

# **Protein- und Naturstoffstruktur als Leitprinzipien für die Entwicklung von Verbindungsbibliotheken**

Zur Erlangung des akademischen Grades eines  
Doktors der Naturwissenschaften  
(Dr. rer. nat.)  
vom Fachbereich Chemie  
der Universität Dortmund  
angenommene

## **Dissertation**

von Apotheker

**Marcus Alexander Koch**

aus Albstadt

**1. Gutachter:** Prof. Dr. H. Waldmann

**2. Gutachter:** Prof. Dr. C. M. Niemeyer

**Tag der mündlichen Prüfung:** 25. Mai 2005







## Abkürzungen und Definitionen

2D	zweidimensional
3D	dreidimensional
ACE	Angiotensin-Conversionsenzym
AChE	Acetylcholinesterase
ACN	Acetonitril
ATC	Acetylthiocholiniodid
11 $\beta$ HSD1	11 $\beta$ -Hydroxysteroid-Dehydrogenase Typ 1
11 $\beta$ HSD2	11 $\beta$ -Hydroxysteroid-Dehydrogenase Typ 2
Cdc25A	Cdc25A-Phosphatase
CDK	Cyclin-dependent Kinase
CE	Combinatorial Extension
CPK	Corey-Pauling-Koltun
DC	Dünnschichtchromatographie
DFMO	<i>DL</i> - $\alpha$ -Difluormethylornithin
DHB	2,5-Dihydroxybenzoesäure
DHPS	Dihydropteroatsynthetase
DMF	<i>N,N</i> -Dimethylformamid
DMSO	Dimethylsulfoxid
DTE	Dithioerythritol
DTNB	5,5'-Dithiobis-(2-nitrobenzoesäure)
DTT	Dithiothreitol
EI	Electron Impact
ER	Estrogenrezeptor
ESI	Electron Spray Ionization
FAB	Fast Atom Bombardment
FC	Flashchromatographie
FSSP	Fold Classification Based on Structure-Structure Alignment of Proteins
FXR	Farnesoid-X-Rezeptor
ges.	gesättigt
GFP	grün-fluoreszierendes Protein

---

GR	Glucocorticoidrezeptor
HCl	Hydrochlorid
Hit	In einem Assay als aktiv identifiziertes Molekül, das unterhalb einer bestimmten Ausschlußkonzentration eine Targetfunktion gezielt moduliert
HPLC	High Performance Liquid Chromatography
HR	High Resolution
HTS	High Throughput Screening
HSD	Hydroxysteroid-Dehydrogenase
IBX	1-Hydroxy-1,2-benziodoxol-3(1 <i>H</i> )-on-1-oxid
$K_i$	Hemmkonstante
KD	Katalytische Domäne
KK	Katalytischer Kernbereich
LAH	Lithiumaluminiumhydrid
LD	Ligandenbindungsdomäne
LDA	Lithiumdiisopropylamid
LC-MS	Liquid Chromatography Mass Spectrometry
LK	Ligandenbindender Kernbereich
LT	Leukotrien
LTA <sub>4</sub> H	Leukotrien-A <sub>4</sub> -Hydrolase
LXR	Leber-X-Rezeptor
MALDI	Matrix Assisted Laser Desorption Ionization
MeOH	Methanol
MS	Massenspektrum
n.b.	nicht bestimmt
3-NBA	3-Nitrobenzylalkohol
NR	Nukleärer Rezeptor (Kernrezeptor)
ODC	Ornithindecaboxylase
p56 <sup>lck</sup>	lymphozytenspezifische Kinase
PDB	Protein Data Bank
PEP	Phosphoenolpyruvat
PLP	Pyridoxal-5'-phosphat
<i>p</i> -NPP	<i>p</i> -Nitrophenylphosphat
PPAR $\gamma$	Peroxisome Proliferator-Activated Receptor $\gamma$

PSSC	Protein Structure Similarity Clustering (Proteinstruktur-Ähnlichkeitsclustering)
<i>p</i> TSA	<i>p</i> -Toluolsulfonsäure
ODC	Ornithindecaboxylase
R <sub>f</sub>	Retentionsfaktor
RMSD	mittlere quadratische Abweichung
R <sub>t</sub>	Retentionszeit
RTK	Rezeptor-Tyrosinkinase
SCOP	Structural Classification of Proteins
SCONP	Structural Classification of Natural Products
SDR	Short-Chain Dehydrogenase
SI	Sequenzidentität
Smp.	Schmelzpunkt
Target	Zielstruktur (i.d.R. ein Protein), mit der biologisch aktive Verbindungen interagieren, und über die der biologische Effekt vermittelt wird.
<i>t</i> Bu	<i>tert</i> -Butyl
TFA	Trifluoressigsäure
THF	Tetrahydrofuran
TOF	Time of Flight
Tris	Tris-(2-Hydroxyethyl)amin

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Allgemeiner Teil</b>	<b>3</b>
2.1	Wirkstofffindungsprozeß	3
2.2	Domänenarchitektur von Proteinen	5
2.3	Die ligandenbindenden Kernstrukturen von Proteindomänen	7
2.4	Naturstoffe	8
2.5	Chemische Genetik, Chemische Genomik oder Chemogenomik?	10
<b>3</b>	<b>Ziel der Arbeit</b>	<b>13</b>
<b>4</b>	<b>Spezieller Teil</b>	<b>15</b>
4.1	Proteinstruktur-Ähnlichkeitsclustering (PSSC)	15
4.1.1	Analyse der Literatur – Eine Machbarkeitsstudie	15
4.1.2	Cdc25A-Phosphatase, Acetylcholinesterase und 11 $\beta$ -Hydroxysteroid-Dehydrogenasen – Ein Proteinstruktur-Ähnlichkeitscluster	29
4.1.3	Ornithindecaboxylase und Dihydropteroatsynthetase – Ein Proteinstruktur-Ähnlichkeitscluster	56
4.1.4	Konzeptionelle Arbeiten zu PSSC	69
4.2	Strukturelle Klassifizierung der Naturstoffe (SCONP)	78
4.2.1	Exkurs: Naturstoff-inspirierte Entwicklung von strukturell atypischen Kinaseinhibitoren	79
4.2.2	SCONP – Ein Hypothesengenerator und Navigator durch den Strukturraum der Naturstoffe	87
<b>5</b>	<b>Zusammenfassung und Ausblick</b>	<b>122</b>
<b>6</b>	<b>Experimenteller Teil</b>	<b>130</b>
6.1	Bioinformatik und Molecular Modeling	130
6.1.1	Allgemeine Arbeitsvorschriften (AAV)	130
6.1.2	Retrospektive Analyse von Literaturbeispielen	140
6.1.3	Arbeiten zum Proteinstruktur-Ähnlichkeitsclustering (PSSC)	143
6.1.4	Konzeptionelle Arbeiten	149
6.1.5	Docking	152
6.2	Biochemie und Zellbiologie	154
6.2.1	Inhibition der Cdc25A-Phosphatase	154
6.2.2	Inhibition der 11 $\beta$ -Hydroxysteroid-Dehydrogenasen Typ 1 und Typ 2	154
6.2.3	Inhibition der Acetylcholinesterase	156
6.2.4	Inhibition der Ornithindecaboxylase	157
6.2.5	Translokation des Glucocorticoidrezeptors (GR)	158
6.2.6	GR-abhängige Transaktivierung	158
6.2.7	Inhibition von Kinasen	159
6.3	Chemie	160
6.3.1	Meßgeräte und Hilfsmittel	160
6.3.2	Versuche zur Synthese von Ulocladolderivaten	161
6.4	Chemoinformatik	172
6.4.1	Primäre molekulare Prozessierung des Naturstoff-Datensatzes	172
6.4.2	Extraktion der Grundgerüste	172
6.4.3	In Silico-Deglykosylierung der Naturstoffe	173
6.4.4	Baumdiagramm der Naturstoffgrundgerüste	174
6.4.5	Statistische Auswertung des Strukturbaums	176
6.4.6	Annotation des Strukturbaums	176

---

<b>7</b>	<b>Literaturverzeichnis</b> .....	<b>178</b>
<b>8</b>	<b>Anhang</b> .....	<b>194</b>
8.1	Anhang I: Von Dysidiolid abgeleitete Bibliothek .....	194
8.2	Anhang II: Von Dysidiolid und Glycyrrhetinsäure abgeleitete Bibliothek .....	207
8.3	Anhang III: Kollektion von Benzolsulfonamid-Derivaten .....	229
8.4	Anhang IV: Von Ulocladol abgeleitete Verbindungskollektion .....	231



# 1 Einleitung

Die Identifizierung neuer chemischer Verbindungen mit der Eigenschaft, Proteinfunktionen gezielt zu modulieren, ist das Wesen des Hit- und Leitstrukturfundungsprozesses und stellt gleichzeitig einen wichtigen Engpaß bei der Entwicklung neuer Arzneistoffe dar.

