

Entwicklung eines Pseudorezeptormodells für das virtuelle Screening

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Biowissenschaften
der Johann Wolfgang Goethe-Universität
Frankfurt am Main

von
Yusuf Tanrikulu
aus Frankfurt am Main

Frankfurt 2008
(D30)

vom Fachbereich Biowissenschaften der
Johann Wolfgang Goethe-Universität Frankfurt am Main als Dissertation angenommen.

Dekan:	Prof. Dr. Volker Müller
Erster Gutachter:	Prof. Dr. Gisbert Schneider
Zweiter Gutachter:	Prof. Dr. Joachim Engels
Datum der Disputation:	29.04.2009

Inhaltsverzeichnis

Zusammenfassung	5
Summary	6
Einleitung	7
Das Pseudorezeptorkonzept.....	9
Voraussetzungen zum Erstellen eines Pseudorezeptors	11
Bekannte Ansätze zur Generierung von Pseudorezeptoren.....	13
Hypothese	18
Daten, Methoden und Algorithmen	19
Bearbeitete Zielproteine	19
Implementierung eines Pseudorezeptorkonzeptes.....	28
Verwendete Methoden und Programme	38
Moleküldatenbanken	43
Ergebnisse und Diskussion	44
Kompatibilität zwischen kristallisierten Bindetaschen und Pseudorezeptoren	44
Konstruktion eines Pseudorezeptors für Antagonisten des H4R.....	51
Evaluation von potentiellen H4R Bindetaschen.....	54
Virtuelles Screening nach H4R Liganden	59
Vergleich mit bisherigen Anwendungen	62
Schlusswort und Ausblick	68
Literaturverzeichnis	70
Eidesstattliche Erklärung	83
Danksagung	84
Lebenslauf	85
Publikationsliste	86
Anhang	89

Abkürzungsverzeichnis

2D	Zweidimensional
3D	Dreidimensional
5-HT _{1A}	5-Hydroxytryptamin (Serotonin) Rezeptor 1A
CDK	<i>Chemistry Development Kit</i>
COX-1	Cyclooxygenase-1
COX-2	Cyclooxygenase-2
DBD	DNA-Bindedomäne
GPCR	G-Protein gekoppelte Rezeptor(en)
H4R	Histamin H ₄ Rezeptor
IUPAC	<i>International Union of Pure and Applied Chemistry</i>
LBD	Ligandenbindedomäne
MD	Molekulare Dynamik(simulation)
PDB	<i>Protein Data Bank</i>
PPAR	Peroxisom Proliferator aktivierter Rezeptor
QSAR	quantitativen Struktur-Aktivitätsbeziehung
ROC	<i>Receiver operator characteristic</i>
RXR	Retinoid-X-Rezeptor
VCAM1	<i>Vascular Cell Adhesion Molecule 1</i>
VLA4	<i>Integrin Very Late Antigen 4</i>
ZNS	Zentrales Nervensystem

Zusammenfassung

Im Rahmen dieser Arbeit wurde die Eignung von Pseudorezeptoren im virtuellen Screening untersucht. Hierzu wurde nach intensiver Auseinandersetzung mit bisher bekannten Konzepten ein neues Computerprogramm zur automatischen Konstruktion von Pseudorezeptormodellen entwickelt. Das Ziel von Pseudorezeptoren ist die Konstruktion eines alternativen, artifiziellen Wirtssystems aus bekannten Liganden eines Zielproteins, dessen dreidimensionale Struktur unbekannt ist.

Der generierte Pseudorezeptor ist zu verstehen als die Menge aller Pseudoatome, die um die Ausgangssubstanz(en) projiziert werden. Bei multiplen Referenzliganden wird eine Gewichtung der Pseudoatome durchgeführt. Zudem wird ausschließlich von Distanz- und Winkelparametern Gebrauch gemacht, die aus Untersuchungen von Kokristallstrukturen gewonnenen wurden.

Eine abschließende Kodierung generierter Pseudorezeptoren als 90-dimensionalen Korrelationsvektor wurde zum virtuellen Screening eingesetzt. In zwei retrospektiven Fallbeispielen wird gezeigt, dass die generierten Pseudorezeptoren für COX-2 und PPAR α mit den realen Zuständen ihrer kokristallisierten Bindetaschen in den PDB Einträge 6cox und 2p54 kompatibel sind.

Im retrospektiven virtuellen Screening in der Wirkstoffdatenbank COBRA (8.311 Moleküle) nach COX-2 Inhibitoren (136 Aktive) konnte eine Anreicherung der aktiven Strukturen in den ersten zwei Perzentilen gezeigt werden (54% der Aktiven). Zudem konnten 80% der aktiven Moleküle bereits nach Vorhersage von 10% Falsch-Positiven gefunden werden. Im Falle des retrospektiven Screenings nach 94 PPAR Liganden konnten 30% der aktiven Moleküle nach der Vorhersage von 10% Falsch-Positiven entdeckt. Nach 20% Falsch-Positiver wurden 46% der PPAR Liganden wieder gefunden.

Weiterhin konnte mit den ligandenbasierten Informationen eines H₄ Pseudorezeptors eine Justierung einer potentiellen Bindetasche des Histamin H₄ Rezeptors aus einer molekularen Dynamiksimulation vorgenommen werden. Schließlich wurde in einem prospektiven virtuellen Screening nach Histamin H₄ Liganden mit einem Pseudorezeptor zwei Strukturen mit unterschiedlichem Grundgerüst und einem $K_i \sim 30 \mu\text{M}$ identifiziert.

Summary

In this thesis, the suitability of pseudoreceptors for virtual screening applications was analyzed. An automated pseudoreceptor construction program was developed after known design principles had been thoroughly studied and compared. The aim of pseudoreceptor modelling is the construction of an alternative host system for known ligands of a given target protein in the absence of three-dimensional structure information.

The constructed pseudoreceptor is represented as the sum of all pseudoatoms, which are projected around reference ligand(s). A weighting scheme is introduced, when pseudoreceptors are generated from multiple reference ligands. For pseudoatom placing distance and angle parameters from a survey of known co-crystal structures were used.

For virtual screening pseudoreceptors were encoded as correlation vectors. It is demonstrated that the generated pseudoreceptors match with their respective co-crystallized binding pockets, taking COX-2 and PPAR-alpha as an example (PDB entries 6cox and 2p54).

In a retrospective virtual screening in the drug collection COBRA (8,311 molecules) for COX-2 inhibitors (136 actives) high enrichment of ligands in the first two percentiles was yielded (54% of the actives). 80% of all active compounds were found after the prediction of only 10% false-positives. In a retrospective screening study for 94 PPAR ligands, 30% of the actives were found together with 10% false-positives. After the prediction of 20% false-positives, 46% of all PPAR ligands could be found.

In addition, a putative binding pocket of the histamine H₄ receptor from a molecular dynamics simulation could be adjusted using ligand-based information of a H₄ pseudoreceptor.

Finally, two micromolar ligands with different scaffolds were identified with a $K_i \sim 30 \mu\text{M}$ by a pseudoreceptor-based prospective virtual screening for novel H₄ ligands.

Einleitung

Rechnergestützte Methoden unterstützen die pharmazeutische Forschung und Entwicklung nunmehr seit drei Jahrzehnten. Hierbei bildeten sich zwei Hauptentwicklungsprinzipien, nämlich rezeptorbasierte (auch „strukturbasierte“ genannt) und ligandenbasierte Ansätze (Congreve *et al.*, 2005; Klebe, 2006; Schneider & Baringhaus, 2008). Diese Strategien werden, insbesondere in Kombination mit fragmentbasierter Leitstrukturfindung (Carr *et al.*, 2005), als Ergänzung zum Hochdurchsatzscreening (*high-throughput screening*, HTS) angesehen, welches trotz immenser Kosten die Hauptinnovationsquelle für viele pharmazeutische Firmen bildet (Bleicher *et al.*, 2003; Alber *et al.*, 2007; Koehn, 2008).

Anwendungsgebiete von ligandenbasierten Methoden, wie die molekulare Ähnlichkeitssuche oder das pharmakophorbasierte virtuelle Screening, sind der Austausch von molekularen Grundgerüsten (*scaffold-hopping*) (Schneider *et al.*, 1999; Schneider *et al.*, 2006; Brown & Jacoby, 2006; Zhao, 2007) und Studien zur Wirkstoffwiederverwendung (*repurposing*) (Carley, 2005a; Carley, 2005b; Bisson *et al.*, 2007). Grundvoraussetzung für die Anwendung ist die Verfügbarkeit von mindestens einem bioaktiven Referenzmolekül. Wenn eine dreidimensionale (3D) Rezeptorstruktur verfügbar ist, können auch rezeptorbasierte Methoden angewendet werden (Mohan *et al.*, 2005; Rester, 2006). Oft findet eine simultane Anwendung beider Techniken statt. Pseudorezeptormodelle, die vor 20 Jahren erdacht wurden (Kato *et al.*, 1987; Kato *et al.*, 1992), können als „Brücke“ zwischen den Ansätzen fungieren, wenn die parallele Anwendung beider Strategien nicht möglich ist.

Zu Beginn eines Wirkstoffentwicklungsprojektes ist die 3D Struktur eines Zielproteins nicht immer verfügbar. Dies liegt an dem Trend zur Suche nach Wirkstoffen für G-Protein gekoppelter Rezeptoren (GPCR), für die es derzeit an kristallisierten Strukturen mangelt. Somit sind diese Projekte auf die Nutzung von ligandenbasierten Methoden eingeschränkt. Gegenwärtig beinhaltet die *Protein Data Bank* (PDB) (Bernstein *et al.*, 1977), seit 1998 unter der Leitung des *Research Collaboratory for Structural Bioinformatics* (RCSB) (Bermann *et al.*, 2000), ungefähr 50.000 Einträge von aufgeklärten Strukturen, deren Zahl jährlich exponential wächst. Nichtsdestotrotz sind es nur 27 Einträge (0,05%), die mit GPCR in Beziehung stehen. Hierunter fallen 23

Strukturen des Rhodopsins (PDB Einträge 1ln6, 2j4y, 1h68, 1jfp, 1jgj, 2i35, 1gue, 1xio, 2f93, 2i36, 2i37, 1f88, 1gu8, 1hzx, 119h, 1u19, 2f95, 2ii7, 2ii8, 2g87, 2hpy, 1gzm und 2ped), drei Strukturen des β_2 adrenergen Rezeptors (PDB Einträge 2rh1, 2r4r und 2r4s) und eine Struktur des β_1 adrenergen Rezeptors (PDB Eintrag 2vt4). Durch die limitierte Zahl von 3D Informationen zu Membranproteinen ist die Anwendbarkeit von rezeptorbasierten Ansätzen für viele pharmazeutisch interessante Zielproteine eingeschränkt (Scapin, 2006; Lacapère, 2007). Weiterentwicklungen der Kristallisierungsmethoden von Membranproteinen in Kombination mit Hochdurchsatzkristallisation sollen diesen Rückstand nachbessern (Granseth *et al.*, 2007). Zwischenzeitlich verschafft man sich Abhilfe durch die Anwendung von rezeptorbasierten Ansätzen auf Homologiemodellen, die von verwandten Proteinen mit hoher Sequenzidentität abgeleitet sind (Costanzi, 2005).

Um diese Kluft zu überbrücken, wurden hybride Ansätze entwickelt. Der prominenteste Ansatz ist die Modellierung von quantitativen Struktur-Aktivitätsbeziehungen (QSAR) zur Korrelation von Bindeaffinitäten bekannter Liganden und deren physikochemischen und topologischen Eigenschaften (Tropsha & Golbraikh, 2007; Lill, 2007). Im Kontext von Pseudorezeptoren sind insbesondere diejenigen QSAR Techniken interessant, welche molekulare Merkmale auf ein Feld (Oberfläche oder Gitter) um die molekulare Überlagerung von Referenzkonformeren generieren (3D QSAR). Einige Beispiele auf diesem Gebiet sind CoMFA (Cramer *et al.*, 1988), GOLPE (Baroni *et al.*, 1993), CoMSIA (Baroni *et al.*, 1993) und QSiAR (Kubinyi, 1998). Diese können zur Erklärung der für eine Interaktion von Ligand und Rezeptor verantwortlichen molekularen Attribute herangezogen werden. Folglich werden 3D QSAR Methoden auch zur Vorhersage von Aktivitäten neuer Screeningsubstanzen eingesetzt. Obwohl diese Fähigkeit bereits in vielen Publikationen demonstriert wurde (Tropsha & Golbraikh, 2007; Lill, 2007), ist die Nutzung von nicht-atombasierten 3D QSAR Methoden in Wirkstoffentwicklungsprojekten unüblich. Ein Grund hierfür ist die bessere Erklärbarkeit von Rezeptor-Liganden Interaktionen bei atombasierten Methoden, die ein direktes Design von chemischen Strukturen erlauben. Pseudorezeptormodelle repräsentieren eine Methode, die solche Vorzüge bietet.

Tabelle 1. Chronologische Übersicht aller Pseudorezeptorentwicklungen.

Programm	Jahr	Typ	Literatur
RECEPS	1987	gitterbasiert	(Kato <i>et al.</i> , 1987 & 1992)
HASL	1988	gitterbasiert	(Doweyko, 1988)
MORPHEUS	1989	gitterbasiert	(Andrews <i>et al.</i> , 1989)
VOROM	1991	partitionsbasiert	(Crippen, 1991; Bradley <i>et al.</i> , 1993)
Yak	1993	peptidbasiert	(Vedani <i>et al.</i> , 1993; Vedani <i>et al.</i> , 1995)
GERM	1994	atombasiert	(Walters & Hinds, 1994)
RECEPTOR	1995	oberflächenbasiert	(Hahn, 1995; Hahn & Rogers, 1995)
QXP	1997	fragmentbasiert	(McMartin & Bohacek, 1997)
EGSITE2	1997	partitionsbasiert	(Crippen, 1997)
SKELGEN	1997	partitionsbasiert	(Todorov & Dean, 1997 & 1998)
PrGen	1998	peptidbasiert	(Zbinden <i>et al.</i> , 1998)
PARM	1998	atombasiert	(Chen <i>et al.</i> , 1998)
Quasar	1998	atombasiert	(Vedani & Zbinden, 1998)
Hostdesigner	2002	fragmentbasiert	(Hay & Firman, 2002)
FLARM	2002	atombasiert	(Pei & Zhou, 2002)
RAPTOR	2004	oberflächenbasiert	(Lill <i>et al.</i> , 2004)
WeP	2004	atombasiert	(Chae <i>et al.</i> , 2004)
FLARM2.0	2005	atombasiert	(Pei <i>et al.</i> , 2005)
ConCept	2007	fragmentbasiert	(Chen <i>et al.</i> , 2007)

Das Pseudorezeptorkonzept

Die Idee von Pseudorezeptormodellen ist die Konstruktion eines Ersatzes für eine 3D Struktur des Zielproteins um Ligandenkonformationen (Tanrikulu & Schneider, 2008). Das Modell soll dabei die Form der Bindetasche und die für eine erfolgreiche Ligandenbindung notwendigen Interaktionspunkte vorzeigen. Pseudorezeptoren beschreiben einen Rezeptorersatz, der ein alternatives Wirtssystem zur nicht-kovalenten Bindung bekannter Liganden mit hoher Affinität darstellt.

Pioniere der Pseudorezeptormodellierung konstruierten diese auf manuelle Art und Weise – zum Beispiel durch die Faltung einer Peptidkette in eine „Bettform“, bestehend aus einem zentralen β -Faltblatt flankiert von zwei α -Helices (Momany *et al.*, 1989). Andere entwickelten einen Pseudorezeptor indem sie um einen GRID-Konsensuspharmakaphor (Davis *et al.*, 1994) mit Hilfe des Programms SYBYL (Tripos Inc., St. Louis, USA, www.tripos.com) die zur Ligandeninteraktion identifizierten Aminosäureketten positionierten (Höltje & Anzali, 1992). Zeitgleich wurden interaktive Pseudorezeptormodellierungsprogramme entwickelt. Obwohl Vedani *et al.* (Snyder *et al.*, 1993) eine Definition zur Unterscheidung von „Minirezeptoren“ (ein nicht

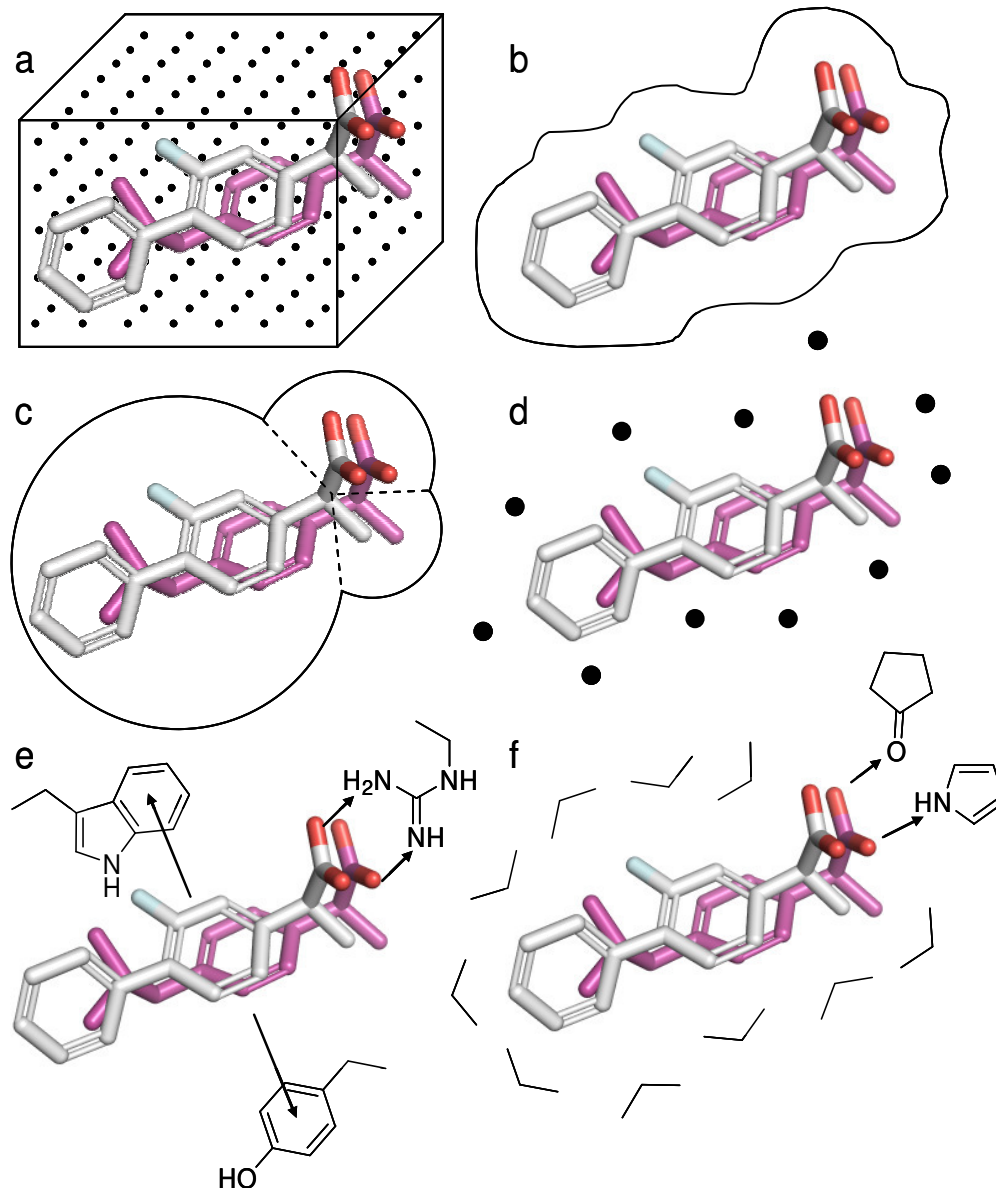


Abbildung 1. Konzeptskizzen eines a) gitter-, b) oberflächen-, c) partitions-, d) atom-, e) peptid- und f) fragmentbasierten Pseudorezeptors um das Strukturalignment von Flurbiprofen (hellgrau) und Ibuprofen (violett) (PDB Einträge 1e9g und 1e9h).

verbundener Satz von kleinen Molekülen, die die Rezeptorseitenketten um einen Liganden simulieren) und „Pseudorezeptoren“ (voll verbundener Satz von peptidischen Molekülen, die eine Bindetasche formen) einführt, sollen in dieser Arbeit alle Variationen als Pseudorezeptoren (bzw. -modelle) bezeichnet werden. Tabelle 1 fasst alle Entwicklungen von Pseudorezeptoren in chronologischer Reihenfolge zusammen. Die Methodik der Pseudorezeptorgenerierung beinhaltet drei fundamentale Schritte (Botta *et al.*, 2002):

1. Definition der essentiellen Rezeptor-Liganden Interaktionspunkte.

2. Konstruktion eines Pseudorezeptors um die Interaktionspunkte.
3. Optimierung der Modellkoordinaten.

Eine breite Auswahl von Techniken zur Pseudorezeptorgenerierung wurde entwickelt, welche hier in sechs Kategorien (gitter-, oberflächen-, partitions-, atom-, peptid- und fragmentbasierte Ansätze) eingeteilt werden sollen (Abbildung 1). Das Hauptziel ist die Generierung eines Pseudorezeptors um Ligandenkonformationen, welcher die Bindetasche des Zielproteins widerspiegelt.

Voraussetzungen zum Erstellen eines Pseudorezeptors

Die Verfügbarkeit mindestens einer aktiven Molekülstruktur ist die Grundvoraussetzung zum Erstellen eines Pseudorezeptors. Bei Verfügbarkeit von mehreren Molekülen müssen diese mittels molekularem Alignment übereinander gelagert werden. Wie in jedem QSAR Ansatz sollten die Ligandenaktivitäten drei bis vier Größenordnungen aufspannen, um die Entwicklung von statistisch plausiblen und generalisierenden Modellen zu ermöglichen. Geht man von einem gemeinsamen Bindemodus der verfügbaren Referenzstrukturen aus, sollten funktional äquivalente molekulare Substrukturen, die für die nicht-kovalenten Wechselwirkungen eines Liganden verantwortlich sind (Schlüsselgruppen), in gleiche Orientierung gebracht werden. Diese Annahme ist jedoch nicht immer zutreffend, da ein Zielprotein mehrere Möglichkeiten zur Ligandenbindung besitzen kann (Kim, 2007a; Kim, 2007b). Deswegen ist die Auswahl von Referenzmolekülen, deren Bindemodus experimentell durch Mutationsstudien bestimmt wurde (Evers *et al.*, 2005), anzuraten.

Die räumliche Anordnung der funktionellen Gruppen eines Moleküls bezeichnet man als „Pharmakophor“. Er bildet die Grundlage zur Bindung eines Liganden am Zielprotein. Die *International Union of Pure and Applied Chemistry* (IUPAC) Definition lautet (Wermuth *et al.*, 1998):

Pharmacophore: The ensemble of steric and electrostatic factors, which are necessary to insure supramolecular interactions with a specific biological target structure.

Die Verwendung von pharmakophorbasierten Ansätzen in der rechnergestützten Wirkstoffforschung ist eine weit verbreitete und erfolgreiche Methode zum Auffinden von aktiven Molekülen (Güner, 2000). Die Anwendung dieser Technik wurde 1996 von Sheridan *et al.* eingeführt (Sheridan *et al.*, 1996). Drei Jahre später zeigten Schneider *et al.* mit dem Programm CATS, dass sich diese Methode für das Scaffold-Hopping eignet, indem Pharmakophorpunkte unter Anbetracht ihrer topologischen Distanz auf dem Molekülgraphen analysiert wurden (Schneider *et al.*, 1999). Um den Verlust von Informationen der 3D Molekülkonformationen zu verringern, wurde das Programm CATS3D entwickelt, welches die Euklidische Distanz zwischen den Pharmakophorpunkten berechnet (Fechner *et al.*, 2003). In dem Programm SQUID wurde hiernach eine Gruppierung der Pharmakophorpunkte erreicht, die von nun an als univariate Gaußverteilungen im Raum kodiert wurden (Renner & Schneider, 2004). Jüngst wurde auch diese Methode um eine trivariate Darstellung der Gaußschen Pharmakophorpunkteverteilungen erweitert (Tanrikulu *et al.*, 2007).

Das Konzept der Überlagerung von Pharmakophorpunkten (auch „Pharmakophoralignment“) ist eine der Grundlagen für ein molekulares Alignment. Ein Problem zeigt sich durch die besonders hohe Konformationsfreiheit flexibler Moleküle. So treten Fälle auf, die keine Überlagerung der Schlüsselgruppen ermöglichen, oder nach einem Alignment die Strukturen zwar überlagert sind, obwohl sie tatsächlich keinen gemeinsamen Bindemodus haben.

Ein Lösungsansatz ist die Auswahl der potentesten Referenzmoleküle zur Pseudorezeptorgenierung, die nur einem Bindemodus entsprechen könnten. Anders geht der *Active Analogue Approach* (Marshall *et al.*, 1979; Marshall, 1993) vor, welcher eine Reihe von Molekülkonformationen mit verschiedenen Pharmakophoren vorschlägt. Hierbei nimmt man an, dass die bioaktive Konformation im Vergleich zur freien, energieniedrigsten Konformation des Moleküls nicht völlig unterschiedlich ist (Pei *et al.*, 2005). Ein anderer Ansatz ist die initiale Generierung eines Pseudorezeptors mit dem aktivsten Liganden und die iterative Addition und Optimierung der Verbleibenden (Zbinden *et al.*, 1998).

Falls Pseudorezeptoren ergänzend zu rezeptorbasierten Methoden erstellt werden, kann man erste Koordinaten (z.B. von Aminosäureseitenketten) aus Röntgenkristallstrukturen ableiten. Zudem kann das Alignment der Referenzmoleküle aus vorhergesagten

Bindeposen von Dockingexperimenten gewonnen werden, welches vergleichbar zu Ligandenüberlagerungen durch Proteinalignments aus Kristalldaten ist (Pei *et al.*, 2005).

Der initiale Schritt der Überlagerung von Pharmakophorpunkten ist der wichtigste Schritt, weil es den Startzustand der Pseudorezeptorgenerierung bestimmt. Kleine Veränderungen in der Überlagerung können zu stark unterschiedlichen finalen Modellen führen (Walters & Hinds, 1994; Chen *et al.*, 1998; Pei & Zhou, 2002; Pei *et al.*, 2005).

Bekannte Ansätze zur Generierung von Pseudorezeptoren

Gitterbasierte Ansätze

Erste Ansätze zur Pseudorezeptorgenerierung folgten der Idee, das Molekülalignment mit einem kubischen Gitter zu umspannen (Abbildung 1a) (Kato *et al.*, 1987; Doweiko, 1988; Andrews *et al.*, 1989; Kato *et al.*, 1992). Vergleichbar zu bekannten 3D QSAR Methoden wurden Maschenweiten von 1,0-2,0 Å verwendet. An jedem Gitterpunkt wurden verschiedene physikochemische Eigenschaften der Liganden aufgenommen, wie z.B. das elektrostatische Potenzial, die Ladungsverteilung und Wasserstoffbrückenfähigkeit. Bei einer korrekten Generierung stellen diese Modelle ein abstraktes Abbild der darunter liegenden Moleküle dar, ohne auf deren Topologie oder atomare Struktur, jedoch auf deren Funktion einzugehen. Die Verfeinerung solcher Modelle wurde durch iteratives Hinzufügen weiterer Referenzmoleküle realisiert.

Oberflächenbasierte Ansätze

Um gitterbasierte Probleme, wie die Wahl der korrekten Maschenweite, zu umgehen, wurden oberflächenbasierte Ansätze entwickelt (Abbildung 1b). Fortgeschrittene Methoden stellen die aufgenommenen Eigenschaften als eine analytische Oberfläche mittels dem *Marching Cube* Algorithmus (Delibasis *et al.*, 2001) dar. Durch die Verwendung von van der Waals oder Wyvill Funktionen (Wyvill *et al.*, 1986) werden ästhetisch ansprechende 3D Modelle der möglichen Bindetasche generiert (Hahn, 1995; Hahn & Rogers, 1995; Hahn & Rogers, 1998).

Die Darstellung der Bindetaschenoberflächen wird meist durch Triangulation erreicht (Heiden *et al.*, 1990) – ihre Form stellt ein Abbild der Bindetasche dar, welches zum

Ligandendesign nach dem Schlüssel-Schloss-Prinzip (Fischer, 1894) herangezogen werden kann. Diese bildliche Darstellung repräsentiert die Definition der expliziten Oberflächen (Hahn, 1995). Es wird angenommen, dass diese Darstellungsform eher der physikalischen Natur von Bindetaschen entspricht als es implizite Oberflächen (Ergebnis einer CoMFA Berechnung (Cramer *et al.*, 1988)) tun. Implizite Oberflächen basieren im Allgemeinen aus mehreren Oberflächen, welche die räumliche Ausdehnung von bestimmten Moleküleigenschaften erfassen. Oberflächen können nicht nur als Flächen modelliert werden, sondern auch punktuell als Oberflächendekomposition erfasst werden (Proschak *et al.*, 2008).

Durch die Einführung von „offenen“ Pseudorezeptoren wurde die Generierung von hypothetischen Bindetaschen möglich, die „Löcher“ auf ihrer Oberfläche beinhalten. Diese Bereiche kennzeichnen Lösungsmittel zugängliche oder unbekannte Bereiche der Bindetasche. Sie machen gerade bei denjenigen Liganden Sinn, die ins Solvent ragen oder nur einen Teil der Proteintasche ausfüllen.

Jüngere Weiterentwicklungen heben den Nutzen von multiplen Oberflächen hervor (Lill, 2007). Hierbei resultiert eine innere Oberfläche aus den Erkenntnissen der potentesten Referenzliganden, wohingegen eine äußere Schicht die der Restlichen umfasst. Das Konzept soll es ermöglichen, auf induzierte Passformphänomene (*induced-fit* (Koshland, 1958)) einzugehen, in der sich die Rezeptorstruktur der Topologie des Liganden anpasst (Tropsha & Golbraikh, 2007; Lill, 2007).

Partitionsbasierte Ansätze

Konzeptuell unterschiedliche Methode gehen weg von gitterbasierten und oberflächenbasierten Ansätzen. Diese Gruppe generiert Pseudorezeptoren durch Partitionierung des Raumes um die Referenzmoleküle in polyhedrale Voronoi-Flächen (Voronoi, 1908), um distinkte Subtaschen zu modellieren (Abbildung 1c) (Crippen, 1991; Bradley *et al.*, 1993). Weiterentwicklungen dieses Ansatztyps erlauben die automatisierte Gruppierung von Referenzmolekülen anhand von sphärischen Partitionen mit verschiedenen Radii um deren Pharmakophorpunkte. Eine ähnliche Idee wurde in dem Programm SKELGEN verfolgt, welches seinen Ursprung in dem rezeptorbasierten *de novo* Design hat (Todorov & Dean, 1997; Todorov & Dean, 1998). Dean *et al.* führten den „Ligandenmodus“ in SKELGEN ein, welcher die Eingabe von Gaußschen

Formen erlaubt, die zwar keine chemische Funktion modellierten, aber den verfügbaren Raum für das Design eingrenzen.

Atombasierte Ansätze

Im Jahr 1994 entwickelten Walters *et al.* eine Methode, bei der Pseudorezeptoren durch eine Menge von Atomen repräsentiert wurden (Abbildung 1d) (Walters & Hinds, 1994). Atome werden dabei auf eine Kugel projiziert, die den größten Liganden im Alignment einkreist. Bildlich kann man sich dies wie einen um die Liganden aufgeblasenen Ballon vorstellen, auf dem Punkte aufgemalt wurden. Durch das Ablassen der Luft verringert sich das Volumen des Ballons, wobei dieser am Ende die Umrisse der Liganden in ihm widerspiegelt. Schließlich positioniert man Atome an die markierten Stellen – üblicherweise 40-60 im gesamten Modell. Die Atomtypen werden dabei zufällig ausgewählt. Hieraus ergibt sich ein multidimensionales Optimierungsproblem, weil eine enorme Anzahl von Modellen generiert werden kann. Zu dessen Lösung kann unter anderem eine evolutionäre Strategie mit genetischem Algorithmus verwendet werden, welche Pseudorezeptoren als eine lineare Aneinanderreihung von Atomen repräsentiert (Walters & Hinds, 1994).

Chen *et al.* (1998) veränderten diesen Prozess, indem sie die Wahl der initialen Atomtypen von der elektrostatischen Ladung der Liganden abhängig machten. Andere Weiterentwicklungen sind flexible atombasierte Pseudorezeptormodelle, um Fehler in der initialen Atompositionierung zu umgehen (Chen *et al.*, 1998). Im Laufe des Verfeinerungsprozesses wird den Atomen ein Spielraum von 0,1 Å gewährt. Im Jahr 2005 ermöglichten Pei *et al.* noch flexiblere Modelle, indem sie Konformationsänderungen der Liganden während der Optimierung zuließen (Pei *et al.*, 2005).

Ein weiteres Mitglied der atombasierten Ansätze ist das Programm Quasar (Vedani & Zbinden, 1998). Es unterstützt die Verwendung von Flip-Flop Wasserstoffbrückenatomtypen, wodurch Pseudoatome gleichzeitig als Wasserstoffbrückenakzeptoren und –donoren gekennzeichnet werden können. Jüngst wurde eine weitere Methode entwickelt, die nicht nur die Auswahl der korrekten Atomtypen optimiert, sondern auch eine Gewichtung der Pseudoatome vollzieht (Chae *et al.*, 2004).

Peptidbasierte Ansätze

Abgesehen von der Positionierung einzelner Atome an dedizierten Positionen wurden ebenfalls Ansätze zur Platzierung kompletter Aminosäureseitenketten entwickelt (Abbildung 1e) (Vedani *et al.*, 1993). Anders als bei den bisher beschriebenen Modellen wird hierbei die räumliche Ausdehnung durch vorgegebene Aminosäuren angegeben. Die exakten Aminosäurekoordinaten werden ausgehend von denjenigen Ligandenatomen identifiziert, die unter anderem an einer Wasserstoffbrücke beteiligt sein könnten. Die Distanzen und Winkel zu diesen Atomen werden aus einer Kristallstrukturanalyse entnommen. Folgende Vektoren werden zu Hilfe gezogen:

- Wasserstoffbrückenvektoren, die von Ligandendonatoratomen (-NH-, -OH und -SH) zu Rezeptorakzeptorpositionen zeigen.
- Freies-Elektronenpaarvektoren, die von Ligandenakzeptoratomen (N, O und S) auf Rezeptordonoratome zeigen.
- Hydrophobe Vektoren, die auf Interaktionen zwischen lipophilen Bereichen des Liganden und des Rezeptors hinweisen.

Die Wichtigkeit eines Interaktionsbereiches wird durch ein Gewicht angegeben, welches sich aus der Anzahl, der Varianz und der Orientierung der Vektoren in diesem Bereich errechnet. Passende Aminosäuren werden ausgehend von einer Rotamerbibliothek, welche multiple Konformationen der Seitenketten beinhaltet, an die entsprechenden Endpunkte der Vektoren platziert (Ponder & Richards, 1987). Optional können Sekundärstrukturen des Zielproteins hinzugefügt werden, um eine höhere Ähnlichkeit zu modellieren.

Im Jahr 1998, wurde dieses Konzept weiterentwickelt (PrGen (Zbinden *et al.*, 1998)). Pseudorezeptormodelle die mit Hilfe von PrGen generiert wurden, sind unabhängig von der initialen Überlagerung der Referenzmoleküle. Liganden werden auf den Zustand eines primären Pseudorezeptors angepasst, welcher selbst durch den potentesten Liganden generiert wird. Zudem ist die Optimierungsphase so erweitert, dass nacheinander der Pseudorezeptor und ein Molekül aus der Menge der Referenzmoleküle energieminiert werden. Optional kann eine Monte Carlo Konformationsuche eingeschaltet werden, die ein globales Minimum der Ligandenkonformationen findet (Metropolis *et al.*, 1953). Resultierende Pseudorezeptoren werden anhand der Korrelation zwischen berechneten und experimentell beobachteten Aktivitäten bewertet.

Fragmentbasierte Ansätze

Die letzte Klasse der Pseudorezeptoren baut auf niedermolekularen Fragmenten auf, die Interaktionen zu den Referenzmolekülen ausbilden können (Abbildung 1f). McMartin *et al.* publizierten ein Softwarepaket, welches mit einem Subprotokoll (namens PSEUD) Pseudorezeptoren komplementär zu einem Pharmakophortemplat aufbauen kann (McMartin & Bohacek, 1997). Der Algorithmus funktioniert nicht vollautomatisch, weshalb ein Nutzer des Programms die Fragmente, die an einer Wasserstoffbrücke beteiligt sein können, manuell positionieren muss. Üblicherweise werden Pyrrol- oder Cyclopentanonringe als Aminosäuren-äquivalente Fragmente ausgewählt. Sie werden nahe an den Referenzmolekülen platziert und mit besonderen Bindungen verbunden. Im Nachhinein füllt der Algorithmus den verbleibenden Raum um die Liganden mit Propan, welches eine Kohlenstoffdichte ähnlich dem Valin oder Phenylalanin besitzt. Schließlich ergibt dies eine hydrophobe Tasche mit isolierten Wasserstoffbrückenbereichen, welche zum automatischen Ligandendocking herangezogen werden kann.

Eine völlig unterschiedliche Intention hatte die Entwicklung des Programms ConCept (Chen *et al.*, 2007). Hier entsprach das Ziel eher dem Design einer Rezeptorsatzstruktur, welche das eigentliche Protein im biologischen Experiment tatsächlich ersetzt. Einzig ein weiterer ähnlicher Ansatz ist bekannt, der jedoch auf Metallionen komplexierende Rezeptoren spezialisiert ist (Hay & Firman, 2002). Obwohl die Motivation der letzten beiden Methoden anders ist als die in dieser Arbeit besprochenen Pseudorezeptoren, ähneln sich die *in silico* Prozeduren stark. Offensichtlich müssen aber bei der Entwicklung von künstlichen Hostsystemen weitere Einschränkungen, wie die synthetische Erreichbarkeit, beachtet werden.

Pseudorezeptormodelle gehören in erster Linie zu den 3D QSAR Methoden in der Chemieinformatik. Ihre Anwendung zielte bisher auf folgende Aspekte hin:

- Untersuchung der Rezeptor-Liganden Interaktionsbereiche.
- Bewertung von neuen Screening- und Leitstrukturkandidaten.
- Faltungsvorhersage einer Proteinsequenz.
- Energieminimierung von Liganden.
- Ideengenerator für die Entwicklung von neuen Wirkstoffen.

Hypothese

In dieser Arbeit wird die Hypothese überprüft, ob sich Pseudorezeptoren zum virtuellen Screening im Sinne des Scaffold-Hoppings eignen.

Hierzu ist die Entwicklung einer Methode zur automatisierten Konstruktion von Pseudorezeptoren aus einer gegebenen Überlagerung von Referenzmolekülen notwendig. Ein generiertes Pseudorezeptormodell sollte den Vergleich mit Kokristallstrukturen bestehen und die Justierung von theoretischen Proteinmodellen erlauben.

Entsprechend der Hypothese sollte in einem retrospektiven virtuellen Screening die Anreicherung von bekannten aktiven Strukturen beobachtet werden können. Im prospektiven Experiment sollten neuartige Moleküle mit ähnlicher Bioaktivität vorgeschlagen werden.

Daten, Methoden und Algorithmen

Bearbeitete Zielproteine

Im Folgenden werden die drei Zielproteine Cyclooxygenase-2 (COX-2), Peroxisom Proliferator aktivierter Rezeptor α (PPAR α) und Histamin H₄ Rezeptor (H4R) beschrieben.

COX-2

COX-2 gehört mit COX-1 (Xie *et al.*, 1991) zu den Arachidonsäure verarbeitenden Enzymen (Hla & Neilson, 1992). Sie produzieren Prostaglandine und Thromboxane durch ihre katalytischen Fähigkeiten zur Cyclooxygenase- und Peroxidasereaktion (Romano & Claria, 2003). COX-2 wird im Zentralen Nervensystem (ZNS) und den Nieren konstitutiv exprimiert (Rocca *et al.*, 2001). Zusätzlich kann es durch sekundäre Induktion mit Wachstumsfaktoren überall im Körper (z.B. Leber, Blutkörperchen, Magen-Darm Trakt) aufgebaut werden (Morita, 2002). Zuletzt wurde eine Splicingvariante von COX-1 mit unbekannter Funktion namens COX-3 bestätigt (Chandrasekharan *et al.*, 2002).

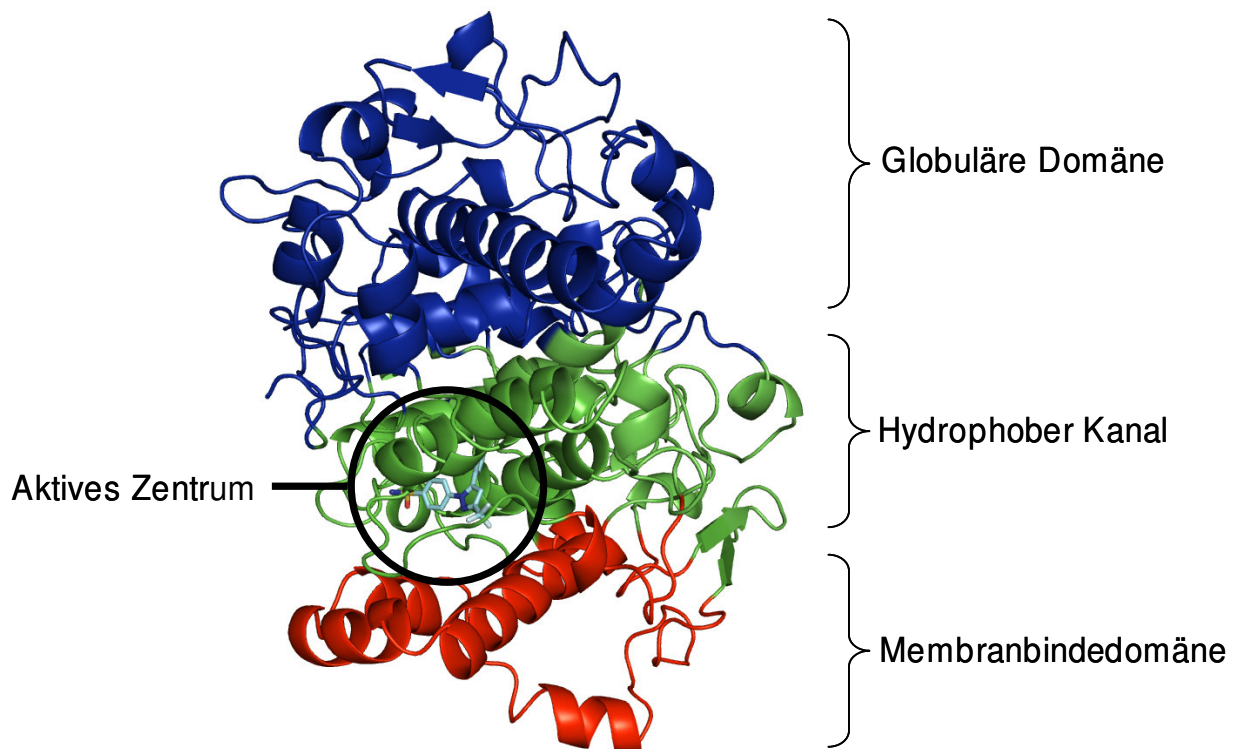


Abbildung 2. COX-2 Proteinstruktur mit dem Inhibitor SC-558 (6cox (Kurumbail *et al.*, 1996)).

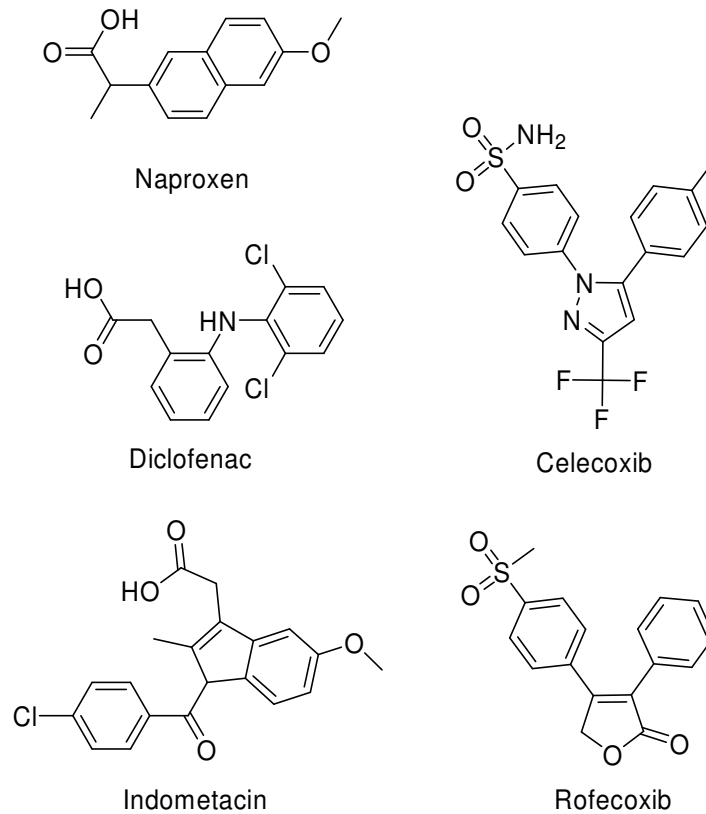


Abbildung 3. Ausgewählte Cyclooxygenase Inhibitoren. Naproxen, Diclofenac und Indometacin (Rowlinson *et al.*, 2003) sind COX-1 und COX-2 Liganden. Celecoxib und Rofecoxib sind selektive COX-2 Inhibitoren (McAdam *et al.*, 1999; Catella-Lawson *et al.*, 1999).

COX-2 hat eine Sequenzlänge von 587 Aminosäuren und bildet eine globuläre Domäne, einen hydrophoben Kanal und eine Membranbindedomäne aus (Abbildung 2). Das aktive Zentrum sitzt innerhalb des hydrophoben Kanals und wird von den Aminosäuren Arg120, Val349, Tyr355, Tyr385, Val434, Arg513, Val523 und Ser530 gebildet (Steinhilber *et al.*, 2005).¹ Die COX-2 3D Proteinstruktur ist seit 1996 bekannt (PDB Eintrag 6cox, Auflösung: 2,8 Å (Kurumbail *et al.*, 1996)).

Cyclooxygenasen sind die klassischen Zielproteine bei der Therapie von Arthritis und Rheuma. Eine Schmerzlinderung wird durch die Verabreichung von COX Inhibitoren erreicht, die eine entzündungshemmende Wirkung haben (Brooks & Day, 1991; Hawkey, 1999; Vane & Botting, 1998; Van Ryn & Pairet, 1999). COX-1 und COX-2 unterscheiden sich im aktiven Zentrum durch drei Aminosäuren, weshalb eine Großzahl an Inhibitoren an beiden Cyclooxygenasen gleichzeitig wirken. COX-2 besitzt je ein Valin an den Positionen 434 und 523, was aufgrund der geringeren Atomanzahl im

¹ Angabe der Aminosäuren im Drei-Buchstabencode mit der zugehörigen Sequenzposition.

Vergleich zu Isoleucin (bei COX-1) zu einer Vergrößerung des aktiven Zentrums führt (Bertolini *et al.*, 2001). An der Sequenzposition 513 ist zudem anstatt eines Histidins (bei COX-1) ein Arginin zu finden. In der Wirkstoffforschung führten diese Kenntnisse zur Entwicklung von selektiven COX-2 Inhibitoren (Wong *et al.*, 1997; Marnett & Kalgutkar, 1999), deren Nachfrage aufgrund der starken Nebenwirkungen im Magen-Darmtrakt von COX nicht-selektiven Inhibitoren ansteigt (Tanaka *et al.*, 2002).

