dem Protein befindet, er auch noch geraume Zeit benötigt, um sich in das Protein einzupassen. Emil Fischer hat vor etwa 100 Jahren dafür das Bild der Schlüssel-Schloß-Erkennung geprägt. Dieser Vorgang wird durch Computersimulation in eine Bildersprache übersetzt, die Analogien aus dem täglichen Leben verwendet, wie beispielsweise das Bild von Schlüssel und Schloß oder Hand und Handschuh. Da die Komplexität der eigentlichen Reaktion jedoch sehr groß ist, enthalten die bildlichen Darstellungen einen hohen Abstraktionsgrad und trotzdem noch sehr viel Detailinformation. Mit Hilfe der modernen Computergraphik kann der mechanistische Ablauf der Arzneistoff-Rezeptor-Wechselwirkung in farbige, dreidimensionale und sogar bewegte Bilder umgesetzt werden. Die Technologie wird als „Molecular Modeling“ oder „Computer Aided Molecular Design“ (CAMD) bezeichnet. Der besondere Vorteil liegt aber nicht in der bildlichen Darstellung, sondern in der Modellierung. Jeder Zustand des Modellsy-

stems ist zu jedem Zeitpunkt im Computer präsent. Auf diese Weise lassen sich virtuelle Experimente durchführen.

Dadurch, daß die Simulation gerechnet wird und nicht als Film existiert, lassen sich wie in einer künstlichen Welt zu jedem Zeitpunkt die Bedingungen ändern. Nichts anderes stellt auch ein naturwissenschaftliches Experiment dar. Die CAMD-Technologie erlaubt also virtuelle Experimente über Arzneistoffwirkungen, die innerhalb des gewählten Systems Vorhersagekraft haben. Sie unterliegen natürlich den gleichen Grenzen der Vorhersagekraft, die unsere reduktionistische Beschreibung der Realität erzeugt.

Die Experimente erlauben auch eine Eigenschaftsbeschreibung der interagierenden Spezies. Der Ligand (Arzneistoff) kann in seinen chemischen Eigenschaften, wie Struktur, Ladung etc., charakterisiert werden. Dies ist nach dem, was oben gesagt wurde, dann auch für den virtuellen Liganden, ein noch nicht existierendes Molekül, möglich. Damit lassen sich neue Liganden für ein Protein im Computer entwerfen.

Eigenschaften können in sehr komplexen Oberflächen dargestellt werden. Es ergibt sich eine neue

Dimension in der Beschreibung molekularer Vorgänge, die sich eher von der Struktur löst und die Änderung von Eigenschaften als Maß benutzt. Dies ist der eigentliche, über die bisherigen Möglichkeiten hinausgehende Lerneffekt bei der Verwendung von CAMD-Techniken.

2 Eigenschaftsbeschreibung von Molekülen

Für Arzneimittelmoleküle gilt überwiegend eine Eigenschafts-Wirkungsbeziehung. Gleiche Eigenschaften sollten eigentlich reproduzierbar die gleichen Wirkungen erzeugen.

Darauf beruht die Möglichkeit einer medikamentösen Therapie. Für die gezielte Entwicklung eines neuen Wirkstoffs wäre es deshalb ideal, seine Eigenschaften und die daraus zu erwartenden Wirkungen möglichst genau kennenzulernen. Für eine Modellierung würde dieser Ansatz bedeuten, daß eine völlig deterministische Sicht der molekularen Abläufe in einer Zelle existiert. Systeme, wie die biochemischen Zyklen in einer Zelle, sind aber fast immer dissipative Systeme. Damit existieren hochgeordnete Strukturen fern vom thermodynamischen Gleichgewicht abgeschlossener Systeme. Da deshalb ein völlig deterministisches Bild eines lebendigen Organismus kaum erwartet werden kann, macht es Sinn, in Näherungen diejenigen Teile in der Funktionskette zu beschreiben, die einer chemischen, biochemischen, physikalischen und pharmakologischen Messung zugänglich sind. In vielen Fällen läßt sich daraus nach Ableitung allgemeiner Gesetzmäßigkeiten eine Simulation der wichtigen Vorgänge an theoretischen Modellen erzeugen. Methoden der modernen Molekularbiologie erlauben den Aufbau ganzer Signalketten von Rezeptoren in Mikroorganismen, wie z.B. Hefen. In dieser „artifizialen Rezeptorwelt“ lassen sich Schritt für Schritt die Signaltransduktionsvorgänge unter verschiedenen Bedingungen studie-

ren. Dies geht von gezielter Mutagenese bis hin zur Bereitstellung ganz anderer Rezeptorproteine.

3 Eigenschaftsbeschreibung von Arzneistoff-Rezeptor/Enzym-Komplexen

Bei der eigentlichen Wirkung eines Arzneimittels, also dem Informationsaustausch zwischen Arzneistoff und biologischer Struktur im Organismus, kommen viele solcher Eigenschaften gleichzeitig oder in bestimmter zeitlicher Abfolge zum Tragen.

Grundsätzliche Überlegungen zur Simulation eines solchen Vorgangs stammen aus der Ableitung der bereits erwähnten Eigenschafts-Wirkungsbeziehungen. Wenn ich eine Reihe von Strukturen, wie oben dargestellt, in ihren Eigenschaften beschrieben habe und aus pharmakologischen Versuchen ihre biologische Wirkung kenne, bin ich in der Lage, die für die Wirkung wichtigen Eigenschaften herauszukristallisieren und sie jeder einzelnen Struktur zuzuordnen.

Die Graphikcomputer erlauben es nun, diese Eigenschaften in Form molekularer Felder zu verdeutlichen. Sie bestehen nicht aus abstrakten Zahlenreihen, sondern aus farbcodierten, dreidimensionalen Gebilden, die auf den ersten Blick als charakteristisch für eine Substanz erkannt werden. Eigenschaften können sein: Größe, Chiralität, Ladung usw.

Das menschliche Fehlererkennungssystem, trainiert auf parallelen Input, ist unmittelbar in der Lage, aufgrund der graphischen Zuordnung der Eigenschaften zwischen „guten“ und „schlechten“ Molekülen zu differenzieren. Perfektioniert wird diese Differenzierung durch die Betrachtung eines Wechselwirkungskomplexes, bestehend aus Arzneistoff und biologischer Erkennungsstruktur. Die Ausbildung dieses Komplexes und eine funktionierende Informationsübertragung beruhen nämlich auf gegenseitiger Erkennung nach dem oben erwähnten Schlüssel-Schloß Prinzip. Die Eigenschaften beider

Partner müssen also zueinander komplementär sein. Diese Komplementarität oder „Paßform“ wird auf 3D-Graphiksystemen sehr gut erkannt, in der zweidimensionalen Zeichenebene ist sie unmöglich abzuschätzen.

Eine Modellierung des Protein-Ligand Komplexes als Basis einer rationalen Entwicklung eines Wirkstoffs (Drug Design Studie) ist demnach auf verschiedenen Wissensebenen möglich.

Im folgenden sollen fünf Hierarchie- oder Wissens Ebenen diskutiert werden, von denen aus ein Prozeß der rationalen Arzneistoffentwicklung beginnen kann.

Jede Ebene ermöglicht einen Entwurf, aber mit sehr unterschiedlicher Qualität bezüglich der Vorhersagekraft. Der hierarchische Unterschied zwischen den Ebenen ist durch die Integration geeigneter experimenteller Verfahren gekennzeichnet, die vorzugsweise sehr nahe am molekularen Ereignis lokalisiert sind. Spektroskopische und molekularbiologische Verfahren stehen dabei im Vordergrund.

4 Die fünf verschiedenen Hierarchie- oder Wissens Ebenen

4.1 Substrate, Inhibitoren, (Ant-) Agonisten und deren Kinetik

Diese Datenbasis ist nach wie vor fast immer der Ausgangspunkt für die Ableitung von Struktur-Funktionsbeziehungen in der Arzneistoffentwicklung. Dies ist darin begründet, daß die eigentliche Arzneistofffindung trotz aller Kenntnisse von Funktions-Strukturzusammenhängen in der Biologie ein Zufallsprozeß ist.

Kombinatorische Chemie, Strukturbibliotheken, systematische Testung (screening)

Moderne Suchstrategien verlassen sich verstärkt auf systematisches Abtasten multidimensionaler Strukturräume, die über Screening zufällig oder nahezu zufällig entstandener

großer Strukturbibliotheken zugänglich sind. Automatisierte Testverfahren, von einfachen Affinitätsmessungen bis zum funktionellen Rezeptor-assay parallel in bis zu hundert verschiedenen Therapieklassen, erlauben die biologische Analytik großer Zahlen von Strukturvarianten in relativ kurzer Zeit.

Sehr oft setzt man sich zum Ziel, die größte existierende Strukturbibliothek, die Natur, nachzuahmen. Die Anwendung genetischer Strategien in der Synthese erlaubt die Neukombination eines beschränkten Satzes von Strukturelementen, bzw. reaktiven Ausgangskomponenten unter zufällig variierenden Startbedingungen. Die Ausgangskomponenten sind in geeigneter Weise markiert, sodaß eine Rückverfolgung nach der Reaktion möglich wird. Auf diese Weise entsteht durch chemische Synthese oder durch biosynthetische Ansätze eine große Zahl (bis zu mehreren Millionen) neuer Verbindungen. Mit automatisierten *in vitro* Testsystemen lassen sich diese vielen Verbindungen in überschaubarer Zeit auf biologische Effekte testen. Viele Kriterien können schon vordergründig zum Ausscheiden von neuen Substanzen führen. Substanzen, die als Medikamente dienen, sollten beispielsweise eine gewisse Wasserlöslichkeit aufweisen, sie müssen sich in einem bestimmten Fenster des Verteilungskoeffizienten zwischen Fettphase und Wasserphase aufhalten. Häufig weisen sehr lipophile Verbindungen starke biologische Effekte auf. Sie wirken sehr oft zelltoxisch und es gelingt manchmal nicht mehr, sie nachträglich bezüglich ihrer Wasserlöslichkeit in einen physiologischen Bereich zu bringen, ohne daß sie gleichzeitig an Wirksamkeit einbüßen.