Die kombinatorische Chemie hat innerhalb des letzten Jahrzehnts in diesem Prozeß eine entscheidende Stellung eingenommen. Die Entwicklung dieses mächtigen Werkzeugs zur Darstellung großer Verbindungsbibliotheken geht einher mit der anschwellenden Datenflut zu biologischen Systemen, wie sie durch globale Ansätze, die auf Analysen der DNA-Sequenz (Genomik), der Proteinstruktur (Strukturelle Genomik) und der Proteininteraktionen (Proteomik) beruhen, erzeugt wurde. Dies nährte die Hoffnung, daß die so gewonnenen biologischen Basisdaten sehr schnell in Wissen umgesetzt werden könnten, Wissen, das letztlich zum Wohle der Menschen in die Entwicklung neuer Therapien und Arzneistoffe münden sollte.

Die äußerst hohe Produktivität des kombinatorischen Ansatzes bezogen auf die schiere Zahl neuer Moleküle und die enorme technische Aufrüstung der pharmazeutisch-chemischen Forschungsabteilungen haben jedoch nicht zu der ersehnten deutlich höheren Anzahl neuer Wirkstoffkandidaten geführt, die die versiegenden Forschungspipelines füllen könnten. Im Gegenteil, die Anzahl neu zugelassener Arzneistoffe fiel auf ein Zwanzigjährestief von 37 im Jahr 2001 und eine Trendwende ist noch nicht in Sicht.<sup>[1]</sup> Desillusionierung machte sich breit. Einfach nur mehr Verbindungen zu generieren, zahlte sich nicht aus. So steht nach wie vor eine der zentralsten Fragen der modernen Wirkstoffentwicklung unbeantwortet im Raum: Wo im riesigen chemischen Strukturraum sind die biologisch relevanten Verbindungen zu finden?

Die Antwort auf diese Frage ist mit Sicherheit alles andere als einfach. Es wurden bereits einige Kriterien definiert, die biologisch relevante Substanzkollektionen erfüllen sollten: „Diversität“<sup>[2-5]</sup>, „Wirkstoff-Ähnlichkeit“<sup>[6-10]</sup> und „biologische Relevanz“<sup>[11-13]</sup>. Das letztgenannte Kriterium kann erfüllt werden, wenn sogenannte „privilegierte“<sup>[14]</sup> Strukturen oder Naturstoffe als strukturelle Leitprinzipien verwendet werden. Tatsächlich steigen die Hitraten in biologischen Screenings signifikant, wenn Naturstoffe oder ihre Analoga in ‚High Throughput Screenings‘ (HTS) miteinbezogen werden.<sup>[11, 12, 15, 16]</sup> Dies legt nahe,

Eigenschaften und Struktur von Naturstoffen zu untersuchen, um etwaige, diesen innewohnende Prinzipien erkennen zu können.

Auf der anderen Seite stellt sich die Frage, wie Eigenschaften potentieller Zielproteine für das zielgerichtete Design von Substanzkollektionen genutzt werden können. Hier wurden bereits Konzepte vorgestellt, die auf einer Gruppierung von Zielproteinen entsprechend ihrer evolutionären Verwandtschaft oder Funktion in sogenannte Target-Familien beruhen. Der wesentliche Nachteil dieser Methoden liegt darin, daß sie allesamt lediglich nahverwandte Sequenzhomologe berücksichtigen. Eine der wesentlichen Erkenntnisse, nämlich daß die prinzipielle Struktur von Proteinen evolutionär stärker konserviert ist als ihre Sequenz,<sup>[17]</sup> wurde hingegen bislang nur wenig genutzt. Es liegt daher nahe, diese Erkenntnis bei der Entwicklung neuer Leitprinzipien für das Design biologisch relevanter Verbindungsbibliotheken stärker als bisher zu nutzen.

Ziel dieser Dissertation war es daher, neue Werkzeuge zu schaffen, die die Welt der Proteine und die Welt der Naturstoffe miteinander verbinden. Auf der Grundlage einer rein strukturellen Sicht auf Proteine sollte ein abstrahierendes Leitprinzip entwickelt werden, das die Fesseln der bisherigen überwiegend evolutionären und funktionalen Betrachtungsweise sprengt und zu Proteinstruktur-Ähnlichkeitsclustern als Ausgangspunkten für eine biologisch relevante Konzipierung von Verbindungsbibliotheken führt. Eine grundlegende strukturelle Betrachtung von Naturstoffen sollte darüber hinaus zu einer Klassifizierung führen, die als Navigator durch die Strukturvielfalt der Naturstoffe und als pragmatischer Hypothesen- und Ideengenerator für die Auswahl von biologisch prävalidierten Ausgangspunkten für den gezielten Aufbau einer Substanzkollektion dienen kann. Beide Konzepte erwiesen sich als nützlich, zeigten sich jedoch vor allem durch ihren synergistischen Gebrauch besonders vielversprechend.

## 2 Allgemeiner Teil

### 2.1 Wirkstofffindungsprozeß

Die im Rahmen des Wirkstofffindungsprozesses zum Einsatz kommenden Methoden können in wissensbasierte Technologien einerseits und solche, die kein Wissen über das Target und seine Liganden erfordern, andererseits eingeteilt werden.

Um einen Hit für ein bestimmtes Target zu erhalten, werden in einem klassischen Ansatz im Wirkstofffindungsprozeß der pharmazeutischen Industrie große Kollektionen von Verbindungen Assay-Verfahren unterzogen, die meist im HTS-Format robotergestützt durchgeführt werden. In der Regel kommen hierzu die über Jahrzehnte gewachsenen „historischen“ Verbindungskollektionen, in die Zwischenprodukte und Verbindungen aus früheren medizinischen, agrochemischen oder anderen Forschungsprogrammen einfließen, zum Einsatz. Typische Verbindungsbibliotheken großer Firmen umfassen ca. eine Million Verbindungen.<sup>[18]</sup> Darüber hinaus wurden zur Ergänzung dieser Verbindungskollektionen seit Beginn der frühen 90er Jahre große Verbindungsbibliotheken mittels Hochdurchsatz-Synthesetechnologien erstellt. Die Getriebenheit durch schiere Zahlen begann mit Bibliotheken, bei denen die einzelnen Verbindungen als Mischungen vorlagen, und ging über zu Verbindungskollektionen von Reinverbindungen. Als Leitprinzip diente zunächst die einfache chemische Zugänglichkeit der Verbindungen, was vor allem an der beschränkten Verfügbarkeit effizienter chemischer Methoden zur Darstellung einer Vielzahl komplexer Verbindungen lag. Es ist daher nicht verwunderlich, daß gerade die frühen kombinatorischen Bibliotheken aus Verbindungen bestanden, die kaum als arzneistoffähnlich bezeichnet werden konnten. Da die Hitraten den hohen Erwartungen nicht entsprachen, wurde dieser Ansatz wieder verlassen, um die kombinatorische Methodik auf kleinere, fokussierte Verbindungskollektionen anzuwenden, die um wenige, vielversprechende Grundgerüste aufgebaut wurden. Damit wurde ein strategischer Wechsel des Einsatzes der kombinatorischen Chemie vom Leitstrukturfindungsprozesses hin zur Anwendung in der Leitstrukturoptimierung vollzogen.<sup>[19]</sup> Aussichtsreiche Grundgerüste können sogenannte privilegierte Strukturen<sup>[14]</sup> sein, die sich z.B. im Bereich der Entwicklung von Liganden für G-Protein-gekoppelte Rezeptoren als zielführende Strategie erwiesen haben.<sup>[18]</sup> Insgesamt wurde die Entwicklung von Verbindungskollektionen auf ausgewählte Zielproteinklassen ausgerichtet.<sup>[20]</sup> Hier werden Strukturinformationen des biologischen Targets und Kenntnisse