Abbildung 3 zeigt die Wirkstoffe Naproxen ($IC_{50} = 18 \text{ nM}$), Diclofenac ($IC_{50} = 77 \text{ nM}$) und Indomethacin ($IC_{50} = 207 \text{ nM}$), welche nicht-selektive COX Inhibitoren sind (Rowlinson *et al.*, 2003). Zusätzlich werden die Strukturen Celecoxib ($IC_{50} = 0,5 \text{ nM}$) und Rofecoxib ($IC_{50} = 2 \text{ nM}$) als selektive COX-2 Inhibitoren gezeigt (McAdam *et al.*, 1999; Catella-Lawson *et al.*, 1999). Charakteristische Merkmale der letzten beiden Moleküle sind die Sulfonyl- bzw. Sulfonamidgruppen und die „Mickey Mouse“ Struktur, als welches das Molekülgerüst, bestehend aus einem fünfgliedrigen Heterozyklus mit flankierenden Benzolgruppen („Mickey Mouse Ohren“), bezeichnet wird.

PPAR α

Im zweiten Fallbeispiel wurde ein Pseudorezeptor für PPAR α konstruiert, welcher den intrazellulären, nukleären Rezeptoren angehört (Rau *et al.*, 2008). Mit den zwei weiteren Subtypen, PPAR δ und PPAR γ , spielen sie eine große Rolle in der Lipid- und Glukoseregulation. Aktivierte PPAR wandern aus dem Cytosol in den Zellkern, wo es zur Bildung eines Heterodimers mit dem Retinoid-X-Rezeptor (RXR) kommt (Desvergne & Wahli, 1999). Der Komplex bindet an eine spezifische DNA-Sequenz (*DNA response element*) und steigert die Gentranskription, die zu einer starken Peroxisomenproduktion führt (Steinhilber *et al.*, 2005). Peroxisomen sind die für den Transport und den Metabolismus von Fettsäuren zuständige Zellorganellen (Desvergne

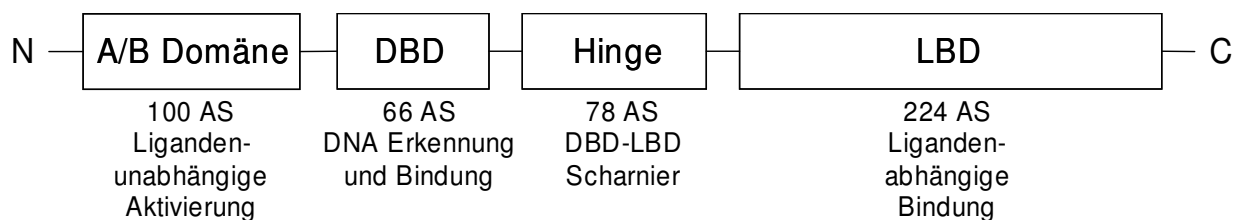


Abbildung 4. Funktionsübersicht der PPAR α Sequenz (AS=Aminosäuren).

& Wahli, 1999). Je nach Subtyp werden PPAR an unterschiedlichen Stellen im Körper exprimiert: PPAR α wird in der Leber, den Nieren und dem Herzen produziert. PPAR δ kommen, neben den Lokalisationen von PPAR α , auch im Darm vor. PPAR γ wird in Fettgeweben exprimiert und steigert dort die glykolytische Aktivität.

Die PPAR α Proteinsequenz lässt sich in vier Bereiche mit unterschiedlicher Funktion einteilen (Abbildung 4). Am C-terminalen Ende befindet sich die Ligandenbindedomäne (LBD), welche für die Erkennung und Bindung von PPAR Modulatoren zuständig ist. Sie macht den Hauptteil der Gesamtsequenz aus (224 von 468 Aminosäuren). Die LBD ist über eine Scharniereinheit mit der DNA-Bindedomäne (DBD), welche zur Identifikation des DNA Responselements fähig ist, verbunden. Die LBD und die DBD sind fundamentale Bereiche der PPAR, weil die gesamte

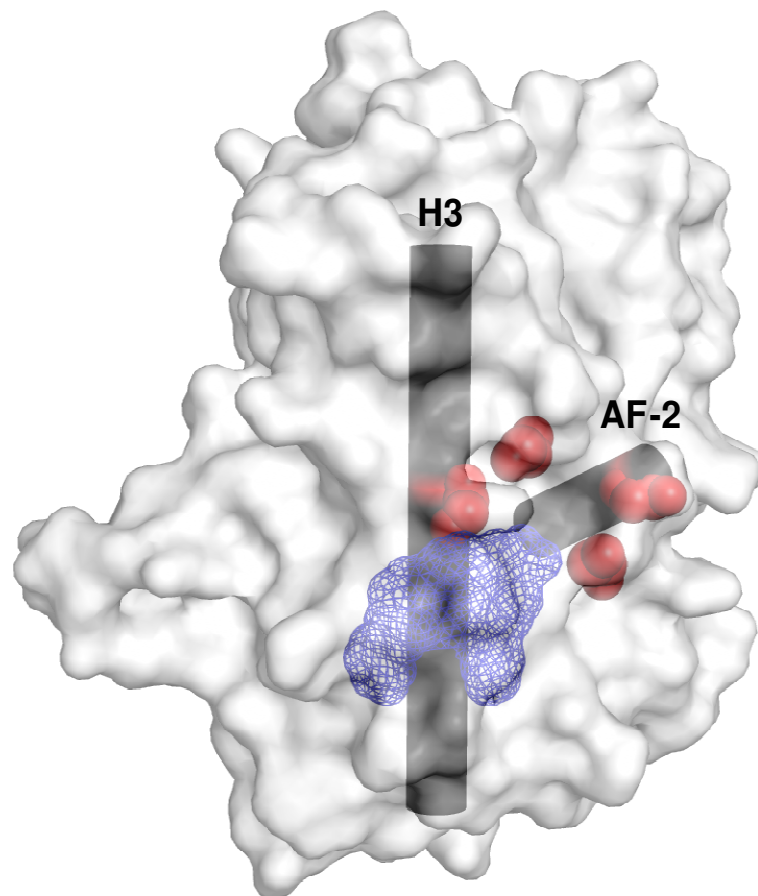
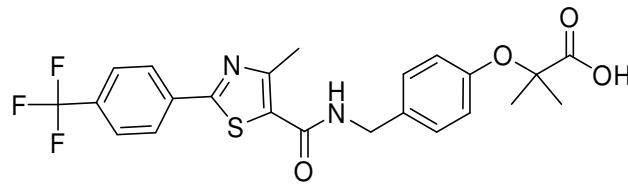
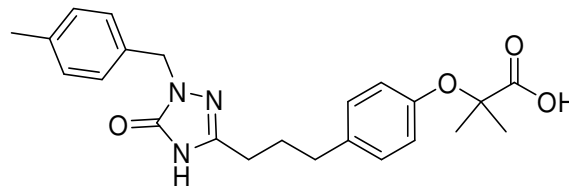


Abbildung 5. 3D Struktur von PPAR α (PDB Eintrag 1k7l). Gezeigt ist die Oberfläche des Proteins mit den wichtigen Helices H3 und AF-2 (schwarze Zylinder). Zusätzlich sind die U-förmige Bindetasche (blaues Gitter) und die Rückgrat-Atome der vier Schlüsselaminosäuren (rote Kugeln) gezeigt.



GW590735



LY518674

Abbildung 6. Die PPAR α Agonisten GW590735 (Sierra *et al.*, 2007) und LY 518674 (Singh *et al.*, 2005) mit dem typischen 2-Hydroxy-2-Methylpropionsäurerüst (nach rechts orientiert).

Proteinfunktion hier manifestiert ist. Am N-terminalen Ende sitzt eine zusätzliche proteinaktivierende Domäne, die von einer Ligandenbindung unabhängig ist.

Die dreidimensionale Proteinstruktur von PPAR α wurde im Jahre 2001 mit Hilfe der Röntgenkristallographie aufgeklärt (PDB Eintrag 1k7l (Auflösung: 2,5 Å, Xu *et al.*, 2001), Abbildung 5). Das zentrale Element der Struktur ist die Helix H3, an der potenzielle PPAR Modulatoren in einer U-förmigen Konformation binden. Dabei ist eine Bindung an den Schlüsselaminosäuren Ser280, Tyr314, His440 und Tyr464 für die Stabilisierung der AF-2 Helixkonformation und die damit einhergehende PPAR α Aktivierung verantwortlich (Sierra *et al.*, 2007). PPAR Bindetaschen besitzen durchschnittlich ein Volumen von 1300-1400 Å³ (Zoete *et al.*, 2007). Die unterschiedliche Bindetaschengröße und -oberflächenformen der drei PPAR Subtypen dienen als Selektivitätskriterien bei der Entwicklung von neuen Modulatoren (Markt *et al.*, 2007).

Aufgrund ihrer Funktionen besteht ein großes pharmazeutisches Interesse an PPAR Liganden. So dienen sie in erster Linie als Zielproteine für die Erforschung von therapeutischen Maßnahmen zur Behandlung von Diabetes mellitus Typ-2 und Fettstoffwechselstörungen (Gelman *et al.*, 2007; Duval *et al.*, 2007). Zudem wurde eine Rolle in Entzündungskrankheiten festgestellt (Széles *et al.*, 2007).

Abbildung 6 zeigt zwei selektive PPAR α Liganden, die der Klasse der Fibrate angehören. GW590735 ($EC_{50} = 4$ nM, Sierra *et al.*, 2007) befindet sich zur Zeit dieser Arbeit in der klinischen Entwicklungsphase II, welche LY518674 ($EC_{50} = 24,5$ nM, Singh *et al.*, 2005) vor kurzem nicht erfolgreich verließ (<http://www.medscape.com/viewarticle/554105>, 10.08.2008).

H4R

Im Jahr 2000 wurde der H4R entdeckt (Oda *et al.*, 2000; Nakamura *et al.*, 2000). Er gehört, wie die Histamin Rezeptoren H₁, H₂ und H₃, zur Klasse A der GPCR. Diese werden in sechs Klassen unterteilt, wobei nur die ersten drei im Menschen vorkommen (Horn *et al.*, 2003):

- Klasse A: rhodopsinähnliche Rezeptoren
- Klasse B: sekretinähnliche Rezeptoren
- Klasse C: metabotrope Glutamatrezeptoren
- Klasse D: Pheromonrezeptoren der Pilze
- Klasse E: cAMP Rezeptoren
- Klasse F: „Frizzled/Smoothened“ Rezeptoren

Die pharmazeutische Behandlung an den Histaminrezeptoren H₁, H₂ und H₃ ermöglicht die Steuerung von Schlaf-Wach Zuständen und allergischen Erkrankungen (bei H₁ (Simons, 2004)), Herzschlagsteuerung und Magensäureregulation (bei H₂ (Valle & Gantz, 1997)), sowie die Regulation der Neurotransmitterausschüttung im zentralen Nervensystem (bei H₃ (Leurs *et al.*, 2005)). Aufgrund der Expression des H4R in dem Knochenmark, der Milz, dem Darm und der Thymus wird eine regulatorische Rolle in der Entzündungs- und Immunantwort vermutet (Lim *et al.*, 2006; Akdis & Simmons, 2006). Er wird vor allem in Neutrophilen, Basophilen, Eosinophilen, Mastzellen und T-Zellen produziert (Gantner *et al.*, 2002; Hofstra *et al.*, 2003). Eine funktionale Beteiligung an der Chemotaxis von Mastzellen und Eosinophilen, sowie an der Freisetzung von Interleukin-16 von T-Zellen wurde bereits erwiesen (Gantner *et al.*, 2002; Hofstra *et al.*, 2003). Jüngste Arbeiten berichten zusätzlich von einem geweckten therapeutischen Interesse in der Abwehr von Brustkrebs (Medina *et al.*, 2008).

Histaminrezeptoren liegen in der Plasmamembran und besitzen sieben Transmembranhelices (Uniprot Eintrag Q9H3N8, Abbildung 7). Sie besitzen wie alle aminergen GPCR sowohl das DRY Motiv in der dritten und das NPxxY Motiv in der siebten Transmembranhelix, als auch eine Cysteinbrücke im Bereich der extrazellulären Matrix (Jacoby *et al.*, 2006).

Grundsätzlich liegt die Aufgabe des H4R in der Wechselwirkung mit dem Botenstoff Histamin. Es aktiviert daraufhin das G-Protein und führt die Signaltransduktion aus dem extrazellulären Raum in das innere der Zelle fort (De Esch *et al.*, 2005). Histamin kommt als in Vesikeln gespeicherter Botenstoff besonders in Granulozyten und Mastzellen vor – bei einer allergischen Reaktion wird es durch Exozytose freigesetzt und von den Histaminrezeptoren gebunden (Akdis & Blaser, 2003). Als Konsequenz findet eine schnelle Zellantwort in Form von Ionenkonzentrations- und Enzymaktivitätsveränderungen, oder eine langsame Zellantwort in Form von Regulationen des Genexpressionsapparats statt (Pierce *et al.*, 2002).

Die endogenen Liganden der Histamin- und Serotoninrezeptoren werden auch unter

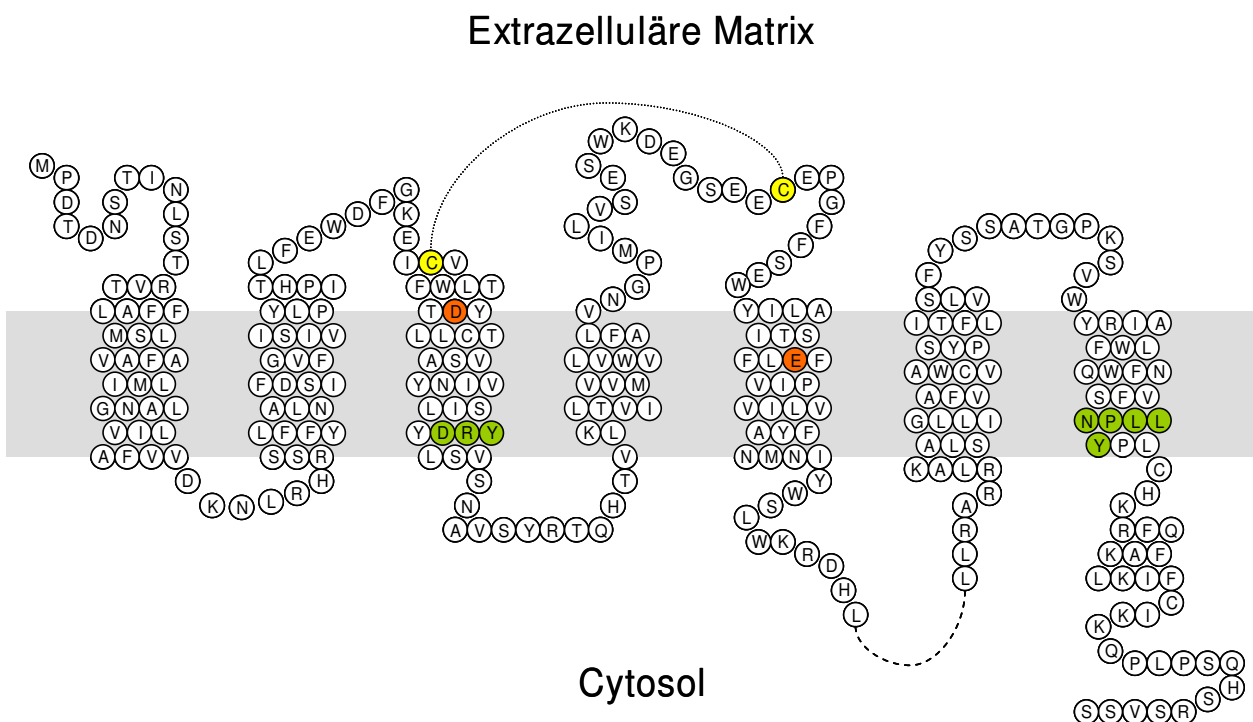


Abbildung 7. Sequenz des Histamin H₄ Rezeptors mit allen 390 Aminosäuren. Die sieben Transmembranomänen befinden sich in der Membran (graues Band). Farblich hervorgehoben sind die Cysteinbrücke in gelb, die Aminosäuren Asp98^{3,32} und Glu182^{5,46} in orange und die konservierten Motive DRY und NPxxY in grün. Der Übersicht halber wurden die Aminosäuren zwischen Lys207 und Lys295 nicht dargestellt (gestrichelte Linie).

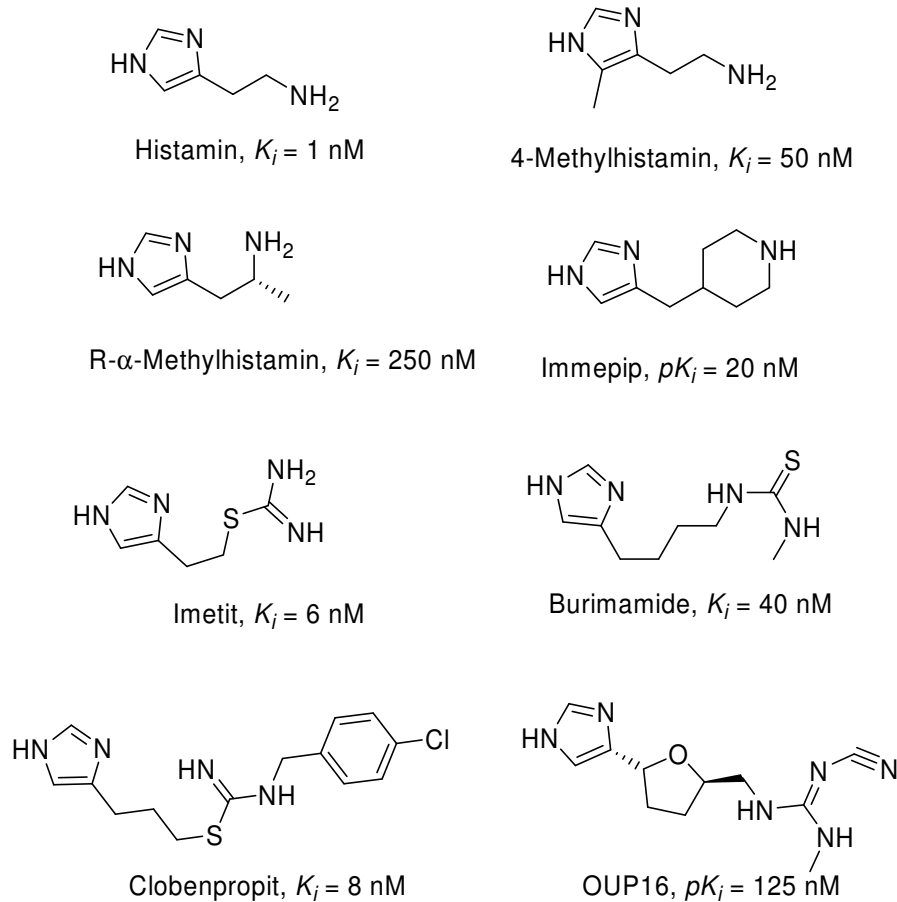
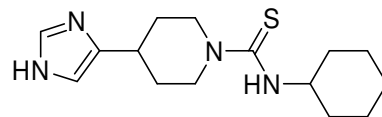


Abbildung 8. H4R Agonisten und ihre Bindeaffinitäten.



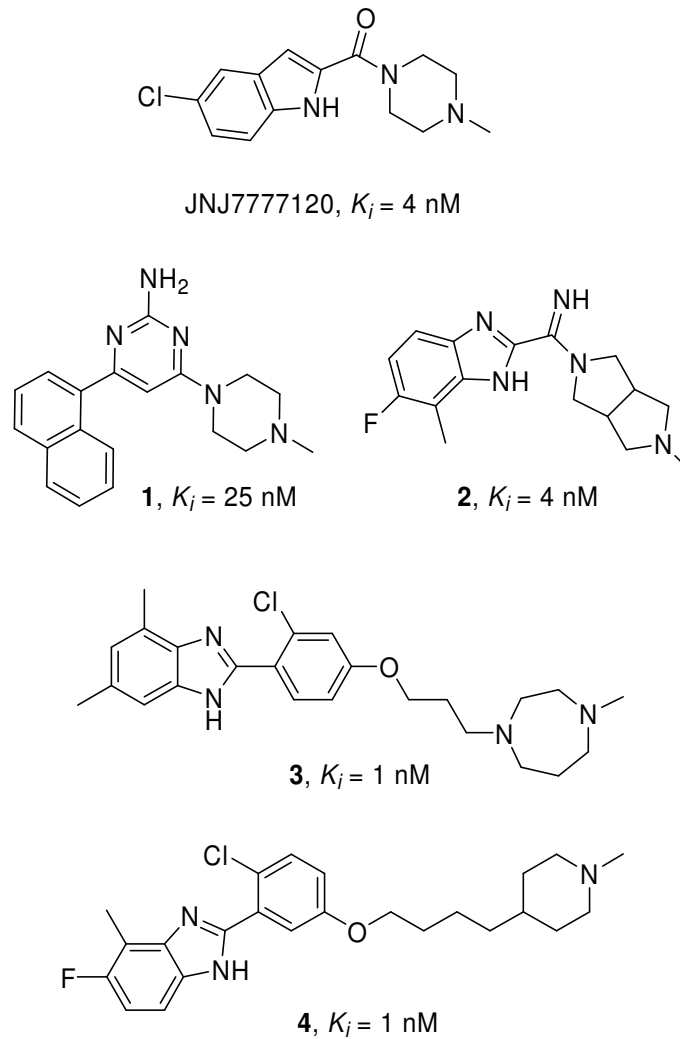
Thioperamid, $K_i = 125 \text{ nM}$

Abbildung 9. Ein inverser Agonist des H4R.

dem Oberbegriff „biogene Amine“ zusammengefasst - sie entstehen aus der Decarboxylierung von Tryptophan und Histidin. Durch diese Gemeinsamkeit ist eine Nebenwirkung der Liganden der Histaminrezeptoren auf Serotoninrezeptoren nicht ausgeschlossen (Steinhilber *et al.*, 2005).

Agonisten und Antagonisten des H4R

Die Funktionen des H4R beruhen auf einer Konformationsänderung durch die Bindung von Histamin, der als Agonist wirkt. Liganden werden aufgrund der implizierten Aktivität auf dem Zielprotein in Agonisten (auch „volle“ Agonisten genannt), partielle

Abbildung 10. Ausgewählte H_4 Antagonisten.

Agonisten, inverse Agonisten, kompetitive Antagonisten und nicht-kompetitive Antagonisten klassifiziert (Neubig *et al.*, 2000; Jacoby *et al.*, 2006).

Im Allgemeinen können sich GPCR auch ohne jeglichen Liganden in einem aktiven Zustand (konstitutive Aktivität oder „Basalaktivität“) befinden, bei der eine Signalprozessoren durchgeführt wird (Seifert & Wenzel-Seifert, 2002).

- Bei der Bindung eines **vollen Agonisten** wird die Leistungsfähigkeit auf das Maximum gesteigert.
- **Partielle Agonisten** bewirken auch eine Steigerung der Gesamtaktivität, jedoch steigt diese nicht bis zur maximalen Fähigkeit an – auch nicht bei einer Konzentrationserhöhung des Liganden.

- **Inverse Agonisten** steigern die Aktivität des Zielproteins nicht, sondern setzen diese unter die Basalaktivität herab.
- **Kompetitive Antagonisten** bewirken eine Blockierung der Agonistenbindung, indem sie an derselben Bindetasche binden und den Agonisten verdrängen.
- Anders ist die antagonistische Wirkung des **nicht-kompetitiven Antagonisten** zu erklären, der durch Bindung in einer anderen Bindetasche eine Konformationsänderung des Zielproteins bewirkt. Durch die implizite allosterische Veränderung der Agonistenbindetasche wird der aktuelle Agonist entlassen und eine weitere Bindung wird unmöglich.

Beispiele für H4R Agonisten, inverse Agonisten und Antagonisten sind der Abbildung 8, Abbildung 9 und Abbildung 10 zu entnehmen. Für die Wirkstoffforschung am Zielprotein H4R sind insbesondere Antagonisten von Interesse, weil sie den endogenen Liganden und die damit verknüpfte Grundfunktion des Rezeptors blockieren.

Obwohl sich kein Wirkstoffkandidat in der klinischen Phase befindet, wurde die antagonistische Wirkung des ersten selektiven Antagonisten JNJ7777120 an Tiermodellen bestätigt (Jablonowski *et al.*, 2003; Thurmond *et al.*, 2004). Die Moleküle **1** (WO 2005/014556), **2** (WO 2006/056848 A1), **3** und **4** (WO 2005/044807 A2) wurden der Patentliteratur entnommen.

Implementierung eines Pseudorezeptorkonzeptes

Im Rahmen dieser Arbeit wurde eine automatische Pseudorezeptorkonstruktion implementiert. Das Hauptaugenmerk wurde auf die Eignung als virtuelle Screeningmethode in der Wirkstoffforschung gelegt. Die Basis des Pseudorezeptors bildet ein einzelner Ligand oder ein Ligandenensemble (Referenzliganden). Die vorgestellte Methode projiziert anhand eines Gitters Pseudoatome um Substrukturen der Referenzliganden, die eine Wasserstoffstoffbrückeninteraktion eingehen können. Pseudoatome geben mögliche Positionen von Rezeptoratomen an. Die gesamte Population der Pseudoatome ergibt schließlich den Pseudorezeptor. Zur Positionsberechnung eines Pseudoatoms wurde auf eine Sammlung von Wasserstoffbrückengeometrien zurückgegriffen, welche aus einer Kristallstrukturuntersuchung (Mills & Dean, 1996) abgeleitet sind.

Abbildung 11 verschafft einen Überblick der Konstruktionsschritte:

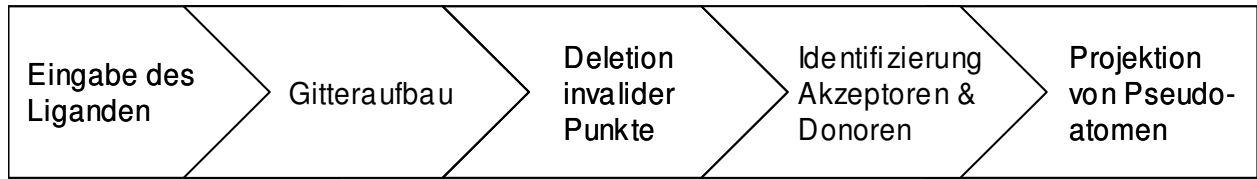


Abbildung 11. Überblick der einzelnen Schritte zum Aufbau eines Pseudorezeptors.

Schritt 1: Als Referenzmolekül muss mindestens eine Konformation eines bekannten Liganden des Zielproteins deklariert werden. Bei einer Anzahl von mehr als einem Liganden müssen diese strukturell übereinander gelagert sein.

Schritt 2: Um die Referenzliganden wird ein Gitter mit einer Maschenweite von $0,5 \text{ \AA}$ aufgebaut. Die Maschenweite wurde kleiner als die mittlere C-C Bindungslänge gewählt, um hinreichend genau Pseudoatome platzieren zu können (üblicherweise beträgt die Maschenweite bei 3D QSAR Ansätzen $1-2 \text{ \AA}$ (Schneider & Baringhaus, 2008)). Die Ausmaße des Gitters werden automatisch so eingerichtet, dass die Referenzliganden zentriert mit einem Abstand von 5 \AA zum Gitterrand positioniert wird. Jeder Gitterpunkt kann im späteren Verlauf zu einem Pseudoatom umgewandelt werden.

Schritt 3: In diesem Schritt werden alle Gitterpunkte gekennzeichnet, die zu nahe an Referenzliganden sind, weil sie aufgrund von van der Waals Potentialen nicht als Pseudoatom in Frage kommen. Hierzu werden die Distanzen aller Gitterpunkte zu allen Schweratomen der Referenzliganden berechnet. Falls ein Gitterpunkt eine Distanz kleiner als den van der Waals-Radius eines Molekülatoms hat, wird dieser für „ungültig“ erklärt. Tabelle 2 gibt einen Überblick der verwendeten Radien.

Schritt 4: Alle Atome der Referenzliganden, welche eine Wasserstoffbrückeninteraktion eingehen können (Wasserstoffbrückenakzeptoren und -donoren), werden identifiziert.

Schritt 5: Ausgehend von den Atompositionen der Akzeptoren und Donoren werden nun diejenigen Gitterpunkte als Pseudoatome deklariert, welche in einer passenden Wasserstoffbrückengeometrie stehen und somit mögliche Positionen von Akzeptoren und Donoren auf der Rezeptorseite simulieren.

Schritt 6: Jedem Pseudoatom wird ein Pharmakophortyp zugewiesen.

Tabelle 2. Überblick der verwendeten van der Waals Radien zur Kennzeichnung ungültiger Gitterpunkte (Schneider & Baringhaus, 2008).







Elementsymbol	van der Waals-Radius (Å)
C	1,85
O	1,4
N	1,54
P	1,9
S	1,85
F	1,35
Cl	1,18

Als Resultat erhält man eine Verteilung von Pseudoatomen um Referenzliganden, die die Positionen von möglichen Rezeptoratomen repräsentieren. Die Gesamtheit dieser Punkte wird als Pseudorezeptor bezeichnet, weil es alle Bestandteile für Wasserstoffbrückeninteraktion zum Liganden beinhaltet.

Algorithmen zur Modellierung von Wasserstoffbrücken

Zur Implementierung von Wasserstoffbrückengeometrien wurde ein geometrisches Modell entwickelt. Ausgehend vom Liganden und nach vorliegender Substruktur lassen sich Wasserstoffbrücken in vier Geometrien einteilen (Tabelle 3).

Tabelle 3. Übersicht der Wasserstoffbrückengeometrien und angewendeten Algorithmen.

Geometrie	Substrukturbeispiel	Bezeichnung	Algorithmus
—	—OH	Linear	Linear
		Planar mit 4 Atomen	Planar
		Planar mit 3 Atomen	Planar
		Tetrahedral	Linear

Eine lineare Geometrie findet sich beispielsweise bei Hydroxylgruppen, bei welchen das Sauerstoffatom die Wasserstoffbrücke eingeht, aber an ein Kohlenstoffatom verankert ist. Für diesen Fall wird der *lineare* (nicht im Sinne der Algorithmenlaufzeit, sondern der Wasserstoffbrückengeometrie) Algorithmus zur Bestimmung der möglichen Positionen des gegenüberliegenden Rezeptoratoms angewendet (Abbildung 12a). Ausgehend vom Ankeratom A und der Geraden durch Atom B, welches an der Wasserstoffbrücke beteiligt ist, wird eine gültige Region für Pseudoatome definiert. Hierzu werden Minima und Maxima für die Variablen r (Distanz) und φ (Winkel) festgesetzt. Befinden sich Gitterpunkte innerhalb dieser Region, werden sie zu einem Pseudoatom mit komplementärer Wasserstoffbrückeneigenschaft ernannt.

Andere Substrukturen werden mit dem *planaren* Algorithmus modelliert, wobei zwei unterschiedliche Fälle auftreten können. Zum einen kann eine planare Substruktur aus vier Atomen bestehen. Dies trifft insbesondere im Falle der Carbonylgruppe zu, wo ein doppelt gebundener Sauerstoff auf der Ebene liegt, welche von seinem Ankeratom und dessen Nachbaratomen aufgespannt wird (sp² Hybridisierung). Die Region für gültige Wasserstoffbrückenpartner liegt ebenfalls an der Ebene und wird daher mit einer zusätzlichen Winkelvariable ϕ festgesetzt (Abbildung 12b). Zum anderen bestehen

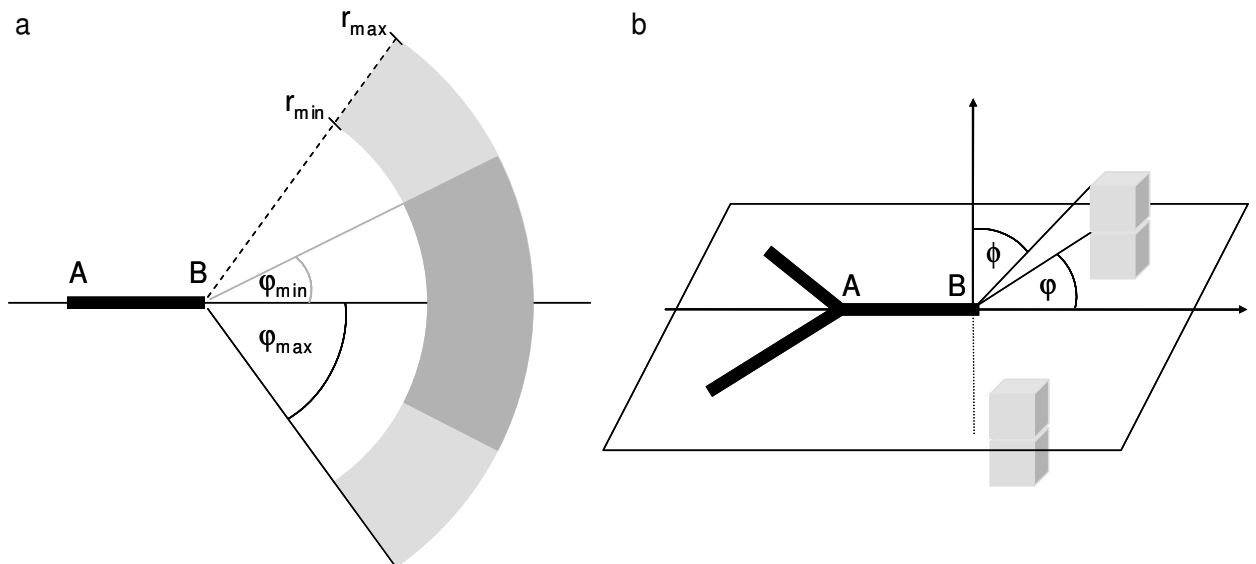


Abbildung 12. a) Aufbau des linearen Algorithmus. Alle Gitterpunkte innerhalb der hellgrauen Region um Atom B haben eine gültige Wasserstoffbrückengeometrie. b) Aufbau des planaren Algorithmus. Gitterpunkte in den beiden grauen Zonen werden zu Pseudoatomen ernannt.

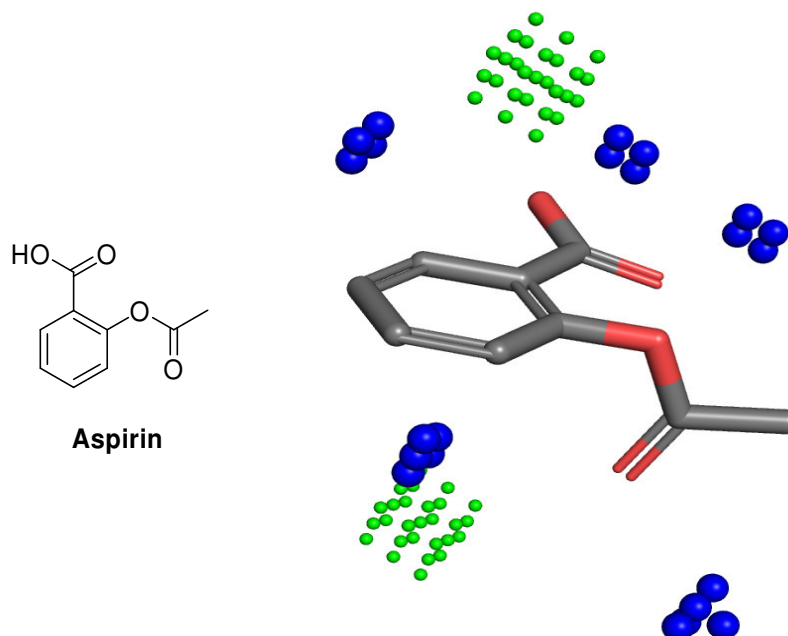


Abbildung 13. Pseudorezeptor um Aspirin. Blaue Kugeln stellen Donorpseudoatome, und grüne Kugeln aromatische Pseudoatome dar.

planare Gruppen aus nur drei Atomen, wie es zum Beispiel in der Ethergruppe der Fall ist. Hier wird ebenfalls auf den planaren Algorithmus zurückgegriffen, wobei ein virtuelles Ankeratom A zwischen die beiden realen Ankeratome projiziert wird.

Der letzte Fall tritt auf, wenn das an der Wasserstoffbrücke beteiligte Atom drei Ankeratome besitzt (*tetrahedrale* Geometrie, sp^3 Hybridisierung). Hier wird analog zum ersten Fall der lineare Algorithmus angewandt, nachdem ein virtuelles Ankeratom in der Mitte der drei realen Ankeratome projiziert wurde, um eine Gerade mit dem interagierenden Atom zu bilden.


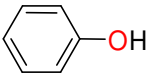




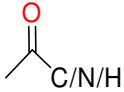




Substrukturen und Pharmakophoreigenschaften


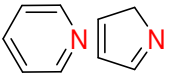
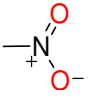
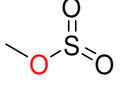
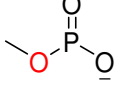
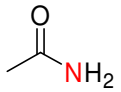
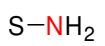
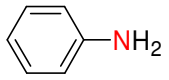
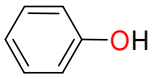




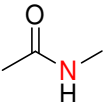
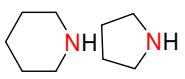

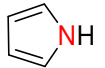
Um Festzustellen, welche Ligandenatome zur Bildung von Wasserstoffbrücken geeignet sind, wird eine Substruktursuche mit Hilfe der Beschreibungssprache MQL (Proschak *et al.*, 2007) durchgeführt. Tabelle 4 zeigt alle hier verwendeten Substrukturen, ihre Pharmakophoreigenschaft (Akzeptor oder Donor), die Schwellenwerte für die Variablen (r , φ , ϕ) und den verwendeten Algorithmus an. Zusätzlich wurde eine Definition von gerichteten aromatischen Interaktionen eingeführt, welche die Erkennung von Benzolringen ermöglicht. Pseudoatome werden in diesem Fall in einer Distanz von

3,5-4,5 Å parallel zur Ebene des Ringes generiert (McGaughey *et al.*, 1998). Es werden keine *edge-to-face* und Kation- π Wechselwirkungen modelliert, ebenso werden weit reichende aromatische Interaktion (6-11 Å) in dieser Studie ignoriert (McGaughey *et al.*, 1998).

Abbildung 13 zeigt am Beispiel des Aspirins, wie eine resultierende Verteilung von Pseudoatomen um erkannte Substrukturen aussieht. Der Pseudorezeptor umgibt das Molekül. Die verschieden farbigen Pseudoatome sind vorhergesagte Positionen von Rezeptoratomen mit der jeweiligen Pharmakophoreigenschaft. Man beachte die Vorhersage für aromatische Pseudoatome parallel zur Ebene des aromatischen Systems und den Brückensauerstoff des Esters, welcher keiner Substruktur entspricht und somit zur Projektion von Pseudoatomen nicht herangezogen wird.

Tabelle 4. Zur Erkennung von Substrukturen verwendete Parameter, unter Angabe des angewendeten Algorithmus und den Distanz- bzw. Winkelparametern. Das rote Atom zeigt jeweils das zur Pseudoatomprojektion herangezogene Ankeratom der erkannten Substruktur mit der in der Spalte „Pharmakophor“ annotierten Pharmakophoreigenschaft. Projizierten Pseudoatomen wird die komplementäre Pharmakophoreigenschaft zugeteilt (MQL Notation in Anhang I).

Substruktur	Algorithmus	Pharmakophor	Distanz (Å)	φ	ϕ
	Linear	Akzeptor	2,6-3,3	141°-180°	-
	Linear	Akzeptor	2,6-3,17	100°-153°	-
	Linear	Akzeptor	2,6-3,3	110°-130°	-
	Linear	Akzeptor	2,6-3,03	100°-161°	-
	Linear	Akzeptor	2,6-3,03	110°-130°	-
	Linear	Akzeptor	2,6-3,03	109°-180°	-
	Planar	Akzeptor	2,6-3,3	110°-130°	80°-100°
	Planar	Akzeptor	2,6-3,17	110°-130°	80°-100°
	Planar	Akzeptor	2,6-3,17	110°-130°	80°-100°
	Planar	Akzeptor	2,6-3,42	140°-160°	80°-100°
	Planar	Akzeptor	2,6-3,42	150°-170°	80°-100°

	Planar	Akzeptor	2,6-3,73	90°-110°	80°-100°
	Planar	Akzeptor	2,6-3,24	145°-180°	80°-100°
	Planar	Akzeptor	2,6-3,48	100°-120°	80°-100°
	Planar	Akzeptor	2,6-3,17	103°-180°	80°-100°
	Planar	Akzeptor	2,6-3,03	103°-146°	80°-100°
	Linear	Donor	2,6-3,3	109°-141°	-
	Linear	Donor	2,8-3,2	140°-180°	-
	Linear	Donor	2,6-3,3	103°-146°	-
	Linear	Donor	2,6-3,03	103°-136°	-
	Linear	Donor	2,6-3,23	140°-180°	-
	Linear	Donor	2,6-3,42	90°-124°	-
	Linear	Donor	2,6-3,1	90°-141°	-
	Linear	Donor	2,6-2,92	140°-180°	-
	Planar	Donor	2,6-3,23	140°-180°	80°-180°
	Planar	Donor	2,6-3,09	140°-180°	75°-105°
	Planar	Donor	2,6-3,09	150°-180°	80°-100°
	Planar	Donor	2,6-3,09	140°-180°	70°-110°

Gewichtete Pseudorezeptoren bei multiplen Referenzliganden

Ist die Basis eines Pseudorezeptors nicht ein einzelnes Molekül, sondern eine Ensemble aus übereinander gelagerten Referenzmolekülen, findet eine Gewichtung der generierten Pseudoatome nach Gleichung (1) statt und resultiert in einem gewichteten Pseudorezeptor.

$$w_{PA} = \frac{u}{v} \quad (1)$$

Dementsprechend ist das Gewicht eines Pseudoatoms PA der Quotient aus der Anzahl der Moleküle, die diesen Punkt bilden, u , und der Gesamtzahl der Moleküle in der Überlagerung v . Somit haben Pseudoatome, welche von allen Molekülen gebildet werden, immer ein Gewicht von 1 und Pseudoatome, welche nur von einem Molekül aus projiziert sind, ein Gewicht von $1/v$. Abbildung 14 zeigt ein Beispiel eines Pseudorezeptors, welches um eine Überlagerung aus Flurbiprofen und Ibuprofen gebildet wurde. Da zwei Referenzmoleküle Ausgangspunkt der Pseudorezeptorgenerierung sind, können Pseudoatome mit einer Gewichtung von 0,5 und 1,0 auftreten. Man achte auf die aromatischen Pseudoatome: die grünen Punkte auf der rechten Seite resultieren aus dem zweiten aromatischen System des Flurbiprofens, welches nicht im Ibuprofen vorkommt. Diese Punkte haben ein Gewicht von 0,5. Dahingegen haben die grünen Punkte in der Mitte ein Gewicht von 1,0, weil sie aus dem gemeinsamen aromatischen System hervorkommen.

Kodierung von Pseudorezeptormodellen als Autokorrelationsvektoren

Zur Durchführung einer Ähnlichkeitssuche wird ein Pseudorezeptor in einen 90-dimensionalen Autokorrelationsvektor übersetzt. Die Nutzung von korrelationsbasierten Deskriptoren ist einer der ältesten Ansätze in der Chemieinformatik zur mathematischen

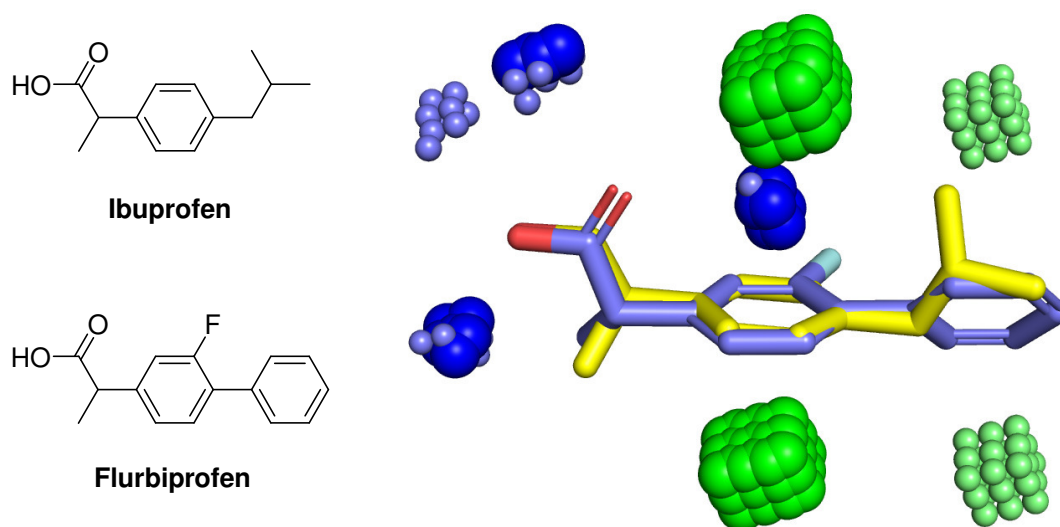


Abbildung 14. Pseudorezeptor aus einer Überlagerung von Ibuprofen und Flurbiprofen. Pseudoatome mit größerem Radius werden von beiden Molekülen generiert.

Beschreibung von Molekülen (Moreau & Broto, 1980).

Prinzipiell kodiert der aus einem Pseudorezeptor gewonnene Autokorrelationsvektor die Häufigkeiten der Pseudoatomdistanzen verschiedenen Typs, welche in einem Pseudorezeptor vorkommen (Abbildung 15). Da die in dieser Arbeit implementierte Pseudorezeptorgenerierung drei Typen von Pseudoatomen unterscheidet, nämlich Akzeptoren, Donoren und aromatische Pseudoatome, können Distanzen folgender Pseudoatompaare observiert werden:

1. Akzeptor zu Akzeptor (AA)
2. Akzeptor zu Donor (AD)
3. Akzeptor zu Aromatisch (AR)
4. Donor zu Donor (DD)
5. Donor zu Aromatisch (DR)
6. Aromatisch zu Aromatisch (RR)

Es werden nur Distanzen von 0-15Å notiert, welche in 1 Angström große Abschnitte eingeteilt sind. Es werden die Distanzen aller Pseudoatome zueinander berechnet.

Schließlich entsteht ein Deskriptor mit einer Länge von 90 Dimensionen (6×15). Formal wird die Häufigkeit einer Distanz d zwischen Pseudoatomen der Typen X und Y mit Gleichung (2) berechnet.

$$Freq_d(X, Y) = \sum_i^X \sum_j^Y \delta_d^{i,j}, \quad (2)$$

wobei δ das Kroneckerdelta bezeichnet und genau dann eins ergibt, wenn die jeweilige Distanz d zwischen den Pseudoatomen i und j existiert. Durch die Partitionierung des

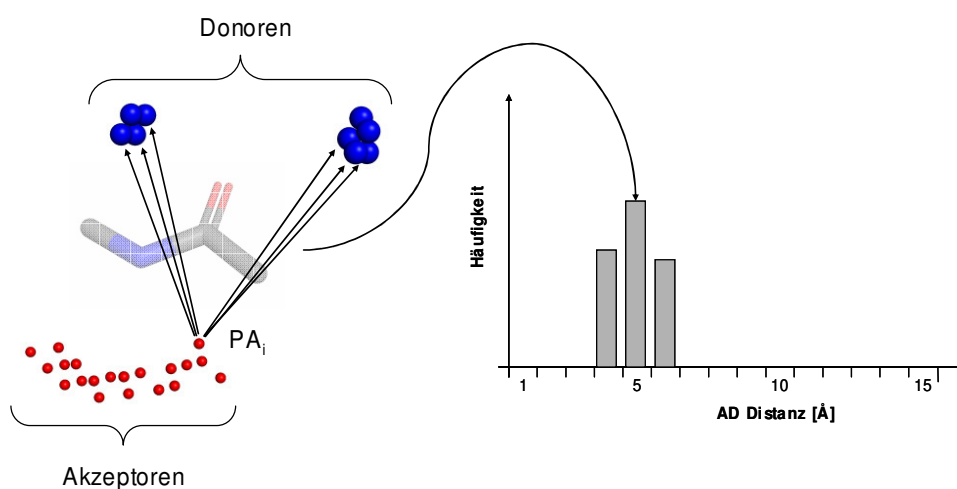


Abbildung 15. Kodierungsprinzip des Pseudorezeptordescriptors.