Vielfach setzt man wieder verstärkt auf „natürliche“ Ressourcen. Nahezu zwei Drittel der Pflanzen auf unserer Erde sind bezüglich ihrer biologischen Wirkungen noch nicht charakterisiert. Im Meer wurden Syntheseleistungen entdeckt, die in ihrer strukturellen Exotik Chemiker begeistern.

Modellierung von Protein-Ligand-Komplexen I

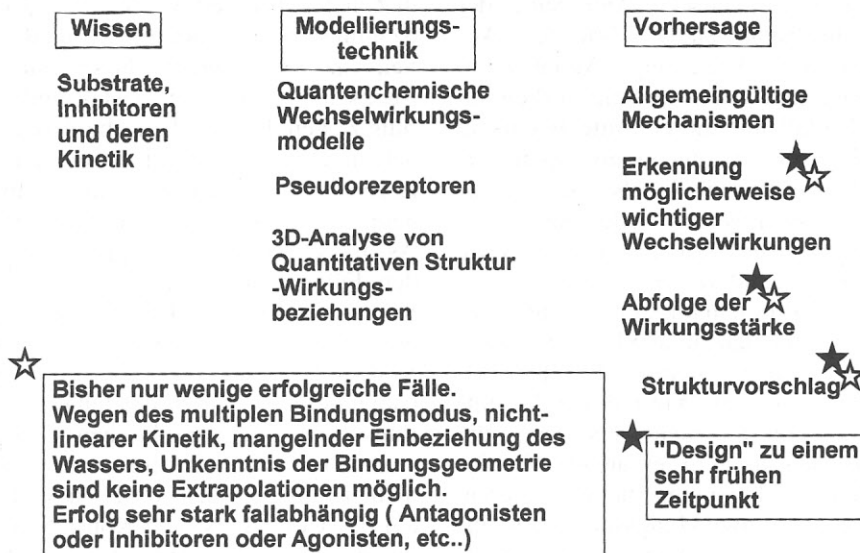


Abbildung 3: Auf der Wissenssebene der Substrat/Inhibitor- und (Ant-)Agonistkinetiken stehen im wesentlichen drei Modellierungstechniken zur Auswahl, die im Erfolgsfall ein Design auf sehr früher Stufe, ohne erhebliche Zusatzinvestitionen möglich machen. Multiple oder alternative Bindungsmodi strukturanaloger Derivate, Unkenntnis von Lösungsmiteleinflüssen und nicht zulässige, lineare Extrapolation reduzieren oft die Erfolgsquoten. Erfolgspositionen der jeweiligen Modellierungstechniken sind mit gefüllten, Caveats mit offenen Symbolen gekennzeichnet.

Modellierung von Protein-Ligand-Komplexen II

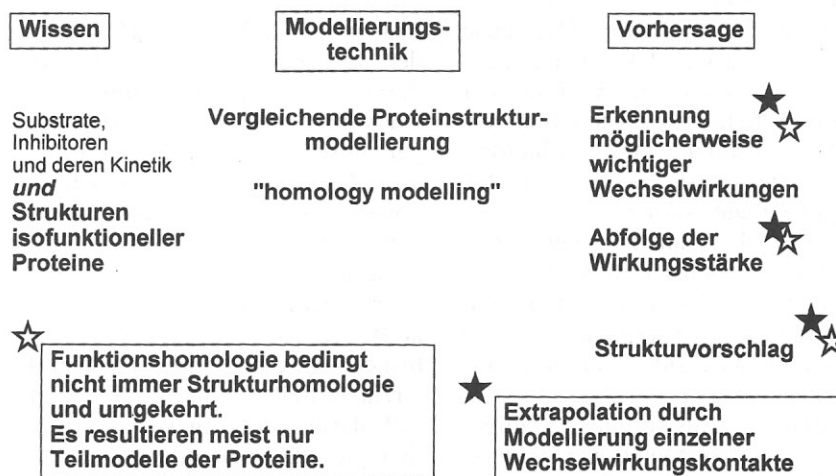


Abbildung 4: Vergleichende Strukturanalyse und deren Übertragung auf Strukturmodelle basieren auf der Kenntnis der Struktur isofunktionaler Proteine. Funktionsessentielle Elemente innerhalb der Proteinsequenzen wurden in der Evolution konserviert. Sie lassen sich durch Sequenzvergleiche erkennen und auf das unbekannte Protein übertragen.

Ein „Treffer“, der damit eine Leitsubstanz charakterisiert, ist statistisch sehr wahrscheinlich. Ab dieser Stufe beginnt meist das Nachdenken

über rationale Verfahren zur Optimierung der gefundenen Leitstruktur, aus der schließlich ein Arzneimittel entstehen soll. Hier leistet

CAMD entscheidende Arbeit, um die physiologische Anpassung der Leitsubstanz zu gewährleisten. Absorption, Verteilung, Metabolisierung und Ausscheidung sollten im Idealfall auf dieser Stufe bereits in die Strukturplanung einbezogen werden, ein Prozeß, der als „*pharmaceutical engineering*“ oder als „*drug engineering*“ zu bezeichnen wäre (Folkers und Merkle, 1995).

Die Problematik der Zukunft wird eher sein, daß zuviele Leitsubstanzen aus solchen systematischen Testserien auftauchen. Wie müssen die Entscheidungskriterien aussehen, die aus den als aktiv erkannten Verbindungen diejenigen heraussortieren, die als Entwicklungssubstanzen ihren sehr teuren Weg zum Arzneimittel gehen?

Molekularbiologie

Hier versucht man über eine molekulare Definition der Erkrankung, beispielsweise über gezielte Mutagenese und Expression von Rezeptoren, die biologische Zielstruktur eines Arzneistoffs zu definieren und dann gezielt oder in Kombination mit Strukturbibliotheken einen geeigneten Liganden zu finden. Der letzte Ansatz ist stark molekular- und zellbiologisch orientiert, während der erstere chemisch ausgerichtet ist. Vom molekular- und zellbiologischen Ansatz ist zu erwarten, daß sich in Zukunft stärker die Elemente des rationalen Entwicklungsprozesses auf der höchsten Wissenssebene, nämlich die Abstimmung von Modellierung und Experiment in jedem Entwicklungsschritt, verwirklichen lassen.

Durch die Auswertung zufälliger oder gezielter Mutationen gelingt nicht nur die analytische Beschreibung eines Proteins. Die Techniken des „Protein Designs“ erlauben auch den synthetischen Schritt zur kreativen Gestaltung von Proteineigenschaften. Dort ist durchaus auch der Zufall wieder der Mittler. Dennoch konnten beispielsweise von an G-Protein gekoppelten Rezeptoren – nach heutigen Erkenntnissen ist dies die am häufigsten vertretene Rezep-

torklasse – daueraktivierte Mutanten in Zellen exprimiert werden. Aus der Kenntnis der Änderungen in der Sequenz des Proteins lassen sich Rückschlüsse auf die Gestaltsänderung ziehen. Der synthetische Ansatz geht noch einen Schritt weiter. Die Rezeptoren können nun mit ihren natürlichen oder auch artifiziellen Assoziaten zusammengruppiert werden. Führt man dies für menschliche Rezeptoren in einer Hefezelle durch, erhält man zwar ein nicht-natürliches System, kann aber diese Zusammenhänge mit dem Vorteil des fehlenden Rauschens der menschlichen Zelle studieren. Die Situation erinnert nicht zufällig an die virtuellen Welten des Computers. Die Konstellation ist sehr ähnlich. Das ist auch der Grund, warum CADD/CAMD und Molekularbiologie ein großes synergistisches Potential aufweisen.

Es stellt sich die Frage nach der Behandlung der Komplexität. Die Tatsache, daß es sich bei alldem um komplexe Systeme handelt, wurde bereits mehrfach erwähnt. Die Reduktion auf handhabbare Szenarien, sowohl im Computer als auch in der Molekularbiologie oder in den automatisierten Reihentestverfahren, ist der bisherige Ausweg. Die sich durch den Reduktionismus möglicherweise auftuenden Verfälschungen unserer Weltsicht sind in epischer Breite in der neueren fach- und populärwissenschaftlichen Literatur dargelegt.

Tatsächlich gewinnt der umgekehrte Ansatz nur äußerst langsam an Boden. Komplexitätsforschung ist bisher kein eingeführtes Fach in der Arzneimittelforschung. Je weiter sich diese jedoch in Richtung Gentherapie und Verständnis der Genregulation entwickelt, umso eher bedarf es der Anwendung der Theorien komplexer Systeme. Gentherapie wird theoretisch einen völlig individuellen Therapieansatz ermöglichen. Damit entfällt der statistische Ansatz, die „Sicherheit der großen Zahl“, die bisher Extrapolationen erlaubte. Der Technik des CADD/CAMD kommen solche individualisierten Ansätze eher näher. Ein für

einen einzelnen Patienten individuell zugeschnittenes synthetisches Therapeutikum wäre zwar der extremste Ansatz, aber auch der beste. Denn gerade in der Leitstrukturoptimierung anhand eines bekannten Proteins postulieren wir, ohne es zu wissen, daß sich der Ligand-Protein Komplex in jeder Zelle eines jeden beliebigen Individuums gleich verhält.

Leitstrukturoptimierung

Eine Optimierung der Leitstruktur auf der Basis von enzymatischen oder Rezeptorgewebdaten ist im wesentlichen nach drei Verfahren möglich (Abb. 3).

Klassische oder 3D-QSAR Methoden stehen sicher aufgrund ihrer historischen Erfolge im Vordergrund (Kubinyi, 1994). Pseudorezeptoren (Vedani, 1994) oder quantenchemische Interaktionsmodelle (Höltje und Kier, 1975) sind aufgrund der wachsenden Kenntnis über die Gesetzmäßigkeiten der Protein-Ligand Wechselwirkung eine zukünftige Alternative. Die Etablierung eines quantenchemischen Interaktionsmodells für die Acetylcholinesterase (Höltje und Kier, 1975) gibt ein Beispiel für die mögliche Vorhersagekraft. Der Vorschlag einer Ammonium-Aromaten Wechselwirkung wurde in der Röntgenstruktur (Sussman et al., 1991) bis in quantitative Details, wie z.B. die Abstände der beiden Gruppen, bestätigt.