über den biologischen Wirkmechanismus genutzt, um das Design einer Verbindungsbibliothek zu steuern. Auch werden rechnergestützte Verfahren des virtuellen Screenings eingesetzt, wobei zwischen virtuellen Filterungsprozessen und virtuellem Screening unterschieden wird. Beim virtuellen Filtern werden Verbindungen eliminiert, die allgemein definierten Anforderungen nicht genügen. Dazu gehören Richtlinien wie z.B. Lipinskis „Rule of Five“,<sup>[21]</sup> die es auf der Grundlage einfacher Deskriptoren erlauben, die Wahrscheinlichkeit einer intestinalen Resorption und damit der peroralen Bioverfügbarkeit abzuschätzen. Aus einer Analyse von 2245 Stoffen des „World Drug Index“ (Thomson Derwent, <http://thomsonderwent.com>) resultierte die Regel, daß eine schlechte Resorption wahrscheinlicher ist, wenn eine Verbindung mehr als 5 Wasserstoffbrückendonatoren, mehr als 10 Wasserstoffbrückenakzeptoren, ein relatives Molekulargewicht  $\geq 500$  und einen berechneten logP-Wert (CLogP, Octanol/Wasser-Verteilungskoeffizient) von  $\geq 5$  aufweist. Des Weiteren werden substrukturbasierte Filter für die Analyse des toxikologischen Profils, möglicher Cytochrom-P450-Interaktion<sup>[22]</sup> oder des Interaktionspotentials mit dem kardialen hERG-(,human ether-a-go-go-related gene’)-Kaliumkanal eingesetzt.<sup>[23]</sup> Für das virtuelle Screening stehen Algorithmen zur Verfügung, die auf zweidimensionalen topologischen Deskriptoren beruhen und ein schnelles Durchforsten von Datenbanken erlauben. Das virtuelle Screening mit dreidimensionalen Pharmakophoren gewinnt zunehmend an Bedeutung.<sup>[24]</sup> Die vierte Dimension stellt die Zielproteinstruktur dar. Das dem biologischen Screening am nächsten kommende virtuelle Verfahren ist das virtuelle ‚Docking‘ und ‚Scoring‘, bei dem vielversprechende Verbindungen nach der Qualität ihrer Interaktion mit einer Bindungstasche ausgewählt werden. Dieses Verfahren wird aber stark limitiert durch die Verfügbarkeit von Kristallstrukturen (insbesondere solcher mit gebundenen Liganden) und durch noch ungelöste Probleme im Bereich der Bewertungsfunktionen für Bindungsaffinität, v.a. im Hinblick auf das Screening großer virtueller Bibliotheken.<sup>[25]</sup> Diese virtuellen Methoden können insgesamt allenfalls als Werkzeuge für die Vorhersage angesehen werden und erfordern eine kontinuierliche Validierung der vorgeschlagenen aktiven Verbindungen durch chemische Synthese und biologische Testung.

Eine wichtige Stellung im Wirkstofffindungsprozeß nimmt eine Vorgehensweise ein, die unter dem Begriff „Chemogenomik“ subsumiert wird. Hier werden Verbindungsbibliotheken dazu genutzt, neue Targets unter Verwendung arzneistoffartiger Moleküle in einem stark integrierten Prozeß zu identifizieren und zu validieren. Rechnergestützte Methoden spielen hier nicht nur beim Aufbau der Verbindungsbibliotheken eine große Rolle, sondern auch bei

der Identifizierung und dem Clustering biologischer Targets. Der Target-Raum wird meist aufgrund der Sequenz der Proteine definiert. Sequenzähnlichkeiten werden genutzt, um Aufschlüsse über die phylogenetische Verwandtschaft von Protein-Targets zu erhalten und um diese letztlich entsprechend ihrer sequentiellen Ähnlichkeit in sogenannten Target-Familien zusammenzufassen. Dies erlaubt die Entwicklung von Liganden ausgehend von auf Target-Familien zugeschnittenen fokussierten Verbindungsbibliotheken. Der erste Schritt in einem Chemogenomik-Ansatz ist daher, die DNA- oder Aminosäuresequenz neuer Target-Gene bzw. -Proteine aufzuspüren und diese (bekannten) Target-Familien zuzuordnen. Das enorme Wachstum verfügbarer Daten über potentielle Targets aus Genomik- und Proteomik-Projekten erfordert die Anwendung und Entwicklung effizienter bioinformatischer Methoden, um diese für die Wirkstoffentwicklung nutzbar machen zu können. Festzustellen ist, daß bei den gegenwärtig verfolgten Chemogenomik-Ansätzen die Proteinstruktur gegenüber der Sequenz eine untergeordnete Rolle spielt. In dieser Arbeit soll dieser Aspekt der Proteinstruktur beleuchtet werden, nicht zuletzt aufgrund der Tatsache, daß im Rahmen der Evolution die Struktur von Proteinen stärker konserviert ist als ihre Sequenz.<sup>[17]</sup>

Insgesamt bleibt festzuhalten, daß sich nur eine verstärkte Integration virtueller Methoden, der Synthese und der biologischen Testung zu einem überdurchschnittlich erfolgversprechenden Wirkstofffindungsprozeß führt.

## 2.2 Domänenarchitektur von Proteinen

Proteine sind modular aufgebaute Biomoleküle, die aus sogenannten Domänen als Baugruppen bestehen. Eine Domäne ist als ein diskreter Teil eines Proteins definiert, der sich unabhängig von der übrigen Struktur zu einer kompakten Anordnung von Sekundärstrukturelementen faltet, die über mehr oder weniger komplexe Verbindungspeptide miteinander verknüpft sein können. Verschiedene Domänen können den gleichen Faltungstyp aufweisen. Ursächlich hierfür kann eine konvergente Entwicklung aufgrund funktionaler oder physikalischer Zwänge sein, da es nur eine begrenzte Anzahl energetisch möglicher Anordnungen von Sekundärstrukturelementen gibt. Als weitere Ursache kommt eine divergente evolutionäre Entwicklung ausgehend von einem gemeinsamen Vorfahren in Betracht, die so weit gehen kann, daß eine Sequenzverwandtschaft nicht mehr erkennbar ist.<sup>[26, 27]</sup> Proteindomänen können daher als strukturell konservierte, jedoch genetisch mobile Einheiten verstanden werden.<sup>[28]</sup> So ist es möglich, daß eine ähnliche Proteinarchitektur in

mehr oder weniger modifizierter Form in unterschiedlichen funktionalen Kontexten in der Natur wiederverwendet wird. Sehr häufig sind dann die Ligandenbindungs- oder katalytischen Stellen strukturell divers, während die prinzipielle Architektur ähnlich ist. Obwohl die Zahl verschiedener Proteine im Menschen auf 100 000 bis 450 000 geschätzt wird, ist man sich einig, daß die Zahl der Domänenfamilien oder gar topologisch unterschiedlicher Faltungstypen deutlich geringer ist. Zur Zeit sind der SCOP-(Structural Classification of Proteins)-Datenbank<sup>[29]</sup> zufolge ca. 800 Faltungstypen bekannt,<sup>[30]</sup> die aus einer Klassifizierung aller strukturell charakterisierter Proteine nach ihrer 3D-Struktur und evolutionären Argumenten resultieren. Daten aus den laufenden Projekten zur Genomsequenzierung erlauben eine Abschätzung der Anzahl der in der Natur vorkommenden Faltungstypen und Proteinfamilien. Aktuelle Schätzungen zufolge liegt die Zahl der Faltungstypen im Bereich 1 000 bis 10 000 und die Zahl der Sequenzfamilien zwischen 4 000 und 50 000.<sup>[27, 31-38]</sup> Es gilt als wahrscheinlich, daß sich die große Mehrzahl der Proteinfamilien ca. 1 000 Faltungstypen zuordnen läßt.<sup>[38]</sup> In diesem Zusammenhang sollte nicht unerwähnt bleiben, daß die Definition des Faltungstyps in vielen Fällen eine empirische Näherung darstellt und daß sich auch Experten über die Zuordnungen vieler Proteine zu einem Faltungstyp nicht einig sind. Häufig fließen nicht nur Strukturinformationen, sondern auch funktionelle und evolutionäre Erwägungen in die Zuordnung mit ein. Mit großer Wahrscheinlichkeit wäre jedoch eine Kategorisierung nach ausschließlich strukturellen Aspekten zielführender, da Proteine des gleichen Faltungstyps nicht notwendigerweise einen gemeinsamen Vorfahren haben oder eine ähnliche physiologische Rolle spielen.