Vektors in je 15 Distanzbereiche pro Pseudoatompaar wird eine Abstraktion vorgenommen – der Distanzbereich a enthält immer die Häufigkeit der auftretenden Distanzen $(a-1), 1$ bis $a, 0$ (Beispiel $a=3$, Distanzen von 2,1 bis 3,0Å).

Wurde ein Pseudorezeptor aus einer Molekülüberlagerung generiert und somit eine Gewichtung der einzelnen Pseudoatome durchgeführt, werden gewichtete Häufigkeiten durch Gleichung (3) angegeben.

$$Freq_d(X, Y) = \sum_i^X \sum_j^Y \delta_d^{i,j} w_i w_j, \quad (3)$$

wobei w_i und w_j die Gewichte der jeweiligen Pseudoatome i und j sind (Gleichung (1)).

Virtuelles Screening mit Pseudorezeptoren

In einem virtuellen Screening mittels eines Pseudorezeptors wird eine Moleküldatenbank so sortiert, dass die auf den Pseudorezeptor passenden Moleküle auf besseren Rängen (an vorderen Positionen) zu finden sind. Hierzu ist es erforderlich, dass die Moleküle in der Datenbank ebenfalls in den 90-dimensionalen Autokorrelationsvektor übersetzt werden. Die Umsortierung der Datenbank erfolgt nach Berechnung der paarweisen Euklidischen Distanz (4) des Pseudorezeptorvektor p und eines der Moleküle x in der Datenbank.

$$d(p, x) = \sqrt{\sum_i^{90} (p_i - x_i)^2}, \quad (4)$$

wobei p_i und x_i der i -te Eintrag im Deskriptorvektor sind. Umso kleiner die Distanz, desto mehr ähnelt der Pseudorezeptor der Screeningsubstanz dem gesuchten Pseudorezeptor.

In einem retrospektiven Screening sind die biologische Aktivität und die Molekülstruktur der Wirkstoffe einer Datenbank bekannt. Eine Analyse der sortierten Liste kann aufgrund der Aktivitätsinformation der Screeningsubstanzen zur Berechnung der Screeningperformanz herangezogen werden.

Im prospektiven Screening werden neue aktive Moleküle gesucht. Auch hierbei wird die paarweise Ähnlichkeit zum Pseudorezeptordescriptor ermittelt und daraus eine Rangliste aufgebaut. Üblicherweise wird nach einer manuellen Selektion innerhalb der ersten Perzentile (*cherry-picking*) eine Validierung im experimentellen Versuch (*assay*) durchgeführt.

Verwendete Methoden und Programme

„Waschen“ von Molekülen

Die Moleküle der Screeningdatenbank wurden mit dem Programm MOE V2007.09 (Chemical Computing Group, Montreal, Kanada, www.chemcomp.com) so vorbereitet, dass azide Gruppen deprotoniert und basische Gruppen protoniert vorliegen. Diesen Vorgang bezeichnet man als „Waschen“.

Konformergenerierung

Für jedes Molekül der Screeningdatenbank wurde mit Hilfe des Programms Corina V3.2 (Molecular Networks GmbH, Erlangen, www.molecular-networks.com) 3D Atomkoordinaten generiert und Wasserstoffe hinzugefügt.

Handelte es sich um eine Konformergenerierung von Referenzmolekülen für die Konstruktion eines Pseudorezeptors, wurde dies mittels einer stochastischen Konformersuche in MOE V2007.09 durchgeführt. Alle Konformationen mit einer Energieabweichung von 7 kcal zur energieniedrigsten Konformation wurden in einer Datenbank gespeichert.

Pharmakophoralignment von Molekülen

Molekülüberlagerungen wurden mit dem flexiblen Pharmakophoralignment-Verfahren in MOE V2007.09 durchgeführt. Die energieniedrigste Konformation des potentesten Referenzmoleküls diente als Templat. Alle weiteren Referenzmoleküle wurden auf das fixierte Templat überlagert.

Molekulare Dynamiksimulation des H4R

Die molekulare Dynamiksimulation des H4R basierend auf einem Homologiemodell mit dem $\beta 2$ adrenergen Rezeptortemplat (Rasmussen *et al.*, 2007) wurde von Tim Werner (Arbeitsgruppe Prof. Schneider, Goethe-Universität Frankfurt a.M.) durchgeführt. Die Sequenzähnlichkeit vom H4R und dem $\beta 2$ Adrenorezeptor in den Transmembrandmänen ist mit 28% höher als die 18% H4R-Rhodopsin Identität (Palczewski *et al.*, 2000).

Das Homologiemodell wurde mit dem Programm Modeller V9.3 erstellt (Šali & Blundell, 1993; Eswar *et al.*, 2000). Eine Überprüfung erfolgte mit Procheck und

WhatIf (Laskowski *et al.*, 1993; Vriend & Sander, 1993). Der durchschnittliche WhatIf Qualitätswert betrug -0,536. Nur vier Aminosäuren (in der intrazellulären Schleife) unterschritten einen Wert von -5.

Das H4R Homologiemodell wurde in ein Palmitoyoleoylphosphatidylcholin (POPC) Doppellipidschicht eingefügt, sodass die achte α -Helix parallel zur Membranoberfläche platziert war. Ein TIP3P Wassermolekülmodell wurde um den Protein-Membran Komplex gelegt, welches zu einem System mit 10.558 Wassermolekülen und 183 Lipiden führte. Die Simulation beinhaltete eine Gesamtzahl von 60.976 Atomen und wurde mit NAMD V2.6 in einem CHARMM27 Kraftfeld durchgeführt (Brooks *et al.*, 1983; Phillips *et al.*, 2005). Die Simulationsschrittweite betrug 12 Femtosekunden. Die elektrostatischen Interaktionen wurden mit der *Particle-Mesh-Ewald* Methode modelliert (Darden *et al.*, 1993). Der Gesamtdruck wurde auf 1,01325 bar, und die Systemtemperatur auf 310 K gesetzt.

Die molekulare Dynamiksimulation verlief in den drei Phasen: Energieminimierung, Erwärmung und Equilibrierung. Die Simulationsdauer im letzten Schritt betrug 1 ns, um einen Verlust der inaktiven Konformation durch eine zu lange Simulation zu vermeiden.

Bewerten von Proteinbindetaschen mittels geometrischem Hashing

Auf Datenbasis einer molekularen Dynamiksimulation (MD) wurden die Aminosäurekoordinaten der H4R Bindetasche (nach (Jongejan *et al.*, 2008)) extrahiert, welches für jede Standaufnahme der MD eine Punkteverteilung liefert. Mit Hilfe von geometrischem Hashing wurde daraufhin nach einer ähnlichen Punkteverteilung wie in dem gesuchten Pseudorezeptor gesucht (Norel *et al.*, 1994; Wolfson & Rigoutsos, 1997). Die Methode wurde von Tim Geppert (Arbeitsgruppe Prof. Schneider, Goethe-Universität Frankfurt a.M.) implementiert. Die Funktionsweise des geometrischen Hashing ist in Abbildung 16 an einem 2D Beispiel erklärt:

- a) Prinzipiell wird innerhalb von zwei Punkteverteilungen (einer Referenzverteilung, z.B. einem Pseudorezeptor, und einer Kandidatverteilung, z.B. extrahierte Aminosäurekoordinaten) nach Übereinstimmungen gesucht. Hierzu ist die Konstruktion einer Hashtabelle erforderlich, die zu einem gegebenen Hasheintrag immer einen bestimmten Wert speichert.

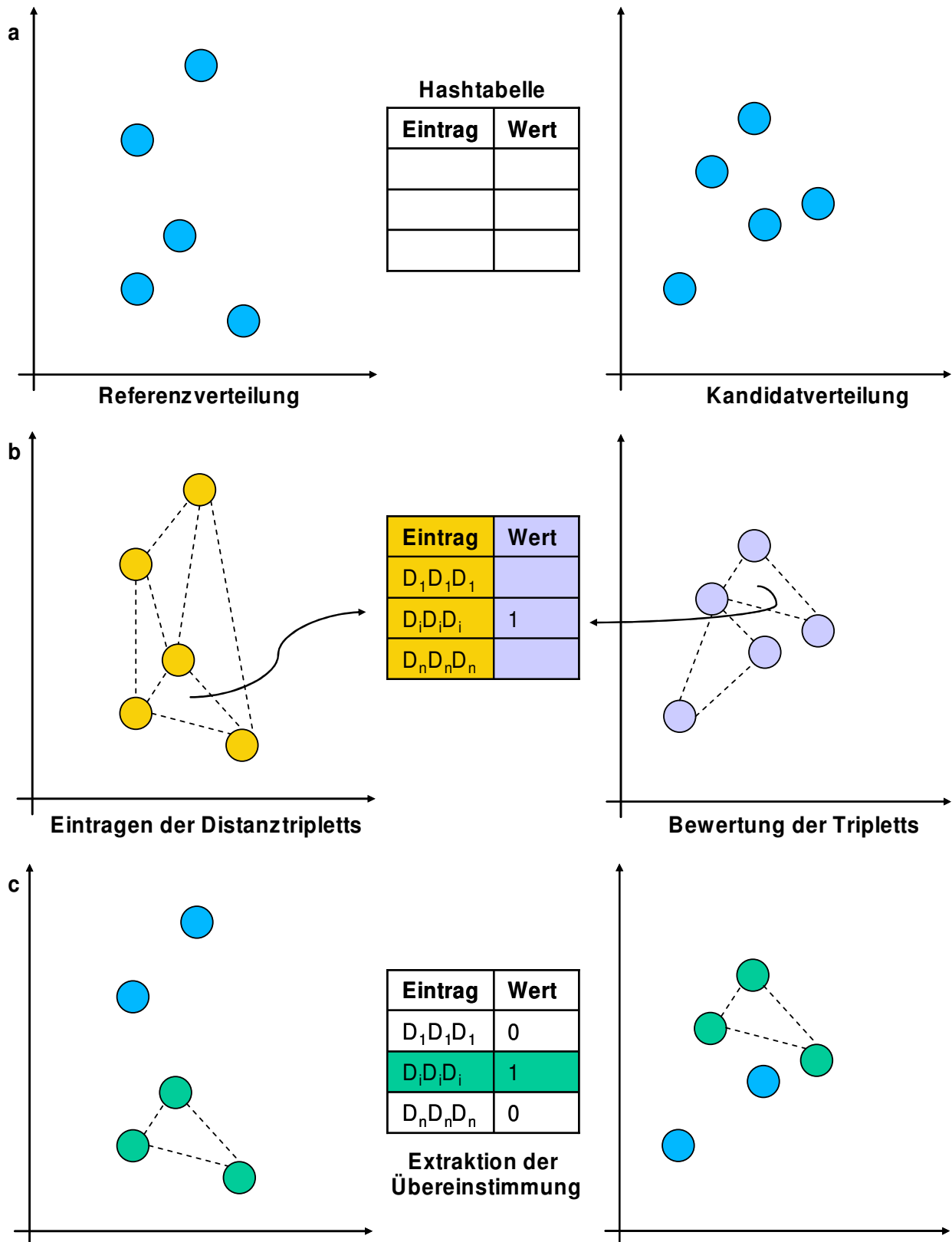


Abbildung 16. Funktionsprinzip des geometrischen Hashing.

- b) Aus der Referenzverteilung werden alle Triplets von Punkten mit den drei jeweiligen Distanzwerten und ihren Pharmakophortypen zur Generierung der Hasheinträge genutzt. Danach werden Triplets aus der Kandidatverteilung generiert. Falls ein Triplet mit übereinstimmenden Distanzen auftaucht, wird der Wert des jeweiligen Eintrags in der Hashtabelle um eins inkrementiert. Alle nicht übereinstimmenden Triplets (respektive Einträge) erhalten keine Bewertung.
- c) Aus der finalen Hashtabelle lässt sich die Übereinstimmung beider Punkteverteilungen extrahieren, indem man die Hashtabelle nach markierten Einträgen durchsucht. Dies geschieht in linearer Zeit, weil die Suchzeit nur von der Größe der Hashtabelle abhängt. Bewertete Distanztriplets sind in beiden Verteilungen präsent, wobei unbewertete nur in der Referenzverteilung vorkommen.

Bewertung der Güte eines retrospektiven virtuellen Screenings

Die Fähigkeit eines Deskriptors zum virtuellen Screening wird mit Hilfe von Anreicherungsuntersuchungen in der Rangliste aus einem retrospektiven Experiment ermittelt. Prinzipiell wird angegeben, wie viele korrekt Positive sich in den vorderen Rangpositionen befinden. Es gibt verschiedene Methoden zur Anreicherungsuntersuchung, wie z.B. die Berechnung des *enrichment factor* (Xu & Agrafiotis, 2002), die Berechnung des *BEDROC*-Wertes (Truchon & Bayly, 2007) oder *receiver operator characteristic* (ROC) Kurven (Fawcett, 2006). In dieser Arbeit wurde letztere Methode verwendet, um die Anreicherung in den vorderen Perzentilen der Screeningdatenbank zu veranschaulichen. Die Methode zur Darstellung der ROC Kurven entstammt der Elektrotechnik und wurde im Jahre 1982 erstmals in der medizinischen Radiologie publik (Hanley & McNeil, 1982). In der Chemieinformatik wurde es in der Qualitätsbewertung von maschinellen Lernalgorithmen eingesetzt und findet auch eine breite Anwendung im retrospektiven virtuellen Screening (Bradeley, 1997; Givehchi & Schneider, 2005);

Eine ROC Kurve ist ein Graph, der die Effizienz eines virtuellen Screenings in Abhängigkeit zu der Fehlerrate angibt. In einem retrospektiven virtuellen Screening, in dem die Aktivität der Moleküle bekannt ist, wird mit Hilfe einer Screeningmethode für

jedes Molekül entschieden, ob es zu der Klasse der Aktiven gehört oder nicht. Hierbei können folgende Zugehörigkeiten auftreten:

- *True-positive* (TP): Ein aktives Molekül wurde zu den Aktiven klassifiziert.
- *False-positive* (FP): Ein inaktives Molekül wurde für aktiv befunden.
- *True-negative* (TN): Ein inaktives Molekül wurde zu den Inaktiven klassifiziert.
- *False-negative* (FN): Ein aktives Molekül wurde für inaktiv befunden.

Die Ordinate der ROC Kurve gibt die Effizienz des Screenings an, indem der Anteil korrekt vorhergesagten Aktiven (üblicherweise als Sensitivität bezeichnet) berechnet wird (Gleichung (5)).

$$\text{Sensitivität} = \frac{TP}{TP + FN} \quad (5)$$

Die Fehlerrate wird auf der Abszisse durch die Berechnung der Spezifität (üblicherweise 1-Spezifität) angegeben (Gleichung (6)).

$$\text{Spezifität} = \frac{TN}{TN + FP} \quad (6)$$

Anhand von ROC Kurven kann erkannt werden, welcher Anteil an korrekt vorhergesagten Aktiven in der Rangliste eines retrospektiven Experiments relativ zum beobachteten Anteil der falsch Vorhergesagten gefunden wurde.

Eine effiziente Screeningmethode ist durch eine ROC Kurve mit großer Fläche (unter der Kurve) erkennbar. Je weiter die Kurve in den oberen linken Bereich des Graphen ragt, desto besser die Performanz.

Programmierungsumgebung

Als Programmiersprache wurde Java (www.sun.com) in der Version 5 benutzt. Zusätzlich wurde die frei verfügbare Bibliothek *Chemistry Development Kit* (CDK) (www.cubic.uni-koeln.de/cdk) in der Version vom 24.04.2008 verwendet (Steinbeck *et al.*, 2003, Steinbeck *et al.*, 2006).

Visualisierungen

Alle 3D Visualisierungen dieser Arbeit erfolgten mit PyMOL (<http://pymol.sourceforge.net/>) in der Version 1.1. Chemische Strukturen wurden mit ChemDraw 7.0 (CambridgeSoft, Cambridge, England, www.cambridgesoft.com) gezeichnet.

Moleküldatenbanken

In dieser Arbeit wurden retrospektive virtuelle Screenings mit Hilfe der Wirkstoffdatenbank COBRA in der Version 6.1 durchgeführt (Schneider & Schneider, 2003). Diese Datenbank enthält aus wissenschaftlichen Publikationen 8.311 annotierte Wirkstoffe unter Angabe ihrer Zielproteine.

Prospektive virtuelle Screenings wurden mit einer zusammengesetzten Datenbank (im Folgenden „Screeningdatenbank“) aus den Substanzen der Anbieter Asinex (Moskau, Russland, www.asinex.com) und Specs (Delft, Niederlande, www.specs.net) durchgeführt, welche käuflich erwerbbar Moleküle beinhalten. Details sind der Tabelle 5 zu entnehmen.

Tabelle 5. Übersicht über die Screeningdatenbank.

Name	Version	Anzahl Moleküle
Asinex Gold	November 2007	233.554
Asinex Platinum	November 2007	126.616
Specs	April 2008	202.681
Gesamt		562.851

Ergebnisse und Diskussion

In Rahmen dieser Arbeit wurde die Eignung von Pseudorezeptoren zum virtuellen Screening analysiert. Hierzu wurde eine Methode zur automatisierten Generierung von Pseudorezeptoren entwickelt. Zur Konstruktion ist die Verfügbarkeit mindestens eines Liganden erforderlich. Bei einer Menge von Liganden müssen diese überlagert vorliegen.

Zur Überprüfung des Konstruktionsprinzips wird an Fallbeispielen die Kompatibilität je eines COX-2 und PPAR α Pseudorezeptors demonstriert. In beiden Fällen lag bereits eine Kristallstruktur vor, welches einen Vergleich der vorhergesagten Interaktionspunkte (Pseudoatome) und den im Kristallkomplex vorkommenden Aminosäurepositionen ermöglicht.

Schließlich wurde die Pseudorezeptorgenerierung auf Liganden des H4R angewendet, dessen Proteinstruktur nicht bekannt ist. Die Plausibilität des Pseudorezeptors wird an der Bewertung von möglichen H4R Bindetaschen erklärt. Der finale Pseudorezeptor wurde zur prospektiven Suche nach neuen Liganden im virtuellen Screening eingesetzt.

Kompatibilität zwischen kristallisierten Bindetaschen und Pseudorezeptoren

Fallbeispiel A - ein COX-2 Pseudorezeptor

Zur Bestätigung der Pseudorezeptorkonstruktion wurde die Kokristallstruktur von COX-2 und dem selektiven Inhibitor SC-558 ausgewählt (PDB Eintrag 6cox). SC-558 bietet dieselbe Struktur wie Celecoxib, bis auf eine durch Brom ersetzte Methylgruppe.

Nachdem die energieniedrigste Konformation des Referenzliganden berechnet (52,6 kcal/mol) wurde, wurde ein COX-2 Pseudorezeptor mit dem in dieser Arbeit entwickelten Programm automatisch generiert.

Abbildung 17 zeigt einen Ausschnitt des aktiven Zentrums aus der Kristallstruktur und dem Pseudorezeptor (nicht alle Aminosäuren des aktiven Zentrums sind dargestellt). Der Inhibitor SC-558 wird neben seiner aufgeklärten, bioaktiven Konformation (hellgrau) auch in der berechneten, energieniedrigsten Konformation (grau) gezeigt. Letztere diente als alleinige Referenzstruktur zur Gewinnung des Pseudorezeptors,

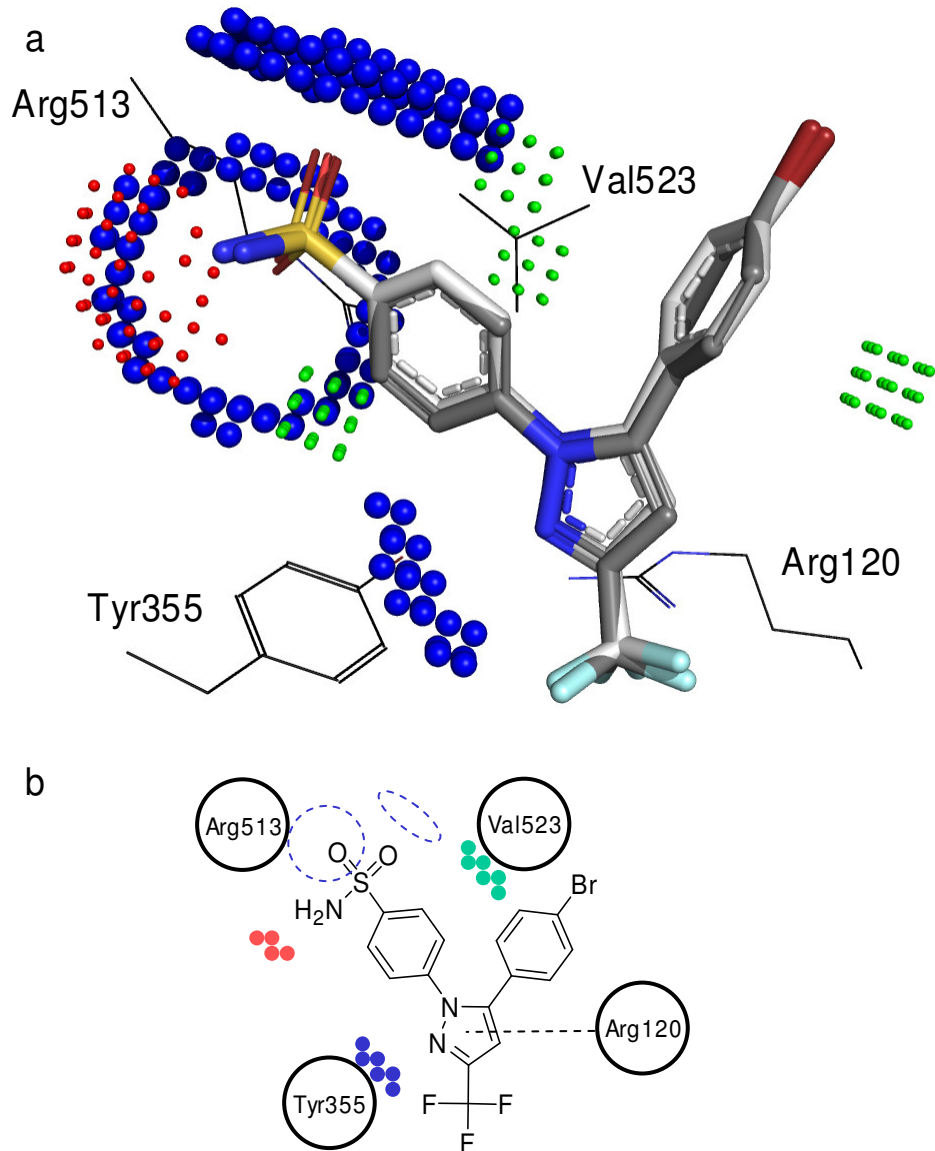


Abbildung 17. a) Abbildung des komplexierten SC-558 aus der PDB Struktur 6cox (hellgrau) mit der energieniedrigsten Ligandenkonformation (dunkelgrau) und die berechneten Pseudorezeptorpunkte (rot: Pseudoakzeptoratome, blau: Pseudodonoratome, grün: aromatische Pseudoatome). b) 2D Schema aus a).

welcher die Positionen für putative Rezeptoratome (Pseudoatome) vorhersagt. Die erkannten Substrukturen führen zur Berechnung folgender möglicher Pseudoatompositionen:

- Pseudoakzeptoratome ausgehend vom Stickstoff der Sulfonamidgruppe (rote Punktwolke im linken oberen Bereich).
- Pseudodonoratome ausgehend vom Stickstoff des Pyrazolringes (längliche blaue Punktwolke im linken unteren Bereich).

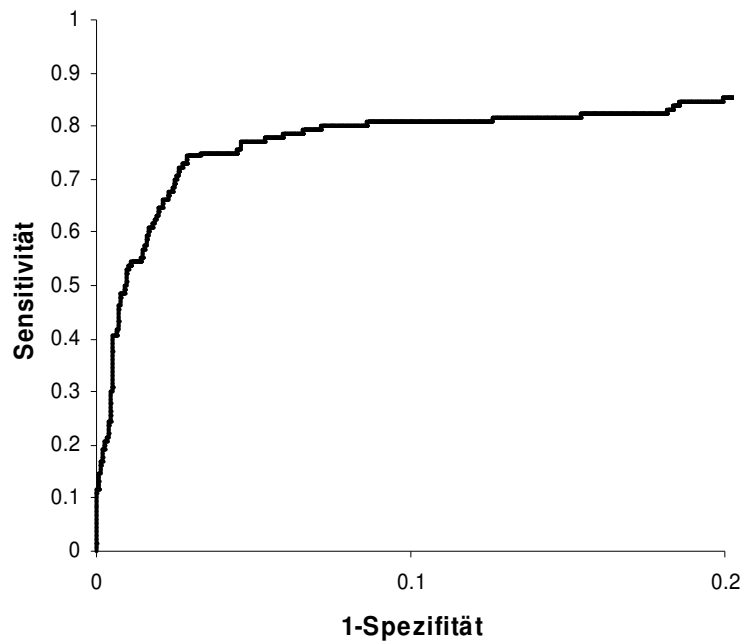


Abbildung 18. ROC Kurve der Anreicherung von COX-2 Inhibitoren nach einem virtuellen Screening mit einem COX-2 Pseudorezeptor. Der Graph gibt den Anteil der Richtig-Positiven (Sensitivität) in Relation zu den Inaktiven (1-Spezifität) in der sortierten Rangliste an.

- c) Pseudodonoratome um die Sauerstoffatome der Sulfonamidgruppe (blaue Ringe im linken oberen Bereich; der zweite Ring ist von der Seite abgebildet und als Ebene sichtbar).
- d) Aromatische Pseudoatome jeweils planar über und unter den Benzolgruppen (grüne Punktwolken).

Aus der Überlagerung des Pseudorezeptors mit der in der Kristallstruktur beobachteten Situation erkennt man mehrere Übereinstimmungen. Zum einen wird deutlich, dass die berechnete energieniedrigste Konformation von SC-558 mit seiner bioaktiven Konformation übereinstimmt. Zum anderen erkennt man, dass einige Pseudoatome auf die Positionen der tatsächlichen Aminosäurepositionen zeigen. Die Aminosäuren Arg513, Tyr355 und Val523 befinden sich exakt auf vorhergesagten möglichen Rezeptoratompositionen. Bei den ersten beiden Aminosäuren wurde auch die Funktion der wasserstoffbrückenfähigen Atome richtig vorhergesagt, welche für eine Wasserstoffbrückenausbildung zum Liganden nötig ist. Das Val523 befindet sich in dem für aromatische Rezeptoratome vorhergesagten Bereich. Es besitzt keine aromatische, aber eine lipophile Eigenschaft.

Weiterhin ist zu erkennen, dass nicht jedes Pseudoatom einem Pendant in der Kristallstruktur entspricht. So befindet sich zum Beispiel keine Aminosäure auf den Positionen für Pseudoakzeptoratome, welches an einer Missinterpretation der Kristallstrukturdaten im PDB Eintrag 6cox liegt (PDB Eintrag 4cox zeigt ein Histidin (Pos. 90) an exakt den roten Pseudoatompositionen). Außerdem konnte der Pseudorezeptor nicht die essentielle Interaktion vom Arg120 zum Pyrazol vorhersagen, welche laut Kristallstruktur (Abbildung 17, rechts unten, hintere Ebene) stattfindet.

Die Güte des COX-2 Pseudorezeptors ist nicht nur anhand seiner Korrespondenz zur Kristallstruktur, sondern auch mit der Fähigkeit zum virtuellen Screening zu beurteilen. Um dies zu untersuchen, wurde ein auf Korrelationsvektoren basierender Deskriptor des Pseudorezeptors berechnet und die Wirkstoffdatenbank COBRA (8.311 Moleküle) nach COX-2 Inhibitoren (136 Aktive) durchsucht. Abbildung 18 zeigt das Ergebnis des retrospektiven Screening nach COX-2 Inhibitoren mit dem generierten COX-2 Pseudorezeptor an. Die Abbildung 18 veranschaulicht die Ergebnisse in den vorderen Perzentilen der sortierten Datenbank (maximal 20% Falsch-Positive).

Eine Screeningmethode sollte in den ersten Perzentilen eine hohe Anreicherung der aktiven Moleküle in der sortierten Wirkstoffdatenbank aufweisen. Üblicherweise werden nur Substanzen in diesem Bereich für weitere Forschungszwecke in Betracht gezogen. Die Abbildung 18 zeigt, dass 80% der aktiven Moleküle bereits nach Vorhersage von nur 10% Falsch-Positiven gefunden werden. Bereits in den ersten zwei Perzentilen der sortierten Datenbank (166 Moleküle) befanden sich 74 selektive COX-2 Inhibitoren, fünf nicht-selektive COX Inhibitoren und zwei Dualinhibitoren der Cyclooxygenase und 5-Lipoxygenase (nicht gezeigt). Obwohl der COX-2 Pseudorezeptor nur auf einem einzelnen Referenzmolekül basiert, ist er in der Lage, die charakteristische Konstellation der Pharmakophorpunkte der Coxibe in einer Moleküldatenbank effizient wieder zu finden. Dies zeigt, dass der Pseudorezeptor des Fallbeispiels A zum virtuellen Screening nach COX-2 Inhibitoren geeignet ist.

Fallbeispiel B - ein PPAR α Pseudorezeptor

Im Jahr 2007 wurde eine hochaufgelöste, komplexierte 3D Struktur von PPAR α und GW590735 publiziert (PDB Eintrag 2p54, Auflösung: 1,8 Å (Sierra *et al.*, 2007)), welche in diesem Fallbeispiel als Kontrollstruktur dient.

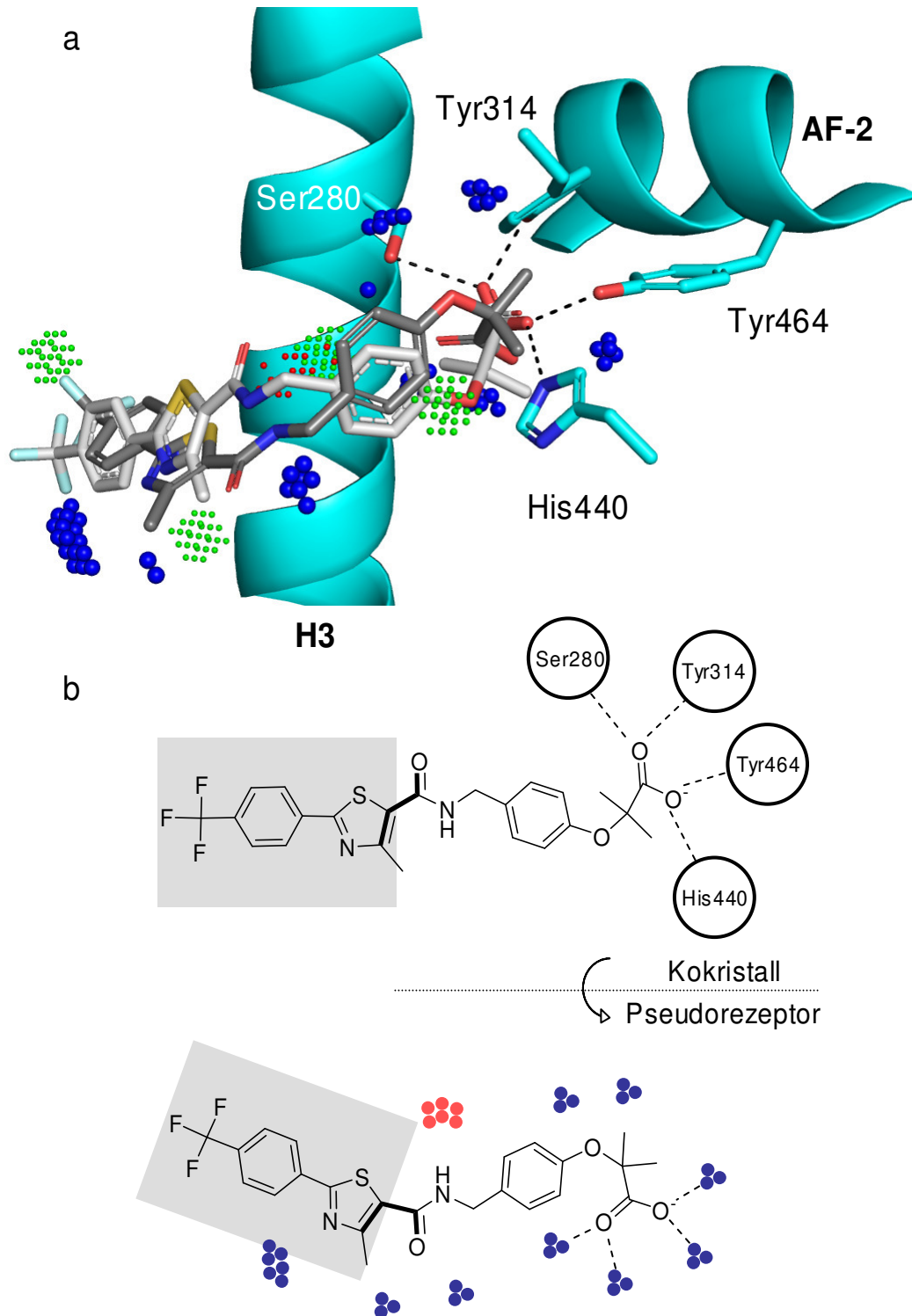


Abbildung 19. a) 3D Vergleich des PPAR α Pseudorezeptors, aufbauend auf der energieniedrigsten Konformation des GW590735 (dunkelgrau), mit der Situation in der Kokristallstruktur. Gezeigt ist das kristallisierte GW590735 (hellgrau), die Helices H3 und AF-2 und die Schlüsselinteraktionen (gestrichelte Linien) zu den Schlüsselamino säuren b) 2D Schema des Vergleichs aus a) mit dem Kristallzustand (oben) und dem Pseudorezeptor (unten). Hervorgehoben sind die Konformationen des GW590735 und die realen sowie vorhergesagten Interaktionen der Carboxylgruppe.

Wie im Fallbeispiel A, wurde die energieniedrigste Konformation von GW590735 berechnet (88,8 kcal/mol). Sie diente als Referenzstruktur zur automatischen Konstruktion eines PPAR α Pseudorezeptors. Der Vergleich zwischen dem Pseudorezeptorkonstrukt und der Kokristallstruktur von PPAR α mit dem Liganden GW590735 (PDB Eintrag 2p54) ist in Abbildung 19 verdeutlicht. Abbildung 19a zeigt eine Überlagerung des Pseudorezeptors und der kokristallisierten PPAR α Bindetasche, welche um die Helix H3 positioniert ist. Die kristallisierte, aktive Konformation des Liganden (hellgrau) baut Wasserstoffbrückeninteraktionen (gestrichelte Linien) zu den vier Aminosäuren Ser280, Tyr314, Tyr464 und His440 auf. Diese Interaktionen werden als Schlüsselinteraktionen für eine Bindung an PPAR α genannt (Sierra *et al.*, 2007). Zusätzlich ist die AF-2 Helix dargestellt, weil sich hier, unter anderem, eine der Schlüsselaminosäuren befindet (Tyr464). Die Pseudoatome, die in ihrer Gesamtheit den Pseudorezeptor bilden, wurden um die Referenzstruktur, der energieminierten Konformation des GW590735 (dunkelgrau), projiziert. Auf den ersten Blick sind keine Gemeinsamkeiten zwischen der aktiven und berechneten, energieniedrigsten Konformation festzustellen: sie unterscheiden sich im 3D Raum scheinbar stark. Um die Ursache hierfür zu erörtern, ist eine vereinfachte Darstellung in 2D hilfreich (Abbildung 19b). Im oberen Teil der Abbildung 19b ist die Situation aus der Kokristallstruktur ersichtlich. GW590735 interagiert mit seiner Carboxylgruppe mit allen vier Schlüsselaminosäuren (Kreise mit Indices). Im unteren Teil der Abbildung 19b ist die Situation im Pseudorezeptor dargestellt. Die energieniedrigste Konformation des GW590735 ist nicht dieselbe wie in der Kristallstruktur. Durch eine Rotation an der Einfachbindung vor der Amidgruppe (dicke Bindungen) liegt der Teil von der Amidgruppe bis zur Carboxylgruppe in einem „gespiegelten“ Zustand (Spiegelebene = gepunktete Linie) vor. Der grau hinterlegte Teil beider Konformationen ist von dieser Rotation nicht betroffen.

Obwohl die Referenzstruktur zur Konstruktion des PPAR α Pseudorezeptors nicht der aktiven Konformation entspricht, stellt es die Distanzrelationen insbesondere für die Schlüsselinteraktionen korrekt dar. Die gestrichelten Linien im oberen Teil der Abbildung 19b zeigen die beobachteten Wasserstoffbrücken in der Kristallstruktur. Im unteren Teil geben die gestrichelten Linien die vorhergesagten Interaktionen der Carboxylgruppe an. Hiernach werden Rezeptoratome mit einer Wasserstoffbrücken-

donoreigenschaft an vier Positionen um die Carboxylgruppe erwartet, die in ihren Distanzen und Winkeln zueinander eben der Situation im Kristall entsprechen.

Wie im Fallbeispiel des COX-2 Pseudorezeptors wurde die Güte des Pseudorezeptors hinsichtlich seiner Performanz zur Anreicherung von PPAR Liganden in einem retrospektiven virtuellen Screening untersucht. Hierzu wurde ein Korrelationsvektor-basiertes Screening der Wirkstoffdatenbank COBRA mit dem PPAR α Pseudorezeptor durchgeführt. In der COBRA sind insgesamt 94 PPAR Liganden annotiert. Diese setzen sich aus 16 PPAR α , 44 PPAR γ , 7 PPAR δ und 24 Dualliganden (gleichzeitige Liganden der Subtypen α und γ oder δ und γ) zusammen.

Abbildung 20 zeigt die Leistung des Pseudorezeptors im virtuellen Screening nach PPAR Liganden. 30% der aktiven Moleküle wurden bereits nach der Vorhersage von 10% Falsch-Positiven entdeckt. Nach 20% Falsch-Positiver wurde fast die Hälfte der PPAR Liganden (46%) entdeckt. Obwohl der PPAR α Pseudorezeptor auf einem einzelnen PPAR α Referenzliganden basiert, ist er offensichtlich in der Lage, die charakteristische Konstellation der Carboxylgruppeninteraktionen der Fibrate in einer Moleküldatenbank wieder zu finden. Hierbei ist die Performanz niedriger als die des COX-2 Pseudorezeptors, weil die Carboxylgruppe kein konserviertes Strukturkriterium

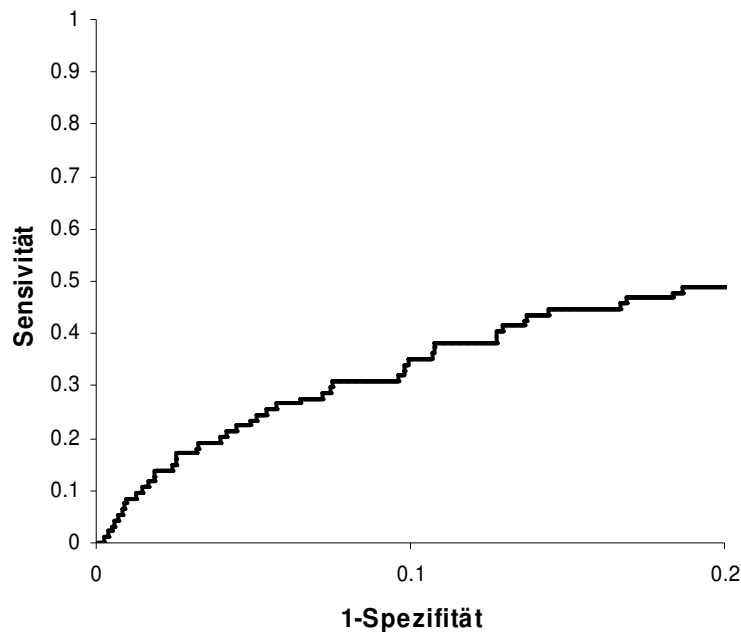


Abbildung 20. ROC Kurve der Anreicherung von PPAR Liganden nach einem virtuellen Screening mit einem PPAR α Pseudorezeptor (Sensivität = Richtig-Positive, 1-Spezifität = Anteil Inaktive).

(„privilegierte“ Substruktur) für PPAR Liganden ist.

Neben dem Auffinden von allgemeinen PPAR Modulatoren ist die Leistung im selektiven Screening von Bedeutung. In dem retrospektiven Screening nach PPAR Liganden (Abbildung 20) wurden nach der Vorhersage von 20% falsch-positiver Strukturen insgesamt 43 PPAR Liganden gefunden. Davon sind 15 PPAR α , 18 PPAR γ , 3 PPAR δ und 7 Dualliganden. In den ersten zwei Perzentilen werden 12 PPAR Liganden entdeckt, wobei 4 PPAR α und 8 PPAR γ Liganden sind. Es werden keine PPAR δ Liganden und keine Dualliganden gefunden. Dies zeigt, dass der im Fallbeispiel B entworfene PPAR α Pseudorezeptor zum virtuellen Screening, auch hinsichtlich Selektivitätskriterien, geeignet ist.

Konstruktion eines Pseudorezeptors für Antagonisten des H4R

Alle Antagonisten aus Abbildung 10 wurden zur Konstruktion eines H4R Pseudorezeptors herangezogen, der den „antagonistischen Zustand“ der Bindetasche modellieren soll. JNJ7777120 dient als Überlagerungstemplat für die restlichen Ausgangssubstanzen. Es besitzt nicht die höchste Affinität, jedoch weist es zum einen eine höhere Affinität als **1** und **2**, und zum anderen eine rigidere Struktur als **3** und **4** auf.

Nachdem die energieniedrigste Konformation von JNJ7777120 berechnet wurde (50,9 kcal/mol), wurden die restlichen Verbindungen **1**, **2**, **3** und **4** anhand ihrer gemeinsamen Pharmakophorpunkte übereinander gelagert. Abbildung 21 stellt eine schematische Darstellung der 3D Strukturüberlagerung in 2D dar, in denen die Ausgangssubstanzen anhand ihrer Gemeinsamkeiten übereinander positioniert wurden:

- Wie alle biogenen Amine besitzen diese Verbindungen eine Wasserstoffbrückendonorfunktion (basischer Stickstoff der Methypiperazinkopfgruppe des JNJ7777120, erstes blaues Band von rechts).
- Die zweite Gemeinsamkeit ist eine Wasserstoffbrückenakzeptorfunktion, die am Amidsauerstoff des JNJ7777120 ausgerichtet ist (rotes Band)
- Eine zweite Wasserstoffbrückendonorfunktion am Indolstickstoff bildet den dritten gemeinsamen Pharmakophorpunkt (drittes blaues Band von rechts)

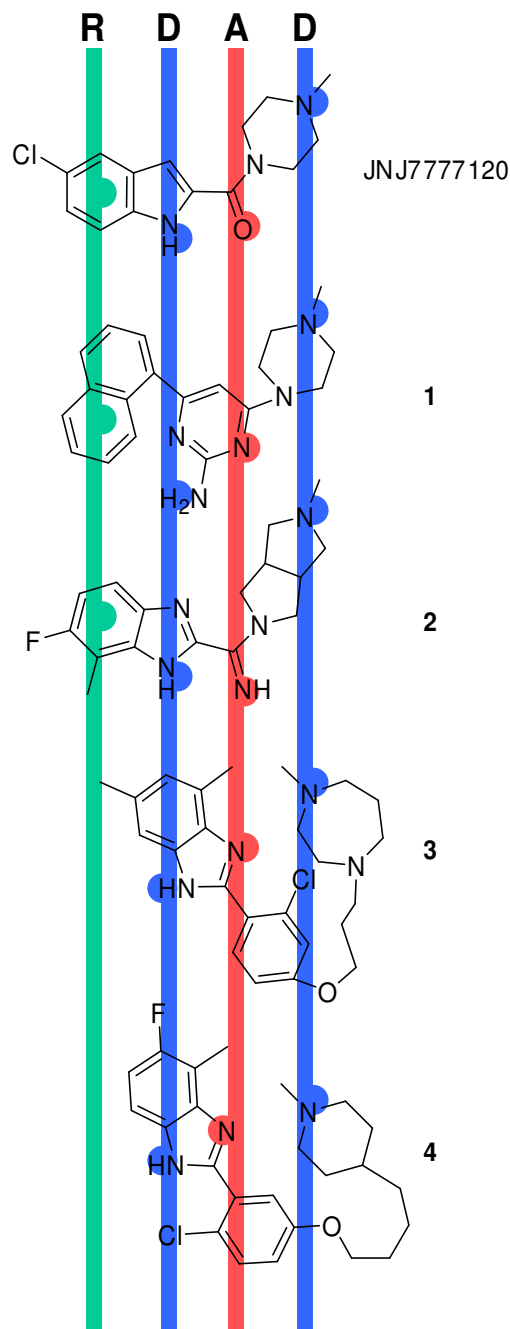


Abbildung 21. Gemeinsame Pharmakophorpunkte (A = Akzeptor, D = Donor und R = aromatisch) der Ausgangsverbindungen zur H₄ Pseudorezeptorkonstruktion.

- Eine weitere Gemeinsamkeit ist eine aromatische Gruppe (grünes Band). Die aromatischen Ringe der Verbindungen **3** und **4** liegen sowohl in 3D als auch in 2D nicht exakt auf dieser Bandenposition.