Pseudorezeptoren stellen dagegen keine Einzelinteraktionsmodelle dar, sondern betten die zu untersuchenden Derivate der Leitstruktur in ein Miniprotein ein. Dieses Miniprotein dient als Rezeptorsurrogat und wird aus den bisher bekannten Regeln einer Protein-Ligand Wechselwirkung abgeleitet. Substituenten am Wirkmolekül, die Wasserstoffdonatoren sind, erhalten eine entsprechende, Wasserstoffbrücken akzeptierende Aminosäure als Wechselwirkungspartner. Das Abtasten des gesamten Moleküls erzeugt so eine Gruppe von Aminosäuren, die den Wirkstoff umgibt. Über die Berechnung von Solvatationsenergien ge-

lingt in einigen Fällen die Vorhersage der Aktivitätsreihe innerhalb des Trainingssatzes (Greenidge et al., 1995; Vedani, 1994). Alle Verfahren haben den Vorteil, daß sie im Erfolgsfall bereits in einem sehr frühen Stadium einen Strukturvorschlag für einen neuen Wirkstoff ermöglichen. Zumindest aber ist auf dieser Stufe bereits die Identifizierung bindungsessentieller Gruppen im Rezeptorprotein machbar. Diese sehr wichtige Information kann auf einer nächsten Wissens Ebene in Mutageneseexperimente einfließen.

Die Erfolgsquote der genannten Verfahren wird durch das Phänomen des multiplen oder alternativen Bindungsmodus selbst nah verwandter Strukturen, durch Unkenntnis von Lösungsmiteleinflüssen und durch nicht-zuläßige lineare Extrapolation, bzw. mangelhafte Trainingssätze stark beeinträchtigt. Multipler oder alternativer Bindungsmodus (Meyer et al., 1986; Mattos und Ringe, 1993) bezeichnet das Phänomen, daß Liganden, obwohl sie sich kompetitiv verhalten, nicht exakt an die gleiche Stelle im Protein oder nicht exakt in der gleichen Geometrie an die Bindungsstelle im Protein binden. Aus der biochemischen Interpretation der Kinetik als „kompetitiv“ darf nach unseren heutigen Erkenntnissen nicht geschlossen werden, daß die als kompetitiv erkannten Liganden auch gleiche geometrische Bindungseigenschaften entwickeln. Für Inhibitoren ist dies unmittelbar einsichtig. Es geht darum, eine Reaktion im Enzym zu blockieren. Dabei spielt es kaum eine Rolle, in welcher Art und Weise das geschieht. Eine vom natürlichen Liganden strukturell völlig abweichende Verbindung kann beispielsweise den Zugang zur aktiven Stelle im Enzym verstopfen. Die Reaktion erstirbt und die Enzymblockade löst einen bestimmten Effekt aus, sie unterbricht beispielsweise die Vermehrung eines pathogenen Keimes. Für den Inhibitor besteht keine Notwendigkeit, die gleichen Bindungspunkte zu verwenden wie der natürliche Ligand. Die alternierende Bindungsgeometrie des

Inhibitors ist aber nur in den Fällen bekannt, in denen eine Röntgenstruktur eines Ko-Kristallits aus Enzym und Inhibitor vorliegt. Dies ist zwar zunehmend, aber dennoch insgesamt äußerst selten der Fall.

Sehr viel komplizierter wird die Situation an Rezeptoren. Dort läßt sich zeigen, daß multiple Bindungsmodi auch für natürliche aktivierende Liganden unterschiedlicher Struktur existieren müssen. Der Schluß auf einen optimierten Liganden aus der Kenntnis einiger anderer Liganden und ihrer biochemischen Bindungskonstanten muß demnach nicht zu einer eindeutigen Lösung führen. Die verschiedenen Verfahren sind damit automatisch abhängig vom bearbeiteten Problem. Daß kompetitiver Antagonismus, der in einem *in vitro* Rezeptorassay gemessen wurde, nicht unbedingt bedeutet, daß der Antagonist eine dem Agonisten verwandte Struktur aufweisen muß, zeigen beispielsweise die Antagonisten der Neurokinine am NK-1 Rezeptor oder anderer G-Protein gekoppelter Rezeptoren (Fong et al., 1993; Schwartz, 1994). Die kompetitive Inhibition eines Enzyms durch einen Hemmstoff des Übergangszustands (*transition state inhibitor*) erfordert dagegen eine möglichst identische Paßform von Substrat und Inhibitor. Die Überlagerung der Strukturen und der Entwurf sterisch identischer, aber reaktiv unterschiedlicher Substituenten hat in Studien, die auf der Kenntnis des molekularen Mechanismus basieren, eine große Erfolgchance. Beispiele hierfür sind Phosphonoamidate bei Proteaseinhibitoren. Flexible Liganden erschweren eine Modellierung durch ihre geometrischen Freiheitsgrade. Eine Energieminimierung, selbst unter Lösungsmiteleinfluß liefert nicht notwendigerweise die gewünschte Geometrie, die die am Rezeptor erforderliche, komplementäre Paßform der Bindungsstelle einnimmt. Die einzige Problemlösung stellt die Einführung von experimentellen Templaten dar. Letztere können rigide Wirkstrukturen sein, wie z.B. Morphin, auf deren Grundgerüst die flexiblen

Liganden angepaßt werden können. Diese Struktur Anpassung setzt jedoch voraus, daß die starre Struktur und der flexible Ligand in der gleichen Orientierung mit sehr ähnlicher Geometrie binden. Dies gilt streng eigentlich nur für strukturell kompetitive Antagonisten und Hemmstoffe des Übergangszustandes in Enzymen.

Kernresonanzspektroskopie (NMR)

Der *Nuclear Overhauser Effect* (NOE) bietet eine elegantere Möglichkeit, Bindungsgeometrien flexibler Liganden zu definieren. Das Experiment wird als *transferred nuclear overhauser effect* bezeichnet. Es setzt aber ein rekombinantes oder sehr sauber isoliertes Zielprotein, in einer dem NMR-Experiment entsprechenden Menge, voraus. Der NOE basiert auf dem Phänomen einer nicht bindungsbedingten elektronischen Wechselwirkung, beispielsweise zweier Wasserstoffatome, die sich auf einen bestimmten Abstand einander genähert haben. Die Empfindlichkeit, mit der sich zwei Wasserstoffe im Raum „sehen“, nimmt mit der sechsten Potenz des Abstands ab und ist deshalb nur unterhalb von fünf Ångström verlässlich zu erfassen. Aufgrund der häufig vorkommenden Wasserstoffe in Molekülen läßt sich aus der Messung ihrer räumlichen Abstände ein Muster für die dreidimensionale Gestalt des beobachteten Moleküls ableiten.

Zuerst erfolgt also die Messung der internen Wasserstoffabstände des Liganden über NOEs. Damit wird die Lösungsgeometrie festgelegt. Im folgenden gibt man das bindende Protein hinzu und wartet die Einstellung des Gleichgewichts ab. Eine Inversion der NOE-Signale zeigt an, daß der Ligand gebunden hat. Die Verschiebung entlang der Frequenzskala gibt die Information, wie seine Geometrie im Übergang zur gebundenen Form geändert wurde (Feng, 1994). Die gemessenen NOE-Distanzen lassen sich als Randbedingungen in eine Moleküldynamiksimulation oder eine Distanzgeometrie-Rechnung des Liganden einset-

zen, aus der dann die gebundene Geometrie bestimmbar ist. Diese experimentelle Distanzmatrix kann dann flexiblen Analogen überlagert werden.

Für komplexere Systeme, wie beispielsweise die Wechselwirkung eines Peptids oder eines Proteins mit einem anderen Protein, kommen heute sogenannte heteronukleare mehrdimensionale NMR Techniken zum Einsatz. Dies hat experimentelle Hintergründe, da oft einzelne Resonanzen von Wasserstoffkernen bei den hohen Peakdichten in Proteinen nicht mehr zugeordnet werden können. Die Techniken der multidimensionalen NMR Spektroskopie setzen allerdings ^{13}C markierte Liganden und im Normalfall auch ^{13}C und/oder ^{15}N markierte Rezeptorproteine voraus. Elegante Anwendungen dieser Methodik finden sich in den Arbeiten zur Strukturaufklärung des Immunsuppressivums Cyclosporin und seiner Wechselwirkung mit seinem Rezeptorprotein Cyclophilin, das Prolinisoimerase-Aktivität aufweist (Weber et al., 1991).

Die klassischen QSAR-Verfahren sind am wenigsten empfindlich gegen Störungen, die aus Bindungsgeometrie-problemen resultieren. Alle stark strukturabhängigen Methoden müssen dagegen sorgfältig auf ihre Eignung für das jeweilige Problem überprüft werden. Letzteres ist nur durch Integration weiterer experimenteller Resultate möglich.

4.2 Strukturen isofunktionaler Proteine

Informationen über die Bindungsstelle spielen naturgemäß die größte Rolle, um die Kenntnis des Protein-Ligand Komplexes zu erhöhen und die rationale Wirkstoffentwicklung zu vervollkommen. In den meisten Fällen wird jedoch die Struktur des Zielproteins unbekannt sein. Dessen Mechanismus ist aber häufig bekannt, besonders dann, wenn die Erkrankung molekularbiologisch und biochemisch (z.B. Proteinssequenz, Kinetik etc.) definiert wurde. In solchen Fällen gelingt es in zuneh-

mendem Maße, aus der Zahl der strukturell bekannten Proteine isofunktionelle Vertreter des Zielproteins zu finden. Aufgrund evolutionärer Entwicklung ist anzunehmen, daß in der Mehrzahl der Fälle isofunktionelle Proteine auch eine sehr ähnliche Architektur, zumindest der funktionell bedeutsamen Bauelemente, besitzen. Vergleichende Strukturanalyse und die Ableitung generell gültiger Bauprinzipien für bestimmte Funktionsklassen innerhalb der Proteine („*homology modeling*“, bzw. in korrekterer Bezeichnung „*comparative modeling*“) können in dieser Situation eingesetzt werden, um aus den erkannten Ähnlichkeitsbeziehungen isofunktionelle Strukturelemente in gleicher Geometrie für das unbekannte Zielprotein zu übernehmen und auf diese Weise eine Bindungsstelle zu modellieren. Anhand des Modells lassen sich aus der Kenntnis der Sequenz wiederum essentielle Bindungsamino-säuren identifizieren. Diese Information fließt sowohl in Mutagenesestudien wie auch in das Modeling der Liganden ein (Abb. 4).