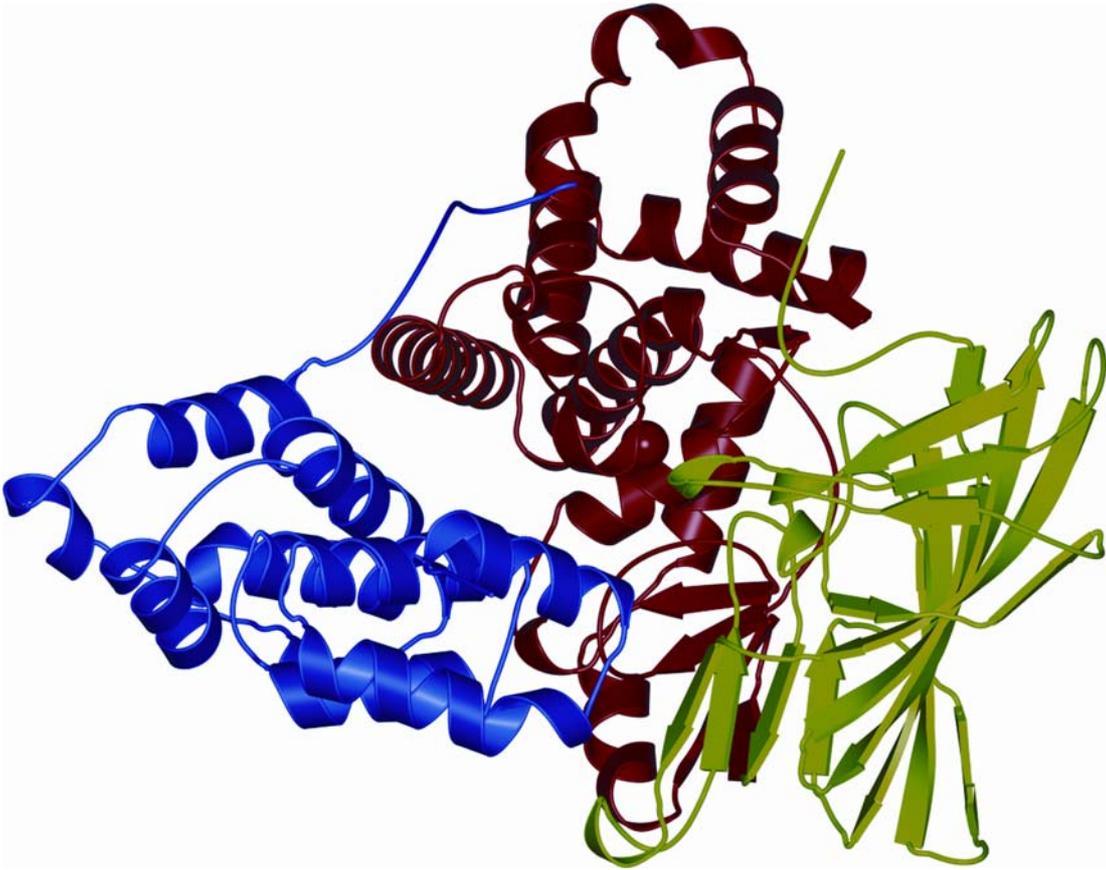


Abb. 1: Domänenarchitektur der Leukotrien-A<sub>4</sub>-Hydrolase (PDB-Code 1HS6). Dieses Enzym besteht aus drei verschiedenen Domänen, einer N-terminalen Domäne (P1-A208, grün, Faltung nach SCOP: ‚LTA<sub>4</sub>H N-terminal domain‘), einer C-terminalen Domäne (D461-D610, blau, Faltung:  $\alpha$ - $\alpha$ -Superhelix) und einer katalytischen Domäne (L209-Y460, rot, in CPK-Darstellung das katalytische Zinkion, Faltung: Zinkin-artig).

### 2.3 Die ligandenbindenden Kernstrukturen von Proteindomänen

Für die Entwicklung niedermolekularer Proteinliganden sind die Ligandenbindungsstellen oder katalytisch aktiven Zentren räumlich gesehen die wichtigsten Bereiche einer Proteindomäne. Diese Bindungsstellen sind in der ligandenbindenden Kernstruktur einer Proteindomäne lokalisiert, in der die katalytische Umsetzung im Falle von Enzymen oder ganz allgemein das Bindungsereignis stattfindet. In ihrer Untersuchung der Beziehung zwischen Sequenz und Struktur von Proteinen definierten Chothia und Lesk strukturell konservierte Kernbereiche und variable Bereiche. Verwandte Proteine weisen in ihrem ‚common core‘ ein ähnliches Faltungsmuster auf, das die wesentlichen um die Bindungsstelle herum angeordneten Sekundärstrukturelemente umfaßt.<sup>[39]</sup>

Die topologisch einzugrenzenden Kernbereiche einer Proteindomäne zu definieren, ist umso wichtiger, je größer die Proteindomäne im Vergleich zum Kernbereich wird. Erwägungen zur

strukturellen Ähnlichkeit bei der Entwicklung von Liganden sollten auf diese ligandenbindenden Kernstrukturen begrenzt werden. Dies erweitert die strukturbasierte Sichtweise von Proteinen, weil damit auch Teilelemente eines zuvor definierten Faltungstyps berücksichtigt werden können. So ist es möglich, daß sich gewisse strukturelle Faltungscharakteristiken in einem anderen Faltungstyp wiederholen können. Wenn diese eine ligandenbindende Kernstruktur beschreiben, so ist eine umfassende Kategorisierung möglich, die ausschließlich auf strukturellen Argumenten basiert und von zuvor definierten Faltungstypen weitgehend unabhängig ist.