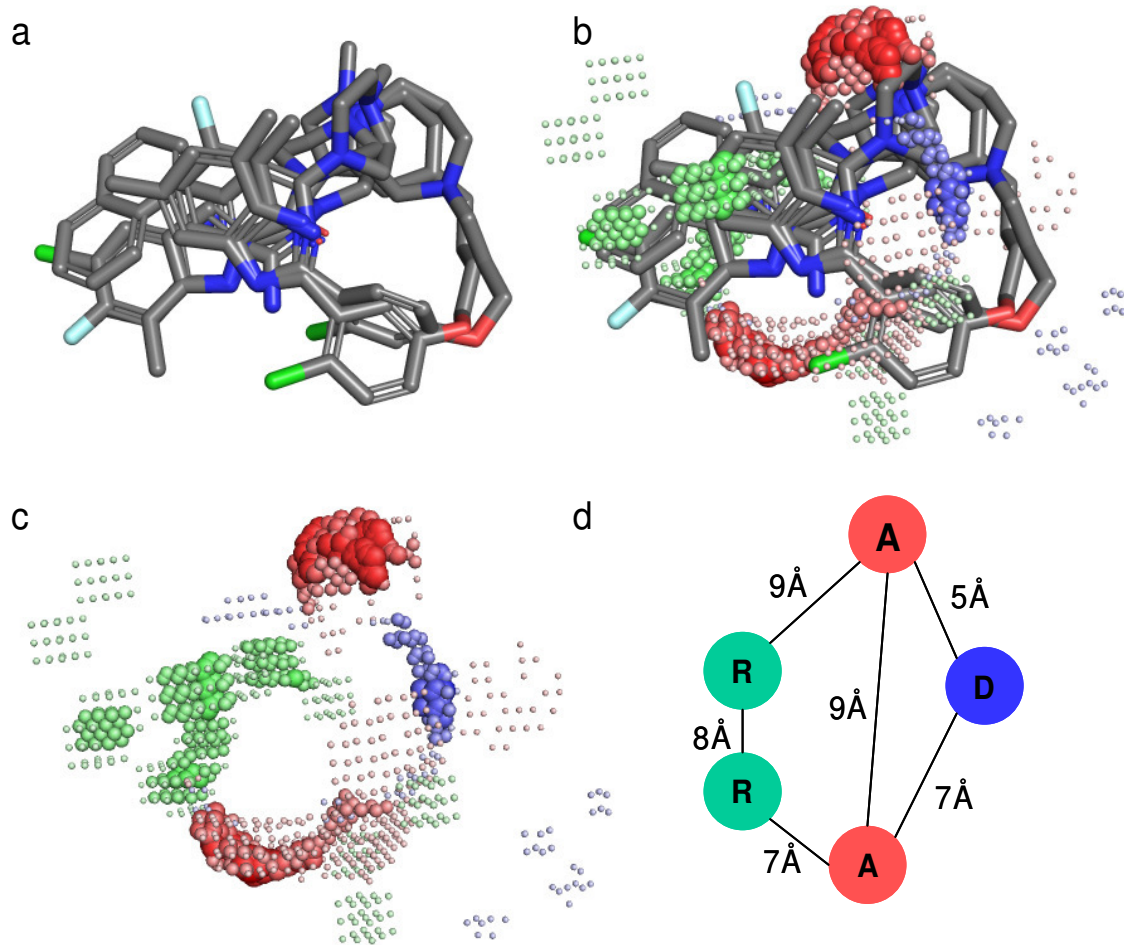


Abbildung 22. Ligandenalignment (a) mit paralleler Darstellung aller generierten Pseudoatome (b), die in der Summe den H₄ Pseudorezeptor (c) bilden. Zusätzliches 2D Schema zur Verdeutlichung aller Pseudoatome und ihren mittleren Distanzen zueinander, die von mindestens der Hälfte aller Referenzliganden gebildet wurden (d).

Das resultierende Ligandenalignment wurde zur automatischen Konstruktion eines H₄ Pseudorezeptors mit dem in dieser Arbeit entwickelten Programm genutzt. Abbildung 22 zeigt den resultierenden Pseudorezeptor mit gewichteten Pseudoatomen. Je größer der Radius der einzelnen Pseudoatome ist, umso mehr Referenzliganden haben zu deren Ausbildung beigetragen. Die automatische Gewichtung der Pseudoatome ist insbesondere bei der Konstruktion von Pseudorezeptoren auf Basis multipler Liganden hilfreich. Sie zeigt, welche Pseudoatome von mehreren Liganden projiziert wurden.

Zusätzlich ist ein 2D Schema (Abbildung 22d) dargestellt, in dem die Komplementarität des Pseudorezeptors zu den gemeinsamen Pharmakophorpunkten der Referenzverbindungen aus Abbildung 21 ersichtlich werden. Zu jedem „Pharmakophorband“ aus dem Ligandalignment gibt es einen korrespondierenden Pseudoatomcluster mit komplementärer Wasserstoffbrücken-bildender Eigenschaft im Pseudorezeptor. Entsprechend finden sich zwei aromatische Cluster, die um das gemeinsame aromatische System vom Großteil der Ausgangssubstanzen angeordnet wurden.

Evaluation von potentiellen H4R Bindetaschen

Der erzeugte H₄ Pseudorezeptor wurde anschließend zur Evaluierung von potentiellen Bindetaschen des H4R genutzt. Dieser Ansatz erfüllt zwei Zwecke: zum einen stützt ein erfolgreiches Auffinden von möglichen Rezeptorbindetaschen die Hypothese des konstruierten Pseudorezeptors. Zum anderen zeigt es eine Möglichkeit der Verknüpfung von ligandenbasierten Informationen in rezeptorbasierten Methoden. Das gefundene Arrangement von Bindetaschenamino säuren könnte folglich in der Lage sein alle Referenzliganden, die zur Konstruktion des Pseudorezeptors beigetragen haben, zu binden.

Die Aminosäuren der H4R Bindetasche

Derzeit existiert keine aufgeklärte 3D Struktur des H4R, welche eine direkte Identifizierung der an der Bindetasche beteiligten Aminosäuren erlauben würde. Infolgedessen wurde mittels selektiven Mutationsstudien die Lage der Bindetasche zwischen den Transmembrandomänen bestätigt (Shin *et al.*, 2002). Zusätzlich wurden Homologiemodelle, basierend auf der Rezeptorstruktur des Rhodopsins, erstellt und für Erklärungen von Bindemöglichkeiten bekannter Agonisten und Antagonisten herangezogen (Jongejan *et al.*, 2008; Jojart *et al.*, 2008; Kiss *et al.*, 2008).

Obwohl die Lage der Bindetasche unter anderem aufgrund der Ähnlichkeiten innerhalb der GCPR Klasse A wohl definiert ist, existieren zwei unterschiedliche Theorien zur Orientierung der Liganden des H4R. Sie unterscheiden sich an der Bindung des positiv geladenen Zentrums der Liganden (z.B. die Methylpiperazinkopfstruktur des JNJ7777120, die Aminogruppe des Histamins) an unterschiedlichen Aminosäuren. Auf der einen Seite wird eine Bindung des geladenen Amins mit dem Aspartat an der

Sequenzposition 94^{3.32} behauptet (Jongejan *et al.*, 2008).² Diese Hypothese wird durch die homologe Position des Aspartats in den Histaminrezeptoreng efestigt (Jacoby *et al.*, 2006). Auf der anderen Seite wurde die konträre Orientierung, nämlich die Bindung des aminergen Zentrums zum Glu182^{5.46} gezeigt (Kiss *et al.*, 2008). Die Aminosäuren Asp94^{3.32} und Glu182^{5.46} befinden sich an den beiden Polen der Bindetasche, und sind die einzigen polaren Aminosäuren, die eine Bindung mit dem geladenen Stickstoff eingehen können (orange Aminosäuren in Abbildung 7). Fakt ist, dass weitere Aminosäuren an der Bindetasche beteiligt sind, die unter anderem für die Aktivität des Rezeptors verantwortlich sein müssen (Shin *et al.*, 2002). Denn diese wird nicht von einer Wasserstoffbrücke zu einer basischen Molekülstruktur gesteuert, weil dies ein allgemeines Bindungskriterium für alle Liganden der Klasse A GPCR darstellt. Folglich sind ausschließlich experimentelle Ergebnisse aus Mutationsstudien für eine Interpretation der Bindetaschenzustände und Aktivitätsauswirkungen heranzuziehen.

Extraktion der Aminosäurekoordinaten aus der H4R Bindetasche

Eine notwendige Voraussetzung zur Bewertung von möglichen Bindetaschenformen des H4R ist die Identifikation von Interaktionspunkten von den an der Bindetasche beteiligten Aminosäuren. Die Auswahl der Aminosäuren wurde aus den Mutationsstudien von (Shin *et al.*, 2002). und (Jongejan *et al.*, 2008). abgeleitet. Tabelle 6 zeigt eine Zusammenfassung der an der Bindetasche beteiligten Aminosäuren und deren als „kritische Punkte“ definierte Interaktionspunkte (alle Wasserstoffbrückenbildende und aromatische Atome). Die ausgewählten Punkte wurden innerhalb einer molekularen Dynamiksimulation (MD) eines H4R Homologiemodells verfolgt.

² Im Folgenden werden Aminosäuren des Histamin H₄ Rezeptors im 3-Buchstabencode mit der zugehörigen Sequenzposition und einer hochgestellten Positionsangabe nach (Ballesteros & Weinstein, 1995) beschrieben. Dieses Schema gibt vor dem Punkt die Nummer der Helix an, in der sich die jeweilige Aminosäure befindet. Zusätzlich wird nach dem Trennzeichen die relative Position zu der am stärksten konservierten Aminosäure in der jeweiligen Helix, die mit dem Index „50“ bezeichnet wird, angegeben.

Tabelle 6. Übersicht der ausgewählten Aminosäuren und ihren Interaktionspunkten (Wasserstoffbrückendonoren, -akzeptoren, gleichzeitiger Akzeptor und Donor, aromatisch), deren Positionen innerhalb der MD beobachtet wurden.

Aminosäure	Position	Kritische Punkte (Element ^{Atomindex})
Asp	94 ^{3.32}	O ¹³⁶⁸ , O ¹³⁶⁹
Tyr	95 ^{3.35}	O ¹³⁸⁵ , C ¹³⁷⁹ , C ¹³⁸⁰ , C ¹³⁸² , C ¹³⁸⁴ , C ¹³⁸⁷ , C ¹³⁸⁹
Cys	98 ^{3.36}	S ¹⁴³⁸
Thr	178 ^{5.42}	O ²⁶⁹⁹
Ser	179 ^{5.43}	O ²⁷¹⁴
Glu	182 ^{5.46}	O ²⁷⁶⁸ , O ²⁷⁶⁹
Phe	183 ^{5.47}	C ²⁷⁷⁹ , C ²⁷⁸⁰ , C ²⁷⁸² , C ²⁷⁸⁴ , C ²⁷⁸⁶ , C ²⁷⁸⁸
Trp	241 ^{6.48}	C ³⁷⁶⁹ , C ³⁷⁷⁰ , N ³⁷⁷² , C ³⁷⁷⁴ , C ³⁷⁷⁵ , C ³⁷⁸⁰ , C ²³⁷⁸² , C ³⁷⁷⁸ , C ³⁷⁷⁶
Tyr	244 ^{6.51}	O ³⁸²³ , C ³⁸¹⁷ , C ³⁸¹⁸ , C ³⁸²⁰ , C ³⁸²² , C ³⁸²⁵ , C ³⁸²⁷
Ser	245 ^{6.52}	O ³⁸³⁸
Gln	272 ^{7.42}	O ⁴²⁸⁹ , N ⁴²⁹⁰

Das Modell basiert auf der Struktur des β_2 adrenergen Rezeptors über die Dauer von 1 ns mit 525 Standbildaufnahmen bei einer Temperatur von 310 K (durchgeführt von Tim Werner, Arbeitsgruppe Prof. Schneider, Goethe-Universität Frankfurt a.M.). Die zeitliche Bewegung der Punkte spiegelt mögliche H4R Bindetaschenformen wieder, die eine zur Ligandenbindung fähige Konstellation annehmen kann. Die Position des Trp241^{6.48} wird als besonders wichtig angesehen, da es als molekularer Schalter für eine agonistische und antagonistische Aktivität des Rezeptors durch eine Konformationsänderung an der Helix 6 verantwortlich sein kann (Shin *et al.*, 2002).

Bewertung der möglichen Bindetaschenformen

Mit Hilfe des geometrischen Hashings wurden die in einer MD vorkommenden Arrangements der in Tabelle 6 genannten kritischen Punkte auf ihre Übereinstimmung mit dem H₄ Pseudorezeptor hin bewertet. Dabei wurden nur Pseudoatome mit einem Gewicht über 0,5 in Betracht gezogen, da diese von mindestens der Hälfte der Referenzliganden projiziert werden.

Abbildung 23a zeigt die resultierende H4R Bindetasche, die die Situation des Pseudorezeptors widerspiegelt. Die Perspektive veranschaulicht die von den sieben Transmembranhelices geformte Bindetasche mit Blick aus der extrazellulären Matrix. Neben den an der Bindetasche beteiligten Aminosäuren aus Tabelle 6 ist der Pseudorezeptor aus Abbildung 22 abgebildet (weiße Kugeln stellen kritische Punkte mit exakter Pseudoatomeigenschaft und Koordinatenübereinstimmung dar).

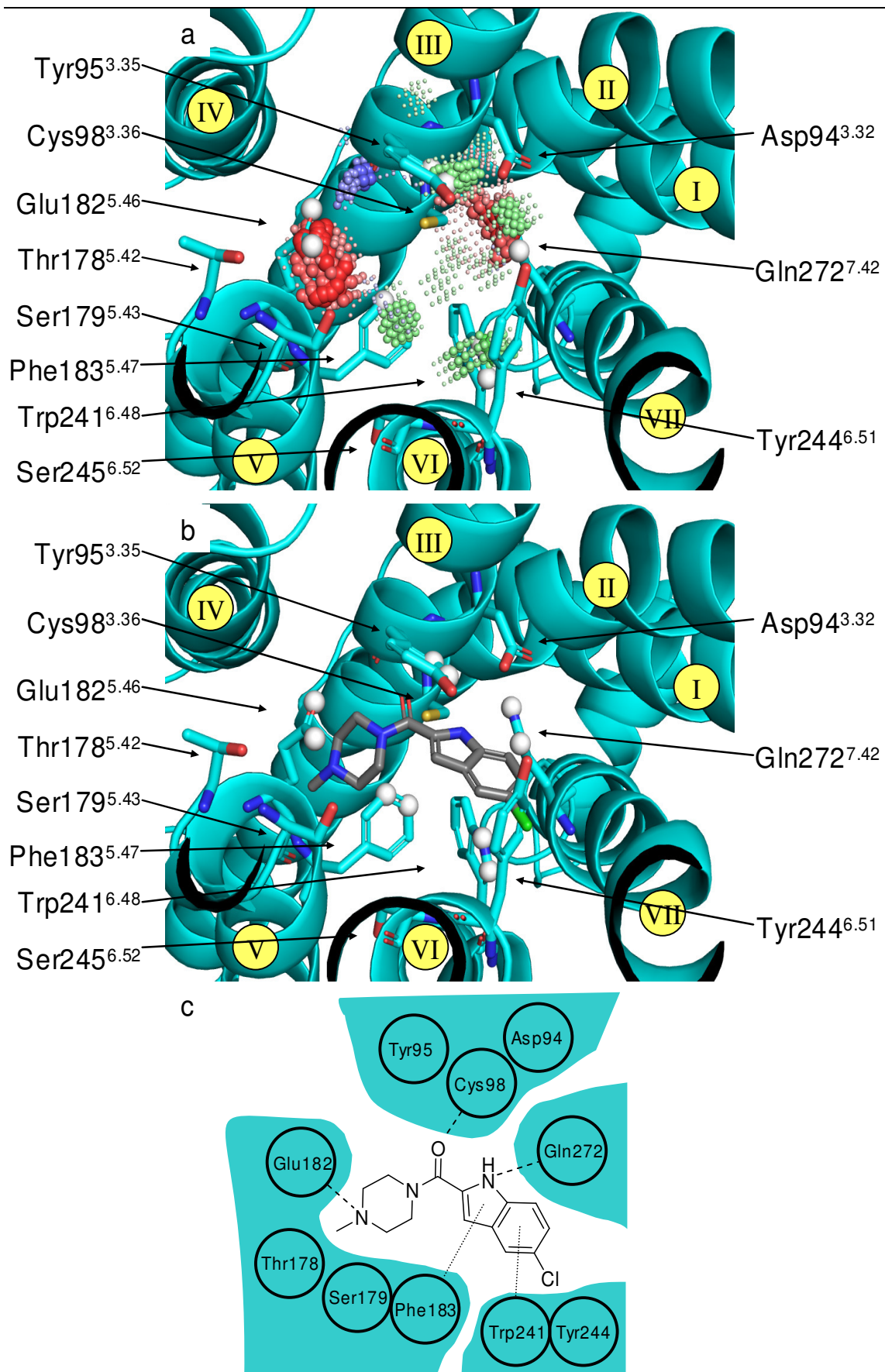


Abbildung 23. Blick aus der extrazellulären Matrix auf die Bindetasche mit der Aminosäurenkonstellation, die dem Pseudorezeptor entspricht (a). Die Helices sind in gelben Kreisen mit römischen Ziffern nummeriert. Die resultierende Orientierung des Referenzliganden JNJ777120 (b) und ein 2D Schema der möglichen Interaktionen (c). Wasserstoffbrücken sind in gestrichelten und aromatische Interaktionen in gepunkteten Linien dargestellt.

Die gefundene Tasche zeigt eine partielle Übereinstimmung der Bindetaschenamino­säuren mit dem Pseudorezeptor. Anhand der Übereinstimmungen ist es möglich, spezifische Interaktionen zwischen Rezeptoratomen und den Referenzliganden in einem gemeinsamen Bindemodus vorherzusagen. Abbildung 23b zeigt die mögliche Orientierung der Referenzliganden am Beispiel des JNJ7777120, welches als Überlagerungstemplat für die restlichen Ausgangssubstanzen fungierte. Neben der 3D Darstellung zeigt Abbildung 23c ein 2D Schema der möglichen Interaktionen.

Im gefundenen Arrangement interagiert das positiv geladene Zentrum des Liganden (Methylpiperazinsubstruktur) mit dem Glu182^{5,46}. Weitere Wasserstoffbrückeninteraktionen sind vom Amidsauerstoff zu Cys98^{3,36} und vom Indolstickstoff zu Gln272^{7,42} möglich. Zusätzlich stabilisieren aromatische Interaktionen zu Trp241^{6,48} und Phe183^{5,47} die Konformation der Helix 6. Abbildung 23c zeigt die sterischen Verhältnisse in der gefunden Bindetasche mittels farblichen Volumen aller elf an der Bindetasche beteiligten Aminosäuren.

Im Verlauf der MD Simulation des H4R Homologiemodells wurde mit Hilfe des Pseudorezeptors eine Aminosäurenkonstellation wieder gefunden, die einen möglichen antagonistischen Zustand des H4R bei Bindung der Referenzliganden repräsentiert. Der vorgeschlagene Bindemodus deckt sich mit der Hypothese aus (Jojart *et al.*, 2008) und (Kiss *et al.*, 2008), die eine Interaktion des basischen Zentrums mit Glu182^{5,46} postulieren. Im Gegensatz zu (Jongejan *et al.*, 2008) zeigt die gefundene Standbildaufnahme kein freies Volumen um Asp94^{3,32}, welches dadurch nicht in der Lage ist, Wasserstoffbrücken auszubilden. Außerdem wird die für eine antagonistische Wirkung der Ligandenbindung essentielle Interaktion zur Helix 6 gezeigt. Sie wird durch aromatische Interaktionen zu dem Trp241^{6,48}, welches in Mutationsstudien von (Shin *et al.*, 2002) als kritische Interaktion annotiert wurde, bestätigt. Somit stellt die gefundene Tasche eine gültige Konstellation von Taschenamino­säuren dar, die die Hypothese des konstruierten Pseudorezeptors stützt.

Mit dem Ergebnis dieses Ansatzes lassen sich folgende Aussagen treffen:

1. Der Ansatz zeigt, dass die Verknüpfung von ligandenbasierten und rezeptorbasierten Informationen mit Hilfe von dem in dieser Arbeit entwickelten Pseudorezeptormodell prinzipiell möglich ist. Erst durch die

- vorgestellte Evaluation der Ausgabe einer MD Simulation wurden die Voraussetzungen für weitergehende rezeptorbasierte Anwendungen geschaffen.
2. Das gefundene Arrangement von Aminosäuren kann nun als Basis weitergehende rezeptorbasierte Ansätze, wie Docking oder rationales Design, dienen. Es ist in der Lage, alle Referenzstrukturen zu binden. Dies stellt den grundsätzlichen Unterschied zu reinen rezeptorbasierten Methoden, wie dem automatischen Ligandendocking, dar. So kann zwar mit Hilfe von Docking ebenfalls ein Bindemodus vorhergesagt werden (Mohan *et al.*, 2005; Rester, 2006), jedoch gilt dieser ausschließlich für den gerade betrachteten Liganden und nicht, wie im Falle des Pseudorezeptoransatzes, für multiple Liganden gleichzeitig. Nichtsdestotrotz kann das hier entwickelte Pharmakophor-abgeleitete Pseudorezeptormodell Dockingprogramme nicht ersetzen. Es integriert keine Energiefunktionen, sodass zur Vorhersage von Bindeaffinitäten Dockingprogramme oder molekulare Mechanikanalysen mit Poisson-Boltzmann Oberflächenmethoden notwendig bleiben (Kuhn *et al.*, 2005).
 3. Das erfolgreiche Auffinden einer gültigen Rezeptorbindetasche stützt die Hypothese des konstruierten Pseudorezeptors, welcher sich demnach zur Suche nach neuen Liganden des H4R eignet.

Virtuelles Screening nach H4R Liganden

Mit dem konstruierten H₄ Pseudorezeptor wurde ein prospektives Screening nach H4R Liganden durchgeführt. Hierbei wurde eine Screeningdatenbank bestehend aus allen Molekülen der drei Datenbanken Specs, Asinex Gold und Asinex Platinum zusammengestellt (insgesamt 562.851 Moleküle). Abbildung 24 zeigt eine Übersicht der zum Screening gehörenden Schritte.

Im ersten Schritt wurde für jedes Molekül eine 3D Struktur (energieniedrigste Konformation) mit dem Programm Corina V3.2 (Molecular Networks GmbH, Erlangen, www.molecular-networks.com) berechnet. Danach wurden im so genannten Waschprozess die Anzahl der Wasserstoffe an Säure- und Basegruppen mit dem Programm MOE korrigiert. Nachdem jedes Molekül in den hier vorgestellten 90-dimensionalen Korrelationsvektor übersetzt wurde, wurde die Substanzsammlung auf Grundlage der Euklidischen Ähnlichkeit der Moleküldeskriptoren und des

Pseudorezeptorkorrelationsvektors sortiert, sodass Moleküle mit geringer Distanz in den oberen Rängen der Datenbank zu finden sind. Im letzten Schritt (*cherry-picking*) wurden unter den ersten 240 Molekülen (0,0004%) der sortierten Screeningdatenbank insgesamt 10 Strukturen nach visueller Inspektion zur Bestellung ausgesucht (Abbildung 25). Dabei wurde insbesondere darauf geachtet, dass die Moleküle neben einem unterschiedlichen Grundgerüst auch eine Gruppe mit basischer Funktion enthalten, um dem basischen Zentrum der Referenzliganden gerecht zu werden.

Die Substanzen wurden in einem Verdrängungsassay auf ihre kompetitive Bindung gegenüber 10nM [3H]Histamin auf 30µg Membransuspension mit humanen H4R getestet (Tim Kottke, Johann Wolfgang Goethe-Universität Frankfurt, Arbeitskreis Prof. Stark, Medizinische Chemie, Tabelle 7).

Zwei der getesteten Strukturen, AG-401/34827007 und AH-262/36083026, zeigen einen $K_i < 40\mu\text{M}$. Der unterschiedliche Chemotyp bei gleicher Bioaktivität zeigt die Generalisierungsfähigkeit der in dieser Arbeit entwickelten Pseudorezeptormethode im virtuellen Screening (*Scaffold-Hopping* (Schneider *et al.*, 1999)).

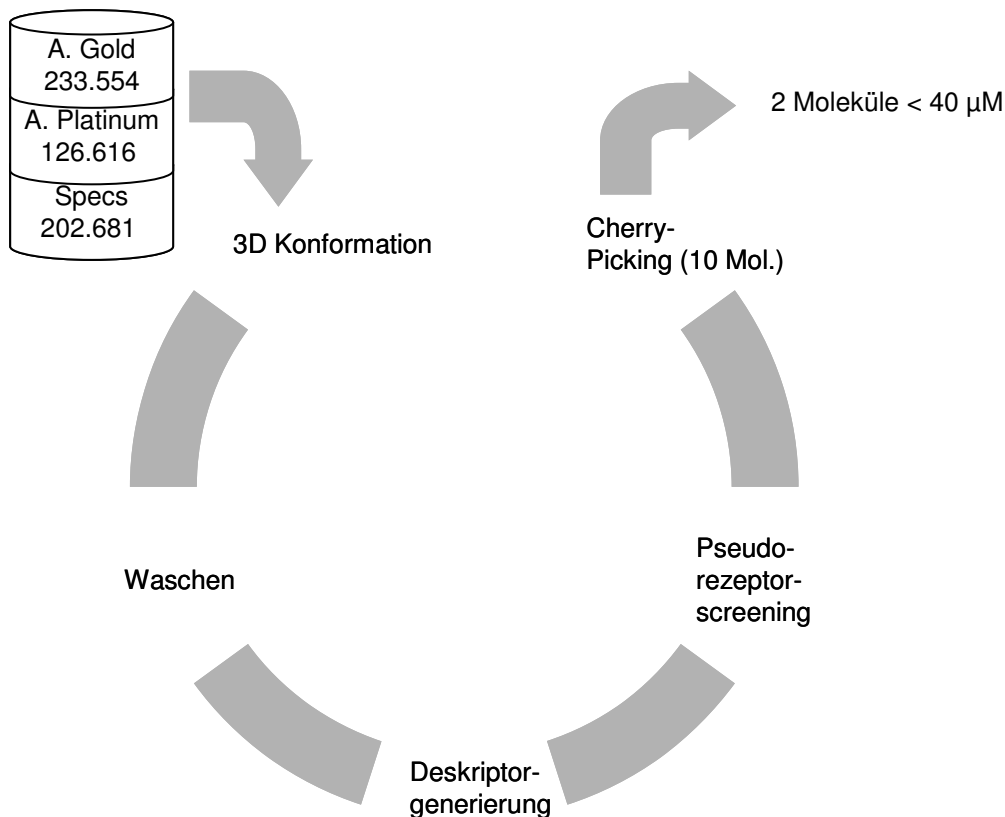


Abbildung 24. Ablaufprotokoll des H₄ Pseudorezeptorscreenings.

Die restlichen Verbindungen zeigen moderate Affinitäten, die teilweise schwanken (BAS 04379911 und BAS 03673618). Generell kann dies auf Löslichkeitsprobleme zurückgeführt werden, was auch bei den Strukturen BAS 00232413 und BAS 03420660 zu fehlenden K_i -Werten geführt haben mag.

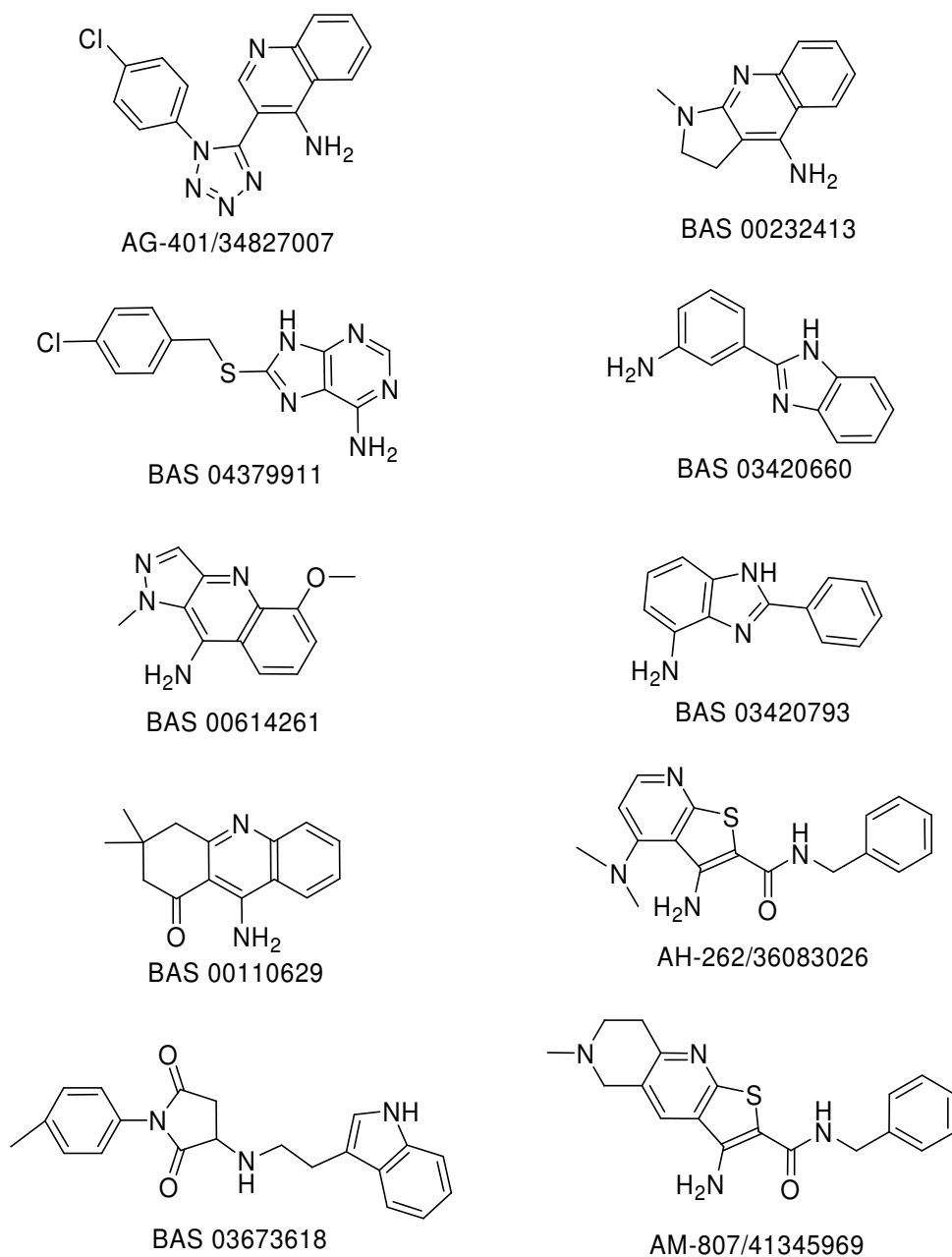


Abbildung 25. Bestellte Moleküle aus dem H_4 Pseudorezeptorscreening.

Tabelle 7. Experimentell bestimmte K_i -Werte der bestellten Moleküle (Abbildung 25).

Molekülname	Datenbank	Messung 1 [μ M]	Messung 2 [μ M]
AG-401/34827007	Specs	35,4	28,5
BAS 04379911	Asinex Gold	140,6	24,5
BAS 00614261	Asinex Gold		<i>nicht geliefert</i>
BAS 00110629	Asinex Gold	194,1	264,5
BAS 03673618	Asinex Gold	109,9	42,9
BAS 00232413	Asinex Gold	91,0	<i>keine Angabe</i>
BAS 03420660	Asinex Gold	113,1	<i>keine Angabe</i>
BAS 03420793	Asinex Gold	190,4	454,1
AH-262/36083026	Specs	33,8	32,8
AM-807/41345969	Specs	<i>keine Angabe</i>	236,6

Vergleich mit bisherigen Anwendungen

Die in dieser Arbeit vorgestellte Methode passt in keine der in Abbildung 1 vorgestellten Pseudorezeptorklassen, wie es in bisher publizierten Anwendungen der Fall ist (Tabelle 1). Es basiert eher auf einem hybriden Konstruktionsprinzip: Zwar bildet, wie in gitterbasierten Ansätzen, ein 3D Gitter das Fundament der automatischen Pseudorezeptorkonstruktion, jedoch werden, analog zu atombasierten Methoden, mögliche Rezeptorelemente als *Pseudoatome* modelliert. Auf der anderen Seite ist die Position der Pseudoatome, parallel zu den peptidbasierten Ansätzen, aus Kristallstrukturuntersuchungen abgeleitet.

Im Unterschied zu allen in Tabelle 1 zusammengefassten Programmen wurde hier eine rein pharmakophorbasierte Pseudorezeptorkonstruktion vorgestellt. Diese ist mit etablierten ligandenbasierten Pharmakophorscreeningmethoden verwandt (Schneider *et al.*, 1999; Fechner *et al.*, 2003; Renner & Schneider, 2004; Tanrikulu *et al.*, 2007). Jedoch wird im Gegensatz dazu kein Pharmakophor der Referenzliganden, sondern eines virtuellen Rezeptors beschrieben. Dadurch findet keine strukturelle Restriktion der untersuchten Moleküle im virtuellen Screening statt. Es wird lediglich nach Substanzen gesucht, die ebenfalls eine Interaktion zu den angezeigten Pharmakophorpunkten auf Rezeptorseite ausbilden können (ohne Pharmakophorpunktpositionen auf der Molekülstruktur vorzugeben), was im Bezug auf Scaffold-Hopping von Vorteil sein könnte.

Durch die unvollständige Umlagerung der Referenzliganden mit Pseudoatomen werden implizit offene Pseudorezeptoren unterstützt. Folglich kommt es zu keiner

Beschränkung des Bindetaschenvolumens; wodurch Liganden gefunden werden können, die nur mit einem Teil ihrer Gesamtstruktur mit den Pseudorezeptor interagieren. In Kombination mit der Kodierung generierter Pseudorezeptoren als Korrelationsvektoren verdeutlicht dies die Ausrichtung der hier vorgestellten Methode für den Einsatz im virtuellen Screening, was in bisherigen Pseudorezeptoranwendungen nicht der Fall war.

Guccione *et al.* nutzten Pseudorezeptoren zur Festlegung der Rezeptor-Liganden Schlüsselinteraktionen und zur Entwicklung von neuen Synthesestrategien für bisher unbekannte Wirkstoffe. Zur Gütebewertung der Vorhersagekraft wurde die q^2 -Statistik verwendet (Golbraikh *et al.*, 2002). Multikonformermodelle seien gerade in gitterbasierten Ansätzen hilfreich, da sie die Generierung von kontinuierlichen Konturkarten ermöglichen. Obwohl in der hier vorgestellten Arbeit keine Anwendung mit mehreren Konformationen eines Liganden demonstriert wurde, ist dies grundsätzlich denkbar: Parallel zur vorgestellten Pseudorezeptorkonstruktion mit mehreren Liganden, kann man einen Pseudorezeptor auf Basis von multiplen Konformationen eines Liganden aufbauen. Die Gewichtung der Pseudoatome würde ebenfalls stark ausgeprägte Interaktionsbereiche aufdecken, die die Schlüsselinteraktionen identifizieren und zur Generierung von Syntheseideen hinzugezogen werden könnten.

Macchiarulo *et al.* haben als eine der Ersten auf die Eignung von Pseudorezeptoren im virtuellen Screening hingewiesen (Macchiarulo *et al.*, 2004), führten aber keine Experimente wie in dieser Arbeit durch. Weiterhin beobachteten sie die Kompatibilität zwischen Pseudorezeptoren und Homologiemodellen. Einen Informationstransfer zwischen Pseudorezeptoren und Proteinmodellen, wie in dieser Arbeit, wurde jedoch ebenfalls nicht präsentiert.

Srivasta und Crippen zeigten wieder, dass sich Pseudorezeptoren zur Erklärung der bekannten Ligandenbindungen eignen (Srivastava & Crippen, 1993). Die Erklärung der Schlüsselinteraktionen von bekannten Liganden stellt eine klassische Anwendung von Pseudorezeptoren dar, die in der hier vorgestellten Arbeit durch den Vergleich von Pseudorezeptoren und Kokristallstrukturen vorgestellt wurde. Sie gingen nicht auf prospektive Fähigkeiten ein.

Lloyd *et al.* bestätigten die Eignung von Pseudorezeptoren für das *de novo* Design am Beispiel neuer nicht-steroidaler Liganden des Östrogenrezeptors (Lloyd *et al.*, 2004). Ein Pharmakophoralignment von fünf bekannten Antagonisten diente zum Aufbau einer Volumenhülle, die mit weiteren fünf Pharmakophorpunkten ausgestattet wurde. Die Platzierung von molekularen Fragmenten in die Hülle führte zu neuen molekularen Grundgerüsten. Die Generierung von völlig neuen molekularen Strukturen mit Hilfe von *de novo* Design stellt besonders im Hinblick auf patentrechtliche Fragestellungen große Vorteile dar. Zwar wurde die hier vorgestellte Methode nicht dafür konzipiert, aber sie könnte zumindest durch die Verwendung als Scoringfunktion innerhalb eines Designprozesses bei der Selektion von geeigneten Kandidaten eingesetzt werden.

Vedani *et al.* erweiterten ihre atombasierte Pseudorezeptorsoftware Quasar (Vedani & Zbinden, 1998), sodass Liganden durch ein Ensemble ihrer Konformationen repräsentiert werden konnten (Vedani *et al.*, 2000a; Vedani *et al.*, 2000b). Später kam die multiple Repräsentation von *induced-fit* Pseudorezeptoren hinzu (Vedani & Dobler, 2002a; Vedani & Dobler, 20002b; Vedani *et al.*, 2005a). Zuletzt wurde die simultane Evaluierung unterschiedlicher Solvatisierungszustände hinzugefügt, welches an der Modellierung eines Östrogenrezeptors veranschaulicht wurde (Vedani *et al.*, 2005b). Sie heben die Notwendigkeit eines Pseudorezeptors hervor, welcher zur Auffassung multipler Liganden- und Bindetaschenkonformationen befähigt ist. Der finale Pseudorezeptor sei in der Lage, die Schlüsselinteraktionspunkte der Kristallstruktur zu erklären, was mit den Ergebnissen dieser Arbeit einhergeht.

Pei *et al.* entwickelten mit einem atombasierten Ansatz ein Pseudorezeptormodell, um 50 bekannte COX-2 Inhibitoren (Liu *et al.*, 2002; Pei *et al.*, 2005). Die Liganden wurden auf das potenteste Molekül (SC-558 (Kurumbail *et al.*, 1996)) aligniert. Mit Hilfe des Modells konnten die Autoren die Aktivitäten von zehn weiteren Inhibitoren, die nicht Teil des Trainingsdatensatzes waren, korrekt vorhersagen. Die Vorhersage von Aktivitäten ist im Sinne klassischer QSAR Anwendungen. In der hier vorgestellten Pseudorezeptormethode zum virtuellen Screening ist die Vorhersage von Aktivitäten mangels Energiefunktion bisher nicht möglich.

Im Jahr 2002 entwickelten Botta *et al.* ein peptidbasiertes Pseudorezeptormodell der Cytochrom P450 abhängigen Lanosterol 14 α -Demethylase (Botta *et al.*, 2002), welches den ersten Schritt der Lanosterolkonversion durch die Entfernung einer Methylgruppe

katalysiert. Sie nutzten 39 Azolantifungizide zur Konstruktion einer Bindetasche, die mit Aminosäuren ausgestattet ist. Das Ligandenalignment wurde aus Dockingexperimenten in ein Homologiemodell gewonnen (Jansen *et al.*, 1997). Der Pseudorezeptor konnte zur Vorhersage von Aktivitäten genutzt werden und erlaubte zudem die Anwendung weiterer rezeptorbasierter Methoden. Zwar erlaubt die hier vorgestellte Methode keine Generierung einer mit Aminosäuren besetzten Bindetasche, jedoch können trotzdem rezeptorbasierte Methoden angewendet werden, wenn ein Informationstransfer in ein Proteinmodell vollzogen wurde.

Bohacek *et al.* generierten einen fragmentbasierten Pseudorezeptor zur Entwicklung von Molekülen, die die Produktion des fetalen Hämoglobins erhöhen (Bohacek *et al.*, 2006). Guanidinium- und Pyrrolfragmente wurden manuell um sechs Referenzmoleküle positioniert, um Wasserstoffbrückeninteraktionen zu modellieren. Die Auffüllung des verbleibenden Raumes mit Propanmolekülen stellte den hydrophoben Charakter der Bindetasche dar. Nach einer Relaxation durch molekulare Dynamiksimulationen des Pseudorezeptors wurde es im automatisierten Docking von 13.000 Molekülen verwendet. Die Bindung von 26 aus den besten 630 Moleküle wurde experimentell überprüft, wo vier Moleküle eine zu >200% erhöhte Aktivität im Vergleich zum Kontrollmolekül Butyrat zeigten. Bohacek *et al.* bestätigten damit die Eignung von fragmentbasierten Pseudorezeptoren als Dockingstrukturen. Aufgrund einer fehlenden Energiefunktion ist ein Docking in die hier vorgestellte Pseudorezeptorkonstruktion nicht möglich. Erst die Platzierung von molekularen Fragmenten würde die Kompatibilität zu verfügbaren Dockingprogrammen und den Einsatz der dort implementierten Funktionen zur Energiekalkulation erlauben.

Es wurde in dieser Arbeit gezeigt, dass Pseudorezeptoren ohne jegliche 3D Strukturinformation erfolgreich im virtuellen Screening eingesetzt werden können und der Transfer von ligandenbasierten Informationen zur Justierung eines rezeptorbasierten Modells eingesetzt werden kann. Ist nun die Strukturaufklärung von Rezeptoren überflüssig? Um diese Frage zu beantworten, muss man sich über die Einschränkungen von Pseudorezeptormodellen bewusst sein. Grundsätzlich repräsentieren Pseudorezeptoren kein Ebenbild der makromolekularen Bindetaschen – sie stellen nur ein Modell, gewonnen aus den Informationen bekannter Liganden, dar.

Es ist insbesondere nicht möglich, neue Wirkstoffe mit einem alternativen Bindemodus für ein Zielprotein zu finden. Es sind Fälle bekannt, in denen zwar mehrere Liganden existieren, diese aber jeweils andere Bindetaschen nutzen (Kim, 2007a; Kim, 2007b). Folglich muss die Annahme eines gemeinsamen Bindemodus aller Referenzliganden nicht erfüllt sein, weshalb bei der Anwendung von Pseudorezeptoren Vorsicht geboten ist. Ein anderer Aspekt betrifft das als Ausgang für die Modellierung dienende Ligandenalignment. In Abbildung 26 sind Überlagerungen von Ibuprofen und Flurbiprofen gezeigt. Beide Substanzen inhibieren die Cyclooxygenase-1 (COX-1), indem sie in dieselbe Bindetasche im gleichen Bindemodus binden. Die Molekülüberlagerung aus Abbildung 26b resultiert aus einem pharmakophorbasierten Alignment multipler Konformationen, wohingegen der Fall in Abbildung 26c aus einer Überlagerung der Proteinkoordinaten aus den Kokristallstrukturen resultiert (PDB Einträge 1eqg (Auflösung: 2,6 Å) und 1eqh (Auflösung: 2,7 Å) (beide aus (Selinsky *et al.*, 2001))). Der Unterschied beider Alignments ist ersichtlich – ein Pseudorezeptor aus dem Pharmakophoralignment muss nicht notwendigerweise der Situation aus Abbildung 26c entsprechen (man beachte die Orientierung der Carboxylgruppen). Fakt ist, dass bei Abwesenheit von 3D Strukturinformationen nur eine ligandenbasierte Überlagerung der Referenzmoleküle zur Verfügung steht, welches von vornherein einen nicht optimalen Startzustand liefern könnte. Aufgrund dieser Problematik wurde in dieser Arbeit die Konstruktion der Pseudorezeptor immer an energieminierten

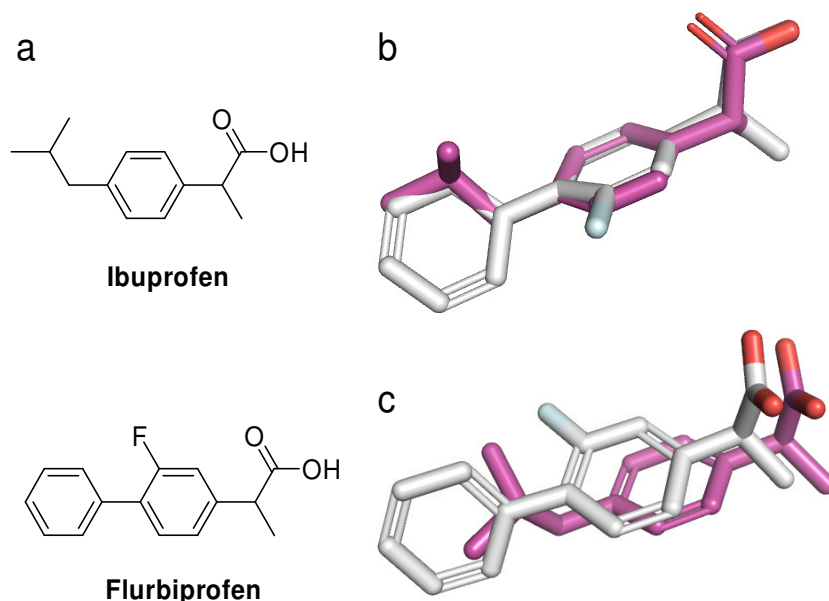


Abbildung 26. a) Strukturen von Ibuprofen und Flurbiprofen und deren b) Pharmakophoralignment und c) Alignment durch Überlagerung der Kokristallstrukturen.

Ligandenkonformationen vorgenommen. Dies folgt der Annahme, dass die bioaktive Konformation im Vergleich zur freien, energieniedrigsten Konformation des Moleküls nicht völlig unterschiedlich ist (Marshall *et al.*, 1979; Marshall, 1993; Pei *et al.*, 2005). Falls 3D Rezeptorinformationen vorhanden sind, kann die Übereinanderlagerung von möglichen Bindepasen aus Dockingstudien ebenfalls als Basis einer Pseudorezeptorkonstruktion dienen.

Schlusswort und Ausblick

Um die Arbeitshypothese der Eignung von Pseudorezeptoren im virtuellen Screening, zu überprüfen, wurde in dieser Arbeit ein Programm zur automatisierten Pseudorezeptorkonstruktion aus einer Molekülüberlagerung entwickelt. Dazu wurden zunächst die beobachteten Distanz- und Winkelparameter von Wasserstoffbrückeninteraktionen aus einer Kristallstrukturanalyse entnommen, um Rezeptoratompositionen zu modellieren. Zur Anwendung im virtuellen Screening wurden die Pseudoatomdistanzen in einem Korrelationsvektor kodiert, die auch berechnete Gewichte beachten kann.

Es wurde gezeigt, dass die konstruierten Pseudorezeptoren mit kokristallisierten Bindetaschen kompatibel sind und sich zum retrospektiven virtuellen Screening eignen. Weiterhin konnte eine Justierung der H4R Bindetasche aus einer molekularen Dynamiksimulation vorgeschlagen werden. Im prospektiven virtuellen Screening konnte gezeigt werden, dass Pseudorezeptoren sich eignen, um neue H4R Liganden zu finden, die ein neues molekulares Grundgerüst besitzen.

Diese Erkenntnisse regen folgende, zukünftige Bearbeitungsmöglichkeiten an:

- Der in dieser Arbeit entwickelte Pseudorezeptorkonstruktor bietet die Typisierung von Pseudoatomen mit Wasserstoffbrückenakzeptor, -donor und aromatischen Pharmakophoreigenschaften. Je nach Zielprotein sind aber auch Eigenschaften, die die Ladung einzelner Ligandenfragmente betreffen, für ein erfolgreiches virtuelles Screening von Bedeutung. Ein Beispiel stellen die Serinproteasen dar, die ein positiv geladenes Fragment in der P1-Tasche benötigen (Wechselwirkung zu einem Aspartylrest des Enzyms) (Böhm *et al.*, 1999). Von daher ist in erster Linie die Erweiterung der Pseudoatomtypisierungen mit elektrostatischen Eigenschaften ein aussichtsreiches Projekt.
- Eine automatische Pseudorezeptorkonstruktion endet mit der Projektion von einzelnen Pseudoatomen um das Ligandenalignment, und stellt ein virtuelles Pharmakophormodell des Rezeptors dar. Dieses ist nicht direkt in rezeptorbasierten Methoden nutzbar. Wenn der Pharmakophor-abgeleitete

Pseudorezeptor mit Aminosäuren aus einer Rotamerbibliothek an den Pseudoatompositionen ausgestattet werden könnte, würde dies die Kompatibilität zu heutigen Dockingprogrammen schaffen. Somit könnten Aktivitätsvorhersagen durch die in Dockingprogrammen integrierte Scoringfunktionen abgeleitet werden.