GPCR Modelle aus vergleichenden Strukturanalysen

Während das Verfahren, historisch gesehen, an Serinproteasen entwickelt wurde, ist das wohl populärste Beispiel für diese Strategie die Modellierung der G-Protein gekoppelten Rezeptoren (GPCR) (Hibert et al., 1991; Baldwin, 1993). Am generellen Aufbau des Rezeptors aus sieben, die Membran durchspannenden Helices besteht kaum Zweifel. Dagegen sind Sequenz und relative Anordnung der Helices nach wie vor umstritten. Das ursprüngliche Modell wurde in Analogie zur Röntgenstruktur des Bakteriorhodopsins erstellt, das als einzige Isofunktionalität zu den Rezeptoren seine Membranständigkeit aufwies. Neue strukturelle Erkenntnisse aus dem Rinder-Rhodopsin, einem „echten“ GPCR, haben zu verbesserten Modellen geführt (Schertler et al., 1993; Baldwin, 1993), die sich zum Teil drastisch von den Modellen unterschei-

den, die aus dem Bakteriorhodopsin abgeleitet wurden. Diese Situation spiegelt die Problematik der vergleichenden Proteinstrukturmodellierung wieder. Es ist bis heute nicht vollständig verstanden, wann eine Aminosäure in einem Protein durch eine andere ersetzbar ist, ohne zum Funktionsverlust zu führen und ob, bzw. wie diese Änderung in der 3D Struktur kompensiert wird.

Eine weitere Schwierigkeit des „*homology modeling*“ ist die Unvollständigkeit der resultierenden Strukturen. Isofunktionalität vorausgesetzt, existieren bei Proteinen aus evolutionär sehr weit voneinander entfernten Organismen dennoch wahrscheinlich große Anpassungen außerhalb des katalytischen oder Bindungszentrums. Eine halbwegs sichere Geometrie ist deshalb nur für einen relativ kleinen inneren Bereich des Zielproteins als Modell zu erhalten. Sollen dort mit CAMD-Methoden Liganden eingepaßt werden, gehen aufgrund der fehlenden äußeren Protein- und Membranteile wesentliche Informationen für die Wechselwirkungsmodelle verloren. Operationen wie Energieminimierung des Liganden in der Bindungsstelle oder Moleküldynamiksimulation werden notwendigerweise durch Artefakte verfälscht, da ein Zerfallen des Proteinteilmodells künstlich verhindert werden muß. Statt des üblichen Fixierens der Proteinketten haben wir ein Verfahren entwickelt, um die Artefakte möglichst gering zu halten. Mit Hilfe von im Computer erzeugten Pseudopartikeln gelingt es an bisher vier verschiedenen Proteinröntgenstrukturen, die auf ihre Bindungsstellen reduziert wurden, deren Ligandkomplexe selbst in der Dynamiksimulation stabil zu halten (Kern, 1995a) (Abb. 7). Die Idee mit den Pseudopartikeln erscheint deshalb als sehr attraktiv, weil sie im Gegensatz zu anderen Verfahren der Selbstorganisation des inneren Bereichs der Proteine eine gewisse Freiheit läßt. Die Kugeln mit einem Durchmesser von etwa sieben Ångström sind gegeneinander beweglich. Sie bilden eine sehr dicke

Schicht, die erst relativ weit außen durch eine (simulierte) harte Wand abgegrenzt und gestützt wird. Die Partikel haben polare Eigenschaften, es ist ihnen erlaubt, sich zu orientieren. Unsere Computereperimente zeigen, daß die Röntgenstrukturen von vier sehr verschiedenen Proteinen, die auf ihre Bindungsstellen zurückgeschnitten wurden, in ausgezeichneter Übereinstimmung reproduziert werden können. Die Pseudopartikel scheinen zu einem Werkzeug entwickelbar zu sein, das einmal Homologiemodellierungsstudien zu einem Routineansatz werden lassen könnte.

Der Vorteil solcher Modelle aus „homology modeling“ Studien, selbst wenn sie aus Partialstrukturen bestehen, ist die Möglichkeit zur Extrapolation. Durch explizite Lokalisation von bindenden Aminosäuren über die Strukturhomologie lassen sich im Gegensatz zu den Verfahren, bei denen ausschließlich Liganden betrachtet wurden, auch solche Verbindungen untersuchen, die nicht mehr über den Trainingssatz abgedeckt sind. Im Sinne einer Optimierungsstrategie lassen sich für noch nicht synthetisierte Derivate einer Leitstruktur an dieser Stelle Einpassungen untersuchen. Es wird beispielsweise klar, ob ein neu eingeführter Substituent in der Bindungstasche Platz findet oder, wenn seine Wirksamkeit bereits bekannt ist, eine alternative Orientierung des Liganden in der Bindungstasche möglich ist.

Eine wesentliche Verbesserung bringt allerdings auch hier die Einbeziehung weiterer experimenteller Verfahren auf der Seite der Bindungsproteine.

4.3 Gezielte Mutation der Bindungsproteine

Molekularbiologische Grundtechniken sind heute im Baukastenverfahren verfügbar. Es gelingt in vielen Fällen, das Zielprotein rekombinant in die Hände zu bekommen. Damit sind auch Mutageneseexperimente möglich (Abb. 5). Die gezielte Mutagenese, ausgehend von Strukturen

Modellierung von Protein-Ligand-Komplexen III

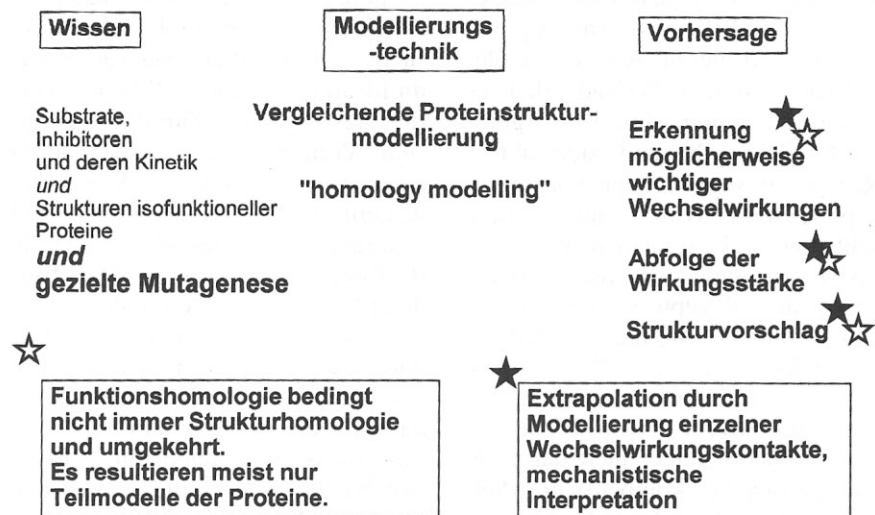


Abbildung 5: Gezielte Mutagenese aus Vorschlägen der Computermodelle trägt wesentlich zur Entstehung der Strukturhypothese des Protein-Ligand Wechselwirkungskomplexes bei. Die analytische Erkenntnis kann mit Hilfe der Molekularbiologie synthetisch geprüft werden.

Modellierung von Protein-Ligand-Komplexen IV

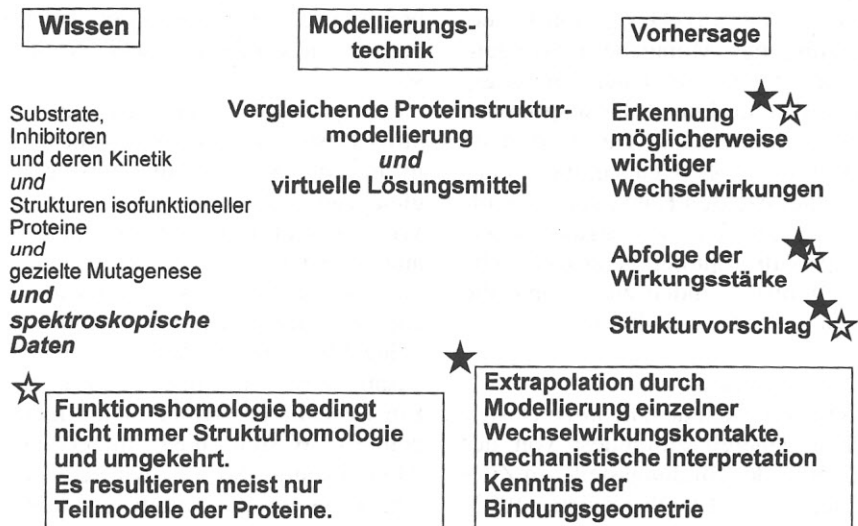


Abbildung 6: Die Integration experimenteller Kartierungs- (mapping) Verfahren erlaubt neben der Verbesserung der mechanistischen Interpretation erste Aussagen über die Orientierung des Liganden in der Bindungsstelle. Die kombinierten Verfahren erhöhen die Auflösung der Homologiemodelle durch experimentelle Sekundärstrukturinformation.

oder Strukturmodellen, gehört zu den elegantesten Verfahren, um nähere Aufschlüsse über die Natur des Protein-Ligand Komplexes zu bekommen. Speziell in der Modellierung der bereits angesprochenen

GPCR konnten Mutagenesestudien wesentlich zur Verfeinerung der Struktur beitragen (Schwartz, 1994). Von großer Bedeutung war dabei die Aufdeckung zusätzlicher Bindungsstellen sowohl für Agonisten als

auch für Antagonisten. Diese Arbeiten haben unter anderem dazu beigetragen, das Bild des (Ant-)Agonismus zu korrigieren. Aus dem molekularbiologischen Befund, daß es aktive Rezeptoren ohne jede Agonistenbindung gibt und sich aktive Rezeptoren sogar mutiert konstitutiv exprimieren lassen, mußte man schließen, daß Agonisten die aktive Form, Antagonisten die inaktive Form des Rezeptors stabilisieren. Damit müssen sie nicht unbedingt die gleiche Bindungsstelle besetzen, was an Neurokininrezeptoren demonstriert wurde (s.o.). Diese Erkenntnis bedingt aber auch, daß nicht automatisch bei kompetitivem Antagonismus auf strukturelle Ähnlichkeit geschlossen werden muß. Sie kann jedoch vorliegen. Dann gelten aber für beide Antagonisten unterschiedliche Modellierungsstrategien. Während im letzten Fall eher eine Blockade, ähnlich wie bei Enzymen, angestrebt wird, muß im ersten Fall quasi ein *de novo* Design mit Unterstützung der Mutagenese erfolgen, da bei unterschiedlicher Bindungsstelle die Struktur des Antagonisten sehr variieren kann. Der Vorteil ist jedoch die große Chance, die Strategie einer exakten Einpassung verfolgen zu können und beispielsweise bei Peptidrezeptoren heterozyklische Antagonisten finden zu können, die quasi agonistisch inaktivieren.