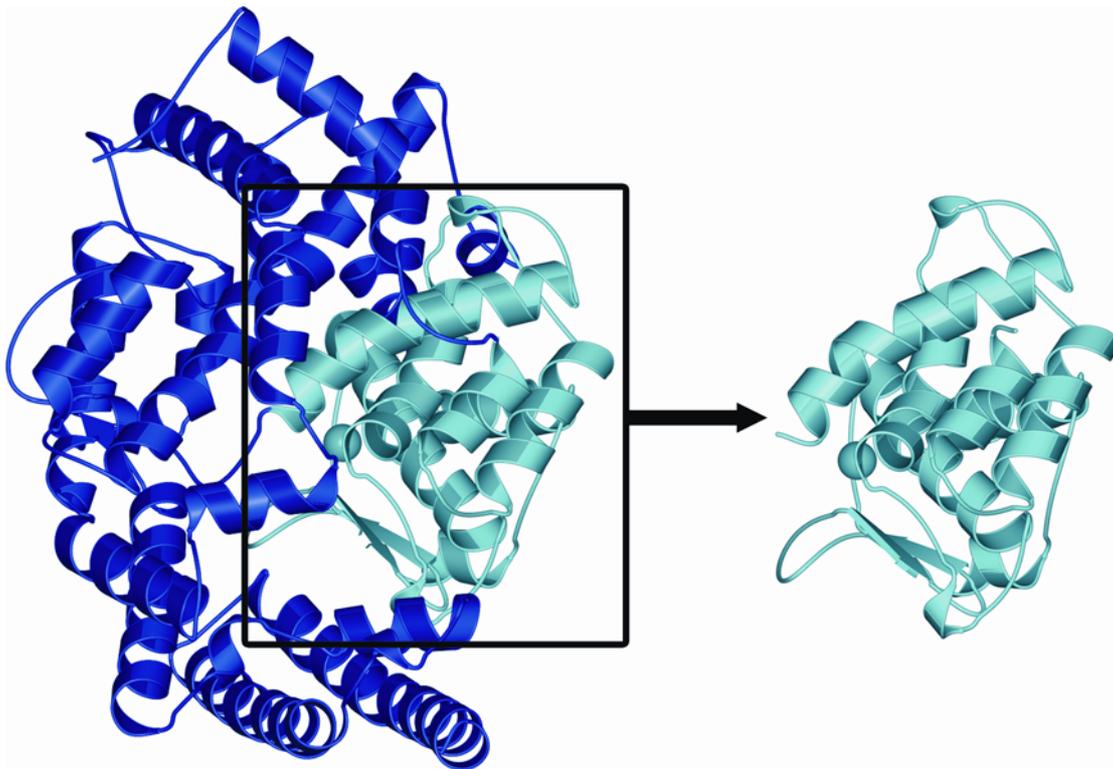


Abb. 2: Definition des katalytischen Kernbereichs am Beispiel des Angiotensin-Conversionsenzyms (ACE, PDB-Code 1O86). Beim ACE handelt es sich um ein Eindomänenprotein (D40-P617), bei dem ausgehend vom katalytischen Zentrum (in CPK-Darstellung das katalytische Zinkion) der für die katalytische Umsetzung relevante Bereich (P312-P424, I521-G583) definiert werden kann.

## 2.4 Naturstoffe

Naturstoffe im engeren Sinne sind niedermolekulare organische Verbindungen, die von Organismen (Tieren, Pflanzen und Mikroorganismen) produziert werden und im wesentlichen ihrem Sekundärstoffwechsel entstammen. Sie nehmen am Primärmetabolismus nicht aktiv teil. Naturstoffe werden daher auch Sekundärstoffe genannt, deren Auftreten sporadisch und bisweilen spezifisch für ein bestimmtes Taxon ist. Sie lassen sich (a) nach ihrer chemischen

Struktur (Tannine, Alkaloide, Amine, Aminosäuren, Flavonoide, Steroide etc.), (b) nach ihrer biogenetischen Herkunft (Polyketide, Isoprenoide, Shikimisäureabkömmlinge etc.) und (c) nach ihrer Funktion bzw. Wirkung (Antibiotika, Farbstoffe, Pheromone etc.) einteilen.

In der Medizin wurden von Anfang an Naturstoffe zur Behandlung menschlicher Krankheiten eingesetzt. Naturstoffe spielten auch in der modernen westlichen pharmazeutischen Industrie als Arzneistoffquellen eine bedeutende Rolle. Einer Analyse zufolge waren 48% von 877 neuen chemischen Substanzen, die zwischen 1981 und 2002 als Arzneistoffe zugelassen wurden, entweder Naturstoffe, Naturstoff-Mimetika oder von Naturstoffen abgeleitete Verbindungen.<sup>[40]</sup> Obwohl Naturstoffe offensichtlich für die Wirkstoffentwicklung gut geeignet sind, haben viele pharmazeutische Firmen ihre Forschungsaktivitäten in der Naturstoffforschung zurückgefahren. Zur Begründung wird häufig vorgebracht, daß Naturstoffe und ihre Derivate synthetisch im allgemeinen zu schwer zugänglich sind, um große Verbindungsbibliotheken aufzubauen, mit denen die immer hungrigen HTS-Roboter gefüttert werden könnten. Naturstoffe könnten jedoch eine Renaissance erleben, da sie für die gezielte Entwicklung von Verbindungskollektionen biologisch relevante Ausgangspunkte im chemischen Strukturraum darstellen, die vergleichsweise hohe Hitraten bei vergleichsweise kleiner Bibliotheksgröße liefern.<sup>[11-13, 41]</sup>

Naturstoffe sind bezüglich ihrer Bindungseigenschaften an biologische Makromoleküle evolutionär erprobt. Sie interagieren im Zuge ihrer Biosynthese mit verschiedenen Proteinen und gehen Wechselwirkungen mit weiteren Proteinen ein, wenn sie z.B. im Rahmen von Kommunikation oder chemischer Verteidigung ihre biologische Funktion erfüllen. Meist werden sie von ihren Erzeugerorganismen, die sich grob in Tiere, Pflanzen, Pilze und Mikroorganismen einteilen lassen, als Sekundärmetabolite produziert, d.h., sie werden in der Regel nicht für lebensnotwendige Stoffwechselprozesse benötigt. Verschiedene Naturstoffklassen weisen darüber hinaus unterschiedliche biologische Aktivitäten auf. Ihre spezielle biologische Prävalidierung und evolutionäre Erprobung machen Naturstoffe zu besonders gut geeigneten Leitmotiven für die Konzipierung biologisch relevanter Verbindungskollektionen. Dieses Kriterium der biologischen Prävalidierung erfüllen jedoch nicht nur Naturstoffe. Verschiedene nicht-natürliche Substanzklassen, die vor allem in medizinisch-chemischen Programmen in einer Art beschleunigter künstlicher Evolution hervorgebracht wurden, weisen ähnliche Eigenschaften auf.

Naturstoffe wurden bereits mehrfach statistisch untersucht.<sup>[10, 42, 43]</sup> Im Vergleich mit Arzneistoffen ist festzustellen, daß ihr durchschnittliches Molekulargewicht fast identisch ist. Interessant ist auch, daß die übrigen Lipinski-Parameter im Mittel in beiden Kollektiven relativ ähnlich sind. Im Vergleich zu Naturstoffen enthalten Arzneistoffe im Schnitt die doppelte Anzahl von Stickstoffatomen, wohingegen erstere mehr Sauerstoffatome enthalten. Unter Verwendung einer Selbstorganisierenden Karte (SOM, self-organizing map) zeigte ein Vergleich topologischer Pharmakophore von Naturstoffen und Arzneistoffen viele ähnliche Konstellationen generalisierter Atomtypen. Daraus kann geschlossen werden, daß es ein vielversprechender Ansatz sein könnte, bei der Entwicklung von Arzneistoffen auf Naturstoffen aufzubauen.<sup>[10]</sup>

## 2.5 Chemische Genetik, Chemische Genomik oder Chemogenomik?

Die Begriffe Chemische Genetik, Chemische Genomik oder Chemogenomik werden parallel gebraucht, um einen neuen Weg für die Erforschung der Biologie eines Targets unter Verwendung chemischer Substanzen zu beschreiben.<sup>[44]</sup>

Die chemische Genetik beschreibt ganz allgemein den Gebrauch niedermolekularer organischer Moleküle, um Veränderungen im Phänotyp eines Organismus zu induzieren.<sup>[45]</sup> Stuart Schreiber beschreibt die chemische Genetik oder Genomik als einen systematischen Ansatz, der neue Entwicklungen in der Robotik, der kombinatorischen Chemie, des HTS und der Bioinformatik nutzt.<sup>[46]</sup> Diese Methode ergänzt den klassischen Weg der Target-Identifizierung und -Validierung über Mutagenese-Studien (klassische Genetik) oder „Knock-out“-Experimente (reverse Genetik) mit anschließender Identifizierung, Reinigung und Charakterisierung endogener Liganden, aus denen dann in mehr oder weniger klassischen Strategien Wirkstoffkandidaten entwickelt werden.<sup>[44]</sup> Eine entscheidende Voraussetzung dafür, solche Gen-„Knock-out“-Experimente unter Verwendung chemischer Liganden durchführen zu können, ist, daß die dafür benutzten kleinen molekularen Substanzen selektiv an ein bestimmtes Genprodukt binden und gezielt dessen Funktion beeinflussen.