- Eine weitere Möglichkeit zur Vorhersage von Ligandenaktivitäten ist die Integration von Energiefunktionen, die entweder mit Pseudoatomen oder mit Aminosäuren an Pseudoatompositionen arbeiten können. Damit wären Energieminimierungen von möglichen Liganden innerhalb des Pseudorezeptors möglich. Dies würde die Berechnung von Aktivitäten basierend auf Energiefunktionen implizieren.

Literaturverzeichnis

A

- Alber JS *et al.* An integrated approach to fragment-based lead generation: philosophy, strategy and case studies from AstraZeneca's drug discovery programmes. *Curr Top Med Chem* **2007**; 7: 1600-1629.
- Alber JS, Blomberg N, Breeze AL, Brown AJH, Burrows JN, Edwards PD, Folmer RHA, Geschwindner S, Griffen EJ, Kenny PW . An integrated approach to fragment-based lead generation: philosophy, strategy and case studies from AstraZeneca's drug discovery programmes. *Curr Top Med Chem* **2007**; 7: 1600-1629.
- Abraham WM, Sielczak MW, Ahmed A, Cortes A, Lauredo IT, Kim J, Pepinsky B, Benjamin CD, Leone DR, Lobb RR. Alpha 4-integrins mediate antigen-induced late bronchial responses and prolonged airway hyperresponsiveness in sheep. *J Clin Invest* **1994**; 93: 776-787.
- Akdis CA, Blaser K. Histamine in the immune regulation of allergic inflammation. *J Allergy Clin Immunol* **2003**; 112: 15-22.
- Akdis CA, Simons FER. Histamine receptors are hot in immunopharmacology. *Eur J Pharmacol* **2006**; 533: 69-76.
- Andrews PR, Quint G, Winkler DA, Richardson D, Sadek M, Spurling TH. Morpheus: a conformation-activity relationships and receptor modeling package. *J Mol Graph* **1989**; 7: 138-145.

B

- Ballesteros JA, Weinstein H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G-protein coupled receptors. *Meth Neurosci* **1995**; 25: 366-428.
- Baroni M, Costantino G, Cruciani G, Riganelli D, Valigi R, Clementi S. Generating Optimal Linear PLS Estimations (GOLPE): An advanced chemometric tool for handling 3D-QSAR problems. *Quant Struct Act Relat* **1993**; 12: 9-20.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucl Acids Res* **2000**; 28: 235-242.
- Bernstein FC, Koetzle TF, Williams GJB, Meyer Jr EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol* **1977**; 112: 535-542.
- Bertolini A, Ottani A, Sandrini M. Dual acting anti-inflammatory drugs: a reappraisal *Pharmacol Res* **2001**; 44.

- Bisson WH, Cheltsov AV, Bruey-Sedano N, Lin B, Chen J, Goldberger N, May LT, Christopoulos A, Dalton JT, Sexton PM, Zhang XK, R. Abagyan R. Discovery of antiandrogen activity of nonsteroidal scaffolds of marketed drugs. *Proc Natl Acad Sci USA* **2007**; 104: 11927-11932.
- Bleicher KH, Böhm HJ, Müller K, Alanine AI. Hit and lead generation: beyond high-throughput screening. *Nat Rev Drug Discov* **2003**; 2: 369-378.
- Blier P, Montigny C. Current advances and trends in the treatment of depression. *Trends Pharmacol Sci* **1994**; 15: 220-226.
- Bohacek R, Boosalis MS, McMartin C, Faller DV, Perrine SP. Identification of Novel Small-molecule Inducers of Fetal Hemoglobin using Pharmacophore and 'PSEUDO' Receptor Models. *Chem Biol Drug Des* **2006**; 67: 318-328
- Botta M, Corelli F, Manetti F, Tafi A. Molecular modeling as a powerful technique for understanding small-large molecules interactions. *Farmaco* **2002**; 57: 153-165.
- Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recog* **1997**; 30: 1145-1159.
- Bradley M, Richardson W, Crippen GM. Deducing molecular similarity using Voronoi binding sites. *J Chem Inf Comput Sci* **1993**; 33: 750-755.
- Broekkamp CL, Leysen D, Peeters BW, Pinder RM. Prospects for improved antidepressants. *J Med Chem* **1995**; 38: 4615-4633.
- Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J Comp Chem* **1983**; 4: 187-217.
- Brooks PM, Day RO. Nonsteroidal anti-inflammatory drugs: differences and similarities. *New Engl J Med* **1991**; 324: 1716-25.
- Brown N, Jacoby E. On scaffolds and hopping in medicinal chemistry. *Mini Rev Med Chem* **2006**; 6:1217-1229.
- Burkly LC, Jakubowski A, Hattori M. Protection against adoptive transfer of autoimmune diabetes mediated through very late antigen-4 integrin. *Diabetes* **1994**; 43: 529-534.

C

- Carley DW. Drug repurposing: identify, develop and commercialize new uses for existing or abandoned drugs. Part I. *Drugs* **2005**; 8: 306-309.
- Carley DW. Drug repurposing: identify, develop and commercialize new uses for existing or abandoned drugs. Part II. *Drugs* **2005**; 8: 310-313.
- Carr RA, Congreve M, Murray CW, Rees DC. Fragment-based lead discovery: leads by design. *Drug Discov Today* **2005**; 10: 987-992.
- Carrol FL, Lewin AH, Boja JW, Kuhar MJ. Cocaine receptor: Biochemical characterization and structure-activity relationships of cocaine analogues at the dopamine transporter. *J Med Chem* **1992**; 35: 969-981.

- Catella-Lawson F *et al.* Effects of specific inhibition of cyclooxygenase-2 on sodium balance, hemodynamics, and vasoactive eicosanoids. *J Pharmacol Exp Ther* **1999**; 289: 735-741.
- Chae CH, Yoo SE, Shin W. Novel receptor surface approach for 3D-QSAR: The weighted probe interaction energy method. *J Chem Inf Comput Sci* **2004**; 44: 1774-1787.
- Chandrasekharan NV, Dai H, Roos KL, Evanson NK, Tomsik J, Elton TS, Simmons DL. COX-3, a cyclooxygenase-1 variant inhibited by acetaminophen and other analgesic/antipyretic drugs: cloning, structure, and expression. *Proc Natl Acad Sci USA* **2002**; 99: 13926-13931.
- Chen H, Zhou J, Xie G. PARM: A genetic evolved algorithm to predict bioactivity. *J Chem Inf Comput Sci* **1998**; 38: 243-250.
- Chen W, Gilson MK. ConCept: De novo design of synthetic receptors for targeted ligands. *J Chem Inf Model* **2007**; 47: 425-434.
- Congreve M, Murray CW, Blundell TL. Structural biology and drug discovery. *Drug Discov Today* **2005**; 10: 895-907.
- Costanzi S. On the applicability of GPCR homology models to computer-aided drug discovery: A comparison between in silico and crystal structures of the beta2-adrenergic receptor. *J Med Chem* **2008**; in press.
- Cramer RD, Patterson DE, Bunce JD. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J Am Chem Soc* **1988**; 110: 5959-5967.
- Crippen GM. Voronoi binding site models. *NIDA Res Monogr* **1991**; 112: 7-20.
- Crippen GM. Validation of EGSITE2, a mixed integer program for deducing objective site models from experimental data. *J Med Chem* **1997**; 40: 3161-3172.

D

- Darden T, York D, Pedersen L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J Chem Phys* **1993**; 98: 10089-10098.
- Davis AM, Gensmantel NP, Johansson E, Marriott DP. The use of the GRID program in the 3-D QSAR analysis of a series of calcium-channel agonists. *J Med Chem* **1994**; 37: 963-672.
- De Esch IJP, Thurmond RL, Jongejan A, Leurs R. The histamine H4 receptor as a new therapeutic target for inflammation. *Trends Pharmacol Sci* **2005**; 26: 462-469.
- Delibasis KS, Matsopoulos GK, Mouravliansky NA, Nikita KS. A novel and efficient implementation of the marching cubes algorithm. *Comput Med Imaging Graph* **2001**; 25: 343-52.
- Desvergne B, Wahli W. Peroxisome proliferator-activated receptors: nuclear control of metabolism. *Endocr Rev* **1999**; 20: 649-688.
- Doweyko AM. The hypothetical active site lattice. An approach to modeling active sites from data on inhibitor molecules. *J Med Chem* **1988**; 31: 1396-1406.

Duval C, Müller M, Kersten S. PPARalpha and dyslipidemia. *Biochim Biophys Acta* **2007**; 1771: 961-971.

E

Eswar N, Marti-Renom MA, Webb B, Madhusudhan MS, Eramian D, Shen M, Pieper U, Šali A. Comparative protein structure modeling with MODELLER, *Current Protocols in Bioinformatics*, Supplement **2000**; 15: 5.6.1-5.6.30.

Evers A, Hessler G, Matter H, Klabunde T. Virtual screening of biogenic amine-binding G-protein coupled receptors: comparative evaluation of protein- and ligand-based virtual screening protocols. *J Med Chem* **2005**; 48: 5448-5465.

F

Fawcett T. An introduction to ROC analysis. *Pattern Recog Lett* **2006**; 27: 861–874.

Fechner U, Franke L, Renner S, Schneider G. Comparison of correlation vector methods for ligand-based similarity searching. *J Comput Aided Mol Design* **2003**; 17: 687-698.

Fischer E. Einfluss der Configuration auf die Wirkung der Enzyme. *Ber Dtsch Chem Ges* **1894**; 27: 2985-2993.

G

Gantner F, Sakai K, Tusche MW, Cruikshank WW, Center DM, Bacon KB. Histamine H4 and H2 receptors control histamine-induced interleukin-16 release from human CD8 T cells. *J Pharmacol Exp Ther* **2002**; 303: 300-307.

Gelman L, Feige JN, Desvergne B. Molecular basis of selective PPARgamma modulation for the treatment of Type 2 diabetes. *Biochim Biophys Acta* **2007**; 1771: 1094-107.

Givehchi A, Schneider G. Multi-space classification for predicting GPCR-ligands. *Mol Divers* **2005**; 9: 371-83.

Golbraikh A, Tropsha A. Beware of q2! *J Mol Graph Model* **2002**; 20: 269-276.

Granseth E, Seppälä S, Rapp M, Daley DO, Von Heijne G. Membrane protein structural biology - how far can the bugs take us? *Mol Membr Biol* **2007**; 24: 329-332.

Guccione S, Doveyko AM, Chen H, Barretta GU, Balzano F. 3D-QSAR using 'multiconformer' alignment: The use of HASL in the analysis of 5-HT1A thienopyrimidinone ligands. *J Comput Aided Mol Des* **2000**; 14: 647-657.

Güner O. Pharmacophore Preception, Development and Use in Drug Design. LaJolla, International University Line **2000**.

H

Hahn M, Rogers D. Receptor Surface Models. 2. Application to quantitative structure-activity relationship studies. *J Med Chem* **1995**; 38: 2091-2102.

- Hahn M, Rogers D. Receptor Surface Models. In: Kubinyi H, Folkers G, Martin YC Eds, 3D QSAR in Drug Design: Recent Advances. Kluwer, Escom. **1998**; 117-133.
- Hahn M. Receptor surface models. 1. definition and construction. *J Med Chem* **1995**; 38: 2080-2090.
- Hanley JA, McNeil BJ. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology* **1982**; 143: 29-36.
- Hawkey CJ. COX-2 inhibitors. *Lancet* **1999**; 353: 307-14.
- Hay BP, Firman TK. Hostdesigner: A Program for the de novo structure-based design of molecular receptors with binding sites that complement metal ion guests. *Inorg Chem* **2002**; 41: 5502-5512.
- Heiden W, Schlenkrich M, Brickmann J. Triangulation algorithms for the representation of molecular surface properties. *J Comput Aided Mol Des* **1990**; 4: 225-269.
- Hla T, Neilson K. Human cyclooxygenase-2 cDNA. *Proc Natl Acad Sci USA* **1992**; 89: 7384-7388.
- Hofstra CL, Desai PJ, Thurmond RL, Fung-Leung W. Histamine H4 Receptor Mediates Chemotaxis and Calcium Mobilization of Mast Cells. *J Pharmacol Exp Ther* **2003**; 305: 1212-1221.
- Höltje HD, Anzali S. Molecular modelling studies on the digitalis binding site of the Na⁺/K⁽⁺⁾-ATPase. *Pharmazie* **1992**; 47: 691-697.
- Horn F, Bettler E, Oliveira L, Campagne F, Cohen FE, Vriend G. GPCRDB information system for G protein-coupled receptors. *Nucl Acids Res* **2003**; 31:294-297.
- Hynes RO. Integrins: bidirectional, allosteric signaling machines. *Cell* **2002**; 110: 673-687.

J

- Jablonowski JA, Grice CA, Chai W, Dvorak CA, Venable JD, Kwok AK, Ly KS, Wei J, Baker SM, Desai PJ, Jiang W, Wilson SJ, Thurmond RL, Karlsson L, Edwards JP, Lovenberg TW, Carruthers NI. The First Potent and Selective Non-Imidazole Human Histamine H4 Receptor Antagonists. *J Med Chem* **2003**; 46: 3957-3960.
- Jacoby E, Bouhelal R, Gerspacher M, Seuwen K. The 7TM G-Protein-Coupled Receptor Target Family. *ChemMedChem* **2006**; 1: 760-782.
- Jansen JM, Koehler KF, Hedberg MH, Johansson AM, Hacksell U, Nordvall G, Snyder JP. Molecular design using the minireceptor concept. *J Chem Inf Comput Sci* **1997**; 37: 812-818.
- Johart A, Kiss R, Viskolcz B, Keseru G. Activation mechanism of the human histamine H4 receptor - an explicit membrane molecular dynamics simulation study. *J Chem Inf Model* **2008**; 48: 1199-1210.

Jongejan A, Lim HD, Smits RA, de Esch IJP, Haaksma E, Leurs R. Delineation of Agonist Binding to the human histamine H4 receptor using mutational analysis, homology modeling, and ab initio calculations. *J Chem Inf Model* **2008**; 48: 1455-1463.

K

Kato Y, Itai A, Iitaka Y. A novel method for superimposing molecules and receptor mapping. *Tetrahedron* **1987**; 43: 5229-5236.

Kato Y, Inoue A, Yamada M, Tomioka N, Itai A. Automatic superposition of drug molecules based on their common receptor site. *J Comput Aided Mol Des* **1992**; 6: 475-486.

Kim KH. Outliers in SAR and QSAR: is unusual binding mode a possible source of outliers? *J Comput Aided Mol Des* **2007**; 21: 63-86.

Kim KH. Outliers in SAR and QSAR: 2. Is a flexible binding site a possible source of outliers? *J Comput Aided Mol Des* **2007**; 21: 421-435.

Kiss R, Noszál B, Rácz A, Falus A, Erős D, Keserű GM. Binding mode analysis and enrichment studies on homology models of the human histamine H4 receptor. *Eur J Med Chem* **2008**; 43:1059-1070.

Klebe G, Abraham U, Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological potency. *J Med Chem* **1994**; 37: 4130-4146.

Klebe G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov Today* **2006**; 11: 580-594.

Koehn FE. High impact technologies for natural products screening. *Prog Drug Res* **2008**; 65: 175, 177-210.

Koshland DE. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci USA*. **1958**; 44: 98-104.

Kubinyi H, Hamprecht FA, Mietzner T. Three-dimensional quantitative similarity-activity relationships (3D QsiAR) from SEAL similarity matrices. *J Med Chem* **1998**; 41: 2553-2564.

Kurumbail RG, Stevens AM, Gierse JK, McDonald JJ, Stegeman RA, Pak JY, Gildehaus D, Miyashiro JM, Penning TD, Seibert K, Isakson PC, Stallings WC. Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents. *Nature* **1996**; 384: 644-648.

L

Lacapère JJ, Pebay-Peyroula E, Neumann JM, Etchebest C. Determining membrane protein structures: still a challenge! *Trends Biochem Sci* **2007**; 32: 259-270.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* **1993**; 26: 283-291.

- Leurs R, Bakker RA, Timmerman H, de Esch IJP. The histamine H3 receptor: from gene cloning to H3 receptor drugs. *Nature Reviews* **2005**; 4: 107-122.
- Lill MA, Vedani A, Dobler M. Raptor: Combining dual-shell representation, induced-fit simulation, and hydrophobicity scoring in receptor modeling: Application toward the simulation of structurally diverse ligand sets. *J Med Chem* **2004**; 47: 6174-86.
- Lill MA. Multi-dimensional QSAR in drug discovery. *Drug Discov Today* **2007**; 12: 1013-1017.
- Lim HD, Smits RA, Leurs R, de Esch IJP. The emerging role of the histamine H4 receptor in anti-inflammatory therapy. *Curr Top Med Chem* **2006**; 6: 1365-1373.
- Liu H, Huang X, Shen J, Luo X, Bingxlong M. Inhibitory mode of 1,5-diarylpyrazole derivatives against cyclooxygenase-2 cyclooxygenase-1: Molecular docking and 3D-QSAR analyses. *J Med Chem* **2002**; 45: 4816-4827.
- Lloyd DG, Buenemann CL, Todorov NP, Manallack DT, Dean PM. Scaffold hopping in de novo design. Ligand generation in the absence of receptor information. *J Med Chem* **2004**; 47: 493-496.

M

- Macchiarulo A, Costantino G, Meniconi M, Pleban K, Ecker G, Bellocchi D, Pellicciari R. Insights into phenylalanine derivatives recognition of VLA-4 integrin: From a pharmacophoric study to 3D-QSAR and molecular docking analyses. *J Chem Inf Comput Sci* **2004**; 44: 1829-1839.
- Markt P, Schuster D, Kirchmair J, Laggner C, Langer T. Pharmacophore modeling and parallel screening for PPAR ligands. *J Comput Aided Mol Des* **2007**; 21: 575-590.
- Marnett LJ, Kalgutkar AS. Cyclooxygenase-2 inhibitors: discovery and future. *TIPS* **1999**; 20: 465-469.
- Marshall GR, Barry CD, Bosshard HE, Dammkoehler RA, Dunn DA. The conformational parameter in drug design: the active analog approach. In: Olson EC, Christoferson RE Eds, *Computer-Assisted Drug Design*. Washington, American Chemical Society Symposium. **1979**; 112: 205-226.
- Marshall GR. Modeling of unknown receptors. In: Kubinyi H Ed, *3D QSAR in Drug Design: Theory Methods and Applications*. Leiden, Escom. **1993**; 80-116.
- McAdam BF *et al.* Systemic biosynthesis of prostacyclin by cyclooxygenase (COX)-2: the human pharmacology of a selective inhibitor of COX-2. *Proc Natl Acad Sci USA* **1999**; 96: 272-277.
- McGaughey GB, Gagné M, Rappé AR. π -stacking interactions: alive and well in proteins. *J Biol Chem* **1998**; 273: 15458-15463.
- McMartin C, Bohacek RJ. QXP: Powerful, rapid computer algorithms for structure-based design. *J Comput Aided Mol Des* **1997**; 11: 333-344.

- Medina V, Croci M, Crescenti E, Mohamad N, Sanchez-Jiménez F, Massari N, Nuñez M, Cricco G, Martin G, Bergoc R, Rivera E. The role of histamine in human mammary carcinogenesis: H3 and H4 receptors as potential therapeutic targets for breast cancer treatment. *Cancer Biol Ther* **2008**; 7: 28-35.
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. *J Chem Phys* **1953**; 21: 1087–1092.
- Mills JEJ, Dean PM. Three-dimensional hydrogen-bond geometry and probability information from a crystal survey. *J Comput Aided Mol Design* **1996**; 10: 607-622.
- Mohan V, Gibbs AC, Cummings MD, Jaeger EP, DesJarlais RL. Docking: Successes and challenges. *Curr Pharm Des* **2005**; 11: 323-333.
- Momany F, Pitha R, Klimkovsky VJ, Venkatachalam CM. Drug design using a protein pseudoreceptor. In: Hohne BA, Pierce TH Eds, *Expert Systems and Applications in Chemistry*. ACS Symp Ser **1989**; 408: 82-91.
- Moreau G, Broto P. Autocorrelation of Molecular Structures. *Nouveau Journal de Chimie* **1980**; 4: 757-759.
- Morita I. Distinct functions of COX-1 and COX-2. *Prostaglandins Other Lipid Mediat* **2002**; 68-69: 165–175.
- Nakamura T, Itadani H, Hidaka Y, Ohta M, Tanaka K. Molecular cloning and characterization of a new human histamine receptor, HH4R. *Biochem Biophys Res Commun* **2000**; 279: 615–620.

N

- Neubig RR, Spedding M, Kenakin T, Christopoulos A. International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII Update on Terms and Symbols in Quantitative Pharmacology. *Pharmacol Reviews* **2003**; 55: 597-606.

- Norel R, Fischer D, Wolfson HJ, Nussinov R. Molecular surface recognition by a computer vision technique. *Protein Eng* **1994**; 7: 39-46.

O

- Oda T, Morikawa N, Saito Y, Masuho Y, Matsumoto S. Molecular cloning and characterization of a novel type of histamine receptor preferentially expressed in leukocytes. *J Biol Chem* **2000**; 275: 36781-36786.

P

- Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**; 289: 739-745.

- Pei J, Zhou J. Flexible atom receptor model. *Acta Chim Sin* **2002**; 60: 973-979.

- Pei J, Chen H, Liu Z, Han X, Wang Q, Shen B, Zhou J, Lai L. Improving the Quality of 3D-QSAR by Using Flexible-Ligand Receptor Models. *J Chem Inf Model* **2005**; 45: 1920-1933.
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. *J Comput Chem* **2005**; 26: 1781-1802.
- Pierce KL, Premont RT, Lefkowitz RJ. Seven-Transmembrane Receptors. *Nature Reviews Molecular Cell Biology* **2002**; 3: 639-650.
- Ponder JW, Richards FM. Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J Mol Biol* **1987**; 193: 775-791.
- Proschak E, Wegner JK, Schüller A, Schneider G, Fechner U. Molecular query language (MQL)—a context-free grammar for substructure matching. *J Chem Inf Model* **2007**; 13: 295-301.
- Proschak E, Rupp M, Derksen S, Schneider G. Shapelets: possibilities and limitations of shape-based virtual screening. *J Comput Chem* **2008**; 29: 108-114.

R

- Rasmussen SG, Choi H, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC *et al.* Crystal structure of the human β_2 adrenergic G-protein-coupled receptor. *Nature* **2007**; 450: 383-388.
- Rau O, Zettl H, Popescu L, Steinhilber D, Schubert-Zsilavecz M. The treatment of dyslipidemia--what's left in the pipeline? *ChemMedChem* **2008**; 3: 206-221.
- Renner S, Schneider G. Fuzzy Pharmacophore Models from Molecular Alignments for Correlation-Vector-Based Virtual Screening. *J Med Chem* **2004**; 47: 4653-4664.
- Rester U. Dock around the clock - Current status of small molecule docking and scoring. *QSAR Comb Sci* **2006**; 25:605-615.
- Rocca B, Secchiero P, Ciabattini G, Ranelletti FO, Catani L, Guidotti L, Melloni E, Maggiano M, Zauli G, Patrono C. Cyclooxygenase-2 expression is induced during human megakaryopoiesis and characterizes newly formed platelets. *Proc Natl Acad Sci USA* **2001**; 99: 7634-7639.
- Romano M, Claria J. Cyclooxygenase-2 and 5-lipoxygenase converging functions on cell proliferation and tumor angiogenesis: implications for cancer therapy. *FASEB J* **2003**; 17: 1986-95.
- Rowlinson SW, Kiefer JR, Prusakiewicz JJ, Pawlitz JL, Kozak KR, Kalgutkar AS, Stallings WC, Kurumbail RG, Marnett LJ. A novel mechanism of cyclooxygenase-2 inhibition involving interactions with Ser-530 and Tyr-385. *J Biol Chem* **2003**; 278: 45763-45769.

S

- Sagara H, Matsuda H, Wada N, Yagita H, Fukuda T, Okumura K, Makino S, Ra C. A monoclonal antibody against very late activation antigen-4 inhibits eosinophil accumulation and late asthmatic response in guinea pig model of asthma. *Int Arch Allergy Immunol* **1997**; 112: 287-294.
- Šali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **1993**; 234: 779-815.
- Scapin G. Structural biology and drug discovery. *Curr Pharm Des* **2006**; 12: 2087-2097.
- Schneider G, Baringhaus KH. *Molecular Design – Concepts and Applications*. Weinheim, Wiley-VCH **2008**.
- Schneider P, Schneider G. Collection of bioactive compounds for focused library design. *QSAR Comb Sci* **2003**; 22: 713-718.
- Schneider G, Neidhart W, Giller T, Schmid G. “Scaffold-Hopping” by topological pharmacophore search: A contribution to virtual screening. *Angew Chemie Int Ed* **1999**; 38: 2894-2896.
- Schneider G, Schneider P, Renner S. Scaffold-hopping: How far can you jump? *QSAR Comb Sci* **2006**; 25: 1162-1171.
- Seifert R, Wenzel-Seifert K. Constitutive activity of G-protein coupled receptors: cause of diseases and common property of wild-type receptors. *Naunyn Schmiedebergs Arch Pharmacol* **2002**; 366: 381-416.
- Selinsky BS, Gupta K, Sharkey CT, Loll PJ. Structural analysis of NSAID binding by prostaglandin H2 synthase: time-dependent and time-independent inhibitors elicit identical enzyme conformations. *Biochemistry* 2001; 40: 5172-5180.
- Sheridan M, Underwood D, Kearsley S. Chemical similarity using geometric atom pair descriptors. *J Chem Inf Comput Sci* **1996**; 36: 128-136.
- Shin N, Coates E, Murgolo NJ, Morse MJ, Bayne M, Strader CD, Monsma FJ. Molecular modeling and site-specific mutagenesis of the histamine-binding site of the histamine H4 receptor. *Mol Pharmacol* **2002**; 62: 38-47.
- Sierra ML, Beneton V, Boullay AB, Boyer T, Brewster AG, Donche F, Forest MC, Fouchet MH, Gellibert FJ, Grillot DA, Lambert MH, Laroze A, Le Grumelec C, Linget JM, Montana VG, Nguyen VL, Nicodème E, Patel V, Penfornis A, Pineau O, Pohin D, Potvain F, Poulain G, Ruault CB, Saunders M, Toum J, Xu HE, Xu RX, Pianetti PM. Substituted 2-[(4-aminomethyl)phenoxy]-2-methylpropionic acid PPARalpha agonists. 1. Discovery of a novel series of potent HDLc raising agents. *J Med Chem* **2007**; 50: 685-695.
- Simons FE. Advances in H1-antihistamines. *N Engl J Med* **2004**; 351: 2203-2217.
- Singh JP, Kauffman R, Bensch W, Wang G, McClelland P, Bean J, Montrose C, Mantlo N, Wagle A. Identification of a novel selective peroxisome proliferator-activated receptor alpha agonist, 2-methyl-2-(4-{3-[1-(4-methylbenzyl)-5-oxo-4,5-dihydro-1H-1,2,4-triazol-3-yl]propyl}phenoxy)propanoic acid (LY518674), that produces marked changes in serum lipids and apolipoprotein A-1 expression. *Mol Pharmacol* **2005**; 68: 763-768.

- Snyder JP, Rao SN, Koehler KF and Vedani A. Minireceptors and pseudoreceptors. In: Kubinyi H Ed, 3D QSAR in Drug Design: Theory, Methods and Applications. Leiden, Escom. **1993**; 336-354.
- Srivastava S, Crippen GM. Analysis of cocaine receptor site ligand Binding by three-dimensional Voronoi site modeling approach. *J Med Chem* **1993**; 36: 3572-3579.
- Steinbeck C, Han Y, Kuhn S, Horlacher O, Luttmann E, Willighagen E. The Chemistry Development Kit (CDK): an open-source Java library for Chemo- and Bioinformatics. *J Chem Inf Comput Sci* **2003**; 43: 493-500.
- Steinbeck C, Hoppe C, Kuhn S, Floris M, Guha R, Willighagen EL. Recent developments of the chemistry development kit (CDK) - an open-source java library for chemo- and bioinformatics. *Curr Pharm Des* **2006**; 12: 2111-20.
- Steinhilber D, Schubert-Zsilavec M, Roth HJ. *Medizine Chemie - Targets und Wirkstoffe*. Stuttgart, Deutscher Apotheker Verlag **2005**.
- Széles L, Töröcsik D, Nagy L. PPARgamma in immunity and inflammation: cell types and diseases. *Biochim Biophys Acta* **2007**; 1771: 1014-1030.

T

- Tanaka A, Hase S, Miyazawa T, Ohno R, Takeuchi K. Role of cyclooxygenase (COX)-1 and COX-2 inhibition in nonsteroidal anti-inflammatory drug-induced intestinal damage in rats: relation to various pathogenic events. *J Pharmacol Exp Ther* **2002**; 303: 1248-1254.
- Tanrikulu Y, Schneider G. Pseudoreceptor models in drug design: bridging ligand- and receptor-based virtual screening. *Nature Rev Drug Discov* **2008**; 7: 667-677.
- Tanrikulu Y, Nietert M, Scheffer U, Proschak E, Grabowski K, Schneider P, Weidlich M, Karas M, Göbel M, Schneider G. Scaffold hopping by "fuzzy" pharmacophores and its application to RNA targets. *ChemBioChem* **2007**; 8: 1932-1936.
- Thurmond RL, Desai PJ, Dunford PJ, Fung-Leung WP, Hofstra CL, Jiang W, Nguyen S, Riley JP, Sun S, Williams KN, Edwards JP, Karlsson L. A potent and selective histamine H4 receptor antagonist with anti-inflammatory properties. *J Pharmacol Exp Ther* **2004**; 309: 404-413.
- Todorov NP, Dean PM. Evaluation of a method for controlling molecular scaffolds diversity in de novo ligand design. *J Comput Aided Mol Des* **1997**; 11: 175-192.
- Todorov NP, Dean PM. A branch-and-bound method for optimal atom-type assignment in de novo ligand design. *J Comput Aided Mol Des* **1998**; 12: 335-350.
- Tropsha A, Golbraikh A. Predictive QSAR modeling workflow, model applicability domains, and virtual screening. *Curr Pharm Des* **2007**; 13: 3494-3504.
- Truchon JF, Bayly CI. Evaluating Virtual Screening Methods: Good and Bad Metrics for the "Early Recognition" Problem. *J Chem Inf Model* **2007**; 47: 488-508.

V

- Valle JD, Gantz I. Novel insights into histamine H₂ receptor biology. *Am J Physiol* **1997**; 273: 987-996.
- Van Ryn J, Pairet M. Clinical experience with cyclooxygenase-2 inhibitors. *Inflamm Res* **1999**; 48: 247-254.
- Vane JR, Botting R. Mechanism of action of nonsteroidal antiinflammatory drugs. *Am J Med* **1998**; 104: 2-8.
- Vedani A, Dobler M. Multidimensional QSAR: Moving from three- to five dimensional concepts. *Quant Struct Act Relat* **2002**; 21: 382-390.
- Vedani A, Dobler M. 5D-QSAR: The key for simulating induced fit? *J Med Chem* **2002**; 45: 2139-2149.
- Vedani A, Zbinden P. Quasi-atomistic receptor modeling. A bridge between 3D QSAR and receptor fitting. *Pharm Acta Helv* **1998**; 73: 11-18.
- Vedani A, Zbinden P, Snyder JP. Pseudo-receptor modeling: A new concept for the three-dimensional construction of receptor binding sites. *J Recept Res* **1993**; 13: 163-177.
- Vedani A, Zbinden P, Snyder JP, Greenidge PA. Pseudoreceptor modeling: The construction of three-dimensional receptor surrogates. *J Am Chem Soc* **1995**; 117: 4987-4994.
- Vedani A, Briem H, Dobler M, Dollinger K, McMasters DR. Multiple conformation and protonationstate representation in 4D-QSAR: The neurokinin-1 receptor system. *J Med Chem* **2000**; 43: 4416-4427.
- Vedani A, McMasters DR, Dobler M. Multiconformational ligand representation in 4D-QSAR: Reducing the bias associated with ligand alignment. *Quant Struct-Act Relat* **2000**; 19: 149-161.
- Vedani A, Dobler M, Dollinger H, Hasselbach KM, Birke F, Lill MA. Novel ligands for the chemokine receptor-3 (CCR3): a receptor-modeling study based on 5D-QSAR. *J Med Chem* **2005**; 48: 1515-27.
- Vedani A, Dobler M, Lill MA. Combining protein modeling and 6D-QSAR. Simulating the binding of structurally diverse ligands to the estrogen receptor. *J Med Chem* **2005**; 48: 3700-3703.
- Voronoi G. Recherches sur les paralléloèdres Primitives. *J reine angew Math* **1908**; 134: 198-287.
- Vriend G, Sander C. Quality control of protein models: Directional atomic contact analysis. *J Appl Cryst* **1993**; 26: 47-60.

W

- Walters DE, Hinds RM. Genetically evolved receptor models: A computational approach to construction of receptor models. *J Med Chem* **1994**; 37: 2527-2536.
- Wermuth C, Ganellin CR, Lindberg P, Mitscher L. Glossary of terms used in medicinal chemistry (IUPAC recommendations). *Pure Appl Chem* **1998**; 70: 1129-1143.

- Wolfson HJ, Rigoutsos I. Geometric hashing: An overview. *IEEE Computational Science and Eng* **1997**; 11: 263-278.
- Wong E, Bayly C, Waterman HL, Riendeau D, Mancini JA. Conversion of prostaglandin G/H synthase-1 into an enzyme sensitive to PGHS-2 selective inhibitor by a double His513 to Arg and Ile523 to Val mutation. *J Biol Chem* **1997**; 272: 9280-9286.
- Wyvill G, Mc Pheeters C, Wyvill B. Data structures for soft objects. *The Visual Computer* **1986**; 2: 227-234.

X

- Xie W, Chipman JG, Robertson DL, Erikson RL, Simmons DL. Expression of a mitogen-responsive gene encoding prostaglandin synthase is regulated by mRNA splicing. *Proc Natl Acad Sci USA* **1991**; 88: 2692-2696.
- Xu H, Agrafiotis DK. Retrospect and prospect of virtual screening in drug discovery. *Curr Top Med Chem* **2002**; 2: 1305-1308.
- Xu HE, Lambert MH, Montana VG, Plunket KD, Moore LB, Collins JL, Oplinger JA, Kliewer SA, Gampe Jr RT, McKee DD, Moore JT, Willson TM. Structural determinants of ligand binding selectivity between the peroxisome proliferator-activated receptors. *Proc Natl Acad Sci USA* 2001; 98: 13919-13924.
- Yang XD, Karin N, Tisch R, Steinman L, McDevitt HO. Inhibition of insulinitis and prevention of diabetes in nonobese diabetic mice by blocking L-selectin and very late antigen 4 adhesion receptors. *Proc Natl Acad Sci USA* **1993**; 90: 10494-10498.

Z

- Zauhar RJ, Moyna G, Tian L, Li Z, Welsh WJ. Shape signatures: a new approach to computer-aided ligand- and receptor-based drug design. *J Med Chem* **2003**; 46: 5674-90.
- Zbinden P, Dobler M, Folkers G, Vedani A. PrGen: Pseudoreceptor modeling using receptor-mediated ligand alignment and pharmacophore equilibration. *Quant Struct Act Relat* **1998**; 17: 122-130.
- Zhao H. Scaffold selection and scaffold hopping in lead generation: a medicinal chemistry perspective. *Drug Discov Today* **2007**; 12: 149-155.
- Zoete V, Grosdidier A, Michielin O. Peroxisome proliferator-activated receptor structures: ligand specificity, molecular switch and interactions with regulators. *Biochim Biophys Acta* **2007**; 1771: 915-925.

Eidesstattliche Erklärung

Die vorliegende Dissertation wurde selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Alle Stellen die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche gekennzeichnet.

Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung nicht vorgelegt worden.

.....
Ort und Datum

.....
Y. Tanrikulu

Danksagung

Ich danke allen Personen, die es mir ermöglicht haben diese Arbeit in der vorliegenden Form anfertigen zu können, insbesondere Prof. Dr. Gisbert Schneider für die exzellente Betreuung dieser Arbeit, und Prof. Dr. Joachim Engels für die Begutachtung.

Ferner danke ich allen Mitgliedern der Arbeitskreise Schneider, Schubert-Zsilavec, Stark und Steinhilber, die das Anfertigen dieser Arbeit durch ihre hilfreiche Diskussionen und Tipps um einiges erleichtert und durch wunderbares Arbeitsklima verschönert haben:

Eugen Proschak, Martin Weisel, Swetlana Derksen, Betty Hofmann, Norbert Dichter, Matthias Rupp, Markus Hartenfeller, Alexander Klenner, Volker Haehnke, Felix Reisen, Jan Hiss, Natalie Jäger, Nickolay Todoroff, Tim Werner, Tim Geppert, Petra Schneider und Brigitte Scheidemantel-Geiß.

Zuletzt danke ich meinen Eltern und meiner Freundin, Beate Hajduk, für die allgegenwärtige Unterstützung und das Interesse an meiner Promotion.

Lebenslauf

	Yusuf Tanrikulu
Geburtsdatum, -ort	12.06.1981 in Frankfurt am Main
	Schulbildung
08.1992 – 06.2000 Leistungskurse	Helmholtzgymnasium (Abitur: 2,4) Biologie, Französisch, Geschichte und Englisch
	Zivildienst
11.2000 – 08.2001 Einrichtung	Fahrdienst und Haushaltsunterstützung Altersheim Hufelandhaus, Frankfurt
	Studium
10.2001 – 12.2005 Hauptfächer Thema	Bioinformatik (Diplom: 1,2) Informatik, Organische Chemie und Genetik <i>Entwicklung einer „Fuzzy“-Pharmakophormethode für das virtuelle Screening nach neuen TAR RNA Liganden</i>
	Promotion
01.2006 – 11.2008 Arbeitsgruppe Thema	Wissenschaftlicher Mitarbeiter und Doktorand Prof. Dr. Gisbert Schneider, Chemieinformatik <i>Entwicklung eines Pseudorezeptormodells für das virtuelle Screening</i>
	Weiterbildung
02.2007 – 08.2007 Themen	Basiszertifikat für Hochschullehre Didaktische Prinzipien, Entwurf von Lehrmaterialien, Aufbau und Reflexion persönlicher Qualitätskriterien, wissenschaftliches Schreiben, Coaching, Körper und Klang: Präsentationsdramaturgie, Moderation, Zuhöreraktivierung

.....
Ort und Datum

.....
Y. Tanrikulu

Publikationsliste

Artikel in wissenschaftlichen Zeitschriften

Eingerahmte Publikationen sind im Anhang II-IV zu finden.

- 1) Schneider G, Tanrikulu Y, Schneider P. Self-organizing molecular fingerprints: a ligand-based view on druglike chemical space and off-target prediction. *Future Med Chem* **2008**, *submitted*.
- 2) Proschak E, Sander K, Zettl H, Tanrikulu Y, Rau O, Schneider P, Schubert-Zsilavecz M, Stark H, Schneider G, From molecular shape to potent bioactive agents II: Fragment-based de novo design, *ChemMedChem* **2008**, *accepted*.
- 3) Proschak E, Zettl H, Tanrikulu Y, Weisel M, Kriegl JM, Rau O, Schubert-Zsilavecz M, Schneider G, From molecular shape to potent bioactive agents I: Bioisosteric replacement of molecular fragments, *ChemMedChem* **2008**, *accepted*.
- 4) Schneider G, Hartenfeller M, Reutlinger M, Tanrikulu Y, Proschak E, Schneider P. Voyages to the (un)known: Adaptive design of druglike compounds. *Trends in Biotechnology* **2008**; *accepted*.
- 5) Schneider P, Tanrikulu Y, Schneider G. Self-organizing Maps in Drug Discovery: Compound Library Design, Scaffold-hopping, Repurposing. *Curr Med Chem* **2008**; *accepted*.
- 6) Tanrikulu Y, Rau O, Schwarz O, Proschak E, Siems K, Müller-Kuhrt L, Schubert-Zsilavecz M, Schneider G. Structure-based Pharmacophore Screening for Natural Product-derived PPARgamma Agonists. *ChemBioChem* **2008**; *accepted*.
- 7) Hofmann B, Franke L, Proschak E, Tanrikulu Y, Schneider P, Seinhilber D, Schneider G. Scaffold-hopping Cascade Yields Potent Inhibitors of 5-Lipoxygenase. *ChemMedChem* **2008**; 3: 1535-1538.
- 8) Tanrikulu Y, Schneider G. Pseudoreceptor Models in Drug Discovery: Bridging the Domains of Ligand- and Receptor-based Virtual Screening. *Nature Rev Drug Discov* **2008**; 7: 667-677.
- 9) Schüller A, Suhartono M, Fechner U, Tanrikulu Y, Breitung S, Scheffer U, Göbel MW, Schneider G. The Concept of Template-Based *de novo* Design From Drug-Derived Molecular Fragments and its Application to TAR RNA. *J Comp Aid Mol Des* **2008**; 22: 59-68.
- 10) Tanrikulu Y, Nietert M, Scheffer U, Proschak E, Grabowski K, Schneider P, Weidlich M, Karas M, Göbel M, Schneider G. Scaffold-hopping by "Fuzzy" Pharmacophores and Application to RNA Targets. *ChemBioChem* **2007**; 8: 1932-1936.
- 11) Fischer L, Hornig M, Pergola C, Meindl N, Franke L, Tanrikulu Y, Dodt G, Schneider G, Seinhilber D, Werz O. The Molecular Mechanism of the Inhibition by Licofelone of the Biosynthesis of 5-Lipoxygenase Products. *Br J Pharmacol* **2007**; 152: 471-480.

- 12) Franke L, Schwarz O, Muller-Kuhrt L, Hoernig C, Fischer L, George S, Tanrikulu Y, Schneider P, Werz O, Steinhilber D, Schneider G. Identification of Natural-Product-Derived Inhibitors of 5-Lipoxygenase Activity by Ligand-Based Virtual Screening. *J Med Chem* **2007**, 50: 2640-2646.
- 13) Schmuker M, Schwarte F, Bruck A, Proschak E, Tanrikulu Y, Givehchi A, Scheiffle K, Schneider G. SOMMER: Self-Organising Maps for Education and Research. *J Mol Model*. **2007**; 13: 225-228.

Buchbeiträge

- 1) Proschak E, Tanrikulu Y, Schneider G. Fragment-based *de novo* design of druglike compounds. In: *Chemoinformatics: An Approach to Virtual Screening* (A. Varnek, A. Tropsha; Eds). RSC Publishing, Cambridge, **2008**; 217-239.
- 2) Schneider G, Proschak E, Grabowski K, Schneider P and Tanrikulu Y. More Than Colored Blobs - "Fuzzy" Descriptions of Potential Pharmacophore Points, In: *Proceedings of the Beilstein Workshop 2006*. Logo-Verlag, Berlin, **2007**; 173-184.

Posterpräsentationen

- 1) Werner T, Geppert T, Tanrikulu Y, Proschak E, Clark T, Schneider G, Pseudoreceptor-based Pocket Selection in Histamine H₄ Receptor Molecular Dynamics Simulation, 4th *German Conference on Chemoinformatics*, 09.-11.11.**2008**, Goslar - Germany.
- 2) Löwer M, Tanrikulu Y, Weisel M, Schneider G, Fuzzy Virtual Ligands for Virtual Screening, 4th German Conference on Chemoinformatics, 09.-11.11.**2008** Goslar - Germany.
- 3) Sander K, Tanrikulu Y, Kottke T, Proschak E, Schneider E, Seifert R, Schneider G, Stark H, Lead Structure Identification by Virtual Screening and Scaffold Modification for Histamine H₄ Receptor Ligands, *XXth International Symposium on Medicinal Chemistry*, 31.08-03.09.**2008**, Vienna - Austria.
- 4) Proschak E, Sander K, Zettl H, Tanrikulu Y, Rau O, Schubert-Zsilavec M, Stark H, Schneider G, SQUIRREL_{novo}: Lead-Finding of PPAR α Agonists by *de novo* Design, *XXth International Symposium on Medicinal Chemistry*, 31.08-04.09.**2008**, Vienna - Austria.
- 5) Proschak E, Tanrikulu Y, Rau O, Zettl H, Weisel M, Schubert-Zsilavec M, Schneider G, The SQUIRREL Hops Over the Scaffolds: Novel PPAR α Agonists, *Frontiers of Medicinal Chemistry*, 02.03.-05.03.**2008**, Regensburg - Germany.
- 6) Klenner A, Becker S, Nietert M, Hartenfeller M, Schüller A, Weisel M, Proschak E, Tanrikulu Y, Schneider G, RNA Ligand Design, *CRC579 RNA-Ligand-Interactions*, 04. & 05.03.**2008**, Frankfurt - Germany.
- 7) Becker S, Klenner A, Nietert M, Hartenfeller M, Schüller A, Weisel M, Proschak E, Tanrikulu Y, Schneider G, The RNA Pocketome, *CRC579 RNA-Ligand-Interactions*, 04. & 05.03.**2008**, Frankfurt - Germany.

- 8) Proschak E, Tanrikulu Y, Hofmann B, Rau O, Zettl H, Weisel M, Kriegl J, Steinhilber D, Schubert-Zsilavec M, Schneider G, The SQUIRREL Hops Over the Scaffolds: Application of a Novel Virtual Screening Tool, *3rd German Conference on Chemoinformatics*, 11.-13.11.2007, Goslar - Germany.
- 9) Grabowski K, Meissner M, Schüller A, Nietert M, Proschak E, Tanrikulu Y, Stauch B, Münk C, Scheffer U, Cichutek K, Schubert-Zsilavec M, Göbel M, Schneider G, Applications of MOE in an Academic Environment, *Fifth European CCG User Group Meeting*, 08.10.-09.10.2007, Berlin - Germany.
- 10) Tanrikulu Y, Proschak E, Nietert M, Grabowski K, Scheffer U, Göbel M, Schneider G, Virtual Screening for Novel TAR RNA Ligands with LIQUID: Scaffold-Hopping Using a Combination of Automated Ligand Docking and Fuzzy Pharmacophore Modeling, *4th Joint Conference on Chemoinformatics*, 18.06.-20.06.2007, Sheffield - United Kingdom.
- 11) Nietert M, Tanrikulu Y, Grabowski K, Scheffer U, Göbel M, Weidlich M, Karas M, Kestner, Gohlke H, Schneider G, Virtual Screening for RNA Ligands - Coping with the Flexibility of Drug Targets, *2nd German Conference on Chemoinformatics*, 12.-14.11.2006, Goslar - Germany.
- 12) Grabowski K, Franke L, Schwarz O, Müller-Kuhr L, Hoernig C, Fischer L, George S, Tanrikulu Y, Schneider P, Werz O, Steinhilber D, Schneider G, Natural Produkt Databases: Sources for Unique Scaffolds and Structures, *2nd German Conference on Chemoinformatics*, 12.-14.11.2006, Goslar - Germany.
- 13) Christian S, Hofmann B, Derksen S, Tanrikulu Y, Rau O, Syha Y, Popescu L, Schubert-Zsilavec M, Rabenstein J, Marschalek R, Scheffer U, von Laer D, Schneider G, Quest for Novel Lead Structures by Virtual Screening: Techniques and Applications, *2nd German Conference on Chemoinformatics*, 12.-14.11.2006, Goslar - Germany.
- 14) Tanrikulu Y, Schneider G, "Flexible" Pharmacophore Points for Alignment-Free Virtual Screening, *1st German Conference on Chemoinformatics*, 13.-15.11.2005, Goslar - Germany.