De Novo Design

Verfahren nach dem *de novo* Design gehen generell zumindest von der Kenntnis der Bindungsstelle im Zielprotein aus. LUDI (Böhm, 1992) gehört zu den am häufigsten verwendeten Programmen. Durch Anwendung von allgemeinen Regeln zur Bindungsart und -geometrie von Liganden, die aus der Analyse von zahlreichen Röntgenstrukturen von Protein-Ligand Komplexen entnommen und auf der Basis von theoretischen Wechselwirkungsmodellen errechnet wurden, lokalisiert das Programm Strukturelemente an den Punkten der Bindungstasche, die für eine Interaktion mit Liganden in Frage kommen. Die Konstellation

der platzierten Strukturelemente, die ja praktisch einem Pharmakophor entspricht, erlaubt durch Datenbank-suche die Identifizierung passender, im Idealfall kommerziell bereits verfügbarer Liganden. Die *de novo* Design Verfahren lassen sich durch einen kombinatorischen Ansatz erweitern, wie er von Caflisch und Mitarbeitern vorgeschlagen wurde (Caflisch et al., 1993). Die Bindungstasche des Zielproteins wird durch einen Satz von für die Bindungsarten charakteristischen, kleinen Liganden – beispielsweise Acetamid für die Bindung an ein Peptidrückgrat oder Methanol für Wasserstoffbrückenbindungen – innerhalb einer Moleküldynamiksimulation „überschwemmt“. Die verbleibenden, gebundenen Fragmente werden lokalisiert. Sie lassen sich zu einem ersten Entwurf eines neuen Liganden zusammensetzen. Die Methode gewinnt sehr viel an Attraktivität durch ihre Schnelligkeit und Variationsbreite, nicht nur wegen ihres Anlehens an die kombinatorische Synthese.

Das prinzipielle Problem der *de novo* Design Verfahren besteht darin, daß der Satz an Bindungsregeln gleichzeitig ihre gesamte erfaßbare Welt darstellt. Liganden, die auf eine andere Art binden könnten, werden nie vorgeschlagen. Aus grundsätzlichen Überlegungen gibt es natürlich keine Möglichkeit, den verfügbaren gesamten Bindungsraum zu erfassen. Ein wesentliches Hindernis ist mangelndes Wissen über molekulare Ähnlichkeiten. So sind wir gezwungen, neue Bindungsmodi aus Röntgenstrukturen abzuwarten, wie beispielsweise die der Acetylcholinesterase, die Aromat-Ammonium-Bindung als eine Wechselwirkungsmöglichkeit von Liganden und Proteinen zeigte.

Analyse von gezielten Mutationen

Die Kombination der gezielten Mutagenese mit den Verfahren der vergleichenden Proteinstrukturmodellierung hat eine synergistische Wirkung für die Strukturvorhersage (Michael et al., 1994). Es erlaubt zum

einen die Prüfung und die Unterstützung oder die Widerlegung von Strukturhypothesen. Andererseits macht es aber auch die Mutagenese besser interpretierbar. Die statistische Mutagenese einzelner Aminosäurereste ist in der Vergangenheit wesentlich überinterpretiert worden. Beim Ersatz einer Aminosäure aus einem konservierten Gebiet nur aufgrund von Sequenzstudien kann nicht ausgeschlossen werden, daß eine Rückfaltung des Proteins eine entfernt liegende, nichtkonservierte Aminosäure in die Nähe des Substitutionsortes bringt und sie dort die Mutation zufällig kompensiert. So resultiert ein nach wie vor aktives Protein, obwohl vielleicht eine sehr empfindliche Position getroffen wurde. Ebenso ist es unmöglich auszuschließen, daß beispielsweise sterische Effekte die Struktur einer Bindungsstelle nach Mutation zusammenbrechen lassen und deshalb eine inaktive Proteinmutante erhalten wird. Der Hintergrund dieser Schwierigkeit liegt in unserer nach wie vor großen Unkenntnis über die Ähnlichkeitsbeziehungen von Molekülen. Daß allein strukturelle Ähnlichkeitsbeschreibungen nicht ausreichend sind, gilt inzwischen als trivial.

Für die uns interessierende herpesvirale Thymidinkinase konnten wir mit der Strategie der gezielten Mutagenese kürzlich zeigen, daß das in Homologiemodellen lokalisierte Substrat mit der Position des Co-Substrats vertauscht werden muß (Kern, 1995b).

4.4 Peptidsynthese und Circular-dichroismus als Kartierungstechnik (*mapping*) für das Zielprotein

Parallel zu den molekularbiologischen Arbeitstechniken bieten die neuen Entwicklungen der automatischen Peptidsynthese in Kombination mit Spektroskopie und immunologischen Verfahren faszinierende Möglichkeiten zur Verfeinerung der Strukturhypothesen (Abb. 6).

CD-Spektroskopische Sekundärstrukturvorhersage

Die Circular dichroismus-Spektroskopie kann dazu verwendet werden, Lösungsmittelgeometrien von Peptiden und Proteinen zu untersuchen. Bekanntlich sind die auftretenden Konformationsfamilien speziell im Fall der Peptide aber sehr stark abhängig von der Umgebung. Dieser Effekt läßt sich ausnützen. Durch Messungen in variierenden Lösungsmitteln, wie beispielsweise der Titration in Trifluoroethanol/Wasser-Gemischen, kann für Peptide mit einer Kettenlänge zwischen 15 und 20 Aminosäuren eine Tendenz zur Ausbildung eines bestimmten Sekundärstrukturtyps beobachtet werden (Zimmermann et al., 1992). Mit der automatischen Peptidsynthese läßt sich in vernünftiger Zeit und guter Ausbeute ein homologes Protein in Form von Peptiden der Kettenlänge 15 bis 20 nachsynthetisieren. Ein Teil der Peptide wird je nach Syntheseproblem auch überlappend sein, es sollten im Schnitt jedoch nicht mehr als 25 synthetische Peptide für ein cytosolisches Protein resultieren. Im Fall der Adenylatkinase, deren Struktur bekannt ist, gelang es auf diese Weise mit einem Satz von 17 synthetischen Peptiden, die gesamte Sekundärstruktur des Proteins mit einer Genauigkeit von mehr als 90% vorherzusagen (Beck-Sickinger, 1995) (Abb. 8, Abb. 9). Diese Sekundärstrukturinformation kann verwendet werden, um im Homologie-modell durch Vergleich mit der lokalen Sekundärstruktur des Templatproteins weitere Proteinbereiche zu modellieren.

Alaninscan

Eine zweite und bereits wesentlich ausgereifere Technik der Kartierung (*mapping*) ist die systematische Variation von Peptiden, die beispielsweise an Hormonrezeptoren binden. Ersetzt man entlang der Sequenz jede Position durch Alanin und korreliert mit der Affinität zum Rezeptor, erhält man eine erste Kartierung des Rezeptors bezüglich der Sensitivität gegenüber dem Verlust der für

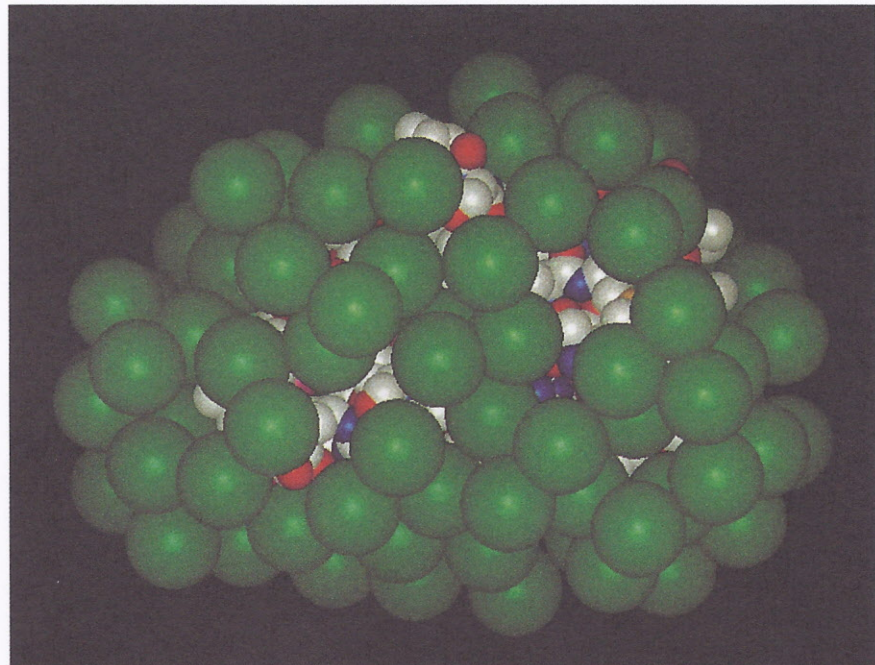


Abbildung 7: Die Pseudopartikel (goldfarben), von denen hier der Übersichtlichkeit wegen nur ein Teil der innersten Schicht dargestellt ist, umgeben das Proteinkonstrukt und betten es ein. So sind Moleküldynamiksimulationen mit eingepaßtem Liganden möglich und erlauben eine Abschätzung über die Affinität des Liganden.