In der klassischen Genetik werden auf eine randomisierte Mutation hin interessante Phänotypen selektioniert, um dann die für die Änderung des Phänotyps verantwortlichen mutierten Gene zu identifizieren. In der chemischen Genetik hingegen werden Zellen einer

Verbindungsbibliothek ausgesetzt und die Moleküle, die die Änderung des Phänotyps bewirken, selektioniert, um schließlich das Protein oder die Proteine, die für den Phänotypwechsel verantwortlich sind, zu identifizieren.<sup>[47]</sup> Die chemische Genetik ist aufgrund von HTS automatisiert durchführbar und verfügt über das Potential, systematisch die Funktion und Rolle von Proteinen in der Zelle zu analysieren. Fernziel der chemischen Genetik ist es daher, für jedes Genprodukt einen niedermolekularen Liganden zu identifizieren.<sup>[48]</sup>

Die reverse Genetik beschreibt das gezielte Ausschalten eines interessanten Gens mit einer „Knock-out“-Mutation und das Studium der sich durch den Verlust des Genprodukts ergebenden Konsequenzen für den Phänotyp. Solche gezielten Mutationen werden traditionell durch homologe Rekombination einer embryonalen Stammzelllinie durchgeführt. Mittlerweile ist aber auch die gezielte Ausschaltung eines Gens durch RNA-Interferenz (RNAi) üblich. Anders als der klassisch-genetische Ansatz ist der reverse spezifisch und reproduzierbar. Der überlebende Mutant ist jedoch gezwungen, den Verlust des Genprodukts zu kompensieren. Dies kann zu einer Reihe von sekundären kompensatorischen Ereignissen führen, die den wahren Effekt der Genausschaltung überlagern können. Im Vergleich dazu wird in der reversen chemischen Genetik ein Protein ausgewählt, gegen das eine Verbindungsbibliothek gescreent wird, um modulierende Liganden für dieses Protein zu identifizieren. Es werden die phänotypischen Konsequenzen untersucht, wenn Zellen diesem Liganden ausgesetzt werden. Ein wesentlicher Vorteil ist, daß das Genprodukt unter hoher zeitlicher Kontrolle direkt angesprochen wird. Durch Entfernung des Moleküls wird die ursprüngliche Funktionsfähigkeit des Genprodukts wiederhergestellt.<sup>[48]</sup>

Um das Ziel der chemischen Genetik oder Genomik, die zelluläre Biologie zu erforschen und Signalwege für bestimmte Genprodukte unter Verwendung kleiner Moleküle als modulierender Liganden zu verstehen, zu erreichen, ist die Verfügbarkeit einer zur biologischen Diversität komplementären chemischen Diversität unabdingbar.

Chemogenomik hingegen beschreibt eher eine Methode, die auf die schnelle Identifizierung neuer Wirkstoffe ausgerichtet ist. Sie umfaßt mehrere Wirkstofffindungs-Technologien von der Target-Identifizierung und -Validierung über Wirkstoffdesign und -synthese bis hin zur biologischen Evaluierung und physikochemischen Profilierung. Diese Technologien werden in einem hoch-integrierten Ansatz nicht nur für einzelne Targets, sondern für ganze Target-Familien eingesetzt.<sup>[44]</sup> Ausgehend vom genetischen Code werden Genprodukte als potentielle

pharmakologische Targets identifiziert. Diese Proteine werden nach ihrer Funktion (Proteasen, Kernrezeptoren, Ionenkanäle, G-Protein-gekoppelte Rezeptoren usw.) klassifiziert und aufgrund ihrer genetischen Sequenz in Unterfamilien gruppiert. Eine ähnliche Abstraktion kann in der „Welt der Chemie“ vorgenommen werden, indem Verbindungen als „Fingerabdrücke“ beschrieben werden, die wiederum in eine chemische Struktur übersetzt werden können. Diese Informationen werden in multidimensionale Pharmakophor-Modelle überführt, nach denen die Moleküle klassifiziert und in Unterbibliotheken nach bestimmten Ähnlichkeitskriterien zusammengefaßt werden. Die resultierenden virtuellen Target- und Ligandenräume müssen dann mit dem Ziel überlagert werden, in der Schnittmenge neuartige und qualitativ hochwertige Liganden für bekannte oder neue Targets zu identifizieren. Das ultimative Ziel ist es, Annotationssysteme zu finden, die eine direkte Übersetzung genetischer Sequenzdaten (Sequenzraum) in chemische Strukturen (topologischer Raum) und umgekehrt erlauben.<sup>[44]</sup> Die Tatsache, daß ähnliche Liganden an ähnliche Targets binden, kann als Grundkonzept wissensbasierter, auf der Chemogenomik aufbauender Strategien im Wirkstofffindungsprozeß angesehen werden.<sup>[49]</sup> Beispielhaft für einen derartigen Ansatz ist das sogenannte „Structure-Activity Relationship Homology“-Konzept, kurz SARAH genannt.<sup>[50]</sup> Dieses Konzept geht davon aus, daß eine Referenzmenge von Liganden für ein bestimmtes Target eine Pharmakophor-Hypothese nicht nur für ein einziges Target, sondern für eine ganze Target-Familie darstellt. Eine Konservierung der Sequenz in den Bindungstaschen innerhalb einer Target-Familie führt zu einer Konservierung ihres Aussehens und ihrer physikochemischen Eigenschaften, was sich wiederum in ähnlichen strukturellen Anforderungen an Liganden widerspiegelt. D.h., daß Moleküle, die als Liganden für ein bestimmtes Target identifiziert wurden, eine gewisse Wahrscheinlichkeit aufweisen, auch Aktivität an anderen Targets derselben Familie zu entfalten. Liganden nahverwandter (homologer) Rezeptoren (v.a. im Bereich der G-Protein- gekoppelten Rezeptoren) werden daher generell im Rahmen von pharmazeutischen Forschungsprogrammen zur Leitstrukturfindung als mögliche Ausgangspunkte für Rezeptoren akzeptiert, für die noch keine spezifischen Liganden identifiziert wurden.

### 3 Ziel der Arbeit

Ziel dieser Arbeit war es, leistungsfähige Konzepte für ein effizientes Design kombinatorischer Verbindungsbibliotheken zu entwickeln und experimentell zu verifizieren. Ausgangspunkt war die Notwendigkeit, im Rahmen der chemisch-biologischen und medizinisch-chemischen Wirkstoffforschung, Bibliotheken chemischer Verbindungen mit biologischer Relevanz zu generieren.<sup>[11-13, 41]</sup> Die Tatsache, daß schiere Bibliotheksgröße allein nicht genügt, um ausreichende Hitraten biologisch aktiver Substanzen hervorzubringen, führte zu der Überlegung, kleinere, fokussierte Bibliotheken ausgehend von biologisch prävalidierten Ausgangspunkten aufzubauen, um höhere Hitraten qualitativ hochwertiger Proteinliganden mit hoher Zuverlässigkeit zu erzielen. Diese Betrachtungsweise schloß nicht nur Erwägungen in bezug auf die Qualität, sondern auch im Hinblick auf die Effizienz eines kombinatorischen Ansatzes ein. Die Frage, die sich dabei stellte, war: Wo sind im chemischen Strukturraum die biologisch relevanten Verbindungen zu finden?

Um diese sicherlich nicht einfache Frage beantworten zu können, sollten sowohl aus der „Welt der Proteine“, d.h. aus der Perspektive der Targets, als auch aus der „Welt der Naturstoffe“, also aus der Perspektive natürlich vorkommender Proteinliganden, Prinzipien abgeleitet werden, die zu einer Ausgangspunkthypothese für das Design fokussierter und biologisch relevanter Verbindungsbibliotheken führen. Diese Prinzipien sollten sowohl bei der Betrachtung von Proteinen als auch von Naturstoffen auf strukturellen Erwägungen beruhen.

Proteine werden im Rahmen der Chemogenomik überwiegend aus der Perspektive ihrer Funktion und ihrer Aminosäureprimärsequenz betrachtet. Diese limitierende Betrachtungsweise sollte im Rahmen dieser Arbeit durch eine rein strukturbasierte, aber dennoch abstrahierende Sicht auf Proteine ergänzt werden. Es sollte untersucht werden, inwieweit eine strukturelle Ähnlichkeit zwischen Proteinen, unabhängig von ihrer Funktion und ihrer Sequenzähnlichkeit, dafür genutzt werden kann, strukturelle Leitprinzipien für fokussierte kombinatorische Verbindungsbibliotheken zu identifizieren. Die aus dieser Untersuchung gewonnenen Erkenntnisse und das daraus resultierende Konzept sollten experimentell in konkreten Liganden-Entwicklungsprojekten verifiziert werden.