Anhang

Ausgewählte Publikationen

- I. MQL Notation der Substrukturen aus Tabelle 4
- II. Pharmakophorsuche nach Naturstoff-abgeleiteten PPAR γ Agonisten
- III. Pseudorezeptormodelle in der Wirkstoffforschung
- IV. Scaffold-Hopping mit „fuzzy“ Pharmakophoren

MQL Notationen der Substrukturen aus Tabelle 4

Akzeptoren

N#C

O[allHydrogens=1]-C[sp2](~*[ring])~*[ring]

O=S(=O)(-*)-C

O[allHydrogens=1]-C[sp3]

O=P(-O[charge=-1])(-*)-*

O=P(-C)(-C)(-C)

O=C(-[C|N|H])-C

O=C(-C)-O[charge=-1]

O=C(-C)-O-C

O[!ring](-C[!bound(=O)])-C[!bound(=O)]

O[ring](-C[!bound(=O)])-C[!bound(=O)]

S=C(~*)~*

N[ring&allHydrogens<1](=*[ring])-*[ring]

O=N[charge=+1](-*)-O[charge=-1]

O(-C)-S(=O)=O

O(-P(=O)(-O[charge=-1]))-*

O=C(-N)-C

Donoren

N[allHydrogens>=2&!ring]-S

N[allHydrogens=2&ring](-C[!bound(=O)])-C[!bound(=O)]

O[allHydrogens=1]-C[sp2&ring](~*[ring])~*[ring]

N[allHydrogens>=1&!ring]=*

N[allHydrogens>=2](-C[!ring])

O[allHydrogens=1&!ring]-C[sp3]

N[allHydrogens=1&sp3](-C)(-C)-C

N[allHydrogens=1](-C=O)-*

N[allHydrogens>=1&ring](-*[!bound(=O)])-C[!bound(=O)]=*

N[allHydrogens>=1&!ring](-C[!bound(=O)])-C[!bound(=O)]

N[allHydrogens>=1&ring](=C[!bound(=O)])-C[!bound(=O)]

Pharmakophorsuche nach Naturstoff-abgeleiteten PPAR γ Agonisten

Structure-based Pharmacophore Screening for Natural Product-derived PPAR γ Agonists

Tanrikulu Y, Rau O, Schwarz O, Proschak E, Siems K,
Müller-Kuhrt L, Schubert-Zsilavec M, Schneider G

ChemBioChem 2008; *accepted*

Peroxisom Proliferator Aktivierte Rezeptoren (PPAR) sind Transkriptionsfaktoren, die eine kritische Rolle in Lipidsignaltransduktion und Immunomodulation haben, und interagieren mit anderen nukleären Rezeptoren der Fettspiegelregulation, wie PXR und NF-kappaB. Daher dienen Agonisten des PPAR α und PPAR γ als therapeutische Agenten in der Behandlung von Dyslipidämie, Diabetes Typ-2 und Artherosklerose, während ihre regulatorische Wirkung auf die Zellproliferation derzeit noch untersucht wird. Verschiedene Naturstoffe wurden identifiziert, die PPAR aktivieren, wie zum Beispiel ein Tetrahydrocannabinol (THC) Säuremetabolit, Carnosolsäure und Carnosol, und Resveratrol. Diese bilden Startpunkte zur kombinatorischen Untersuchung von Naturstoff-abgeleiteten Substanzen zur Leitstrukturfindung und -entwicklung zum Austausch von existierenden PPAR Agonisten mit potentiell sicheren Wirkstoffen mit neuem Strukturgrundgerüst. Hier präsentieren wir ein Protokoll zum virtuellen Screening nach PPAR γ Agonisten aus einer kombinatorischen Substanzbibliothek, welche aus der Grundstruktur des α -Santonin, einem natürlichen Sesquiterpenlacton aus Beifuß, abgeleitet sind. Wir demonstrieren, dass es möglich ist Leitstrukturen aus kleinen kombinatorischen Substanzkollektionen mit minimalem Aufwand durch „fuzzy“ Pharmakophorscreenings zu finden.

DOI: 10.1002/cbic.200((will be filled in by the editorial staff))

Natural product-derived compound screening

Structure-based Pharmacophore Screening for Natural Product-derived PPAR γ Agonists

Y. Tanrikulu,^[a] O. Rau,^[b] O. Schwarz,^[c] E. Proschak,^[a]
K. Siems,^[c] L. Müller-Kuhr,^[c] M. Schubert-Zsilavec,^[b]
G. Schneider^{[a],*}

Peroxisome proliferator-activated receptors (PPARs) are transcription factors that play a critical role in lipid signalling and immunomodulation, and functionally interact with other nuclear receptors in the regulation of lipid metabolism, for example PXR and NF-kappaB.¹ Therefore, agonists of PPAR α and PPAR γ serve as therapeutic agents for treatment of dyslipidaemia, type-II diabetes, and arteriosclerosis, while their effects on the regulation of cell proliferation are under investigation.² Several natural compounds have been identified that activate PPARs, like tetrahydrocannabinol (THC) metabolite THC-11-oic acid,³ carnolic acid and carnosol,⁴ and resveratrol.⁵ These can provide a starting point for combinatorial exploration of natural product-derived compounds for lead discovery and development,⁶ with the aim to substitute existing PPAR agonists by potentially safer drugs containing novel scaffolds.⁷ Here, we present a virtual screening protocol that led to a PPAR γ agonist from a combinatorial compound library which was derived from the scaffold structure of α -santonin, a natural sesquiterpene lactone found in mugwort.⁸ We demonstrate that it is possible to find lead candidates in small combinatorial compound collections with minimal experimental effort by “fuzzy” pharmacophore screening.

For generation of a pharmacophore query, we superimposed four high-resolution X-ray structures of the PPAR γ ligand-binding domain in complex with agonists (PDB⁹ identifiers: 1nyx with Ragaglitazar,¹⁰ 1knu with YPA,¹¹ 1i7i with Tesaglitazar,¹² 1zgy with Rosiglitazone¹³) (Figure 1a). The resulting ligand alignment served as basis for pharmacophoric point assignment by our software LIQUID, as described

previously.¹⁴ Briefly, LIQUID represents potentially pharmacophoric points (lipophilic, hydrogen-bonding) in a molecule by Gaussian densities. These densities are converted to probabilities for pairwise matching of compounds. As a result of LIQUID matching and scoring, a screening library is sorted so that best matching compounds appear at the top of the ranked list. From this list, most promising screening candidates are picked. For generation of the LIQUID descriptors from three-dimensional molecular conformations, we used cluster radii of 1 Å for lipophilic centres and 2 Å for hydrogen-bonding centres. The resulting pharmacophore query is depicted in Figure 1b. This procedure was performed to obtain a “receptor-relevant” pharmacophore model of PPAR γ agonists, instead of using a ligand-based spatial alignment of artificially generated conformers. This structure-based alignment of multiple ligands allowed us to compute “fuzzy” pharmacophoric feature points, so that a probability-weighted consensus model was obtained.

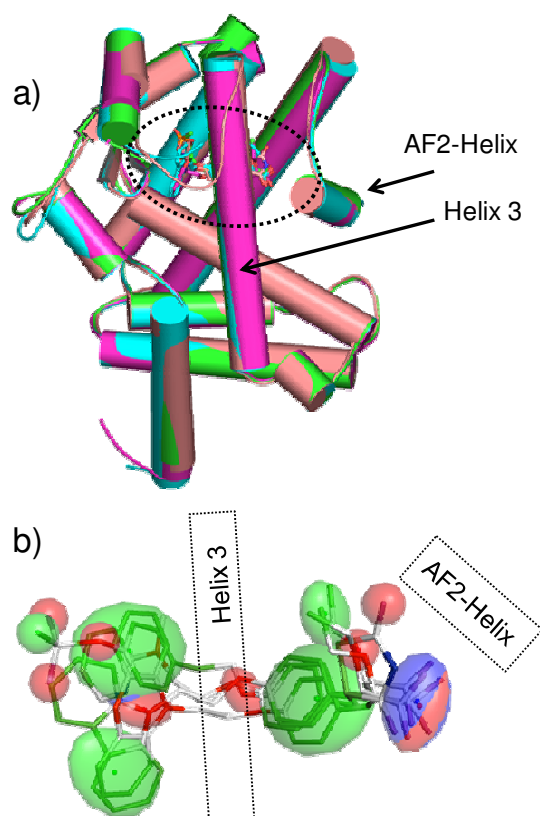


Figure 1. Structural superimposition of four PPAR γ -agonist complexes (a), and the alignment-derived LIQUID pharmacophore query (b). In the pharmacophore model, lipophilic centres are shown in green, potential hydrogen-bond donor sites in blue, potential hydrogen-bond acceptor sites in red. Approximate locations of helix 3 and the AF2-helix are indicated. The trivariate Gaussians are shown with widths of one standard deviation in each direction.

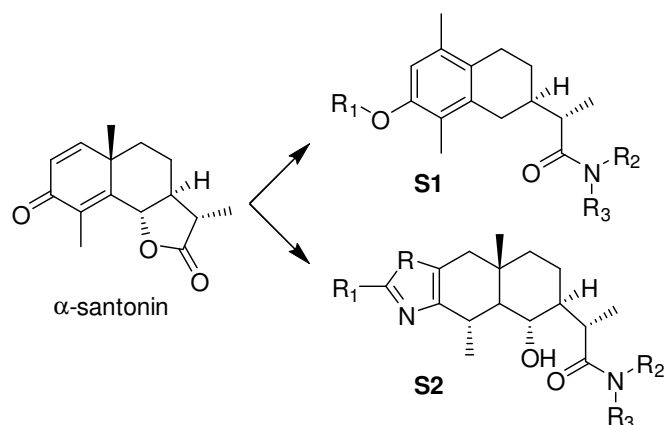
[a] Prof. Dr. G. Schneider, Y. Tanrikulu, E. Proschak
Goethe-Universität, Institut für Organische Chemie und Chemische
Biologie, CMP / ZAFES / LiFF
Siesmayerstr. 70, D-60323 Frankfurt am Main, Germany
Fax: (+) 49 69 798 24880
E-mail: g.schneider@chemie.uni-frankfurt.de

[b] Prof. Dr. M. Schubert-Zsilavec, Dr. O. Rau
Goethe-Universität, Institut für Pharmazeutische Chemie, CMP /
ZAFES / LiFF
Max-von-Laue-Str. 9, D-60438 Frankfurt am Main, Germany

[c] Dr. O. Schwarz, Dr. K. Siems, Dr. L. Müller-Kuhr
AnalytiCon Discovery GmbH
Hermannswerder Haus 17, D-14473 Potsdam, Germany

Then, the AnalytiCon Discovery collection of natural product-derived combinatorial compounds (version 01/2007, 15,590 entries) was searched for hits matching the LIQUID pharmacophore query. A single three-dimensional conformation was computed for each compound using Corina v3.2 (Molecular

Networks GmbH, Erlangen). This concept was shown to be sufficient for first-pass virtual screening in comparison to more demanding multi-conformer approaches.¹⁵ LIQUID descriptors were generated from these conformers and compared to the query descriptor with the Euclidean distance metric. Among the top-scoring semi-synthetic natural product analogues, scaffolds **S1** and **S2** were dominant (in total, 475 compounds; Scheme 1). Both scaffolds had been derived from α -santonin by means of chemical modifications.⁸ Notably, α -santonin itself was inactive in our PPAR assays. We decided to focus on scaffold **S1**, which has the higher ligand efficiency¹⁶ of the two prevalent scaffolds, and manually selected eight compounds from two small series for activity testing (Table 1).



Scheme 1. Chemical structures of two molecular scaffolds derived from α -santonin. Both scaffolds were found among the top-scoring virtual hits obtained by LIQUID pharmacophore searching in the AnalytiCon Discovery compound collection.

Two compounds exhibited PPAR activation in a cell-based reporter gene assay (Table 1). Compound **1** was most potent with an EC_{50} value of approximately 15 μ M at PPAR γ and only slight but measurable activation of PPAR α (Figure 2). Compound **2** stems from the same series, is a weak PPAR γ agonist, but does not activate PPAR α due to the additional piperidine linker. It is noteworthy that compounds containing the tetrahydronaphthole **S1** are known to inhibit 5-lipoxygenase (5-LO).¹⁷ This observation supports the hypothesis that natural product-derived scaffolds might represent “biophores”, that is, preferred molecular frameworks for protein binding.¹⁸ Different side-chain decorations can be explored to yield ligands with an affinity toward different protein targets.

As a final test and to gain a preliminary understanding of the ligand-receptor interaction, we performed automated docking of compound **1** into the binding site of PPAR γ using the software GOLD (version 3.2).¹⁹ Figure 3 shows the best-scoring suggested docking pose of compound **1** ($GoldScore = 59$) in comparison to the co-crystallized binding conformation of Farglitazar ($EC_{50} \approx 5$ nM²⁰), which was chosen as an independent reference (not part of the pharmacophore query). The carboxylic head group forms hydrogen bridges to amino acid side chains that are known to be involved in agonist binding, including Tyr473, which is part of activation function 2 (AF2-helix, helix 12), a critical structural domain for PPAR function.^{1,21} The docking study supports our experimental findings and suggests compound **1** as a full agonist of PPAR γ , as binding to the AF2-helix of the ligand-binding site is required for full agonist activity and partial agonists function via an independent mechanism.²² Due to its comparably small lipophilic tail, compound **1** does not embrace helix 3 (Figure 3),

which might explain its lower potency compared to classical full PPAR γ agonists. Based on EC_{50} values, the computed ligand efficiency¹⁶ of compound **1** is $\Delta g = (\Delta G^{\circ} / \text{number of non-hydrogen atoms}) \approx -1.3$ kJ mol⁻¹ per non-hydrogen atom, and for Farglitazar $\Delta g \approx -1.2$ kJ mol⁻¹ per non-hydrogen atom. This finding leaves room for optimization of compound **1** as a favorable lead structure, bearing in mind that the maximally achievable affinity per non-hydrogen atom has been estimated to be -4.2 kJ mol⁻¹.²³

Summarizing, we demonstrated that natural product-derived combinatorial libraries can be efficiently screened by structure-based pharmacophore matching and scoring, leading to novel bioactive substances. For pharmacophore query generation, we used the alignment of co-crystal structures of PPAR γ with potent agonists, rather than a ligand-based alignment. This virtual concept provides a straightforward route to finding novel bioactive agents with minimal experimental effort. Here, a new chemotype of PPAR agonists was identified, which may serve as a starting point for lead structure generation and optimization.

Table 1. Activity of compounds tested in a cellular reporter gene assay⁷ at a concentration of 30 μ M. Values give PPAR α activation relative to the selective agonist GW7647 ($EC_{50} = 6$ nM), and PPAR γ activation relative to the selective agonist Pioglitazone ($EC_{50} = 0.27$ μ M). Structures are aligned according to the α -santonin derived scaffold.

No.	Chemical structure	PPAR α	PPAR γ
1		16±3%	110±31%
2		0%	33±8%
3		inactive	
4		inactive	
5		inactive	
6		inactive	
7		inactive	
8		inactive	

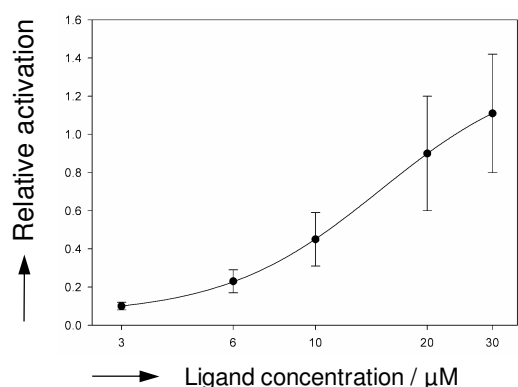


Figure 2. Dose response curve of compound **1** for activation of PPAR γ , determined in a Gal4-dependent luciferase reporter gene assay in Cos7 cells. The plot gives relative activities \pm stdev. compared to the effect of the reference compound Pioglitazone ($n \geq 3$).

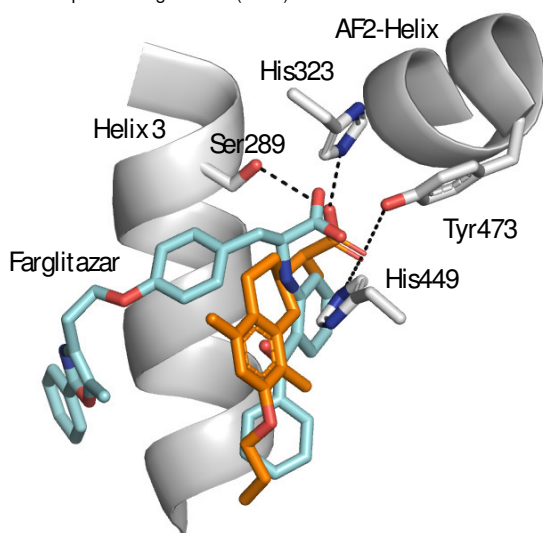


Figure 3. Docking pose of compound **1** (in orange) in a crystal structure of PPAR γ in complex with Farglitazar (PDB identifier: 1fm9²⁴). Potential hydrogen bonds are drawn as dashed lines.

Experimental Section

Receptor activation was tested in a luciferase reporter gene assay as described.⁴ In brief, Cos7 cells were seeded in 96-well plates at a density of 30,000 cells per well (100 μl) and transfected with lipofectamine (Gibco), the Gal4-dependent reporter vector pFR-luc, the Gal4-PPAR-LBD fusions pFA-CMV-PPAR of the respective subtype, and pRL-SV40 for normalization. Four hours after transfection, cells were incubated with the test compound in triplicate wells. Luciferase activity was determined the following day using Dual-glo (Promega). Each experiment was repeated at least three times. Average relative activation was calculated in comparison to the effect of a reference compound (1 μM Pioglitazone for PPAR γ and 1 μM GW7647 for PPAR α). Determination of EC_{50} values was carried out with SPSS2001 (SPSS, Chicago).

Acknowledgements

This research was supported by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften, Frankfurt a.M.

Keywords: Bioisoster · Combinatorial chemistry · Drug discovery · Lipid metabolism · Virtual screening

- [1] a) H. E. Xu, M. H. Lambert, V. G. Montana, D. J. Parks, S. G. Blanchard, P. J. Brown, D. D. Sternbach, J. M. Lehmann, G. B. Wisely, T. M. Willson, S. A. Kliewer, M. V. Milburn, *Mol. Cell.* **1999**, *3*, 397-403; b) V. Arulampalam, *J. Intern. Med.* **2008**, *263*, 607-612; c) O. Rau, H. Zettl, L. Popescu, D. Steinhilber, M. Schubert-Zsilavecz, *ChemMedChem* **2008**, *3*, 206-221; d) W. Wahli, *J. Intern. Med.* **2008**, *263*, 613-619.
- [2] F. Chang, L. A. Jaber, H. D. Berlie, M. B. O'Connell, *Ann. Pharmacother.* **2007**, *41*, 973-983.
- [3] J. Liu, H. Li, S. H. Burstein, R. B. Zurier, J. D. Chen, *Mol. Pharmacol.* **2003**, *63*, 983-992.
- [4] a) O. Rau, M. Wurglics, A. Paulke, J. Zitzkowski, N. Meindl, A. Bock, T. Dingermann, M. Abdel-Tawab, M. Schubert-Zsilavecz, *Planta Med.* **2006**, *72*, 881-887; b) D. Poeckel, C. Greiner, M. Verhoff, O. Rau, L. Tausch, C. Hörnig, D. Steinhilber, M. Schubert-Zsilavecz, O. Werz, *Biochem. Pharmacol.* **2008**, *76*, 91-97.
- [5] S. Ulrich, S. M. Loitsch, O. Rau, A. von Knethen, B. Brüne, M. Schubert-Zsilavecz, J. M. Stein, *Cancer Res.* **2006**, *66*, 7348-7354.
- [6] O. Rau, Y. Syha, H. Zettl, M. Kock, A. Bock, M. Schubert-Zsilavecz, *Arch. Pharm. (Weinheim)* **2008**, *341*, 191-195.
- [7] a) K. Grabowski, K.-H. Baringhaus, G. Schneider, *Nat. Prod. Rep.* **2008**, DOI: 10.1039/b715668p; b) J. M. Rollinger, H. Stuppner, T. Lange, *Prog. Drug Res.* **2008**, *65*, 211, 213-249; b) O. Rau, M. Wurglics, T. Dingermann, M. Abdel-Tawab, M. Schubert-Zsilavecz, *Pharmazie* **2006**, *61*, 952-956.
- [8] O. Schwarz, S. Jakupovic, H. D. Ambrosi, L. O. Haustedt, C. Mang, L. Müller-Kuhrt, *J. Comb. Chem.* **2007**, *9*, 1104-1113.
- [9] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucl. Acids Res.* **2000**, *28*, 235-242.
- [10] S. Ebdrup, I. Pettersson, H. B. Rasmussen, H.-J. Deussen, A. Frost Jensen, S. B. Mortensen, J. Fleckner, L. Pridal, L. Nygaard, P. Sauerberg, *J. Med. Chem.* **2003**, *46*, 1306-1317.
- [11] P. Sauerberg, I. Pettersson, L. Jeppesen, P. S. Bury, J. P. Mogensen, K. Wassermann, C. L. Brand, J. Sturis, H. F. Woldike, J. Fleckner, A.-S. T. Andersen, S. B. Mortensen, L. A. Svensson, H. B. Rasmussen, S. V. Lehmann, Z. Polivka, K. Sindelar, V. Panajotova, L. Ynddal, E. M. Wulff, *J. Med. Chem.* **2002**, *45*, 789-804.
- [12] P. Cronet, J. F. Petersen, R. Folmer, N. Blomberg, K. Sjoblom, U. Karlsson, E. L. Lindstedt, K. Bamberg, *Structure* **2001**, *9*, 699-701.
- [13] Y. Li, M. Choi, K. Suino, A. Kovach, J. Daugherty, S. A. Kliewer, H. E. Xu, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9505-9510.
- [14] Y. Tanrikulu, M. Nietert, U. Scheffer, E. Proschak, K. Grabowski, P. Schneider, M. Weidlich, M. Karas, M. Göbel, G. Schneider, *ChemBioChem* **2007**, *8*, 1932-1936.
- [15] S. Renner, C. H. Schwab, J. Gasteiger, G. Schneider, *J. Chem. Inf. Model.* **2006**, *46*, 2324-2332.
- [16] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Disc. Today* **2004**, *9*, 430-431.
- [17] L. Franke, O. Schwarz, L. Müller-Kuhrt, C. Hoernig, L. Fischer, S. George, Y. Tanrikulu, P. Schneider, O. Werz, D. Steinhilber, G. Schneider, *J. Med. Chem.* **2007**, *50*, 2640-2646.
- [18] a) S. Pickett, in *Protein-Ligand Interactions* (H.-J. Böhm, G. Schneider, Eds), Wiley-VCH, Weinheim **2003**, pp. 73-106; b) M. Kaiser, S. Wetzel, K. Kumar, H. Waldmann, *Cell. Mol. Life Sci.* **2008**, *65*, 1186-1201.
- [19] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, R. D. Taylor, *Proteins* **2003**, *52*, 609-623.
- [20] B. R. Henke, S. G. Blanchard, M. F. Brackeen, K. K. Brown, J. E. Cobb, J. L. Collins, W. W. J. Harrington, M. A. Hashim, E. A. Hull-Ryde, I. Kaldor, S. A. Kliewer, D. H. Lake, L. M. Leesnitzer, J. M. Lehmann, J. M. Lenhard, L. A. Orband-Miller, J. F. Miller, R. A. J. Mook, S. A. Noble, W. J. Oliver, D. J. Parks, K. D. Plunket, J. R. Szewczyk, T. M. Willson, *J. Med. Chem.* **1998**, *41*, 5020-5036.
- [21] a) S. Yu, J. K. Reddy, *Biochim. Biophys. Acta* **2007**, *1771*, 936-951; b) L. Michalik, V. Zoete, G. Krey, A. Grosdidier, L. Gelman, P. Chodanowski, J. N. Feige, B. Desvergne, W. Wahli, O. Michielin, *J. Biol. Chem.* **2007**, *282*, 9666-9677.
- [22] J. B. Bruning, M. J. Chalmers, S. Prasad, S. A. Busby, T. M. Kamenecka, Y. He, K. W. Nettles, P. R. Griffin, *Structure* **2007**, *15*, 1258-1271.
- [23] I. D. Kuntz, K. Chen, K. A. Sharp, P. A. Kollman, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9997-10002.
- [24] R. T. Gampe Jr., V. G. Montana, M. H. Lambert, A. B. Miller, R. K. Bledsoe, M. V. Milburn, S. A. Kliewer, T. M. Willson, H. E. Xu, *Mol. Cell* **2000**, *5*, 545-555.

Pseudorezeptormodelle in der Wirkstoffforschung

Pseudoreceptor Models in Drug Discovery: Bridging the Domains of Ligand- and Receptor-based Virtual Screening

Tanrikulu Y, Schneider G

Nature Review Drug Discovery 2008; 7: 667-677

Rationale Wirkstoffforschung basiert auf expliziten und impliziten Struktur-Aktivitätsbeziehungen. Typischerweise werden rezeptorbasierte oder ligandenbasierte Strategien angewendet, welches von den verfügbaren Informationen zur Liganden und Rezeptorstruktur abhängt. Pseudorezeptormodelle kombinieren die Vorteile beider Strategien und repräsentieren eine Möglichkeit zur Rezeptorkartierung und Ligandensuche. Sie bilden einen Einstiegspunkt für Wirkstoffentwicklungsprojekte, die ohne eine aufgeklärte Zielproteinstruktur auskommen müssen. Dieses Review verschafft einen Überblick über alle entwickelten Pseudorezeptormodelle und jüngsten Erfolgen in der Leitstrukturfindung. Darüberhinaus werden die Voraussetzung zur Pseudorezeptorkonstruktion, die Vorteile und Limitierung verschiedener Ansätze kritisch begutachtet.

Pseudoreceptor models in drug design: bridging ligand- and receptor-based virtual screening

Yusuf Tanrikulu and Gisbert Schneider

Abstract | Rational drug design is based on explicit or implicit structure–activity relationship models. Typically, receptor-based or ligand-based strategies are pursued, depending on the information available about known ligands and the receptor structure. Pseudoreceptor models combine the advantages of these two strategies and represent a unifying concept for both receptor mapping and ligand matching. They can provide an entry point for structure-based modelling in drug discovery projects that lack a high-resolution structure of the target. Here, we review the field of pseudoreceptor modelling techniques along with recent hit and lead finding applications, and critically discuss prerequisites, advantages and limitations of the various approaches.

Fragment-based lead discovery

The identification of bioactive substances by assembling small-molecule fragments. Often, the fragments themselves are building blocks of known drugs.

Molecular similarity searching

These are techniques for searching compound databases with the aim of retrieving molecules that have a similar chemical structure to a query compound. Typically, similarity is expressed on the basis of molecular descriptors; for example substructure fingerprints or pharmacophoric features.

*Institute of Organic Chemistry and Chemical Biology,
Johann Wolfgang
Goethe-University Frankfurt,
Siesmayerstrasse 70,
D-60323 Frankfurt, Germany.
Correspondence to G.S.
e-mail: g.schneider@chemie.uni-frankfurt.de
doi:10.1038/nrd2615
Published online 18 July 2008*

Computer-assisted drug design has supported pharmaceutical research and development for over three decades. Two major design principles have evolved: the receptor-based approach and the ligand-based approach^{1–3}. These strategies, in particular in combination with fragment-based lead discovery⁴, are considered to be complementary to high-throughput screening, which still represents the major source of innovation for many pharmaceutical companies^{5–7}.

Areas of application of ligand-based methods such as molecular similarity searching or pharmacophore-based virtual screening include scaffold-hopping^{8–11} and repurposing^{12–14}. A prerequisite for their use is the availability of at least one bioactive reference structure or search template. When a sufficiently accurate three-dimensional (3D) receptor model is available, receptor-based methods can be used to screen compound collections virtually^{15,16}. Often, both ligand-based and receptor-based techniques may be applied simultaneously. Pseudoreceptor modelling fills the gap between pure ligand-based and receptor-based design, and there is a renewed interest in these techniques, which were conceived approximately 20 years ago^{17,18}. In this Review, we will present the different pseudoreceptor concepts, as well as recent applications in the context of hit and lead finding projects.

3D target (or receptor) information is not always available at the beginning of a drug discovery project. It has been estimated that approximately 50% of early phase drug discovery projects are started without a receptor structure model. This is mainly owing to the pursuit of

G-protein-coupled receptor (GPCR) or other integral membrane protein targets for which structural information is lacking, and so projects are limited to ligand-based design methods¹. At present, the Protein Data Bank (PDB)¹⁹ — since 1998 under the responsibility of the Research Collaboratory for Structural Bioinformatics²⁰ — contains approximately 50,000 entries of target structures that have been determined, and has been increasing exponentially in size. However, notably, only 26 entries (0.05%) are related to GPCRs. These entries include 23 structures of rhodopsin (PDB entries 1ln6, 2j4y, 1h68, 1jfp, 1jgj, 2i35, 1gue, 1xio, 2f93, 2i36, 2i37, 1f88, 1gu8, 1hzx, 1l9h, 1u19, 2f95, 2ii7, 2ii8, 2g87, 2hpy, 1gzm and 2ped), and three β_2 -adrenergic receptor structures (PDB entries 2rh1, 2r4r and 2r4s). The limited number of available membrane protein target structures overall prohibits the applicability of receptor-based programs for many pharmaceutically interesting targets^{21,22}. Advances in crystallization methods for membrane proteins, together with high-throughput crystallography, are anticipated to help remedy this backlog²³, but accurate methods for receptor structure modelling are needed²⁴. Receptor-based design can likewise be used on structures derived from homology-modelled structures of related family members with high sequence identity²⁵.

To meet this challenge, hybrid approaches have been proposed to bridge ligand-based and receptor-based design domains. The most prominent advances use quantitative structure–activity relationship (QSAR) modelling to perceive correlations between the affinity of known

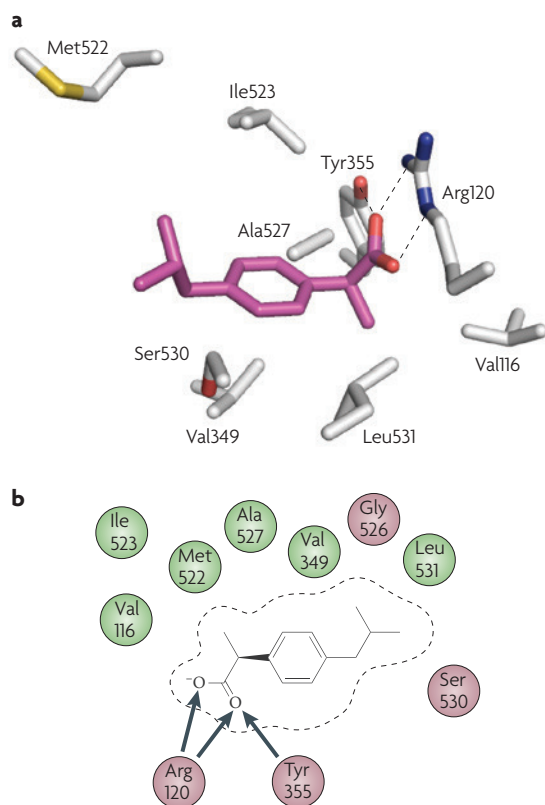


Figure 1 | Aim of pseudoreceptor modelling. **a** | Co-crystal structure of ibuprofen and cyclooxygenase 1 (COX1); Protein Data Bank identifier: [1eqg](#)¹¹⁷. **b** | Two-dimensional sketch of the interaction sites between receptor and ligand (generated with MOE; Chemical Computing Group). Hydrogen bonds are highlighted. A pseudoreceptor model should capture the essential interacting amino-acid side chains, the volume and shape of the ligand binding cavity, and indicate hot spots for receptor–ligand interaction. Polar residues are brown, hydrophobic residues are green.

Pharmacophore

The spatial arrangement of atoms or groups in a molecule known or predicted to be responsible for specific biological activity.

Scaffold-hopping

The identification of isofunctional but structurally different chemotypes.

Repurposing

The use of known leads and approved drugs for new indications. This includes the development of known drugs to preferably interact with so-called off-targets.

Quantitative structure–activity relationship (QSAR)

Mathematical relationships that link chemical structure and pharmacological activity in a quantitative manner for a series of compounds. Methods that can be used in QSAR include various regression and pattern-recognition techniques.

Atomistic

A molecular design concept or molecular representation that considers single atoms.

The idea of pseudoreceptor models is to construct a replacement (surrogate) for the 3D target structure around a single ligand conformation or an ensemble of ligands. The models attempt to capture the shape of the binding site and its interaction points that need to be occupied for successful ligand binding (FIG. 1). Pseudoreceptors can delineate a receptor surrogate that mimics an alternative host system, which is capable of non-covalent interactions with known molecules that have affinity. The pioneers of pseudoreceptor modelling constructed the virtual host systems manually; for example, by folding a peptide chain into a ‘bed’-shaped conformation consisting of a central sheet structure flanked by two helices³². Others³³ developed a pseudoreceptor model of Na⁺/K⁺ ATPase by building a consensus pharmacophore model of reference steroids using GRID³⁴ and positioning appropriate amino acids at the binding subsites with SYBYL (Tripos Inc.). Interactive pseudoreceptor modelling software packages were invented during the same period. Although Vedani and co-workers³⁵ introduced a definition that separates minireceptors (an unconnected set of small molecules imitating receptor side chains encircling a ligand) from pseudoreceptors (fully linked peptidic or non-peptidic molecules forming a binding pocket) we will refer to both types as pseudoreceptors (model) herein.

The main aim of this Review is to provide an overview of pseudoreceptor-generation methods. Unlike previous summaries of the pseudoreceptor field^{35–37}, we will not restrict the overview to the most popular protocols such as Yak^{35,38,39} or PrGen⁴⁰. However, to limit the scope of this article, concepts based on molecular dynamics calculations are not described, and may be found elsewhere^{41,42}. The main strengths and weaknesses of pseudoreceptor modelling are discussed with the aid of selected recently published work. TABLE 1 summarizes the pseudoreceptor software solutions and concepts that are discussed here in chronological order.

The pseudoreceptor concept

The methodology of pseudoreceptor model generation involves the three fundamental tasks shown in FIG. 2. First, the presumed key interaction sites (anchor points) of the ligand–receptor complex are defined; second, the core pseudoreceptor model is assembled around these hypotheses; and last, model coordinates are optimized to gain more accurately calculated binding energies in validation studies⁴³. A wide range of pseudoreceptor construction routes have been developed. For the purpose of this Review, we formed six such categories (FIG. 3): grid-based, isosurface-based, partition-based, atom-based, peptide-based and fragment-based methods, which are each described in more depth later in the article.

The key aim is to generate a pseudoreceptor model in which ligands adopt conformations that are comparable to their binding mode in the real macromolecular target structure. The optimized model can be used for affinity prediction and other receptor-based modelling tasks. It is important to understand that a pseudoreceptor does not necessarily accurately mimic the real receptor

Table 1 | Selected pseudoreceptor approaches in chronological order

Program name	Year	Type of model	Refs
RECEPS	1987	Grid-based	17,18
HASL	1988	Grid-based	57
MORPHEUS	1989	Grid-based	56
VOROM	1991	Partition-based	67,68
Yak	1993	Peptide-based	38,35,39
GERM	1994	Atom-based	51
RECEPTOR	1995	Isosurface-based	60–62
QXP	1997	Fragment-based	81
EGSITE2	1997	Partition-based	50
SKELGEN	1997	Partition-based	69,70
PrGen	1998	Peptide-based	40
PARM	1998	Atom-based	52
Quasar	1998	Atom-based	54
Hostdesigner	2002	Fragment-based	83
FLARM	2002	Atom-based	53
RAPTOR	2004	Isosurface-based	55
WeP	2004	Atom-based	71
FLARM2.0	2005	Atom-based	49
ConCept	2007	Fragment-based	82

structure. Nevertheless, as long as this error is systematic (that is, constant during the modelling study), it is still possible and justified to use such receptor surrogates for virtual screening.

Prerequisites for pseudoreceptor generation

As shown in FIG. 2, a molecular alignment of selected known actives represents the foundation of pseudoreceptor generation. As in any QSAR approach, ligand binding affinities should cover three to four orders of magnitude to allow for statistically plausible model development and testing. According to an assumed common binding mode, ligands are superimposed so that functionally equivalent chemical groups are arranged in a similar orientation. The assumption of a common binding mode is not necessarily true for every ligand of the same target^{44,45}. Therefore, it is advised to select reference molecules for which the binding mode is experimentally supported; for example, by means of site-directed mutagenesis⁴⁶.

The concept of key interacting groups asserted itself as a reliable tool to find such superimpositions. A problem in conformer alignment is the high degree of conformational freedom that flexible ligands have. It can happen that potential pharmacophoric points of reference molecules cannot be arranged according to a consensus bioactive conformation, or alignments are obtained that may have a low root mean square deviation but do not represent a common binding mode^{44,45}.

One approach to address this issue is to restrict pseudoreceptor generation to the most potent compounds only. The Active Analogue Approach, conceived by Marshall

three decades ago^{47,48}, proposes a set of plausible molecular conformers that yield different pharmacophore patterns. Thereby, it is assumed that the receptor-bound (bioactive) conformation of potent ligands is not strikingly dissimilar from the free low-energy state. An alternative approach is to derive the spatial co-location of reference compounds after the generation of a primary pseudoreceptor model obtained from the ligand with the highest affinity⁴⁰. Additional ligands join the model in an iterative generate-and-optimize process.

When pseudoreceptors are generated as a complementary strategy to conventional receptor-based methods, preliminary pseudoreceptor coordinates can be deduced from crystallographic data. Subsequent docking or force-field guided conformational evolution of remaining ligands into a receptor template can sometimes yield molecular alignments, as observed in superimpositions of crystallographic ligand data⁴⁹.

The initial alignment of pharmacophoric points is a crucial step, as it provides the starting point for pseudoreceptor generation, and it sets the course for optimization of the final model in terms of quality and applicability. Small changes of the pharmacophoric point alignment may result in completely different models. So, this initial factor is typically 'softened' either by generating a family of pseudoreceptor models — in which final predictions consist of a range of possible results⁵⁰ or their average^{49,51–53} — or by allowing translational and rotational changes of ligand conformations^{40,54,55}. The latter is usually accompanied with energy minimization of the opponent receptor surrogate during optimization tasks (see later).

Model generation and refinement

Grid-based models. Initial approaches for pseudoreceptor generation followed the intuitive idea of surrounding the molecular alignment of reference compounds with a cubic grid or lattice^{17,18,56,57} (FIG. 3a). Akin to prevalent 3D QSAR techniques, grid points are placed 0.25–2.0 Å from each other, generating a regularly spaced grid. At each grid point, several physicochemical ligand properties are recorded, such as electrostatic potential, partial charge distribution and hydrogen-bonding capabilities. When accurately defined, these models are abstract representations of multiple ligand molecules without reference to their underlying topographical and atomic structure but by exposed chemical features. Further model refinement can be achieved by straightforward superposition of additional reference compounds.

Isosurface-based models. To avoid potential issues that are associated with grid-based models — for example, the question of how to choose an optimal grid-point spacing — isosurface-based methods may be used for pseudoreceptor generation (FIG. 3b). Advanced approaches compile the recorded information to gain an analytically accessible surface by means of computer graphics tools; for instance, the marching cube algorithm⁵⁸. By using van der Waals or Wyvill functions (bounded functions that decay, yielding smoother surfaces)⁵⁹, aesthetic 3D embodiments of the putative binding site can be obtained^{60–62}.

Energy minimization

The optimization of molecule conformation towards idealized geometries by help of an energy function.

Bounded functions

Mathematical functions with a range that has an upper bound and a lower bound.

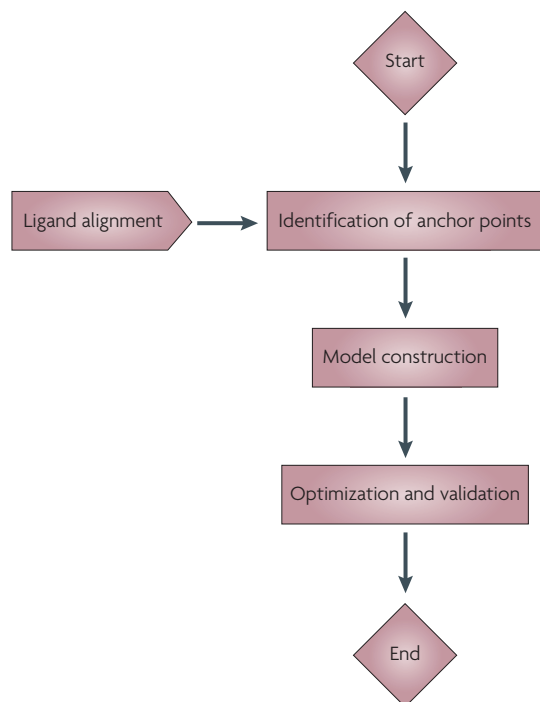


Figure 2 | **Workflow of pseudoreceptor model generation.** The first step is to define the ‘anchor points’ of the ligand–receptor complex. Next, the pseudoreceptor model is assembled around these hypotheses, and then the model coordinates are optimized to obtain more accurately calculated binding energies in validation studies.

Surfaces that envelop the ligand alignment are often represented by triangulated meshes⁶³. Their shapes deliver an imprint of the putative binding site and provide complementary accessible surfaces through which receptor and ligand can interact with each other in terms of the lock-and-key principle⁶⁴. This figurative representation accompanies the definition of explicit surfaces⁶⁰. They are assumed to be more substantial and correspond more closely to the actual physical nature of receptor binding sites as compared with implicit surfaces; for example, from the outcome of a CoMFA calculation²⁸. Implicit surfaces are generally based on multiple fields, which highlight the spatial variation of a feature. Calculation of chemical characteristics at explicit surface points is also possible for isosurface-based approaches⁶⁵.

The introduction of ‘open’ pseudoreceptors allows for discontinuous models: hypothetical receptor models with holes in their surface showing solvent-accessible regions or unknown parts of the binding site. Such models make sense when reference ligands are exposed to the solvent, or when the user wants to create a pseudoreceptor as a set of distinct surface patches. This can be accomplished by cropping areas from the surface by using primitive geometrical objects⁶⁰. More recent developments highlight the need of multiple enclosed surfaces²⁷. Here, an inner shell represents measures of the ligands with highest affinities, and an outer shell introduces ligands with lower affinities. The idea is to gain an approximation of

induced-fit or flexible-fit phenomena that adapt active-site shapes to the topology of the ligand^{27,66}. This strategy augments another prerequisite; that is, the calculation of multiple conformation states of reference molecules. Receptor surfaces that can be altered are represented by generation of an ensemble of receptor models (a pseudoreceptor family) instead of a single solution, and the resultant predictions are computed as averaged values from the individual pseudoreceptors.

Partition-based models. Conceptually different approaches drop the idea of grid-embedded or surface-embedded ligands. They build active-site models by partitioning the ligand environmental space into polyhedral Voronoi patches to represent distinct interaction subsites^{67,68} (FIG. 3c). Enhancements of this approach introduced the automatic clustering of ligand atoms by flowering-out spherical binding regions with variable radii centred at key interacting atoms⁵⁰. A similar idea was transferred into SKELGEN (De Novo Pharmaceuticals), which has its origin in structure-based *de novo* design^{69,70}. Dean and co-workers extended their program with a ‘ligand mode’ of operation that allows the input of Gaussian intrusion shapes. The user must supply a file containing a set of coordinates and radii of ‘featureless spheres’. Sphere borderlines enclose a virtual ligand volume, which can be combined with pharmacophoric constraints from known reference active molecules.

Atom-based models. In 1994, Walters and co-workers developed a technique (FIG. 3d) in which pseudoreceptors are represented as a set of atoms⁵¹. Atoms are placed on a sphere surrounding the largest molecules in the ligand alignment. One can imagine an inflated balloon with the ligand alignment inside and several marked points on its surface. By releasing air, the balloon shrinks around the ligands with increasing tightness, similar to a shrink-wrap. This is continued until the surface of the balloon touches the van der Waals surface of the ligands. Finally, atoms are placed at the marked positions. Usually, such a model consists of 40–60 atoms. Atom types are set in an arbitrary manner (for example, from the CHARMM force-field definitions). As a downside, a multidimensional optimization problem evolves because a vast number of models may be constructed. So, a refinement procedure incorporating a genetic algorithm was used⁵¹ (FIG. 4), in which pseudoreceptors are encoded as linear strings of atom types.

Chen and co-workers modified this procedure by gearing the initial generation of pseudoreceptor individuals⁵². Atom types were selected in a charge-dependent manner, considering the complementary electrostatic field of ligands. More recent developments introduced flexible receptor model atom points to avoid prediction errors of a fixed receptor shape⁵³. During refinement, atoms are allowed to move 0.1 nm from their original position. As many cases are encountered in which ligands are only partly enclosed in the binding site, open receptor models are also allowed. In 2005, Pei and co-workers⁴⁹ strengthened the flexibility and improved the genetic algorithm accordingly. They allowed conformational

Voronoi patches
A set of polygons (polytopes) that ascertain an optimal tessellation (tiling) of the encapsulated space.

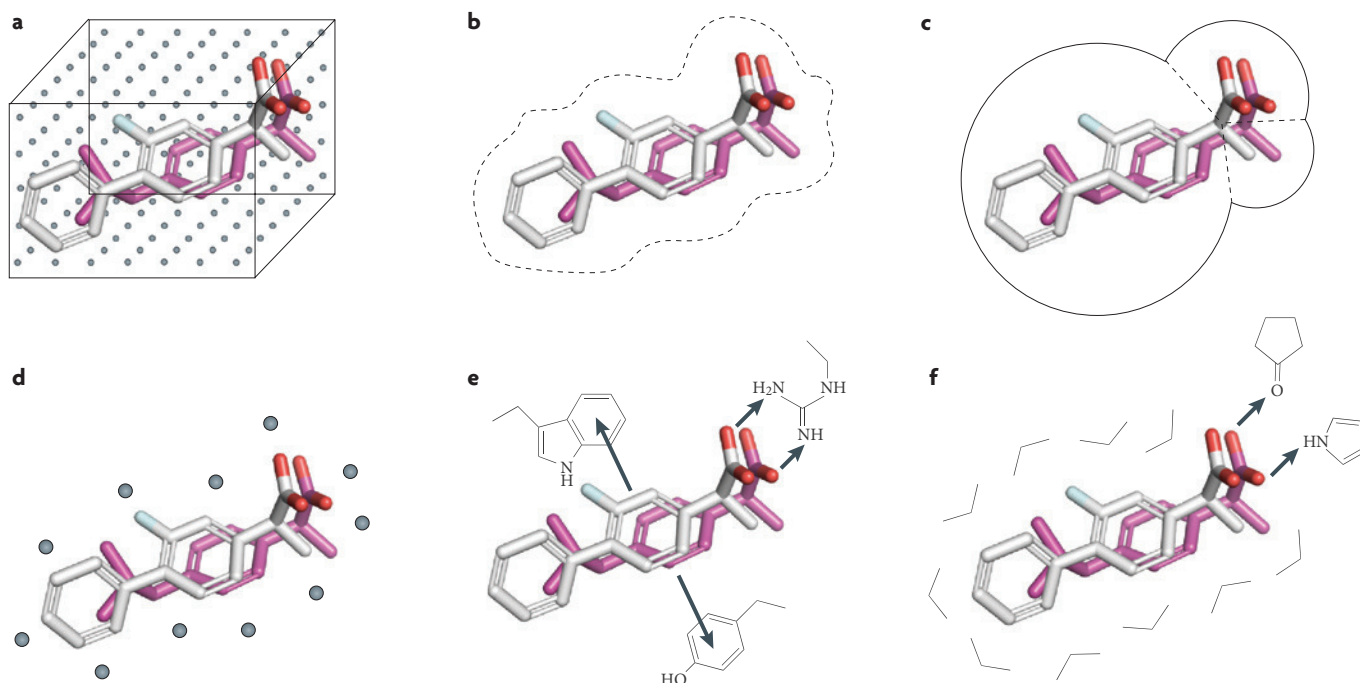


Figure 3 | **Pseudoreceptor construction routes.** Sketches outlining the concepts of grid-based (a); isosurface-based (b); partition-based (c); atom-based (d); peptide-based (e); and fragment-based (f) pseudoreceptor models around a structural alignment of flurbiprofen (light grey) and ibuprofen (magenta).

changes of reference ligands within pseudoreceptor models to be considered while the refinement process proceeds. The whole molecule is free to move in space by translational and rotational shifting. Consequently, ligand locations are rearranged to alternative orientations and placements. This might mimic realistic ligand–receptor binding processes, although mechanisms of flexible-fit phenomena are beyond the modelling capabilities of atomistic pseudoreceptor models. In future developments, one might consider ensembles of virtual receptors and cope with multiscale optimization by parallel computing methods⁴⁹.