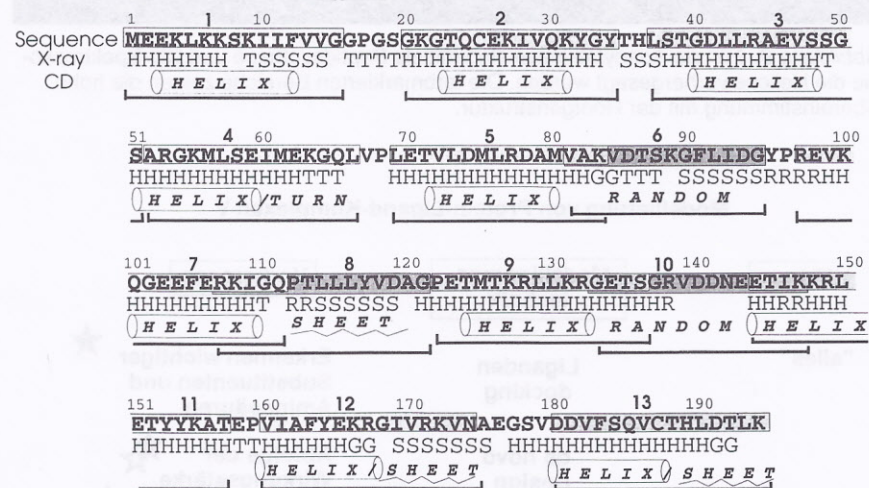


Abbildung 8: Der Vergleich der CD-spektroskopischen Vorhersage der Sekundärstruktur von synthetischen Peptiden der Adenylatkinasesequenz zeigt eine Strukturvorhersage von mehr als 90% des Gesamtproteins. Die Sequenz des Proteins ist in der obersten Zeile dargestellt. Darunter folgt die Röntgenstruktur (H = Helixbereich, S = Faltblattstruktur, T = Turn oder ungeordnet). In der dritten Zeile ist das Ergebnis der CD-Messungen aufgetragen. Die Balken markieren jeweils das synthetische Peptid.

die Bindung essentiellen Aminosäuren. Für die Bindung des Neuropeptidhormons Y (NPY) an seinen Rezeptor wies der Alaninscan auf die Bedeutung zweier C-terminaler Arginine hin (Beck-Sickinger, 1994).

Parallel durchgeführte Mutagenesestudien am NPY-Rezeptor enthüllten zwei korrespondierende Asparaginsäuren, die in idealer Weise als Wechselwirkungspartner in Frage kämen.

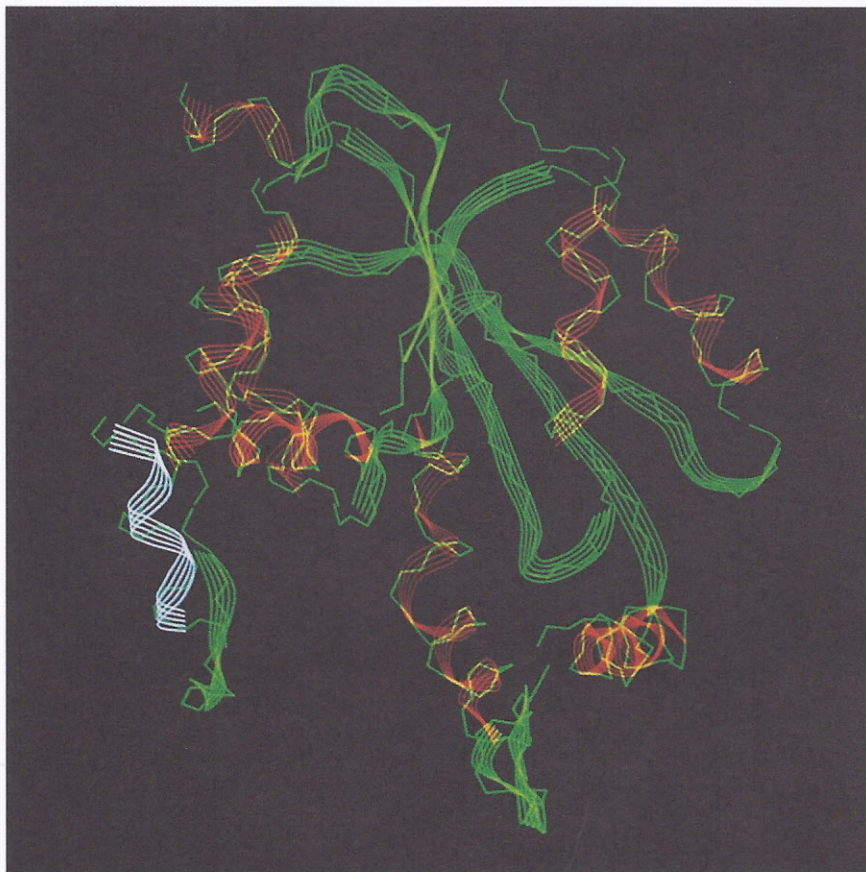


Abbildung 9: Für die Adenylatkinase sind mittels Peptidsynthese und CD-Spektroskopie die Helices vorhergesagt worden. Die farbmarkierten Bereiche zeigen die hohe Übereinstimmung mit der Röntgenstruktur.

Modellierung von Protein-Ligand-Komplexen V

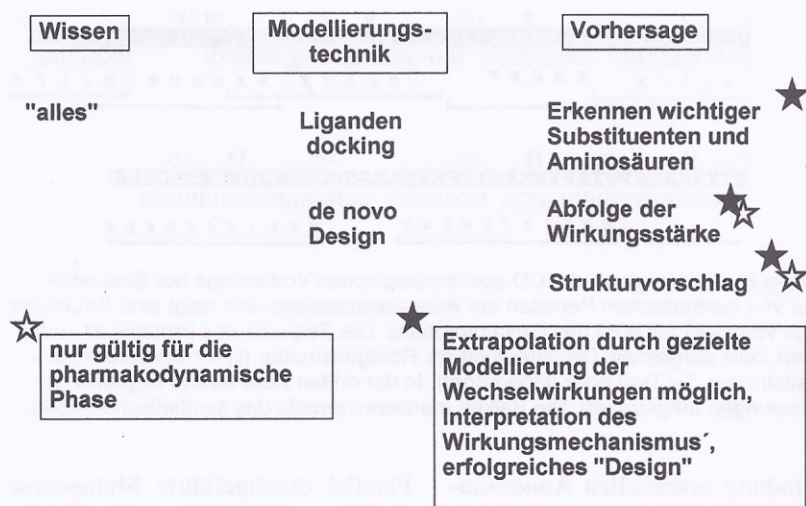


Abbildung 10: Der Einbau und die Simulation der Wechselwirkungen von Liganden in die bekannten Proteinstrukturen gestatten mechanistische Interpretation und Extrapolationen. Der Entwurf neuer, nicht analoger Strukturen (*de novo* Design) wird über die Kenntnis der für die Wirkungsauslösung essentiellen Wechselwirkungen möglich.

Auf dieser Stufe sind Extrapolationen aus verschiedensten Techniken, wie z.B. *Comparative Molecular Field Analysis* (CoMFA), möglich. Durch explizite Modellierung von Interaktionen und deren experimentelle Prüfbarkeit können Arzneistoffklassen mit einbezogen werden, die bisher nicht als potentielle Therapeutika für das Untersuchungssystem angesehen wurden.

4.5 Molekulare Struktur des Wechselwirkungskomplexes

Mit der Kombination von Molekularbiologie, Röntgenstrukturanalyse und Kernresonanzspektroskopie gelingt es heute in zunehmendem Maße, die Zielproteine im Komplex mit dem Wirkstoff bis hin zum atomaren Detail aufzulösen. Damit stehen für die Betrachtung der pharmakodynamischen Phase sehr viele wichtige Daten zur Verfügung. Nach wie vor nicht explizit beschreibbar ist der Zeitverlauf der Assoziation bzw. Dissoziation. Aus der komplexen Hyperfläche können nur kleine Ausschnitte modelliert werden. Strukturübergänge, wie beispielsweise Konformationsänderungen in Rezeptorproteinen, sind bisher nicht genau modellierbar. Allerdings liefert die Zuhilfenahme von spektroskopischen und Mutagenesedaten eine Reihe von stabilen Zwischenstufen, über die man sich dem mechanistischen Ablauf nähern kann (Abb.10). Der Test ist letztendlich die gelungene Wirkungsvorhersage für eine vorher noch nicht synthetisierte Substanz aus der theoretischen Ableitung des Mechanismus.

Ein Beispiel für dieses Vorgehen ist das Design blockierender Peptide, bzw. deren Mimetika für den Haupt-histokompatibilitätskomplex (*major histocompatibility complex*, MHC) der T-Zellantwort in unserem Immunsystem. In diesem Schutzmechanismus unseres Körpers präsentiert eine spezielle Zellart auf ihrer Oberfläche mit Hilfe eines eigens dafür entwickelten Proteins ein Bruchstück eines körperfremden Proteins auf der Oberfläche. Die Kombination aus

peptidischem Bruchstück, präsentierendem Protein und Oberfläche der präsentierenden Zelle wird von einem Rezeptor auf der T-Zelloberfläche erkannt. Er bindet und etabliert auf diese Weise einen ternären Komplex aus Antigen-präsentierendem Protein, Peptid und T-Zellrezeptor. Dieser Erkennungsvorgang löst die physiologische Reaktion der Zerstörung der Fremdeiweiße oder Fremdzellen aus.