Eine analoge Untersuchung sollte für Naturstoffe als natürliche und evolutionär erprobte Liganden für Biomakromoleküle, im speziellen für Proteine, durchgeführt werden. Es sollte

geprüft werden, inwieweit sich die strukturelle Vielfalt der Naturstoffe in einem Ordnungsprinzip erfassen läßt. Eine sich daraus ableitende strukturbasierte Klassifizierung von Naturstoffen könnte dazu dienen, die Prinzipien, die bei der Bindung an Proteine eine Rolle spielen, besser zu verstehen. Das aus diesen Erkenntnissen gewonnene Konzept sollte ebenfalls experimentell verifiziert werden.

Das ultimative Ziel dieser Arbeit war es, die Strukturwelten von Proteinen und Naturstoffen und die abgeleiteten Konzepte zusammenzuführen, um aus diesem Synergismus ein Werkzeug zu entwickeln, das auf hohem Abstraktionsniveau die Identifizierung biologisch relevanter Ausgangspunkte im chemischen Strukturraum ermöglicht. Es sollten Verbindungsbibliotheken entstehen, die trotz ihrer vergleichsweise kleinen Größe idealerweise hohe Hitraten strukturell innovativer, potenter und selektiver Proteinmodulatoren mit hoher Zuverlässigkeit liefern, wobei nach Möglichkeit auch die chemische Zugänglichkeit durch strukturelle Abstraktion und Vereinfachung mitberücksichtigt werden sollte.

Die in dieser Arbeit vorgestellten Konzepte und Ergebnisse erforderten eine Integration von Methoden der Bio- und Chemoinformatik, der organisch-chemischen Synthese, der Biochemie und der Zellbiologie.

## 4 Spezieller Teil

### 4.1 Proteinstruktur-Ähnlichkeitsclustering (PSSC)

#### 4.1.1 Analyse der Literatur – Eine Machbarkeitsstudie

Das Vorhaben, aus der Gruppierung von Proteindomänen nach rein strukturellen Gesichtspunkten ein Konzept für die gezielte Entwicklung von Verbindungsbibliotheken abzuleiten, sollte durch eine retrospektive Analyse von Modellfällen zunächst auf prinzipielle Machbarkeit geprüft werden.

Vor dem Hintergrund, daß in der Evolution die Struktur der Proteine stärker konserviert ist als ihre Sequenz<sup>[17]</sup> und daß Naturstoffe als evolutionär erprobte Proteinliganden die Hitraten steigern lassen, wenn diese im HTS miteinbezogen werden,<sup>[11]</sup> wurden folgende Hypothesen für ein auf Proteinstrukturähnlichkeit basierendes Konzept formuliert:

1. Eine Gruppierung von Proteinen nach ausschließlich strukturellen Gesichtspunkten, auch bei niedriger Sequenzidentität der beteiligten Proteine, führt ähnlich wie beim Clustering von Proteinen in sogenannte Target-Familien, bei dem ein hoher Grad an evolutionärer Verwandtschaft und eine ähnliche Funktion erforderlich sind, zu validen Pharmakophorhypothesen für eine gegebene Gruppe strukturell ähnlicher Proteine.
2. Naturstoffe sind im chemischen Strukturraum besonders gut geeignete Ausgangspunkte für den Aufbau einer Verbindungsbibliothek, die innerhalb eines Clusters strukturell ähnlicher Proteine potente und ggf. selektive Liganden hervorbringt. Diese Hypothese gründet im wesentlichen darauf, daß Naturstoffe evolutionär erprobt sind und inhärent die Eigenschaft besitzen, an Proteine zu binden. Vorläufer der Naturstoffe und sie selbst müssen zunächst an Proteine binden, die ihre Biosynthese katalysieren, und sie binden an weitere Proteine, wenn sie ihre biologische Wirkung, z.B. im Zuge chemischer Kommunikation und Verteidigung, entfalten.
3. Hypothese 2 soll nicht dazu führen, daß nicht-natürliche Verbindungen mit biologischer Relevanz, etwa solche, die „privilegierte“<sup>[14]</sup> Strukturen enthalten oder aus medizinisch-chemischen Programmen hervorgegangen sind, von der Betrachtung ausgeschlossen werden.

Auf der Grundlage dieser Hypothesen wurden aus der Literatur Beispiele für Verbindungskollektionen herausgefiltert, die den in diesen Hypothesen formulierten Kriterien genügen. Auf diese Weise sollte geprüft werden, ob es Anhaltspunkte für die prinzipielle Machbarkeit eines auf Proteinstrukturähnlichkeit beruhenden Clusteringkonzepts gibt, das eine gezielte Entwicklung von Proteinliganden erlaubt. Die identifizierten Beispiele wurden ursprünglich nicht nach diesen Kriterien konzipiert.

### ***Leukotrien-A<sub>4</sub>-Hydrolase und Aminopeptidasen – Entwicklung von Inhibitoren der Leukotrien-A<sub>4</sub>-Hydrolase***

Die Leukotrien-A<sub>4</sub>-Hydrolase (LTA<sub>4</sub>H) ist ein bifunktionelles Zink-Metalloenzym, das LTA<sub>4</sub> in LTB<sub>4</sub> umwandelt und darüber hinaus auch Aminopeptidaseaktivität aufweist. LTB<sub>4</sub> ist ein potenter chemotaktischer Mediator, der z.B. bei Entzündungsreaktionen oder bei der Immunantwort involviert ist. Die kritische Rolle der LTA<sub>4</sub>H bei der Bildung von LTB<sub>4</sub> macht dieses Enzym zu einem attraktiven Target. Die LTA<sub>4</sub>H katalysiert die vinyloge Hydrolyse des Leukotrienepoxids LTA<sub>4</sub> zu LTB<sub>4</sub> in ihrem Zink-enthaltenden aktiven Zentrum.<sup>[51, 52]</sup> Das Zinkion fungiert als Lewis-Säure, indem es den Epoxidring polarisiert und die negative Ladung im Übergangszustand stabilisiert. Im Falle der LTA<sub>4</sub>H stellte das Zinkbindungsmotiv (HEXXH-X<sub>18</sub>-E) den Ausgangspunkt für eine Erforschung der Verwandtschaft dieses Enzyms zu zinkbindenden Metallopeptidasen dar.<sup>[53]</sup> Nach der SCOP-(Structural Classification of Proteins)-Datenbank<sup>[29]</sup> wird dieses Enzym auch dem Zinkin-artigen Faltungstyp und der Superfamilie der Metalloproteasen zugeordnet. Das sowohl bei der LTA<sub>4</sub>H als auch bei Metallopeptidasen konservierte HEXXH-Sequenzmotiv legte nahe, Peptidase-Inhibitoren als potentielle Liganden für die LTA<sub>4</sub>H zu erforschen. In der Tat hemmte Bestatin (**1**), ein Naturstoff und bekannter Aminopeptidasehemmer, auch die LTB<sub>4</sub>-Biosynthese (siehe Abb. 3 und 4B). Dieses Ergebnis und die Tatsache, daß der Angiotensin-Conversionsenzym-(ACE)-Hemmer Captopril (**2**, Abb. 4B) auch die LTA<sub>4</sub>H hemmt,<sup>[53]</sup> haben die Variation dieser Leitstrukturen inspiriert und führten zur Synthese potenter Inhibitoren der Epoxidhydrolase-Aktivität der LTA<sub>4</sub>H. Diese Substanzen wiesen auch Selektivität für LTA<sub>4</sub>H gegenüber anderen Aminopeptidasen auf (siehe Verbindungen **4** and **5**, Abb. 4B).<sup>[54-58]</sup>