Another member of this class of pseudoreceptor-generation programs is Quasar⁵⁴. As a main difference, Quasar supports hydrogen bond flip–flop atom types. These pseudoreceptor atoms have the ability to change their hydrogen-bond propensity depending on the features of the respective ligand. This is considered to be a more realistic description of hydrogen-bonding amino acids, as polar amino acids have variable hydrogen-bond geometries; that is, as hydrogen-bond donors or acceptors. Recently, a novel representative of atom-based pseudoreceptor models has been published, which uses the evolution strategy to find an optimal distribution of atom weights instead of atom types⁷¹.

Peptide-based models. Apart from placing singular atoms in the environment of ligands, approaches to pinpoint amino-acid residues at designated locations were made³⁸ (FIG. 3e). Other than the former models, spatial arrangement of putative receptor sites by a user-selected set of amino acids is based on hydrogen-bonding

geometry, which is derived from X-ray surveys. Vectors are used to delineate interactions from each ligand to suspected receptor sites:

- Hydrogen-extension vectors direct from ligand donor atoms (-NH-, -OH and -SH) towards receptor acceptor coordinates.
- Lone-pair vectors guide paths of potential hydrogen-bond interactions from ligand acceptor atoms (N, O and S) to receptor donor atoms.
- Hydrophobic vectors show interactions of lipophilic ligand and receptor regions.

The importance of the receptor interaction site is computed according to the assigned weight of each vector cluster. The weight depends on the number, variance and orientation of the clustered vectors. Suitable amino-acid conformations that satisfy the respective vector are taken from a rotamer library⁷². Optionally, target residues are added to the model to introduce higher-order structural motifs such as loops, helices or sheets. The pseudoreceptor model is refined by iterative conformational search and directionality-integrated energy minimization of the amino-acid residues.

In 1998, an extension of this concept was implemented, named PrGen⁴⁰. Pseudoreceptor models generated by PrGen are independent from the initial molecular superposition. Ligands are aligned onto the pharmacophore of a generated primordial pseudoreceptor, which was constructed with the most potent ligand. Most importantly, a novel iterative pseudoreceptor calibration algorithm (equilibration protocol) was integrated. Here, the receptor surrogate or a molecule of the ligand set is minimized, whereas all other atoms remain fixed. Optionally,

Rotamer library

A collection of residue conformations, each representing an optimum according to an energy function.

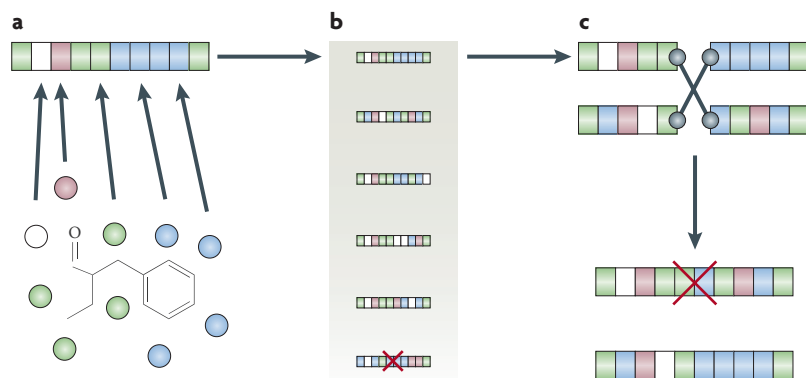


Figure 4 | Pseudoreceptor optimization with a genetic algorithm. Pseudoreceptors are encoded as linear strings of atom types. Every string position corresponds to particular coordinates in Cartesian space (a). A population of possible pseudoreceptors is generated at random, and each model corresponds to a gene/individual (b). From the assumption that bioactivity is proportional to the ligand–receptor interaction energy, each model is analysed to find the best ones. In the original application⁵¹, fitness was calculated as the regression coefficient of $1/\exp(\text{calculated intermolecular van der Waals and electrostatic energies of the model and each ligand})$ versus $\log(\text{bioactivity of each ligand})$. Ideally, model scores increase with a better regression. Pairs of highly ranked individuals are selected as parents to produce a new pair of offspring (c). Offspring individuals are generated by a crossover mutation at a randomly chosen string position. Optionally, a point mutation at a second chosen location can occur. Child models will be deleted if their score is below the worst score in the whole population of models. After a predetermined number of generations or when the average population fitness reaches a threshold, the refinement procedure is stopped, yielding a set of pseudoreceptor models.

a Monte Carlo conformational search can be used to find global minima of ligands⁷³. Resultant pseudoreceptors are scored according to their correlation coefficient of calculated and experimental ligand binding affinities. A noteworthy point is that the score is incorporated into energy calculations to directly affect model refinement.

An alternative way to find suitable receptor atom coordinates was introduced by Dean and co-workers, who performed a crystal survey yielding 3D hydrogen-bonding geometries⁷⁴. They implemented a method to unfurl ‘umbrellas’ around hydrogen-bonding ligand atoms, which displays favoured positions of receptor atoms according to the geometry of putative hydrogen bonds of the respective ligand atom⁷⁵. On this basis, a ligand superposition software package was developed⁷⁶ (application in REF. 77). LUDI^{78,79} and related receptor-based *de novo* design software is grounded on this principle⁸⁰. This method could be used for the development of a pseudoreceptor modelling program.

Fragment-based models. The final class of pseudoreceptor models is based on small-molecule fragments that are capable of interacting with ligand chemical groups (FIG. 3f). McMartin and co-workers published a software suite, which contains pseudoreceptor modelling packages⁸¹. Its module PSEUD is able to model a hypothetical binding site, which is complementary to a given pharmacophore template. The algorithm is not fully automatic, as the user has to place small hydrogen-bonding

fragments manually. Often, pyrrole or cyclopentanone are picked to mimic amino-acid residues. They are positioned close to reference ligands and connected by special bonds. Afterwards, the program fills the residual space with propane to mimic the carbon density of valine or phenylalanine. At the end, this yields a hydrophobic pocket that has isolated hydrogen-bond patches. The pseudoreceptor is energy minimized, in which the hydrogen bonds and the volume field serve as constraints. Final ligands can be used for virtual screening by automated ligand docking.

An entirely different intention was the starting point of the authors of ConConcept⁸², one of the most recent additions to pseudoreceptor research. ConConcept aims to design molecular receptor surrogates, which may replace full target proteins in biological assays. Notably, only one earlier similar approach is known, which was specialized on metal-containing receptors (Hostdesigner⁸³). Although the motivation is different, the *in silico* design process appears to be similar to pseudoreceptor-design algorithms. Obviously, this approach has to take additional restraints into account, such as the synthetic feasibility and energetic suitability (especially considering solvent effects) of designed small biomolecular receptors.

Pseudoreceptors and real binding pockets

A general assumption in receptor-based drug design is that ligand shapes and the corresponding receptor surface are complementary. This statement is valid depending on the level of abstraction one considers^{84,85}. It is important to realize that even for strong complexes, remarkable local discrepancies can be found, such as pocket regions that are unoccupied (voids), pairs of concave–concave and convex–convex surface patches, and gaps between the ligand and the receptor surface (FIG. 5). Such a situation cannot be captured by ligand-driven pseudoreceptor concepts, which attempt to create a virtual receptor surface. The receptor model will inevitably suffer from an optimistic design, in particular in atom-based approaches. When unoccupied regions of the real binding pocket are not captured, virtual screening will prevent the retrieval of compounds that have alternative binding modes. Also, over-detailed pseudoreceptor modelling has the risk of lacking biological relevance and drastically limits the conformational freedom of potential ligands.

FIGURE 6 gives two examples of simplistic pseudoreceptor models together with the experimentally determined receptor structures. In these successful case studies, putative residue positions of carazolol (a ligand of the β_2 adrenergic receptor (a GPCR)) and imatinib (Gleevec, Novartis; a ligand of human ABL kinase) were computed in the absence of the receptor. After superpositioning of predicted amino-acid volumes with the actual co-crystallized structures, a sound occupation of the pseudoreceptor models by amino acids that are essential for ligand binding and activity is observed. This example demonstrates that pseudoreceptor models can capture relevant aspects of receptor–ligand interaction geometry and function.

Monte Carlo conformational search
A technique for random sampling of molecular conformations, which is often combined with the Metropolis criterion for whether or not to accept a conformation.

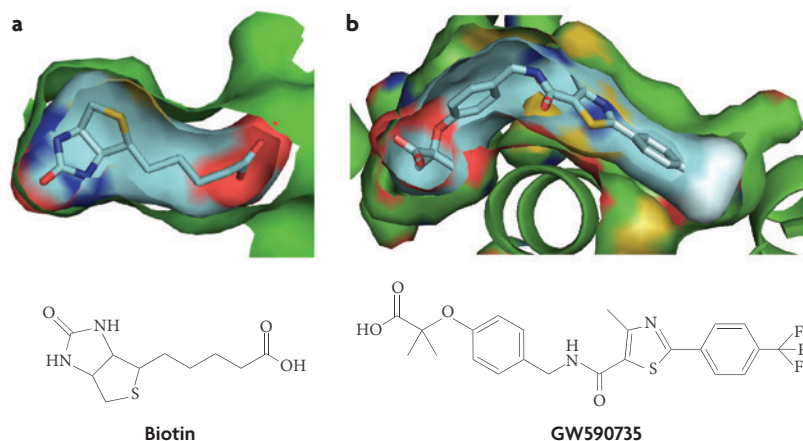


Figure 5 | Shape complementarity of ligands and their receptor binding pockets. **a** | An example of shape complementarity that is almost perfect. The tight fit between the ligand biotin and its receptor streptavidin ($K_d \sim 1 \times 10^{-16}$ M) is indicated by the solvent-accessible surfaces of the ligand and the receptor pocket (Protein Data Bank identifier: [2rtf](#), 1.47 Å resolution¹¹⁸). **b** | The complex formed by the ligand GW590735 and peroxisome proliferator-activated receptor- α ($EC_{50} = 4$ nM) (Protein Data Bank identifier: [2p54](#), 1.79 Å resolution¹¹⁹). Compared with case **(a)**, more gaps between the surfaces are visible in **(b)**.

Pseudoreceptor applications in drug design

Pseudoreceptor models fall into the class of 3D QSAR methods in computational chemistry. Their application has several common goals:

- Exploration of key ligand–receptor interaction sites.
- Evaluation of new candidate compounds.
- Estimation of interaction energies.
- Discerning the way in which a receptor protein sequence folds.
- Energy minimization of ligands in the pseudoreceptor to indicate their active conformation.
- Direct medicinal chemists in the design of novel compounds.

In this section, selected examples of recently published pseudoreceptor modelling projects are presented to demonstrate the utility and applicability of pseudoreceptors.

Analysis of 5-HT_{1A} thienopyrimidinones ligands.

5-hydroxytryptamine (serotonin) receptor 1A (5-HT_{1A}) ligands are thought to have effects on anxiety and depression^{86,87}. Guccione and co-workers investigated a lattice-based pseudoreceptor model (FIG. 3a) of the 5-HT_{1A} binding pocket, based on a set of 23 thienopyrimidinone-type ligands⁸⁸. A low-energy state of the ligand with the highest affinity was found by means of systematic conformational search and energy minimization, and was used as a rigid alignment template for the remaining 22 ligands. The authors generated two pseudoreceptor models. The first contained five superimposed conformations of each ligand found by Monte Carlo rotation sampling, and the second contained the single lowest-energy conformation of each ligand. Each model was validated by leave-one-out cross-validation. The authors used q^2 statistics to estimate the predictive power of their models (a note on this aspect has been published

previously⁸⁹). The resulting contour plots suggested a bulky substituent at the 4-methyl-thiophene, electron-rich or electron-poor atoms at the mercaptopyrimidine, and an electron-rich *ortho*-phenyl substituent at the piperazine. The authors concluded that the averaged multiconformer models comprised a predictive superiority over single conformer models, and stressed that the interpretation of single-molecule-derived pseudoreceptors is limited owing to resultant models that have less filled contours. Multiconformer models, on the other hand, allow the generation of contiguous contour plots (facilitating an interpretation) and average potential conformational states of the ligands. This application showed that pseudoreceptor modelling can be helpful to find key interaction sites and to suggest design strategies for synthesis of novel ligands. In addition, it proposes the consideration of multiple ligand conformations for pseudoreceptor generation.

Investigation of phenylalanine antagonists of VLA4.

The integrin very late antigen 4 (VLA4) is expressed on many types of blood cells⁹⁰, and is a potential target for immunosuppression in asthma^{91,92} and diabetes^{93,94}. The crystal structure of its endogenous ligand, vascular cell adhesion molecule 1 (VCAM1), was used to reveal novel antagonists⁹⁵. Phenylalanine derivatives belong to one class of such antagonists. Their binding mode was investigated by pseudoreceptor modelling with the program RECEPTOR⁹⁶, which belongs to the class of isosurface-based models (FIG. 3b). Site-directed mutagenesis revealed two key interactions of VCAM1 and VLA4, namely the carboxylic group of Asp40 and the centroid of the Leu70 side chain. The carboxylic groups and the centre of the aromatic phenylalanine rings of four phenylalanine derivatives were spatially arranged onto the common pharmacophore. An additional 128 phenylalanine derivatives were added to this primary ligand alignment before a pseudoreceptor model was constructed. Interaction energies were calculated at each surface point, some of which were used in a partial least squares analysis after variable selection. The final model was in agreement with site-directed mutagenesis experiments. Both studies suggested the same arrangement of ligand–receptor interaction sites. The model quality was supported by the docking of phenylalanine derivatives into the binding site of a VLA4 homology model. Proposed binding modes showed common aspects with the pseudoreceptor, and so the pseudoreceptor model may be used to screen for novel antagonists. It is noteworthy how well the pseudoreceptor modelling results complement the docking experiment in a case study by Macchiarulo and colleagues⁹⁶.

Investigation of ligand binding to a cocaine receptor.

Cocaine acts as a local anaesthetic and has multiple effects on the central nervous system. The best-studied site of action of cocaine is at the dopamine transporter protein, which shows two binding sites that differ in ligand affinity⁹⁷. Srivasta and Crippen used a partition approach (FIG. 3c) to build a pseudoreceptor model of the cocaine receptor by means of cocaine derivatives⁹⁸.

Leave-one-out cross-validation

A statistical method to evaluate the performance of prediction tools. A single observation from the original sample is used for validation, and all remaining observations are used as training data. This is repeated such that each observation in the sample is used once in the validation step.

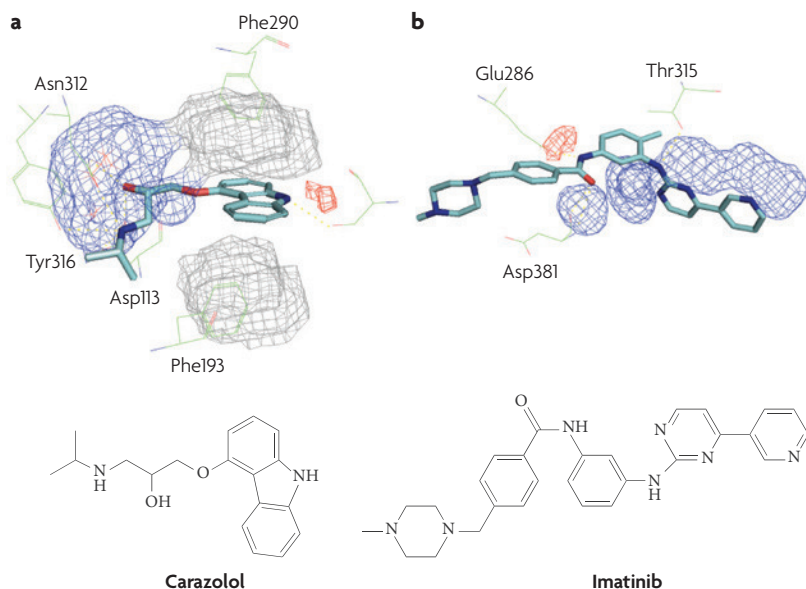


Figure 6 | Two case studies comparing ligand binding to a pseudoreceptor and their true biological receptor. **a** | Superposition of a simplistic pseudoreceptor showing predicted amino-acid positions in space and amino-acid positions of a co-crystallized structure of carazolol and the β_2 -adrenergic receptor (G-protein-coupled receptor subtype A; Protein Data Bank identifier: 2rh1)¹²⁰. Meshed volumes represent computed positions of aromatic (grey), hydrogen-bond donor capable (blue) and hydrogen-bond acceptor capable (red) amino acids. **b** | Superposition of pseudoreceptor and co-crystallized structure of the human ABL kinase ligand imatinib (Gleevec; Novartis) (Protein Data Bank identifier: [2hyy](#))¹²¹.

VOROM⁶⁷ was used to understand the mode of binding of eight cocaine stereoisomers and 12 analogues. Multiple conformations of each ligand were used to deduce a pseudoreceptor. For simplification, pseudo-atoms were placed at mean coordinates of ligand atoms. This yielded an averaged model of multiple ligand states. The resulting pseudoreceptor gives insights into the importance of ligand stereoselectivity and highlights the need for a hydrophobic substituent at the tropane ring of cocaine.

De novo design of non-steroid oestrogen receptor antagonists. Lloyd and co-workers used the *de novo* design tool SKELGEN to generate novel oestrogen receptor antagonists in the absence of structural target information⁹⁹. SKELGEN's ligand-mode demonstrates the utility of partition-based pseudoreceptor approaches for *de novo* design (FIG. 3c). The flexible alignment of five known oestrogen receptor antagonists yielded a hull with five pharmacophoric constraints. This shape was filled with suitable fragments, which were connected to final molecules. After clustering of the candidate molecules by molecular similarity, several non-steroidal scaffolds were suggested. All suggestions were able to mimic the steroid part of antagonists. The scaffolds contain a bicyclic ring connected to another aromatic system by one bond. These core elements are also observed in many known oestrogen receptor inhibitors. Consequently, this pioneering study shows that the use of a pseudoreceptor can be sufficient for successful *de novo* design.

Evaluation of solvation scenarios of the oestrogen receptor. Vedani and co-workers have periodically enhanced their software package Quasar⁵⁴ (FIG. 3d). It was equipped with the possibility to represent ligands by their ensemble of conformations (4D)^{100,101}, and multiple representations of induced-fit pseudoreceptor models (5D)^{102–104}. The most recent addition allows for simultaneous evaluation of different solvation scenarios (6D)¹⁰⁵. The usefulness of this technique was demonstrated by pseudoreceptor model generation of the oestrogen receptor.

The potential binding modes of 88 oestrogen receptor antagonists were analysed by constructing a multidimensional pseudoreceptor model in complement to a co-crystal structure of diethylstilbestrol and the α -domain of the target. The authors highlight the need for a pseudoreceptor that captures multiple ligand and binding-site states, as docking studies showed no consistency in the suggested binding poses. Each ligand was represented by four low-energy conformations from which a family of pseudoreceptor models was constructed. Solvation scenarios were simulated by optimizing each model in terms of its solvent accessibility. The final pseudoreceptor population identified the same key interactions as in the co-crystal structure. Interestingly, the authors show that the predictive power of the model improves with the introduction of additional dimensions, especially when the solvent accessibility of pseudoreceptors is considered.

Pseudoreceptor modelling of COX2. Pei and co-workers followed an atom-based approach (FIG. 3d) with the aim of generating a pseudoreceptor around a set of 50 known cyclooxygenase 2 (COX2) inhibitors^{49,106}. Ligands were aligned with the potent inhibitor SC-558 as a template. The authors note that the lowest-energy conformation of this inhibitor is in agreement with crystallographic data. All ligands were superimposed by starting from arranging two common ring systems, namely the aromatic system at the sulphonamide group and the diazole ring. It is noteworthy that the refined model showed modestly translated ligand coordinates. The model was successfully used to predict pIC_{50} values of ten test compounds that were not contained in the training set. In terms of ligand interacting sites, the pseudoreceptor shows an agreement with the binding pocket of the target protein.

Analysis of azole antifungal agents targeting cytochrome P450-dependent lanosterol 14 α -demethylase. In 2002, Botta and co-workers constructed a peptide-based pseudoreceptor model (FIG. 3e) of the enzyme cytochrome P450-dependent lanosterol 14 α -demethylase⁴³. This enzyme catalyses the first step of lanosterol conversion by removing a methyl group. Owing to its inhibition by antifungals, the concentration of methylated substrates rises and changes the membrane structure of fungi, thereby prohibiting growth.

The authors used 39 azole antifungal agents to construct a peptide binding site. According to another study¹⁰⁷, the ligand alignment was obtained from docking to a subset of 14 residues of a homology model of the

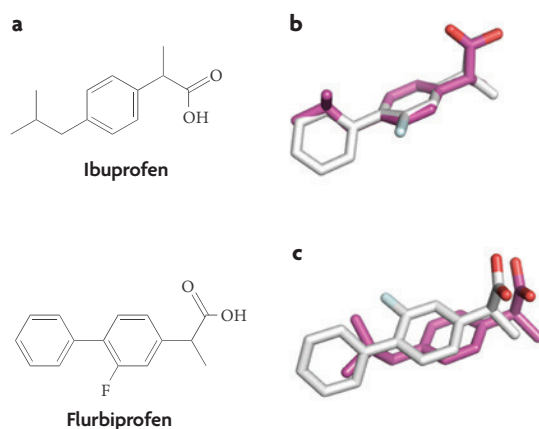


Figure 7 | Differing ligand alignments. This shows the structures of ibuprofen and flurbiprofen (**a**), and their alignment according to a flexible pharmacophore point-matching procedure (**b**), and by rigid superpositioning of the receptor co-crystal structures (**c**). The flexible alignment was performed with MOE (Chemical Computing Group). Receptor structure matching of Protein Data Bank entries [1eqq](#) and [1eqh](#)¹¹⁷ was performed with PyMol (DeLano Scientific).

target. After pseudoreceptor equilibration and relaxation of ligand coordinates by energy minimization, the pseudoreceptor model was able to predict experimental activities of 13 fungicides. It is noteworthy that the authors established a model to characterize the topography of the active site of the enzyme, which may now aid the design of novel inhibitors of pathogenic fungal enzymes.

Pseudoreceptor model of sweet-tasting molecules. Bassoli and co-workers modelled a peptide-based receptor site (FIG. 3e) with sweet-tasting molecules¹⁰⁸. This is a challenging task because the natural binding sites of sweetness taste receptors — TAS1R1, TAS1R2 and TAS1R3 — are unknown. Recent advances led to the hypothesis of a “one protein — multiple binding site” solution, in which one receptor binds molecules with dissimilar scaffolds^{109,110}.

The ligand set contained 24 molecules, which were aligned according to glucophoric functional groups. Thirteen amino acids were chosen for pseudoreceptor construction according to experiments with monoclonal antibodies that yielded information about interacting residues. The pseudoreceptor model was established in two steps. First, a model was deduced from guanidinium-containing compounds providing preliminary positions of amino acids of the pseudoreceptor binding site, and then the aligned ligands were minimized within the pseudoreceptor. The pseudoreceptor model was able to predict free energies of ligand binding for five test ligands. It was successfully validated by cross-validation and Y-scrambling. In summary, the authors were able to create a pseudoreceptor model that can recognize different types of sweet-tasting compounds and predict their sweet-tasting potency.

Y-scrambling

A statistical test of prediction tools, in which models are fitted for randomly reordered property/activity values and compared with the model obtained for the actual property/activity values.

Design of novel gene inducers of fetal haemoglobin.

Bohacek and co-workers used a fragment-based pseudo-receptor approach (FIG. 3f) to find new gene inducers that raise fetal haemoglobin production¹¹¹. Guanidinium and pyrrole fragments were placed around a low-energy pharmacophoric point alignment of six known antibody-inducing molecules. Residual volume was filled with propane to conserve the hydrophobic character of the pockets. After molecular dynamics relaxation and initial validation with an external ligand set, the pseudoreceptor was used for receptor-based screening by automated ligand docking. Property filtering of 13,000 commercially available compounds yielded a subset of 630 molecules that were docked into the model. *In vitro* testing of 26 manually selected molecules generated four novel gene-inducing molecules that had an activity of >200% compared with butyrate. This successful project is an exceptional example of a pseudoreceptor model that bridges ligand-based and receptor-based hit finding. It contains aspects of both ligand-based pharmacophore modelling and receptor-based virtual screening.

Conclusions

Several recent success stories demonstrate that pseudoreceptors can be useful tools in early stage hit and lead finding, thus complementing conventional ligand-based and receptor-structure-based virtual-screening campaigns. Initial applications show their potential in projects that lack an experimentally determined target structure. So, one might pose the question of why would one go to the effort of determining the 3D structures of protein–ligand complexes if all that is needed for drug discovery is a pseudoreceptor model? The answer is manifold, as pseudoreceptors have several limitations. Most importantly, they do not represent real macromolecular binding sites, but only a model that is generated from known ligands. Implicitly, finding novel ligand structures with potentially alternative binding modes in the same receptor pocket is impossible. Pseudoreceptor models should therefore be treated with caution because it is known that multiple ligand modes and binding sites may exist, even in one target. Consequently, the inverse assumption that all ligands that were used to construct the pseudoreceptor model have a common binding mode may not be true^{44,45}.

Another general issue of pseudoreceptors and related 3D QSAR techniques arises from the initial ligand alignment, which is typically performed on the basis of modelled ligand conformations alone. In FIG. 7, the alignment of ibuprofen and flurbiprofen is shown as an example. Both compounds inhibit COX1 by binding to the same protein pocket in an essentially identical binding mode. The ligand alignment shown in FIG. 7b is the result of flexible pharmacophore-based conformer matching (note that many tools are available for both rigid and flexible ligand alignment for pharmacophore modelling^{112,113}), and in FIG. 7c the alignment results from receptor structure superpositioning. The two alignments differ markedly, and a computed pseudoreceptor based on the alignment shown in

FIG. 7b does not, therefore, necessarily represent the situation found in the actual binding pocket (note, for example, the different orientation of the carboxylic acid groups in the two ligand alignments). It is evident that for pseudoreceptor generation in the absence of receptor structure information one can only perform ligand-based conformer alignment (FIG. 7b). We emphasize this potential discrepancy between the computed pseudoreceptor and the real structure of the ligand binding crevice.

Another caveat is related to the validation of QSAR models in general. Several studies have clearly demonstrated that a very careful statistical assessment of a (pseudoreceptor) model is essential, and computing q^2 values alone is insufficient^{89,114–116}. Model validity is particularly required when pseudoreceptors are used for subsequent virtual screening, ligand synthesis and biochemical testing. Keeping these caveats in mind, pseudoreceptor models represent a valuable addition to the toolbox of the drug designer.

1. Schneider, G. & Baringhaus, K. H. *Molecular Design: Concepts and Applications*. 1–277 (Wiley-VCH, Weinheim, 2008).
2. Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov. Today* **11**, 580–594 (2006).
3. Congreve, M., Murray, C. W., Blundell, T. L. Structural biology and drug discovery. *Drug Discov. Today* **10**, 895–907 (2005).
4. Carr, R. A., Congreve, M., Murray, C. W. & Rees, D. C. Fragment-based lead discovery: leads by design. *Drug Discov. Today* **10**, 987–992 (2005).
5. Bleicher, K. H., Bohm, H. J., Muller, K. & Alanine, A. I. Hit and lead generation: beyond high-throughput screening. *Nature Rev. Drug Discov.* **2**, 369–378 (2003).
6. Alber, J. S. et al. An integrated approach to fragment-based lead generation: philosophy, strategy and case studies from AstraZeneca's drug discovery programmes. *Curr. Top. Med. Chem.* **7**, 1600–1629 (2007).
7. Koehn, F. E. High impact technologies for natural products screening. *Prog. Drug Res.* **65**, 175, 177–210 (2008).
8. Schneider, G., Neidhart, W., Giller, T. & Schmid, G. "Scaffold-hopping" by topological pharmacophore search: a contribution to virtual screening. *Angew. Chem. Int. Ed.* **38**, 2894–2896 (1999).
9. Schneider, G., Schneider, P. & Renner, S. Scaffold-hopping: how far can you jump? *QSAR Comb. Sci.* **25**, 1162–1171 (2006).
10. Brown, N. & Jacoby, E. On scaffolds and hopping in medicinal chemistry. *Mini Rev. Med. Chem.* **6**, 1217–1229 (2006).
11. Zhao, H. Scaffold selection and scaffold hopping in lead generation: a medicinal chemistry perspective. *Drug Discov. Today* **12**, 149–155 (2007).
12. Carley, D. W. Drug repurposing: identify, develop and commercialize new uses for existing or abandoned drugs. Part I. *Drugs* **8**, 306–309 (2005).
A comprehensive description of the repurposing principle.
13. Carley, D. W. Drug repurposing: identify, develop and commercialize new uses for existing or abandoned drugs. Part II. *Drugs* **8**, 310–313 (2005).
14. Bisson, W. H. et al. Discovery of antiandrogen activity of nonsteroidal scaffolds of marketed drugs. *Proc. Natl Acad. Sci. USA* **104**, 11927–11932 (2007).
15. Mohan, V., Gibbs, A. C., Cummings, M. D., Jaeger, E. P. & DesJarlais, R. L. Docking: successes and challenges. *Curr. Pharm. Des.* **11**, 323–333 (2005).
16. Rester, U. Dock around the clock — current status of small molecule docking and scoring. *QSAR Comb. Sci.* **25**, 605–615 (2006).
17. Kato, Y., Itai, A. & Iitaka, Y. A novel method for superimposing molecules and receptor mapping. *Tetrahedron* **43**, 5229–5236 (1987).
18. Kato, Y., Inoue, A., Yamada, M., Tomioka, N. & Itai, A. Automatic superposition of drug molecules based on their common receptor site. *J. Comput. Aided Mol. Des.* **6**, 475–486 (1992).
19. Bernstein, F. C. et al. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542 (1977).
20. Berman, H. M. et al. The Protein Data Bank. *Nucl. Acids Res.* **28**, 235–242 (2000).
21. Scapin, G. Structural biology and drug discovery. *Curr. Pharm. Des.* **12**, 2087–2097 (2006).
22. Lacapere, J. J., Pebay-Peyroula, E., Neumann, J. M. & Etchebest, C. Determining membrane protein structures: still a challenge! *Trends Biochem. Sci.* **32**, 259–270 (2007).
23. Granseth, E., Seppala, S., Rapp, M., Daley, D. O. & Von Heijne, G. Membrane protein structural biology — how far can the bugs take us? *Mol. Membr. Biol.* **24**, 329–332 (2007).
24. Orry, A. J., Abagyan, R. A. & Cavasotto, C. N. Structure-based development of target-specific compound libraries. *Drug Discov. Today* **11**, 261–266 (2006).
25. Costanzi, S. On the applicability of GPCR homology models to computer-aided drug discovery: a comparison between *in silico* and crystal structures of the β_2 -adrenergic receptor. *J. Med. Chem.* **51**, 2907–2914 (2008).
26. Tropsha, A. & Golbraikh, A. Predictive QSAR modeling workflow, model applicability domains, and virtual screening. *Curr. Pharm. Des.* **13**, 3494–3504 (2007).
27. Lill, M. A. Multi-dimensional QSAR in drug discovery. *Drug Discov. Today* **12**, 1013–1017 (2007).
28. Cramer, R. D., Patterson, D. E. & Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **110**, 5959–5967 (1988).
29. Baroni, M. et al. Generating Optimal Linear PLS Estimations (COLPE): an advanced chemometric tool for handling 3D-QSAR problems. *Quant. Struct. Act. Relat.* **12**, 9–20 (1993).
30. Klebe, G., Abraham, U. & Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological potency. *J. Med. Chem.* **37**, 4130–4146 (1994).
31. Kubinyi, H., Hamprecht, F. A. & Mietzner, T. Three-dimensional quantitative similarity–activity relationships (3D QSiAR) from SEAL similarity matrices. *J. Med. Chem.* **41**, 2553–2564 (1998).
32. Momany, F., Pitha, R., Klimkovsky, V. J. & Venkatachalam, C. M. in *Expert Systems and Applications in Chemistry* (eds Hohne, B. A. & Pierce, T. H.) 82–91 (American Chemical Society, Washington DC, 1989).
33. Holtje, H. D. & Anzali, S. Molecular modelling studies on the digitalis binding site of the Na⁺/K⁺-ATPase. *Pharmazie* **47**, 691–697 (1992).
34. Davis, A. M., Gensmantel, N. P., Johansson, E. & Marriott, D. P. The use of the GRID program in the 3-D QSAR analysis of a series of calcium-channel agonists. *J. Med. Chem.* **37**, 963–972 (1994).
35. Snyder, J. P., Rao, S. N., Koehler, K. F. & Vedani, A. in *3D QSAR in Drug Design: Theory, Methods and Applications* (ed. Kubinyi, H.) 336–354 (Leiden/Escom, Dordrecht, 1993).
36. Schleifer, K. J. in *Pharmacophores and Pharmacophore Searches* (eds Langer, T. & Hoffmann, R. D.) 117–130 (Wiley-VCH, Weinheim, 2006).
A succinct overview of peptide-based pseudoreceptor applications.
37. Gurrath, M., Muller, G. & Holtje, H. D. in *3D QSAR in Drug Design: Recent Advances* (eds Kubinyi, H., Folkers, G. & Martin, Y. C.) 135–157 (Kluwers/Escom, Dordrecht, 1998).
38. Vedani, A., Zbinden, P. & Snyder, J. P. Pseudo-receptor modeling: a new concept for the three-dimensional construction of receptor binding sites. *J. Recept. Res.* **13**, 163–177 (1993).
39. Vedani, A., Zbinden, P., Snyder, J. P. & Greenidge, P. A. Pseudoreceptor modeling: the construction of three-dimensional receptor surrogates. *J. Am. Chem. Soc.* **117**, 4987–4994 (1995).
40. Zbinden, P., Dobler, M., Folkers, G. & Vedani, A. PrGen: pseudoreceptor modeling using receptor-mediated ligand alignment and pharmacophore equilibration. *Quant. Struct. Act. Relat.* **17**, 122–130 (1998).
41. Galvez-Ruano, E., Iriepa-Canalda, I., Morreale, A. & Lipkowitz, K. B. A computational model of the nicotinic acetylcholine binding site. *J. Comput. Aided Mol. Des.* **13**, 57–68 (1999).
42. Freceer, V., Ho, B. & Ding, J. L. Interpretation of biological activity data of bacterial endotoxins by simple molecular models of mechanism of action. *Eur. J. Biochem.* **267**, 837–852 (2000).
43. Botta, M., Corelli, F., Manetti, F. & Tafi, A. Molecular modeling as a powerful technique for understanding small-large molecules interactions. *Farmaco* **57**, 153–165 (2002).
44. Kim, K. H. Outliers in SAR and QSAR: is unusual binding mode a possible source of outliers? *J. Comput. Aided Mol. Des.* **21**, 63–86 (2007).
45. Kim, K. H. Outliers in SAR and QSAR: 2. Is a flexible binding site a possible source of outliers? *J. Comput. Aided Mol. Des.* **21**, 421–435 (2007).
46. Evers, A., Hessler, G., Matter, H. & Klabunde, T. Virtual screening of biogenic amine-binding G-protein coupled receptors: comparative evaluation of protein- and ligand-based virtual screening protocols. *J. Med. Chem.* **48**, 5448–5465 (2005).
47. Marshall, G. R., Barry, C. D., Bosshard, H. E., Dammkoehler, R. A. & Dunn, D. A. in *Computer-Assisted Drug Design* (eds Olson, E. C. & Christoferson, R. E.) 205–226 (American Chemical Society, Washington DC, 1979).
Pioneering work that represents the basic principles of modern modelling software packages.
48. Marshall, G. R. in *3D QSAR in Drug Design: Theory Methods and Applications* (ed. Kubinyi, H.) 80–116 (Leiden/Escom, Dordrecht, 1993).
49. Pei, J. et al. Improving the quality of 3D-QSAR by using flexible-ligand receptor models. *J. Chem. Inf. Model* **45**, 1920–1935 (2005).
50. Crippen, G. M. Validation of EGSITE2, a mixed integer program for deducing objective site models from experimental data. *J. Med. Chem.* **40**, 3161–3172 (1997).
51. Walters, D. E. & Hinds, R. M. Genetically evolved receptor models: a computational approach to construction of receptor models. *J. Med. Chem.* **37**, 2527–2536 (1994).
Presentation of the first atom-based pseudoreceptor algorithm, including a genetic optimization algorithm.
52. Chen, H., Zhou, J. & Xie, G. PARM: a genetic evolved algorithm to predict bioactivity. *J. Chem. Inf. Comput. Sci.* **38**, 243–250 (1998).
53. Pei, J. & Zhou, J. Flexible atom receptor model. *Acta Chim. Sin.* **60**, 973–979 (2002).
54. Vedani, A. & Zbinden, P. Quasi-atomistic receptor modeling. A bridge between 3D QSAR and receptor fitting. *Pharm. Acta Helv.* **73**, 11–18 (1998).
55. Lill, M. A., Vedani, A. & Dobler, M. Raptor: combining dual-shell representation, induced-fit simulation, and hydrophobicity scoring in receptor modeling: application toward the simulation of structurally diverse ligand sets. *J. Med. Chem.* **47**, 6174–6186 (2004).
56. Andrews, P. R. et al. Morpheus: a conformation–activity relationships and receptor modeling package. *J. Mol. Graph.* **7**, 138–145 (1989).
57. Doweyko, A. M. The hypothetical active site lattice. An approach to modeling active sites from data on inhibitor molecules. *J. Med. Chem.* **31**, 1396–1406 (1988).
58. Delibasis, K. S., Matsopoulos, G. K., Mouravliansky, N. A. & Nikita, K. S. A novel and efficient implementation of the marching cubes algorithm. *Comput. Med. Imaging Graph.* **25**, 343–352 (2001).

59. Wyvill, G., McPheeters, C. & Wyvill, B. Data structures for soft objects. *Vis. Comp.* **2**, 227–234 (1986).
60. Hahn, M. Receptor surface models. 1. Definition and construction. *J. Med. Chem.* **38**, 2080–2090 (1995).
61. Hahn, M. & Rogers, D. Receptor surface models. 2. Application to quantitative structure–activity relationship studies. *J. Med. Chem.* **38**, 2091–2102 (1995).
62. Hahn, M. & Rogers, D. in *3D QSAR in Drug Design: Recent Advances* (eds Kubinyi, H., Folkers, G. & Martin, Y. C.) 117–133 (Kluwer/Escom, Dordrecht, 1998).
63. Heiden, W., Schlenkrich, M. & Brickmann, J. Triangulation algorithms for the representation of molecular surface properties. *J. Comput. Aided Mol. Des.* **4**, 225–269 (1990).
64. Fischer, E. Einfluss der configuration auf die wirkung der enzyme. *Ber. Deutsch Chem. Ges.* **27**, 2985–2993 (1894) (in German).
65. Proschak, E., Rupp, M., Derksen, S. & Schneider, G. Shapelets: possibilities and limitations of shape-based virtual screening. *J. Comput. Chem.* **29**, 108–114 (2008).
66. Kirchmair, J. *et al.* Fast and efficient *in silico* 3D screening: toward maximum computational efficiency of pharmacophore-based and shape-based approaches. *J. Chem. Inf. Model.* **47**, 2182–2196 (2007).
67. Crippen, G. M. Voronoi binding site models. *NIDA Res. Monogr.* **112**, 7–20 (1991).
68. Bradley, M., Richardson, W. & Crippen, G. M. Deducing molecular similarity using Voronoi binding sites. *J. Chem. Inf. Comput. Sci.* **33**, 750–755 (1993).
69. Todorov, N. P. & Dean, P. M. Evaluation of a method for controlling molecular scaffolds diversity in *de novo* ligand design. *J. Comput. Aided Mol. Des.* **11**, 175–192 (1997).
70. Todorov, N. P. & Dean, P. M. A branch-and-bound method for optimal atom-type assignment in *de novo* ligand design. *J. Comput. Aided Mol. Des.* **12**, 335–350 (1998).
71. Chae, C. H., Yoo, S. E. & Shin, W. Novel receptor surface approach for 3D-QSAR: The weighted probe interaction energy method. *J. Chem. Inf. Comput. Sci.* **44**, 1774–1787 (2004).
72. Ponder, J. W. & Richards, F. M. Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* **193**, 775–791 (1987).
73. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092 (1953).
74. Mills, J. E. J. & Dean, P. M. Three-dimensional hydrogen-bond geometry and probability information from crystal survey. *J. Comput. Aided Mol. Des.* **10**, 607–622 (1996).
- An all-embracing crystal survey of hydrogen-bond geometries in protein–ligand complexes.**
75. Mills, J. E., Perkins, T. D. J. & Dean, P. M. An automated method for predicting the positions of hydrogen-bonding atoms in binding sites. *J. Comput. Aided Mol. Des.* **11**, 229–242 (1997).
76. Mills, J. E., de Esch, I. J., Perkins, T. D. & Dean, P. M. SLATE: a method for the superposition of flexible ligands. *J. Comput. Aided Mol. Des.* **15**, 81–96 (2001).
77. De Esch, I. J. *et al.* Development of a pharmacophore model for histamine H₂ receptor antagonists, using the newly developed molecular modeling program SLATE. *J. Med. Chem.* **44**, 1666–1674 (2001).
78. Bohm, H. J. The computer program LUDI: a new method for the *de novo* design of enzyme inhibitors. *J. Comput. Aided Mol. Des.* **6**, 61–78 (1992).
79. Bohm, H. J. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comput. Aided Mol. Des.* **6**, 593–606 (1992).
80. Schneider, G. & Fechner, U. Computer-based *de novo* design of drug-like molecules. *Nature Rev. Drug Discov.* **4**, 649–663 (2005).
81. McMartin, C. & Bohacek, R. J. QXP: powerful, rapid computer algorithms for structure-based design. *J. Comput. Aided Mol. Des.* **11**, 333–344 (1997).
82. Chen, W. & Gilson, M. K. Concept: *de novo* design of synthetic receptors for targeted ligands. *J. Chem. Inf. Model.* **47**, 425–434 (2007).
83. Hay, B. P. & Firman, T. K. Hostdesigner: a program for the *de novo* structure-based design of molecular receptors with binding sites that complement metal ion guests. *Inorg. Chem.* **41**, 5502–5512 (2002).
84. Norel, R., Wolfson, H. J. & Nussinov, R. Small molecule recognition: solid angles surface representation and molecular shape complementarity. *Comb. Chem. High Throughput Screen.* **2**, 223–237 (1999).
85. Zauhar, R. J., Moyna, G., Tian, L., Li, Z. & Welsh, W. J. Shape signatures: a new approach to computer-aided ligand- and receptor-based drug design. *J. Med. Chem.* **46**, 5674–5690 (2003).
86. Blier, P. & Montigny, C. Current advances and trends in the treatment of depression. *Trends Pharmacol. Clin. Sci.* **15**, 220–226 (1994).
87. Broekkamp, C. L., Leysen, D., Peeters, B. W. & Pinder, R. M. Prospects for improved antidepressants. *J. Med. Chem.* **38**, 4615–4633 (1995).
88. Guccione, S., Doveyko, A. M., Chen, H., Barretta, G. U. & Balzano, F. 3D-QSAR using 'multiconformer' alignment: the use of HASL in the analysis of 5-HT_{1A} thienopyrimidinone ligands. *J. Comput. Aided Mol. Des.* **14**, 647–657 (2000).
89. Golbraikh, A. & Tropsha, A. Beware of q²! *J. Mol. Graph. Model.* **20**, 269–276 (2002).
- Fundamental work indicating obstacles in model-validation studies.**
90. Hynes, R. O. Integrins: bidirectional, allosteric signaling machines. *Cell* **110**, 673–687 (2002).
91. Abraham, W. M. *et al.* Alpha 4-integrins mediate antigen-induced late bronchial responses and prolonged airway hyperresponsiveness in sheep. *J. Clin. Invest.* **93**, 776–787 (1994).
92. Sagara, H. *et al.* A monoclonal antibody against very late activation antigen-4 inhibits eosinophil accumulation and late asthmatic response in guinea pig model of asthma. *Int. Arch. Allergy Immunol.* **112**, 287–294 (1997).
93. Yang, X. D., Karin, N., Tisch, R., Steinman, L. & McDevitt, H. O. Inhibition of insulin and prevention of diabetes in nonobese diabetic mice by blocking L-selectin and very late antigen 4 adhesion receptors. *Proc. Natl Acad. Sci. USA* **90**, 10494–10498 (1993).
94. Burkly, L. C., Jakubowski, A. & Hattori, M. Protection against adoptive transfer of autoimmune diabetes mediated through very late antigen-4 integrin. *Diabetes* **43**, 529–534 (1994).
95. Lin, K. C. & Castro, A. C. Very late antigen 4 (VLA4) antagonists as antiinflammatory agents. *Curr. Opin. Chem. Biol.* **2**, 453–457 (1998).
96. Macchiarulo, A. *et al.* Insights into phenylalanine derivatives recognition of VLA-4 integrin: from a pharmacophoric study to 3D-QSAR and molecular docking analyses. *J. Chem. Inf. Comput. Sci.* **44**, 1829–1839 (2004).
97. Carrol, F. L., Lewin, A. H., Boja, J. W. & Kuhar, M. J. Cocaine receptor: Biochemical characterization and structure–activity relationships of cocaine analogues at the dopamine transporter. *J. Med. Chem.* **35**, 969–981 (1992).
98. Srivastava, S. & Crippen, G. M. Analysis of cocaine receptor site ligand binding by three-dimensional Voronoi site modeling approach. *J. Med. Chem.* **36**, 3572–3579 (1993).
99. Lloyd, D. G., Buenemann, C. L., Todorov, N. P., Manallack, D. T. & Dean, P. M. Scaffold hopping in *de novo* design. Ligand generation in the absence of receptor information. *J. Med. Chem.* **47**, 493–496 (2004).
100. Vedani, A., Briem, H., Dobler, M., Dollinger, K. & McMasters, D. R. Multiple conformation and protonationstate representation in 4D-QSAR: the neurokinin-1 receptor system. *J. Med. Chem.* **43**, 4416–4427 (2000).
101. Vedani, A., McMasters, D. R. & Dobler, M. Multiconformational ligand representation in 4D-QSAR: reducing the bias associated with ligand alignment. *Quant. Struct. Act. Relat.* **19**, 149–161 (2000).
102. Vedani, A. & Dobler, M. 5D-QSAR: the key for simulating induced fit? *J. Med. Chem.* **45**, 2139–2149 (2002).
103. Vedani, A. & Dobler, M. Multidimensional QSAR: moving from three- to five dimensional concepts. *Quant. Struct. Act. Relat.* **21**, 382–390 (2002).
104. Vedani, A. *et al.* Novel ligands for the chemokine receptor-3 (CCR3): a receptor-modeling study based on 5D-QSAR. *J. Med. Chem.* **48**, 1515–1527 (2005).
105. Vedani, A., Dobler, M. & Lill, M. A. Combining protein modeling and 6D-QSAR. Simulating the binding of structurally diverse ligands to the estrogen receptor. *J. Med. Chem.* **48**, 3700–3703 (2005).
106. Liu, H. *et al.* Inhibitory mode of 1,5-diarylpyrazole derivatives against cyclooxygenase-2 cyclooxygenase-1: molecular docking and 3D-QSAR analyses. *J. Med. Chem.* **45**, 4816–4827 (2002).
107. Jansen, J. M. *et al.* Molecular design using the minireceptor concept. *J. Chem. Inf. Comput. Sci.* **37**, 812–818 (1997).
108. Bassoli, A., Drew, M. G., Merlini, L. & Morini, G. General pseudoreceptor model for sweet compounds: a semiquantitative prediction of binding affinity for sweet-tasting molecules. *J. Med. Chem.* **45**, 4402–4409 (2002).
109. Temussi, P. The history of sweet taste: not exactly a piece of cake. *J. Mol. Recognit.* **19**, 188–199 (2006).
110. Temussi, P. The sweet taste receptor: a single receptor with multiple sites and modes of interaction. *Adv. Food Nutr. Res.* **53**, 199–239 (2007).
111. Bohacek, R., Boosalis, M. S., McMartin, C., Faller, D. V. & Perrine, S. P. Identification of novel small-molecule inducers of fetal hemoglobin using pharmacophore and 'PSEUDO' receptor models. *Chem. Biol. Drug Des.* **67**, 318–328 (2006).
- Pioneering prospective approach of the pseudoreceptor concept for finding hits.**
112. Khedkar, S. A., Malde, A. K., Coutinho, E. C. & Srivastava, S. Pharmacophore modeling in drug discovery and development: an overview. *Med. Chem.* **3**, 187–197 (2007).
113. Wolber, G., Seidel, T., Bendix, F. & Langer, T. Molecule-pharmacophore superpositioning and pattern matching in computational drug design. *Drug Discov. Today* **13**, 23–29 (2008).
114. Golbraikh, A. *et al.* Rational selection of training and test sets for the development of validated QSAR models. *J. Comput. Aided Mol. Des.* **17**, 241–253 (2003).
115. Baumann, K. & Stiefl, N. Validation tools for variable subset regression. *J. Comput. Aided Mol. Des.* **18**, 549–562 (2004).
116. Gramatica, P. Principles of QSAR models validation: internal and external. *QSAR Comb. Sci.* **26**, 694–701 (2007).
117. Selinsky, B. S., Gupta, K., Sharkey, C. T. & Loll, P. J. Structural analysis of NSAID binding by prostaglandin H₂ synthase: time-dependent and time-independent inhibitors elicit identical enzyme conformations. *Biochemistry* **40**, 5172–5180 (2001).
118. Katz, B. A. Binding of biotin to streptavidin stabilizes intersubunit salt bridges between Asp61 and His87 at low pH. *J. Mol. Biol.* **274**, 776–800 (1997).
119. Sierra, M. L. *et al.* Substituted 2-[[4-(aminomethyl)phenoxy]-2-methylpropionic acid PPAR agonists. 1. Discovery of a novel series of potent HDLc raising agents. *J. Med. Chem.* **50**, 685–695 (2007).
120. Cherezov, V. *et al.* High-resolution crystal structure of an engineered human β_2 -adrenergic G protein-coupled receptor. *Science* **318**, 1258–1265 (2007).
121. Cowan-Jacob, S. W. *et al.* Structural biology contributions to the discovery of drugs to treat chronic myelogenous leukaemia. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 80–93 (2007).