Die Studie begann mit der Modellierung der Peptidwechselwirkung eines Teilstücks aus dem Influenza Virus Matrixprotein der Kettenlänge fünf, gebunden an MHC codiertes humanes Leukozyten Antigen (HLA) A2 Protein. Ausgehend von der Röntgenstrukturanalyse des HLA-A2 wurde nach vergleichender Proteinstrukturmodellierung das Nonapeptid IMP85-66 Aminosäure für Aminosäure in die Bindungstasche des HLA Moleküls eingedockt. Daran schlossen sich Moleküldynamiksimulationen in Wasserumgebung an, mit deren Hilfe die optimalen Bindungswechselwirkungen des Peptides analysiert werden konnten. Auf diese Weise gelang es, mit theoretischen Verfahren eine Bindungsgeometrie für Peptide dieses Typs vorzuschlagen. Aminosäuren, die für die Bindung an die antigen-präsentierende Struktur, das HLA-Molekül, und Aminosäuren, die für die T-Zell Interaktion wichtig waren, konnten so identifiziert werden (Rognan et al., 1992) (Abb. 11). Der Beweis für die Richtigkeit der Strukturvorhersage wurde ein Jahr später über die Röntgenstrukturanalyse des identischen Peptidkomplexes geliefert (Madden et al., 1993). Die RMS Abweichung (*root mean square* ist das statistische Maß der räumlichen Abweichung von Strukturen) zwischen theoretisch vorhergesagter und experimentell im Kristall gefundener Struktur betrug lediglich 1.2 Å. Sechs der insgesamt neun im Kristall sichtbaren und für die Wechselwirkung wichtigen Wasserstoffbrücken sind auch in der theoretischen Struktur vorhanden (Abb. 12). Die Nützlichkeit theoretischer Simulationen

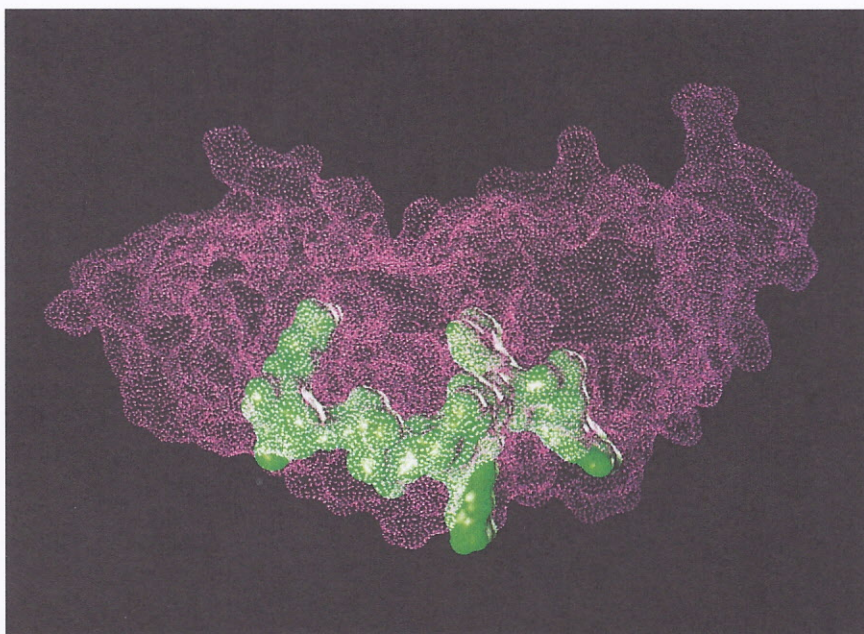


Abbildung 11: Das antigene Peptid (grün) läßt sich eindeutig in einer bestimmten Geometrie (Schlüssel-Schloß-Prinzip) in die Bindungstasche des Antigen präsentierenden Proteins (rote Punkteoberfläche) einpassen.

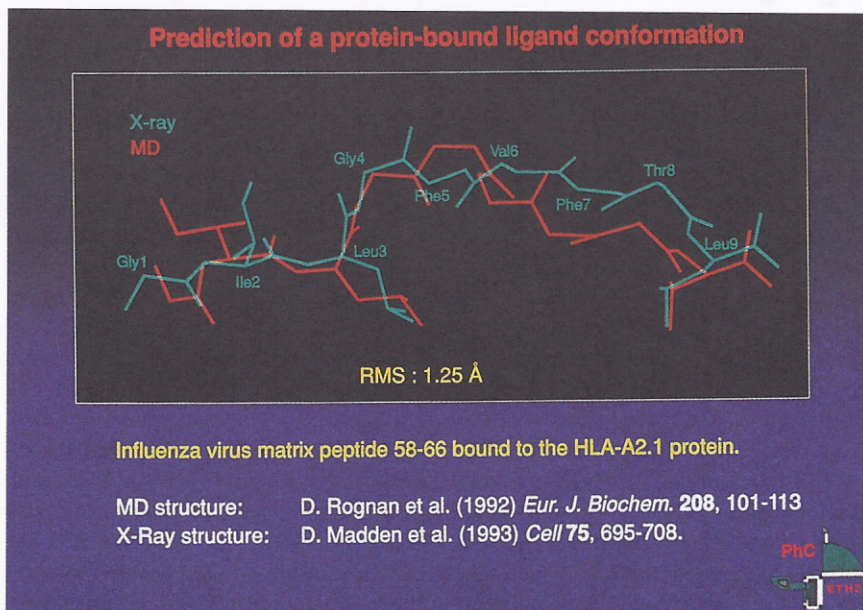


Abbildung 12: Vorhergesagte (rot) und im Kristall gefundene Geometrie (blau) des HLA-A2 gebundenen IMP 58-66 Nonapeptids.

als komplementäres Werkzeug zur Erklärung biologischer Daten wird hierdurch unterstützt und ist bereits mehrfach dokumentiert (van Gunsteren und Mark, 1992). Die aus dieser Analyse entwickelte Designstrategie hatte zum Ziel, die Bindungsenergie der Peptide durch Einbau nicht-na-

türlicher Reste und Spacer so zu erhöhen, daß eine kompetitive Blockade als potentielle Therapie für eine Autoimmunerkrankung erreicht wird. Der Einbau stärker hydrophober und sterisch anspruchsvoller Aminosäuren in diejenigen Positionen die an das HLA Molekül ban-

den, führte unter maximaler Ausnutzung der Bindungstaschen bereits zu einer wesentlich höheren Affinität. Zusätzlich war darauf zu achten, daß keine neue oder weiter erhöhte Antigenität entstand. Mit dem Spacerkonzept, bei dem gezielt einzelne Reste durch möglichst einfache Kohlenwasserstoffe ersetzt wurden, gelang es schließlich, ein Molekül zu entwerfen, das im biologischen Test eine um den Faktor 100 erhöhte Affinität, bei gleichzeitig enorm reduzierter Komplexität aufwies. Der neue Ligand, ein nicht-natürliches Pentapeptid, ist gegenüber dem natürlichen Nonapeptid wesentlich einfacher gebaut (Rognan, 1995).

5 Zusammenfassung und Ausblick

Die Anwendungsbreite bzw. Anwendungsmöglichkeit theoretischer Methoden in der Auffindung und Entwicklung von Arzneimitteln wird stark beeinflusst von der kritischen Kenntnis des Wissensstandes zum Zeitpunkt der Anwendung. Alle Verfahren lassen schon mit sehr ungenügenden Daten eine Analyse zu und provozieren Extrapolationen, deren Güte nur schwer nachzuvollziehen ist. Eine Datenanalyse wird oft zu einem Zeitpunkt verlangt oder vorgenommen, an dem keineswegs die richtigen Daten, sondern nur sehr viele Daten vorliegen. Eine theoretisch arbeitende Gruppe, deren Aufgabe die rationale Entwicklung von Arzneistoffen ist, muß deshalb zum frühest möglichen Zeitpunkt in die Entwicklung miteinbezogen werden. Aus theoretischer Sicht werden sehr oft stärker abweichende Strukturen für ein „*training set*“ und ein „*test set*“ vorgeschlagen, als sie im intuitiven Repertoire des Chemikers vorhanden sind. Strukturen können vor allem auch unter dem Aspekt des „integrierten Ansatzes“ paralleler experimenteller Techniken gewählt werden. Nicht jedes Analogon eignet sich für *transferred*-NOE Experimente oder ist für die *crosslinking* Substitution zugänglich.

Die kritische Auswahl des Datensatzes bestimmt Erfolg oder Mißerfolg der rationalen Entwicklung eines Wirkstoffes, nicht aber die Interpretation einer „historischen“ Datensammlung. „Drug Design Studien“ im Sinne des oben beschriebenen integrierten Ansatzes sind oft sehr erfolgreich, werden aber gleichzeitig aus Schutzgründen selten publiziert.

Ein wesentliches und immer wiederkehrendes Element, das durch theoretische Ansätze erzeugt wird, ist dasjenige des gesteuerten Zufalls, der vielleicht dem im Englischen häufig gebrauchten Ausdruck *serendipity* entspricht. Theoretische Ansätze erzwingen oft Abweichungen von der Routine oder der vermeintlichen Erfahrung und bestätigen das Wort Pasteurs: „Der Zufall begünstigt stets den vorbereiteten Geist.“

Der vorliegende Aufsatz kann die Thematik nur anreißen. Ihm liegen persönliche Erfahrungen des Autors in der Modellierung von Proteinen und im computergestützten Entwurf von Arzneistoffen zugrunde, wobei alle Arbeiten die Zielsetzung hatten, die Grenzen der jeweiligen Methoden kennen zu lernen. Insofern hat jede Arbeitsgruppe ihr eigenes Repertoire an validierten Methoden. Die positiven Erfahrungen mit dem hier vorgestellten integrierten Ansatz zeigen, daß es in Zukunft möglich sein wird, routinemäßig theoretische Verfahren gleichberechtigt in die gesamte Entwicklung eines Arzneimittels einzubeziehen. Die Modellierung von Ligand Komplexen mit dem Zielprotein der Arzneistoffwirkung ist nur einer der Schritte im Arzneistoffschicksal. Jeder Metabolisierung liegen solche Wechselwirkungskomplexe zugrunde, sie stehen damit einer rationalen Optimierung offen. Auch die Membranpassage wird als Phänomen simulierbar. Die hier skizzierte Vorgehensweise sollte damit entlang des gesamten Wegs einer Entwicklung vom Wirkstoff zum Arzneimittel Begleitung bieten.

Wie groß ist die Tierschutzrelevanz der CADD Techniken? Es ist generell festzustellen, daß in der letzten Dekade der „Verbrauch“ von

Versuchstieren drastisch abgenommen hat. Dies ist auf ein zunehmendes Bewußtsein in der Öffentlichkeit und bei den Forschenden, sowie auf die konsequentere Durchsetzung von Gesetzen und Verordnungen zurückzuführen. Der prozentuale Anteil, der einer „rationalen Arzneistoffentwicklung“ dabei zuzuordnen wäre, ist nicht leicht festzustellen und wird kontrovers diskutiert.

Es steht völlig außer Zweifel, daß ein sorgfältig entwickeltes Modell eines Protein-Ligand Komplexes für ein therapeutisch interessantes Zielprotein in der Lage sein wird, Leitstrukturen für ihre therapeutische Anwendung zu optimieren. Das Modell schließt damit eine Reihe von Analogverbindungen, die im Regelfall synthetisiert worden wären, aus dem Tierversuch aus. Jede theoretische Betrachtung von Protein-Ligand Wechselwirkungen birgt damit ein Potential zur Einsparung von Tierexperimenten.

Der oben beschriebene integrierte Ansatz mit der engen Verknüpfung von Biophysik, Biochemie und computerunterstützter Modellierung katalysiert noch viel stärker die *in vitro* Testung anstelle des Tierexperiments.

Allerdings schränkt die notwendige reduktionistische Betrachtung innerhalb des „rationalen Weges“ auch erheblich seine Erfolgsquote ein. Extrapolationen, wie das Überspringen von Therapieklassen und Strukturklassen, sind selten, bzw. nie erfolgreich gewesen. Alternative Bindungsmodi (s.o.), eng mit diesem Problem verknüpft, treten auch bei Agonisten auf und lassen uns eine neue Substanz nur mit einer gewissen Wahrscheinlichkeit ausschließen, nie mit Sicherheit. Unter wissenschaftlichen wie ökonomischen Aspekten bleibt ein „Restrisiko“, eine optimale Substanz verpaßt zu haben. Dem stand oft eine sehr unkritische Überbewertung der „bunten Bilder“ von Seiten der Anwender entgegen und hat zum Teil heftige Ablehnung der Vorgehensweise bei den experimentell tätigen Kollegen erzeugt. Hier kommt umso

stärker der Grundgedanke des Artikels zum tragen, der fordert, möglichst direkt *in vitro* Tests mit der Modellierung zu verknüpfen (Abb. 2). Dies müßte nicht nur für die pharmakodynamische Phase, also die Wirkung des Arzneistoffs am Zielort, sondern ebenso für alle seine pharmakokinetischen und metabolischen Phasen erfolgen.

Die heute versuchte molekulare Definition von Erkrankungen führt zu Tierversuchen. Nach der theoretischen und *in vitro* experimentellen Optimierung eines komplexen Systems, wie z.B. das der T-Zellantwort unseres Immunsystems, ist eine weitere Überprüfung am gesamten System erforderlich. Dies ist im Fall des Immunsystems das (transgene) Tier. Es ist aber auch durchaus umstritten, ob ein synthetischer Ligand des MHC-Komplexes (s.o.) therapeutisch überhaupt sinnvoll ist.

Dies zeigt die Schwierigkeit einer generellen Bewertung der computer-gestützten Arzneimittelentwicklung in Bezug auf Tierversuche. Die Methode trägt sicher dazu bei, den Beginn der Experimente möglichst weit zu verschieben und damit die Tierzahlen zu reduzieren. Dies gilt jedoch nur unter der Voraussetzung, daß diese Methode gleichberechtigt in die Entwicklung eines Arzneimittels einbezogen wird und man die mit ihrer Hilfe abgeleiteten Vorschläge auch akzeptiert.

Literatur:

- Beck-Sickinger, A. G., Wieland, H. A., Wittneben, H., Willim, K.-D., Rudolf, K. und Jung, G. (1994). Complete L-alanine scan of neuropeptide Y reveals ligands binding to Y1 and Y2 receptors with distinguished conformations. *Eur. J. Biochem.* 225, 947-958.
- Beck-Sickinger, A., Kiess, M., Kern, P. und Folkers, G. (1995). Secondary structure prediction of adenylate kinase by circular dichroism spectroscopy of synthetic peptides. *Pharm. Helv. Acta* 70, 33-43.
- Baldwin, M. J., (1993). The probable arrangement of the helices in G-protein coupled receptors. *EMBO J* 12, 1693-1703.
- Blundell, T., Hubbard, R. und Weiss, M. A. (1992). Structural biology and diabetes mellitus, molecular pathogenesis and rational drug design. *Diabetologica* 35, 69-76.
- Böhm, H. J. (1992). The computer program LUDI: A new method for de novo design of enzyme inhibitors. *J. Comput. Aided Mol. Design* 6, 61-78.
- Caffisch, A., Miranker, A. und Karplus, M. (1993). Multiple copy simultaneous search and construction of ligands in binding site: application to inhibitors of HIV-1 aspartic proteinase. *J. Med. Chem.* 36, 2142-2167.
- Caffisch, A. (1995). Computer-aided combinatorial ligand design: application to human thrombin (persönliche Mitteilung, Publikation in Vorbereitung).
- Feng, N. und Scheraga, H. A. (1994). Use of the transferred nuclear overhauser effect to determine the conformations of ligands bound to proteins. *Acc. Chem. Res.* 27, 258-265.
- Fong, T. M., Cascieri, M. A., Hong, Y., Bansal, A. und Swain, C. (1993). Amino-aromatic interaction between histidine 197 of the neurokinine receptor and CP 96345. *Nature* 362, 350-353.
- Folkers, G. und Merkle, H. P. (1995). Technologien für die Medikamente der Zukunft. *ETH Bulletin*, 36-41.
- Greenidge, P. A., Merz, A. und Folkers, G. (1995). A pseudoreceptor modeling study of the Varicella-Zoster Virus and Human Thymidine Kinase Binding Sites. *J. Comput. Aided Mol. Design* (submitted).
- Gunsteren, W. F. van und Mark, A. E. (1992). On the interpretation of biochemical data by molecular dynamics computer simulation. *Eur. J. Biochem.* 204, 947-961.
- Hibert, M., Trumpp-Kallmeyer, S., Bruinvels, A. und Hoflack, J. (1991). Three-dimensional models of neurotransmitter G-protein coupled receptors. *Mol. Pharmacol.* 40, 8-15.
- Höltje, H.-D. und Kier, L. B. (1975). Nature of the anionic or a-site of cholinesterase. *J. Pharm. Sci.* 61, 418-421.
- Kern, P., Rognan, D. und Folkers, G. (1995a). MD simulations in pseudoparticle fluids. *J. of QSAR and Molecular Modeling* (im Druck).
- Kern, P., Blundell, T., Johnson, M. S., Krickl, S., Michael, M. und Folkers, G. (1995b). Knowledge based structure prediction of HSV-1 Thymidine kinase substructures. *Pharm. Helv. Acta* (im Druck).
- Kubinyi, H. (1993). *QSAR: Hansch analysis and related approaches*. Weinheim: VCH.
- Madden, D., Garboczi, D. N. und Wiley, D. C. (1993). The antigenic identity of peptide-MHC complexes: A comparison of the conformations of five viral peptides presented by HLA-A2. *Cell* 75, 695-708.
- Mattos C. und Ringe D. (1993). Multiple binding modes. In H. Kubinyi (Hrsg.), *3D QSAR in Drug Design, Theory, Methods and Applications* (226-254). Leiden: Escom.
- Meyer, E. F., Radhakrishnan, R., Cole, G. M. und Presta, L. G. (1986). Structure of the product complex of acetyl-Ala-Pro-Ala with porcine elastase at 1.65 Å resolution. *J. Mol. Biol.* 189, 533-539.
- Michael, M., Fetzter, J. und Folkers, G. (1994). Site-directed mutagenesis clarifies the substrate position within the 3D model of the active site of HSV-1 Thymidine kinase. *Eur. J. Biochem.* 226, 219-226.
- Rognan, D., Zimmermann, N., Jung, G. und Folkers, G. (1992). Molecular dynamic study of a complex between the human histocompatibility antigen HLA-A2 and the IMP 58-66 nonapeptide from influenza matrix protein. *Eur. J. Biochem.* 208, 101-113.
- Rognan, D., Scapozza, L., Folkers, G. und Daser, A. (1995). Rational design of non-natural peptides as high-affinity ligands for the HLA B2705* human leucocyte antigen. *Proc. Natl. Acad. Sci. USA* 92, 753-757.
- Schertler, G. F., Villa, C. und Henderson, R. (1993). Projection structure of rhodopsin. *Nature* 362, 770-772.
- Schwartz, T. W. (1994). G-Protein Coupled Receptors. *Current Opinion in Biotechnology* 5, 343-444.
- Sussman, J. L., Harel, M., Frolova, F., Oefner, C., Goldman, A., Toker, L. und Silman, I. (1991). Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding protein. *Science*, 253, 872-879.
- Vedani, A. (1994). Das Konzept des Pseudorezeptors für das pharmakologische Screening. *ALTEX* 11, 11-21.
- Vedani, A., Zbinden, P., Snyder, J. P. und Greenidge, P. A. (1995). Pseudoreceptor modeling: The construction of three-dimensional receptor surrogates. *J. Am. Chem. Soc.* 117, 4987-4994.
- Weber, C., Wider, G., von Freyberg, B., Traber, R., Braun, W., Widmer, H. und Wüthrich, K. (1991). The NMR Structure of Cyclosporin A bound to cyclophilin in aqueous solution. *Biochemistry* 30, 6563-6574.
- Zimmermann, N., Beck-Sickinger, N., Müller, I., Folkers, G. und Jung, G. (1992). Conformational and epitope mapping of herpes simplex virus type 1 thymidine kinase using synthetic peptide segments. *Eur. J. Biochem.* 200, 519-528.

Korrespondenzadresse

Gerd Folkers
Departement Pharmazie,
ETH Zürich
Winterthurerstr. 190
CH-8057 Zürich