Acknowledgements

N. Todoroff is thanked for technical support. This work was supported by the Beilstein Institute for the Advancement of Chemical Sciences (Beilstein Institut zur Förderung der Chemischen Wissenschaften), Frankfurt am Main, Germany.

DATABASES

RSCB Protein Data Bank:

<http://www.rcsb.org/pdb/home/home.do?1eqg|1eqh|2hyh|2p54|2rtf>

FURTHER INFORMATION

Laboratory of Gisbert Schneider: <http://www.modlab.de>

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

Scaffold-Hopping mit „fuzzy“ Pharmakophoren

Scaffold-hopping by “Fuzzy” Pharmacophores and Application to RNA Targets

Tanrikulu Y, Nietert M, Scheffer U, Proschak E, Grabowski K,
Schneider P, Weidlich M, Karas M, Göbel M, Schneider G

ChemBioChem 2007; 8: 1932-1936

Virtuelles Screening und *de novo* Design dienen in der Medizinischen Chemie als Ideengeneratoren. Hierbei ist die Fähigkeit des Scaffold-Hoppings ausschlaggebend. Die Aufgabe ist es, Moleküle mit gleicher Funktionalität aber unterschiedlichem Strukturgrundgerüst zu finden. In dieser Arbeit wird die erfolgreiche Anwendung der “fuzzy” Pharmakophormethode LIQUID (LIgand-based QUantification of Interaction Distributions) zur Suche nach neuen HIV RNA Liganden vorgestellt.

DOI: 10.1002/cbic.200700195

Scaffold Hopping by “Fuzzy” Pharmacophores and its Application to RNA Targets

Yusuf Tanrikulu,^[a] Manuel Nietert,^[a] Ute Scheffer,^[a] Ewgenij Proschak,^[a] Kristina Grabowski,^[a] Petra Schneider,^[b] Markus Weidlich,^[c] Michael Karas,^[c] Michael Göbel,^[a] and Gisbert Schneider^{*,[a]}

Ideally, a novel lead structure represents a different chemotype from known ligands of a given target macromolecule. Virtual screening techniques and molecular de novo design are able to perform “scaffold-hopping”, and can thus be used as idea generators for medicinal chemistry.^[1–3] The task is to find iso-functional bioactive molecules with different backbone architectures.^[1] Here, we present the successful application of our “fuzzy” pharmacophore method LIQUID (“ligand-based quantification of interaction distributions”)^[4] to finding novel RNA ligand scaffolds and new inhibitors of in vitro protein expression.

LIQUID is akin to our previously introduced SQUID approach.^[5,6] Both techniques model pharmacophoric features as Gaussian densities. While SQUID operates with univariate distributions, LIQUID employs trivariate models. In a retrospective virtual-screening study, we first compared LIQUID and SQUID, which both work on 3D pharmacophores, with the corresponding 2D CATS approach.^[1] All three methods encode pharmacophoric information as a correlation vector; this allows for rapid alignment-free similarity searching (by pair-wise Euclidian distance) in large compound libraries.^[7] We assessed their ability to retrieve known ligands of six drug targets from a collection of drugs and lead compounds compiled from recent literature.^[8] Overall, the CATS 2D method outperformed the 3D methods in this study, and SQUID retrieved slightly more hits than LIQUID (Table 1). This outcome is not entirely surprising, as we employed only a single calculated conformation for each of the molecules, and owing to the fact that the trivariate pharmacophore points in the LIQUID models are more restrictive than their univariate counterparts in SQUID. Notably, within the error margins, all methods performed comparably

Table 1. Average enrichment of active species expressed as average enrichment factor $\langle ef \rangle = \text{hits}_{\text{found}} / \text{hits}_{\text{expected}} (\pm \text{SD})$,^[25] yielded by retrospective virtual screening.^[a]

Target	CATS (2D)	SQUID (3D)	LIQUID (3D)
ACE (N=34)	26 ± 16	35 ± 21	21 ± 17
DHFR (N=64)	30 ± 10	15 ± 11	10 ± 7
Thrombin (N=183)	17 ± 9	8 ± 6	12 ± 8
COX-2 (N=136)	11 ± 8	14 ± 10	10 ± 8
PPARγ (N=44)	13 ± 8	6 ± 5	5 ± 5
FXa (N=228)	9 ± 6	5 ± 3	5 ± 3

[a] Three correlation-vector methods were compared for their ability to retrieve known active compounds in the first percentile (N=83) of a ranked collection of drugs and lead structures (N=8,311). COX-2: cyclooxygenase 2, ACE: angiotensin converting enzyme, DHFR: dihydrofolate reductase, PPARγ: peroxisome proliferator-activated receptor gamma, FXa: factor Xa.

and clearly demonstrated their ability to retrieve active species from a collection of drug-like compounds.

As all targets employed for this preliminary retrospective assessment were proteins, we checked the scaffold-hopping abilities of SQUID and LIQUID by using an RNA target, namely the *trans*-activation response element (TAR) of human immunodeficiency virus (HIV) 1. AIDS is caused by infection with HIV, which belongs to the class of retroviruses. This viral class is able to transpose its genome and use the host’s replication machinery for proliferation. Regulatory viral units have been selected as targets for blocking viral replication.^[9] Specific interaction of TAR RNA with the Tat protein is essential for virus

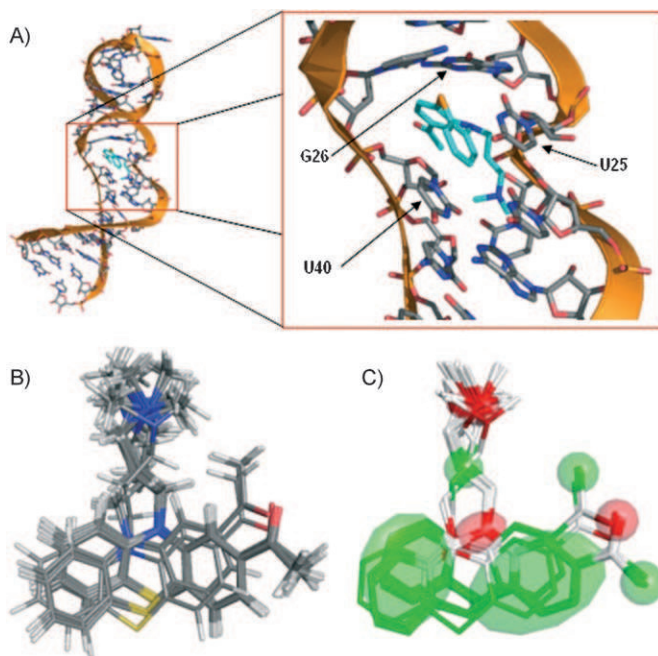


Figure 1. A) Structure of TAR RNA with bound acetylpromazine 1 (model 1, PDB ID: 1lvj).^[16] B) 20 superimposed docking poses of 1. C) “Fuzzy” pharmacophore model of (B). The colored blobs represent trivariate Gaussian ellipsoids indicating potential clusters of lipophilic interactions (green) and hydrogen-bond acceptors (red). For visualization, the shape of the ellipsoids was defined by their standard deviations in each direction.

[a] Y. Tanrikulu, M. Nietert, Dr. U. Scheffer, E. Proschak, K. Grabowski, Prof. Dr. M. Göbel, Prof. Dr. G. Schneider
Institute of Organic Chemistry and Chemical Biology
Johann Wolfgang Goethe University
Siesmayerstrasse 70, 60323 Frankfurt am Main (Germany)
Fax: (+49) 69-798-24880
E-mail: gisbert.schneider@modlab.de

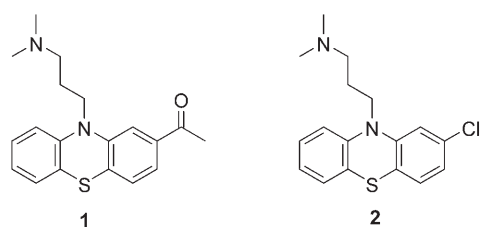
[b] Dr. P. Schneider
Schneider Consulting GbR
George-C.-Marshall Ring 33, 61440 Oberursel (Germany)

[c] M. Weidlich, Prof. M. Karas
Institute of Pharmaceutical Chemistry, Johann Wolfgang Goethe-University
Max-von-Laue-Strasse 9, 60438 Frankfurt am Main (Germany)

Supporting information for this article is available on the WWW under <http://www.chembiochem.org> or from the author.

replication.^[10] The TAR RNA fragment contains a bulge region formed by the nucleosides U25, G26, and U40 (Figure 1 A). This bulge is considered a primary protein–ligand interaction site.^[11] A potent bulge ligand should inhibit the Tat–TAR interaction, thereby potentially interrupting viral replication. Leadoff structure-based molecular modeling approaches had identified acetylpromazine (1; Scheme 1) and chlorpromazine (2) as TAR RNA binders.^[12]

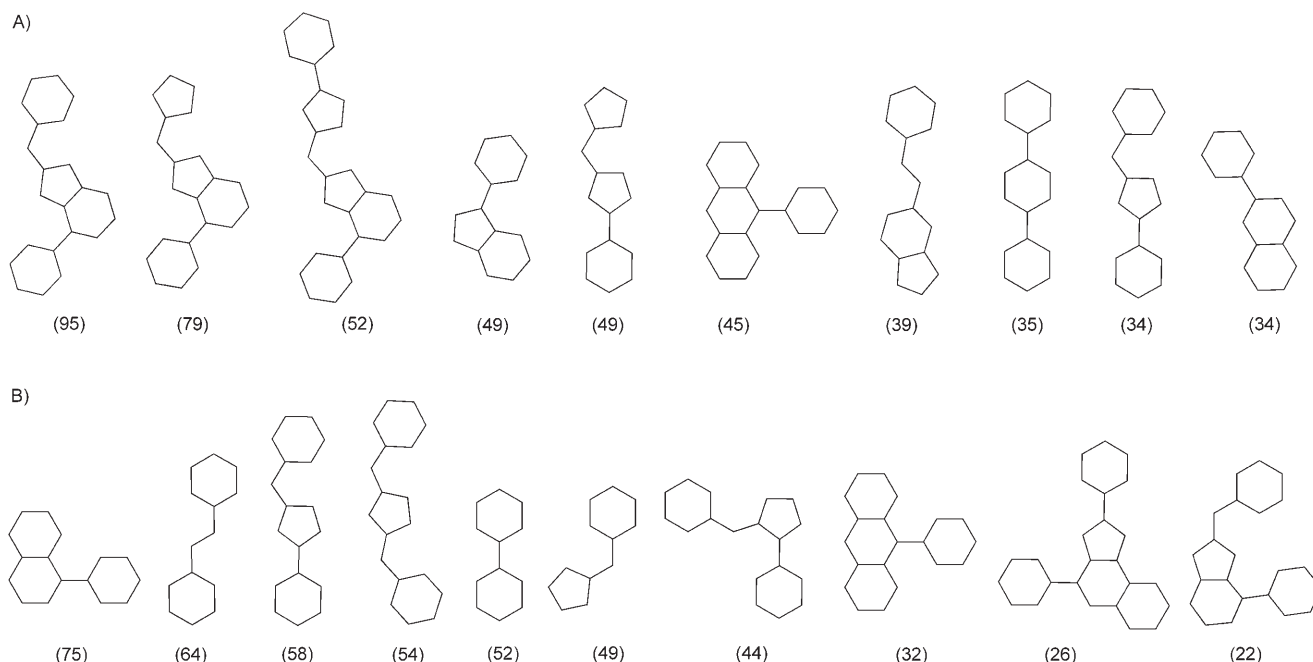
By using acetylpromazine 1 as query, SQUID retrieved 576 different molecular graph scaffolds^[13] among the top-ranking 1960 molecules from the SPECS compound library^[14] (v.2007/05, 196075 compounds). LIQUID retrieved slightly fewer scaffolds (475). Again, this lower number was expected as a consequence of using the more restrictive trivariate Gaussian pharmacophore feature densities (Scheme 2). Notably, there are only a few overlapping scaffolds among the top-scoring ones in both virtual hit lists. This emphasizes that the methods complement each other, despite their seemingly small differences in pharmacophore representation.



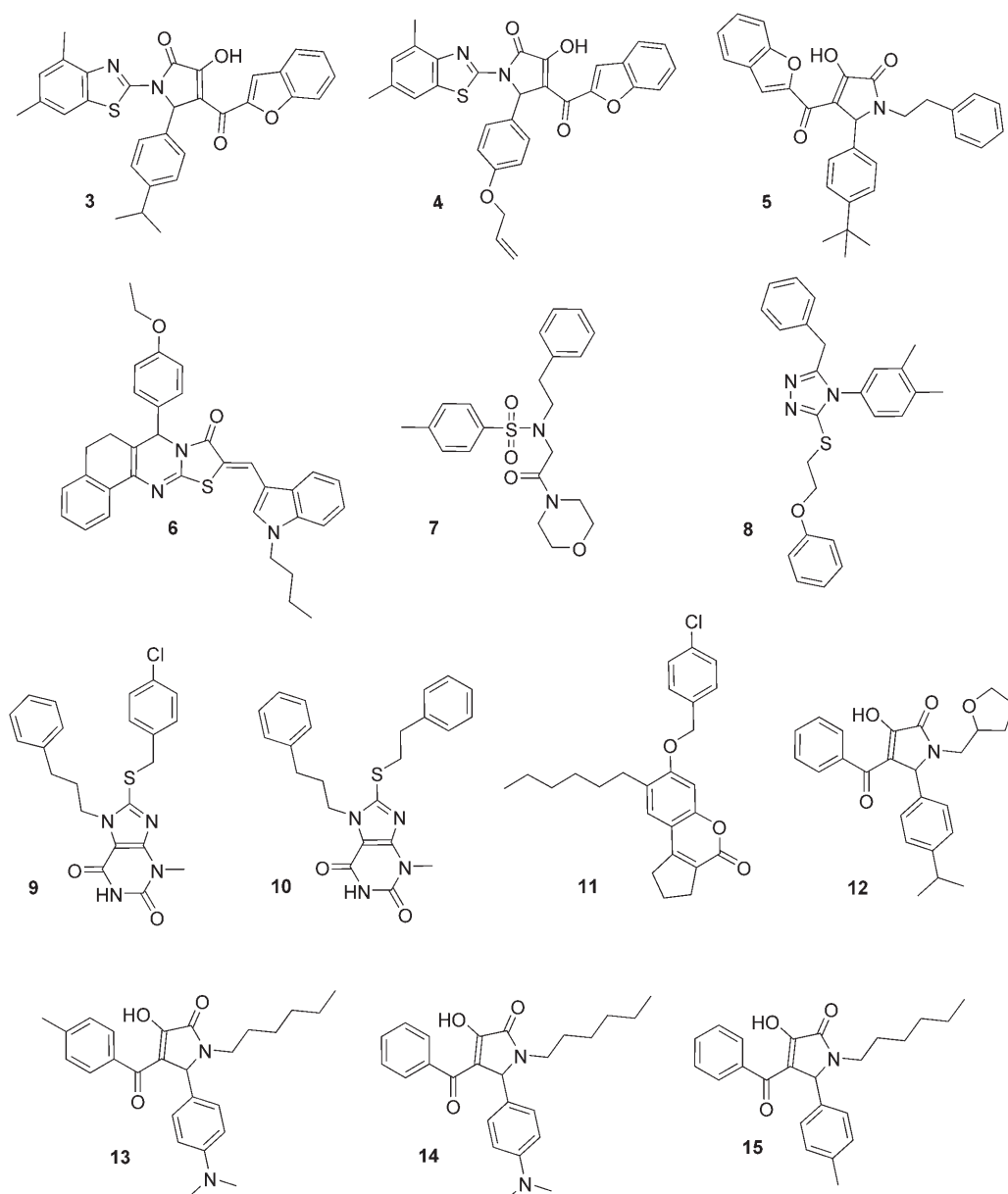
Scheme 1. TAR RNA ligands acetylpromazine 1 and chlorpromazine 2 were both discovered through structure-based approaches.^[12]

A conceptual idea of LIQUID is to combine multiple docking poses of a known inhibitor into a single fuzzy description of potential ligand–receptor interaction sites. The resulting pharmacophore model is then used for virtual screening of a large collection of candidate compounds. 3D conformations of reference ligands can be experimentally determined binding poses, or—as in this first prospective study—stem from computational docking experiments. The idea is to construct a pharmacophore model that accounts for multiple variations of a potential binding mode, rather than using a single docking pose as query. This might also help to avoid artifacts that arise from virtual screening with a single inaccurate docking solution or computed conformation.

In a preliminary prospective virtual screening study, we first applied LIQUID to a particularly difficult task, namely structure-based scaffold-hopping from a poor ligand of a flexible target. Such a scenario is typical for many exploratory projects in drug discovery. In 2002, James and co-workers published an NMR structure of HIV-1 TAR RNA in complex with the bulge ligand acetylpromazine 1 (PDB ID:^[15] 1lvj),^[16] which served as the reference structure in this study (Figure 1 A). TAR RNA is remarkably flexible,^[16,17] and acetylpromazine is a weak binder yielding an IC_{50} of 500 μM in our assay.^[6] Notably, in contrast to earlier publications by Du et al. claiming nanomolar TAR binding of acetylpromazine in a gel-shift assay,^[16] Mayer and James found a K_d of 270 μM in 2004.^[18] This discrepancy coincides with our own observations for acetylpromazine.^[6] Certainly, the use of acetylpromazine as a suitable reference structure is debatable because of its meager blocking of Tat/TAR complex formation. As a consequence, one cannot expect more potent



Scheme 2. Most prevalent molecular graph scaffolds retrieved by A) LIQUID and B) SQUID with acetylpromazine 1 as query structure. Numbers in brackets give the absolute number of a scaffold found among the top-ranking 1960 molecules of the SPECS^[14] (v.2007/05) compound library. The TAR-bound conformation of acetylpromazine was used, and for each library compound a single generated conformation was generated by CORINA (Molecular Networks GmbH, Erlangen).



Scheme 3. Compounds tested for inhibition of Tat-TAR interaction. Molecule **12** yielded the best results in a FRET assay.

compounds to be found by a virtual screening approach *per se*. In this proof-of-principle study we used acetylpromazine as a “tough but typical” example and because of its available NMR reference structure and comparability to previous approaches.

For pharmacophore modeling, we performed 20 redocking experiments of **1** into the bulge of TAR RNA (model 1 from 1lvj) with the docking program GOLD^[19] (v.3.0, the binding site was defined at point $x = -24.82$, $y = 7.96$, $z = 6.48$ with a radius of 10 Å; early docking termination was disabled). This yielded 20 docking poses of acetylpromazine in the TAR bulge region (Figure 1B). LIQUID was employed to compute a pharmacophore model of all 20 docking solutions (poses). Full details of the method are described in the Supporting Information.

Figure 1C depicts the docking solutions and the computed fuzzy pharmacophore model. The model contains two main

lipophilic interaction sites corresponding to the aromatic ring systems of acetylpromazine. Smaller lipophilic pharmacophore points represent the aliphatic linker to the dimethylamine moiety and the carbon of the acetyl group. In our study, all nitrogen atoms and the oxygen of the acetyl group were interpreted as hydrogen-bond acceptors. This pharmacophore model delineates an abstract, generalized representation of all potential interaction sites of the set of acetylpromazine docking solutions. In contrast to many other pharmacophore-modeling techniques, LIQUID does not use hard exclusion spheres to represent potential interaction sites. “Soft” shells are used instead, and yield a probability distribution for each potential ligand–receptor interaction point without the explicit necessity of defining tolerances, “voids”, or “tabu zones”.

The LIQUID model was used to screen for novel TAR RNA ligands in the SPECS database (v.2005/09, 198 999 com-

pounds).^[14] Out of the 100 top-ranked compounds, 13 molecules with a different scaffold from that of acetylpromazine were ordered for experimental testing (Scheme 3). These substances were manually picked under consideration of molecular weight (350–550 Da) and the number of hydrogen-bond donors and acceptors (<5). In addition, all selected molecules were prechecked for a potentially similar TAR-bound conformation as **1** by using GOLD. Finally, a fluorescence-based assay (FRET) was used to measure activity (see the Supporting Information).^[20] This assay had been optimized for Tat/TAR interaction analysis, and was performed as previously described.^[6]

Compounds **12** and **14** were the most potent and yielded an average fluorescence reduction of up to 17% at a concentration of 50 μM ($\text{IC}_{50} > 50 \mu\text{M}$; Table 2). Molecules **7**, **9**, and **10**

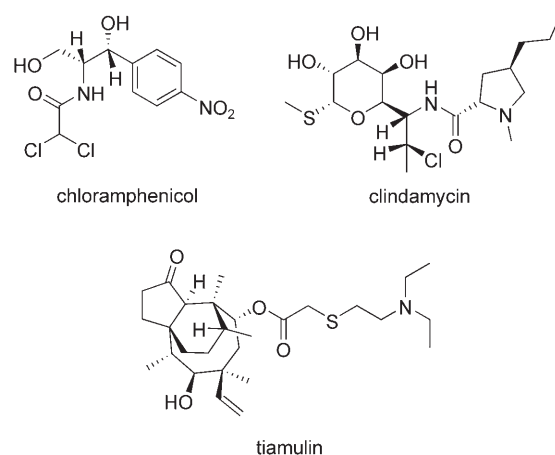
Compound no.	Conc. [μM]	Tat–TAR ^[a] inhibition [%]	Compound no.	Conc. [μM]	Tat–TAR ^[a] inhibition [%]
3	100	0	4	100	0
5	100	0	6	100	0
7	100	7 \pm 10	8	100	3 \pm 15
9	100	11 \pm 5	10	100	8 \pm 4
11	50	0	12	50	17 \pm 10
13	50	0	14	50	12 \pm 4
15	50	1 \pm 4			

also decreased fluorescence; this points to a possible inhibition of the Tat–TAR association. Due to solubility issues of some of the selected substances, molecules **3–10** were tested at a concentration of 100 μM , and molecules **11–15** were tested at 50 μM . For the same reason, determination of IC_{50} values was not possible.

A striking example of the solubility issue encountered in this study is the discrepancy in activity between compounds **13** and **14**. We do not attribute this observation to a steep structure–activity relationship. These substances differ in a single *para*-methyl group, which led to observable precipitation of **13** in the test wells. As a consequence, a lower substance concentration than expected was measured. Therefore, we cannot exclude some binding potential of **13**. This interpretation also holds for compounds **3**, **4**, and **15**. Poor aqueous solubility is a general problem of many substances tested in biological assays, and is particularly crucial for RNA or DNA ligands known to intercalate between the bases.^[21]

To further scrutinize our approach, we applied LIQUID to virtual screening for ligands binding to the ribosomal A site. This time the pharmacophore model was computed from a rigid-body alignment of experimentally determined A-site-bound conformations of the antibiotic agents chloramphenicol (PDB: 1k01),^[22] clindamycin (PDB: 1jzx),^[22] and tiamulin (PDB: 1xbp)^[23] (Scheme 4, Figure 2).

Again, the SPECS catalogue was virtually screened, the 100 top-ranking candidates were docked into the ribosomal A-site structure by using GOLD (v.3.0, binding site of 1xbp at point



Scheme 4. Reference antibiotics binding to the ribosomal A site.

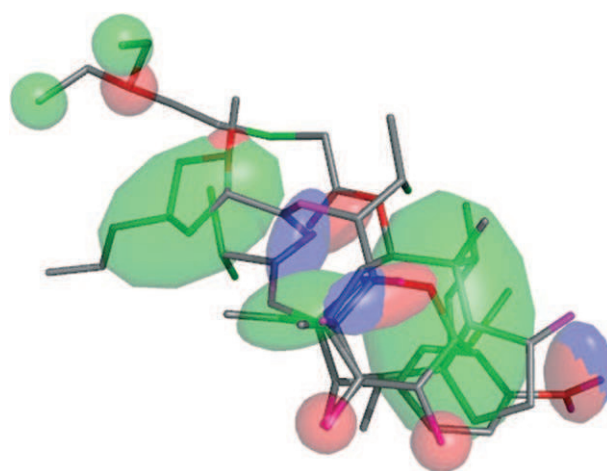
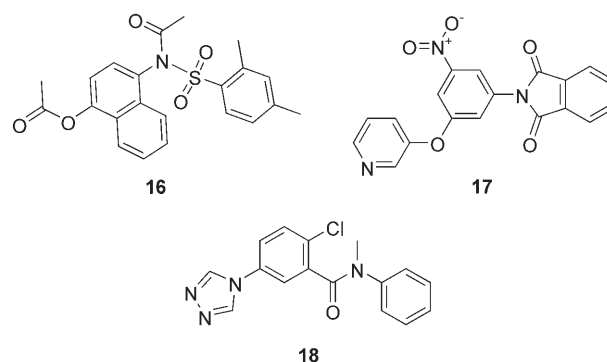


Figure 2. Pharmacophore model of ribosomal A-site ligands.

$x=53.53$, $y=120.79$, $z=113.70$ with a radius of 10 Å; early docking termination was disabled), and 19 molecules were tested in a cell-free transcription/translation assay (Supporting Information). Three compounds **16–18** (Scheme 5), which belong to different chemotypes from the reference compounds, yielded 25 \pm 10% reduction of protein expression at a concentration of 100 μM . Notably, these hits were inactive in the Tat–TAR interaction assay indicating relative selectivity.



Scheme 5. Inhibitors of in vitro transcription/translation found by LIQUID.

Summarizing, these results demonstrate the applicability of our fuzzy pharmacophore concept to virtual screening. In two studies, LIQUID retrieved several new RNA ligands that contain different scaffolds from the reference compounds. These case studies demonstrate how pharmacophore-based virtual screening can be used as an idea generator in early drug discovery projects.^[24]

Acknowledgements

This research was supported by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften and the Deutsche Forschungsgemeinschaft (DFG; SFB 579 "RNA-Ligand Interactions", projects A3 and A11).

Keywords: cheminformatics · drug design · modeling · pharmacophores · ribosomal A site · virtual screening

- [1] G. Schneider, W. Neidhart, T. Giller, G. Schmid, *Angew. Chem.* **1999**, *111*, 3068–3070; *gew. Chem. Int. Ed.* **1999**, *38*, 2894–2896.
- [2] R. D. Cramer, R. J. Jilek, S. Guessregen, S. J. Clark, B. Wendt, R. D. Clark, *J. Med. Chem.* **2004**, *47*, 6777–6791.
- [3] G. Schneider, P. Schneider, S. Renner, *QSAR Comb. Sci.* **2006**, *25*, 1162–1171.
- [4] G. Schneider, E. Proschak, K. Grabowski, P. Schneider, Y. Tanrikulu, *Proceedings of the Beilstein Workshop* (Eds.: M. Hicks, C. Kettner), Logos, Berlin, **2007**; in press.
- [5] S. Renner, G. Schneider, *J. Med. Chem.* **2004**, *47*, 4653–4664.
- [6] S. Renner, V. Ludwig, O. Boden, U. Scheffer, M. Göbel, G. Schneider, *ChemBioChem* **2005**, *6*, 1119–1125.
- [7] a) P. Broto, G. Moreau, C. Vandycke, *Eur. J. Med. Chem. Chim. Ther.* **1984**, *19*, 66–70; b) H. Bauknecht, A. Zell, H. Bayer, P. Levi, M. Wagener, J. Sadowski, J. Gasteiger, *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1205–1213; c) U. Fechner, L. Franke, S. Renner, P. Schneider, G. Schneider, *J. Comput.-Aided Mol. Des.* **2003**, *17*, 687–698.
- [8] P. Schneider, G. Schneider, *QSAR Comb. Sci.* **2003**, *22*, 713–718.
- [9] G. C. Fanning, *Curr. Drug Targets* **2006**, *7*, 1607–1613.
- [10] S. Bannwarth, A. Gatignol, *Curr. HIV Res.* **2005**, *3*, 61–71.
- [11] J. Karn, *J. Mol. Biol.* **1999**, *293*, 235.
- [12] K. Lind, Z. Du, K. Fujinaga, B. Peterlin, T. James, *Chem. Biol.* **2002**, *9*, 185–193.
- [13] G. W. Bemis, M. A. Murcko, *J. Med. Chem.* **1996**, *39*, 2887–2893.
- [14] Specs, Kluyverweg 6, 2629HT Delft, The Netherlands, **2006**, <http://www.specs.net>
- [15] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–242.
- [16] Z. Du, K. E. Lind, T. L. James, *Chem. Biol.* **2002**, *9*, 707–712.
- [17] Y. Mu, G. Stock, *Biophys. J.* **2006**, *90*, 391–399.
- [18] M. Mayer, T. L. James, *J. Am. Chem. Soc.* **2004**, *126*, 4453–4460.
- [19] G. Jones, P. Willet, R. Glen, A. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727–748.
- [20] C. Matsumoto, K. Hamasaki, A. Mihara, A. Ueno, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1857–1861.
- [21] a) L. Di, E. H. Kerns, *Drug Discovery Today* **2006**, *11*, 446–451; b) A. I. H. Murchie, B. Davis, C. Isel, M. Afshar, M. J. Drysdale, J. Bower, A. J. Potter, I. D. Starkey, T. M. Swarbrick, S. Mirza, C. D. Prescott, P. Vaglio, F. Abouela, J. Karn, *J. Mol. Biol.* **2004**, *336*, 625–638; c) C. A. Lipinski, *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235–249.
- [22] F. Schlünzen, R. Zarivach, J. Harms, A. Bashan, A. Tocilj, R. Albrecht, A. Yonath, F. Francesci, *Nature* **2001**, *413*, 814–821.
- [23] F. Schlünzen, E. Pyetan, P. Fucini, A. Yonath, J. M. Harms, *Mol. Microbiol.* **2004**, *54*, 1287–1294.
- [24] a) LIQUID is publicly available as a free PyMOL plug-in for pharmacophore model visualization from our website (<http://www.modlab.de>); b) W. L. DeLano, The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA, USA, **2002**.
- [25] a) H. Xu, D. K. Agrafiotis, *Curr. Top. Med. Chem.* **2002**, *2*, 1305–1320; b) A. N. Jain, *Curr. Opin. Drug Discovery Dev.* **2004**, *7*, 396–403; c) A. Bender, R. C. Glen, *J. Chem. Inf. Model.* **2005**, *45*, 1369–1375.

Received: April 13, 2007

Published online on September 25, 2007

CHEM**BIO**CHEM

Supporting Information

© Copyright Wiley-VCH Verlag GmbH & Co. KGaA, 69451 Weinheim, 2007

Scaffold-Hopping by “Fuzzy” Pharmacophores and Application to RNA Targets

Yusuf Tanrikulu, Manuel Nietert, Ute Scheffer, Ewgenij Proschak, Kristina Grabowski, Petra Schneider, Markus Weidlich, Michael Karas, Michael Göbel, Gisbert Schneider

Defined by the Medicinal Chemistry Section of IUPAC,^[1] a pharmacophore is the “ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response”. According to this concept, ligand-receptor interactions arise by individual functional group contributions. Since we do not know a priori which functional group actually contributes to the interaction they are termed “potential pharmacophore points” (PPPs). Ligand-receptor interactions take place in 3D space. Therefore 3D pharmacophore models represent the most intuitive choice. Noteworthy, in the absence of a receptor-relevant ligand conformation or conformation ensemble, quantitative structure-activity relationship (QSAR) studies that are based on 3D models can still be erroneous. In addition to the problem of conformer generation, an error-prone step in pharmacophore matching methods is the 3D-alignment of molecular features, that is, matching a screening molecule to a given pharmacophore model. To enable rapid database searching the explicit alignment step can be avoided by an alignment-free representation of pharmacophore patterns. One idea is to convert the spatial distribution of PPPs to a vector representation. Such vectors are referred to as “fingerprints”, “bitstrings”, “correlation vectors” (CV), or “spectra” depending on the type of information stored. The trick is to compare these reduced molecular representations instead of explicit 3D feature alignment thus formulating a pharmacophore search as a similarity search.^[2]

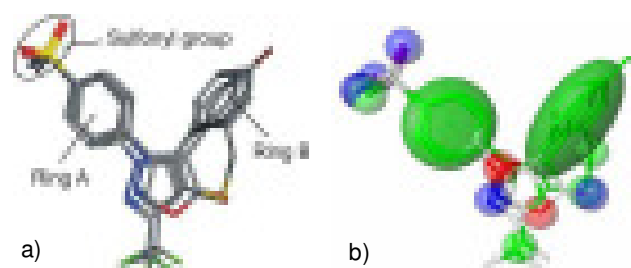


Figure S1. a) Molecular alignment of three COX-2 inhibitors (Rofecoxib, M5, SC-558). The common chemical groups, which are essential interactions for specific COX-2 inhibition, are labeled according to ref. [4]. b) LIQUID Pharmacophore model of a): the visualization shows the ellipsoidal PPP models of both ring systems, which are members of the maximum common substructure of the respective superposition.

In our software LIQUID, PPPs are modeled as trivariate Gaussian distributions and encoded as CV representations. The statistical spread of every PPP reflects the fuzziness of

the pharmacophore model: The probability of a certain interaction decreases with the increasing distance to the PPPs centroid. Three-dimensional visualization tools, e.g. PyMOL^[3], enable us to visually analyze generated pharmacophore models. An example of a LIQUID pharmacophore model is shown in Figure S1, derived from a molecular alignment of the COX-2 inhibitors Rofecoxib, M5, and SC-558.

LIQUID can be used to compute a pharmacophore model of either a single molecular conformation or a conformer ensemble. The first fundamental step is the atom type recognition, where potential pharmacophore features are assigned to the query's atoms. We consider three different interaction types: “lipophilic”, “hydrogen-bond donor” and “hydrogen-bond acceptor”. For example, oxygen atoms are always hydrogen-bond acceptor interaction points, whereas OH groups also possess a hydrogen-bond donor feature. An atom can represent none, one or maximal two (only hydrogen-bond donor + acceptor) of these features. Atoms lacking a pharmacophore feature are not considered for the model. Atom typing results in a transformation of the query molecule(s) into a three-dimensional disposition of interaction points. To gather a fuzzy approximation of this interaction field via Gaussian functions, single interaction points are clustered into PPPs. Every PPP represents a local maximum of the Gaussian distribution of interaction points it contains.

To determine the maxima of the interaction point distribution the cluster radius dependent local-feature-density (LFD) was introduced.^[5] It allows a quantification of common-type atoms in the spatial environment around an atom. Hereby the manually adjustable cluster radius is used to constrain the space around an atom where the maximum is determined. This also enables the user to generate pharmacophore models with varying fuzziness. The fuzzy pharmacophore model of all acetylpromazine docking modes described above was computed with a lipophilic cluster radius of 1.4 Å and 2.0 Å for hydrogen-bond donors and acceptors. The LFD of the k^{th} atom of pharmacophore type T is calculated by Eq. 1.

$$LFD(atom_k^T) = \sum_{i=1}^n \max\left\{0, 1 - \frac{D_2(atom_k^T, atom_i^T)}{r_c}\right\}, \quad (1)$$

where n is the total number of atoms of the pharmacophore type T in the query, D_2 is the Euclidean distance of two atoms, and r_c represents the cluster radius of the specific pharmacophore type. The closer a common-typed atom, the bigger is its impact to the considered atom's LFD. Clustering of the interaction points is done on the basis of the atoms' LFDs via a Union-Find strategy. The following pseudo-code illustrates how the algorithm works:

```
INIT: each atom is a singleton.
FOR each atom  $i$  of type  $T$ 
  FOR each atom  $j$  of type  $T$ 
    calculate Distance( $i, j$ )
    IF Distance  $\leq$  ClusterRadius  $r_c$  THEN
      FIND maxLFD(Cluster $_i$ )
      FIND maxLFD(Cluster $_j$ )
      IF maxLFD(Cluster $_i$ )  $\leq$  maxLFD(Cluster $_j$ ) THEN
        UNION Cluster $_j$  with Cluster $_i$ 
```

Initially, every atom represents a singleton. If a common-typed atom yielding a higher LFD inside the cluster radius of

the considered atom is found, both atoms get “united” into a cluster. Finally, each cluster of interactions points (feature-typed atoms) forms a PPP. One can observe that the number of final clusters depends on the adjusted cluster radius. The centroid of a PPP is calculated as geometric center of its clustered atoms.

We apply the principal component analysis (PCA)^[6] in order to compute the size and orientation of a PPP. The covariance matrix is built up from the Cartesian coordinates of the clustered atoms in relation to the PPPs centroid. The orientation of a PPP is given by the resulting principal components, because their directions span the data space according to the highest variances. Eigenvector approximation is done with the NIPALS algorithm.^[7] The corresponding Eigenvalues provide the distribution of the PPP in the direction of the Eigenvectors.

Having obtained the position, size and orientation of the PPPs, we encode the pharmacophore model as a correlation vector (Eq. 2).^[2,5] LIQUID computes a correlation-vector from the trivariate Gaussian functions, which are used to model the PPPs. Due to pairwise PPP correlation, we encounter six PPP pairs: “lipophilic- lipophilic”, “lipophilic-donor”, “lipophilic-acceptor”, “donor-donor”, “donor-acceptor” and “acceptor-acceptor”. Encoding yields an equally partitioned bin vector. In the case of PPP instances, the correlation vector reduces to scaled occurrence frequencies of atom pairs at distance intervals from 1 to 20 Å.^[2,8] Each bin (vector element) contains a correlated probability, which indicates the presence of a PPP pair at distance d :

$$CV_d^{A,B} = \frac{1}{\# \text{ pairs } (A, B)} \sum_i^A \sum_j^B \frac{1}{2} \{ f_i^V G(\sigma_{1,2,3})_i \cdot f_j^V G(\sigma_{1,2,3})_j \}, \quad (2)$$

for all PPPs i of pharmacophore type A and for all PPPs j of type B. $f_i^V G$ gives the trivariate Gaussian with standard deviations $\sigma_{1,2,3}$, and $\# \text{ pairs}(A, B)$ is the number of pairs of PPPs of type A and B. The result is a 120-dimensional correlation vector-based descriptor designed for fast virtual screening. In this study, the Euclidian distance was used to determine the similarity between vector representations of pairs of molecules.

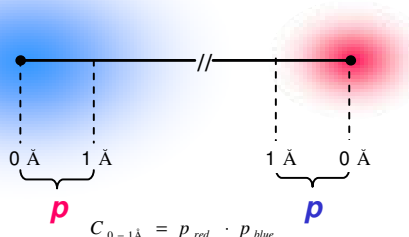


Figure S2. Correlation of probabilities in a certain distance bin

FRET Binding Assay. Fluorescence based binding assays were performed in a microplate reader (Safire²; Tecan; excitation wavelength 489 nm, emission wavelength 590 nm) in 96 well microplates (Corning 6860, black, non binding surface). TAR RNA (Biospring) and Tat peptide (Thermo

Electron Corporation) were both used at concentrations of 10 nM in a final volume of 100 μ L in TK buffer (50 mM Tris-HCl, 20 mM KCl, 0.01% Triton-X 100, pH 7.4) at 37 $^{\circ}$ C. Prior to titration, the RNA (100 nM in 5mM Tris-HCl, pH 7.4) was heated to 90 $^{\circ}$ C for 5 minutes and then immediately placed on ice for additional 2-5 minutes. The fluorescence of the blank (TK buffer only), pure peptide and of the Tat-TAR complex was determined first. Single-point measurements of potential inhibitors were carried out in triplicates at 50 or 100 μ M. Figure S3 shows the TAR model and labeled Tat peptide.

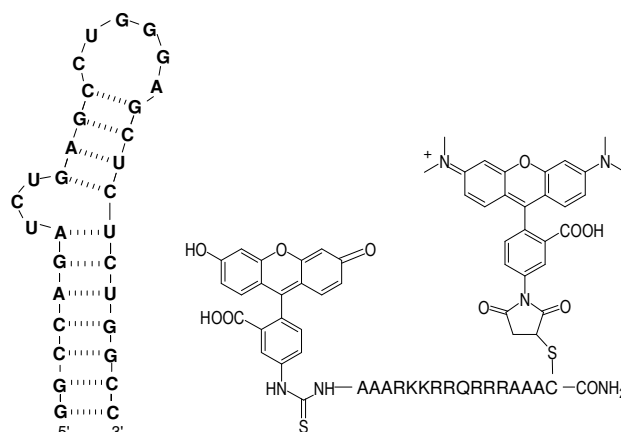


Figure S3. Structures of the HIV-1 TAR model and Tat peptide labeled with fluorescein and rhodamine.

The values delivered by the Safire² spectrometer were transformed to obtain relative fluorescence intensity according to equation (3):

$$x_{ij}^* = \frac{x_{ik} - x_{k, \min}}{x_{k, \max} - x_{k, \min}}, \text{ where} \quad (3)$$

x_{ij}^* = scaled fluorescence intensity in [0,1]

x_{ik} = raw data value

$x_{k, \min}$ = minimum of the measured fluorescence activity (max. quenching value; max. of free peptide)

$x_{k, \max}$ = maximum of the measured fluorescence activity (max. fluorescence value; max. of bound peptide).

Original data from the FRET assay are available from file “Tanrikulu_FRET_primary_data.xls”.

Cell-free Transcription/Translation Assay. S30-extract of *E. coli* A19 was prepared as described previously.^[9] The reaction volume of 25 μ l contained 3.5 mM Tris-acetate (pH 8.2), 220 mM potassium acetate, 13 mM magnesium acetate, 100 mM HEPES-KOH (pH 8.0), 0.7 mM EDTA, 0.2 mM folic acid, 2 mM DTT, 2 % polyethylene glycol 8000, 1.2 mM ATP, 0.8 mM of the NTPs, 20 mM acetyl phosphate (lithium potassium salt), 20 mM phosphoenolpyruvate (trisodium salt), and 0.05 % sodium azide (components purchased from Sigma-Aldrich Chemie GmbH, Schnellendorf, Germany). 500 mg/L of *E. coli* tRNA mixture, 40 mg/L pyruvate kinase, 1 tablet/10 ml complete protease inhibitor (all Roche Diagnostics, Mannheim, Germany), 0.3 U/ μ l RNasin RNase inhibitor (Promega GmbH, Mannheim, Germany), 3 U/ μ l T7-RNA-Polymerase (GE Healthcare Europe GmbH, Munich, Germany), and complete amino acid mixture (final concentration 0.75 – 1mM). 40–60 ng/ μ l of the plasmid pVEX2.3-GFP encoding GFP under the control of a T7-promoter and 30 % *E. coli* A19 S30 extract

were added. Compounds were dissolved with 2% (v/v) dimethyl sulfoxide in the reaction mixture. Reaction mixtures were complemented with 100 μ M of potential inhibitor. Negative inhibition control reactions were complemented with water. CFTT-reactions were run for 5 h at 30 °C. GFP was quantified (1:100 dilutions) by measuring the fluorescence with excitation/emission wavelengths of 395 and 509 nm, respectively using a Hitachi F-4500 Fluorescence Spectrophotometer. Single-point measurements of potential inhibitors were carried out in triplicates.

Original data from the cell-free transcription/translation assay are available from file "Tanrikulu_ivTT_primary_data.xls".

References

- [1] C.G. Wermuth, C.R. Gannelin, P. Lindberg, L.A. Mitschler, *Pure & Appl. Chem.* **1998**, 70, 1129-1143.
- [2] S. Renner, U. Fechner, G. Schneider, In: *Pharmacophores and Pharmacophore Searches* (T. Langer and E. Hoffmann), Wiley-VCH, Weinheim, **2005**, pp. 49-79.
- [3] W.L. DeLano, *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA, USA, **2002**.
- [4] A. Palomer, F. Cabre, J. Pascual, J. Campos, M. Trujillo, A. Entrena, M. Gallo, L. Garcia, D. Mauleon, A. Espinosa, *J. Med. Chem.* **2002**, 45, 1402-1411.
- [5] S. Renner, G. Schneider, *J. Med. Chem.* **2004**, 47, 4653-4664.
- [6] E. Jackson, *A User's Guide to Principal Components*, Wiley, New York, **1991**.
- [7] S. Wold, *Chemica Scripta* **1974**, 5, 97-106.
- [8] R.E. Carhart, D.H. Smith, R. Venkataraghavan R, *J. Chem. Inf. Comput. Sci.* **1985**, 25, 64-73.
- [9] C. Klammt *et al.* *Eur. J. Biochem.* **2004**, 271, 568-580.