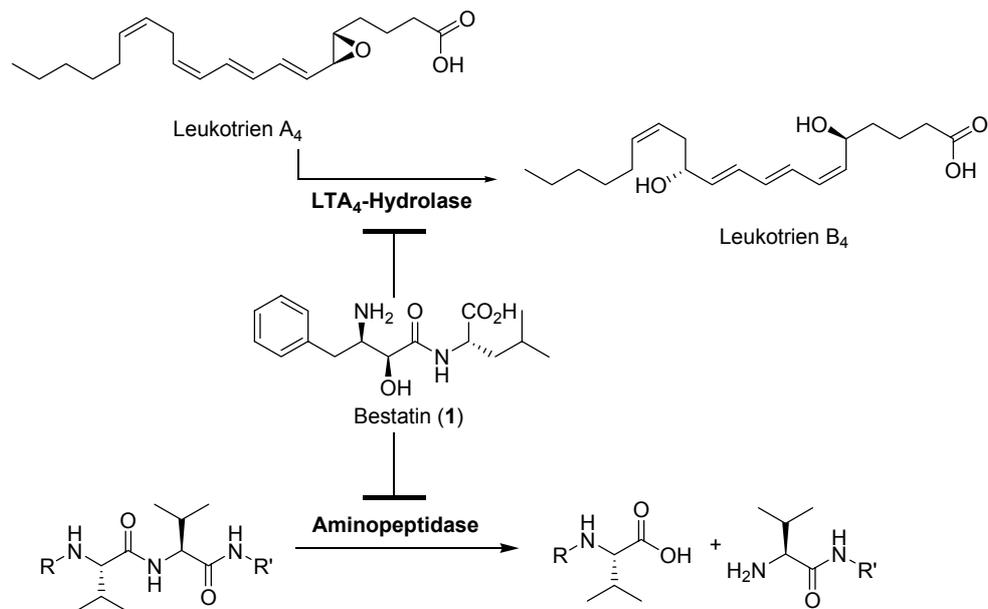


Abb. 3: Von Leukotrien-A<sub>4</sub>-Hydrolase (LTA<sub>4</sub>H) und Aminopeptidasen katalysierte Reaktionen. Trotz der Tatsache, daß diese Enzyme unterschiedliche Reaktionen katalysieren, werden sie durch den Naturstoff Bestatin (**1**), einem bekannten Aminopeptidase-Hemmer, inhibiert. LTA<sub>4</sub>H katalysiert in ihrem Zink-enthaltenden katalytischen Zentrum die vinyloge Hydrolyse des Leukotrienepoxids LTA<sub>4</sub> zu LTB<sub>4</sub>, während Aminopeptidasen Peptidbindungen spalten.

Die katalytische Domäne der LTA<sub>4</sub>H (L209-Y460, Protein bestehend aus drei Domänen und als M1-Metallopeptidase klassifiziert), der katalytische Kernbereich des humanen Angiotensin-Conversionsenzyms (P312-P424 und I521-G583, Eindomänenprotein, klassifiziert als M2-Metallopeptidase) und die M4-Metallopeptidase Thermolysin (I1-K316, Eindomänenprotein) weisen eine geringe Sequenzähnlichkeit auf. Die Sequenzidentität beträgt 8.3% bei einer Alignmentlänge von 216 Aminosäuren, wenn die LTA<sub>4</sub>H und Thermolysin miteinander verglichen werden. Beim Vergleich des strukturabgeleiteten Sequenzalignments des ACE und des Thermolysins ergibt sich eine Sequenzidentität von 14.4% bei einer Alignmentlänge von 90 Aminosäuren (siehe Abb. 4C). Trotz geringer Sequenzähnlichkeit zeigen alle drei Enzyme eine signifikante strukturelle Verwandtschaft in ihren jeweiligen katalytischen Domänen bzw. Kernbereichen (siehe Abb. 4A). Die katalytischen Zentren befinden sich an konservierten Stellen in der Faltung, wie dies die überlappende Position der katalytischen Zinkionen aller drei Enzyme im Raum zeigt. Hervorzuheben ist, daß in dieser konkreten Liganden-Designsituation ein Naturstoff, Bestatin (**1**, siehe Abb. 3 und 4B), ein unspezifischer Aminopeptidasehemmer, das initiale strukturelle Leitmotiv für die Entwicklung von LTA<sub>4</sub>H-Hemmern darstellte.



Interessanterweise zeigen Sulfotransferasen und Nukleotidkinasen, die jeweils aus einer katalytischen Domäne bestehen, eine signifikante strukturelle Ähnlichkeit, wie dies aus der Überlagerung der Uridylatkinase (UK) und der Estrogensulfotransferase (EST) in Abb. 6A zu ersehen ist. Und das, obwohl ihre katalytischen Domänen insignifikante Sequenzähnlichkeit mit einer Sequenzidentität von 9.1% bei einer Alignmentlänge von 143 Aminosäuren aufweisen. In beiden Strukturen werden die Cofaktoren über Wasserstoffbrücken mit den Amidgruppen des Proteinrückgrats gebunden. Das vorletzte Phosphat interagiert jeweils mit einem sogenannten ‚P-Loop‘-Motiv.<sup>[61]</sup> Die spezifischen Interaktionen mit den Aminosäureresten sind jedoch unterschiedlich. Die Phosphatgruppe des Substrats, das von UK phosphoryliert wird, hat die gleiche Orientierung in bezug auf den Cofaktor wie die phenolische Hydroxylfunktion des 17 $\beta$ -Estradiols, die in der von EST katalysierten Reaktion sulfatiert wird. Dies legt nahe, daß der katalytische Mechanismus sowohl der Phosphat- als auch der Sulfatübertragung ähnlich ist. Trotz der hohen Strukturähnlichkeit der katalytischen Domänen beider Enzyme sind nur wenige Aminosäuren konserviert und auch die spezifischen an der Katalyse beteiligten Aminosäuren haben ihren Ursprung an unterschiedlichen Orten im aktiven Zentrum.<sup>[61, 62]</sup> Beide Enzyme sind nach SCOP der Superfamilie der ‚P-Loop‘-enthaltenden Nukleosidtriphosphatase zugeordnet.

Die strukturelle Verwandtschaft der verwendeten Cofaktoren sowie die Ähnlichkeiten bezüglich des Reaktionsmechanismus und der Adenin-Bindungstaschen führten dazu, daß eine Purin-basierte Bibliothek, die ursprünglich für die Entwicklung von Inhibitoren Cyclin-abhängiger Kinasen entwickelt wurde und durch den Naturstoff Olomoucine (**8**, Abb. 6B) inspiriert war,<sup>[63, 64]</sup> auch gegen verschiedene Sulfotransferasen gescreent wurde.

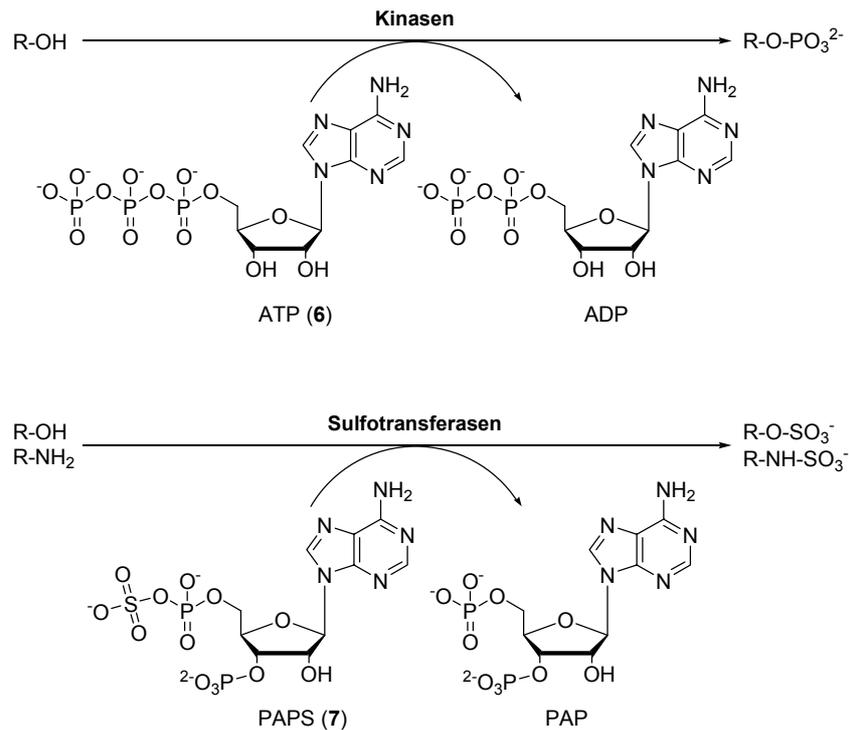
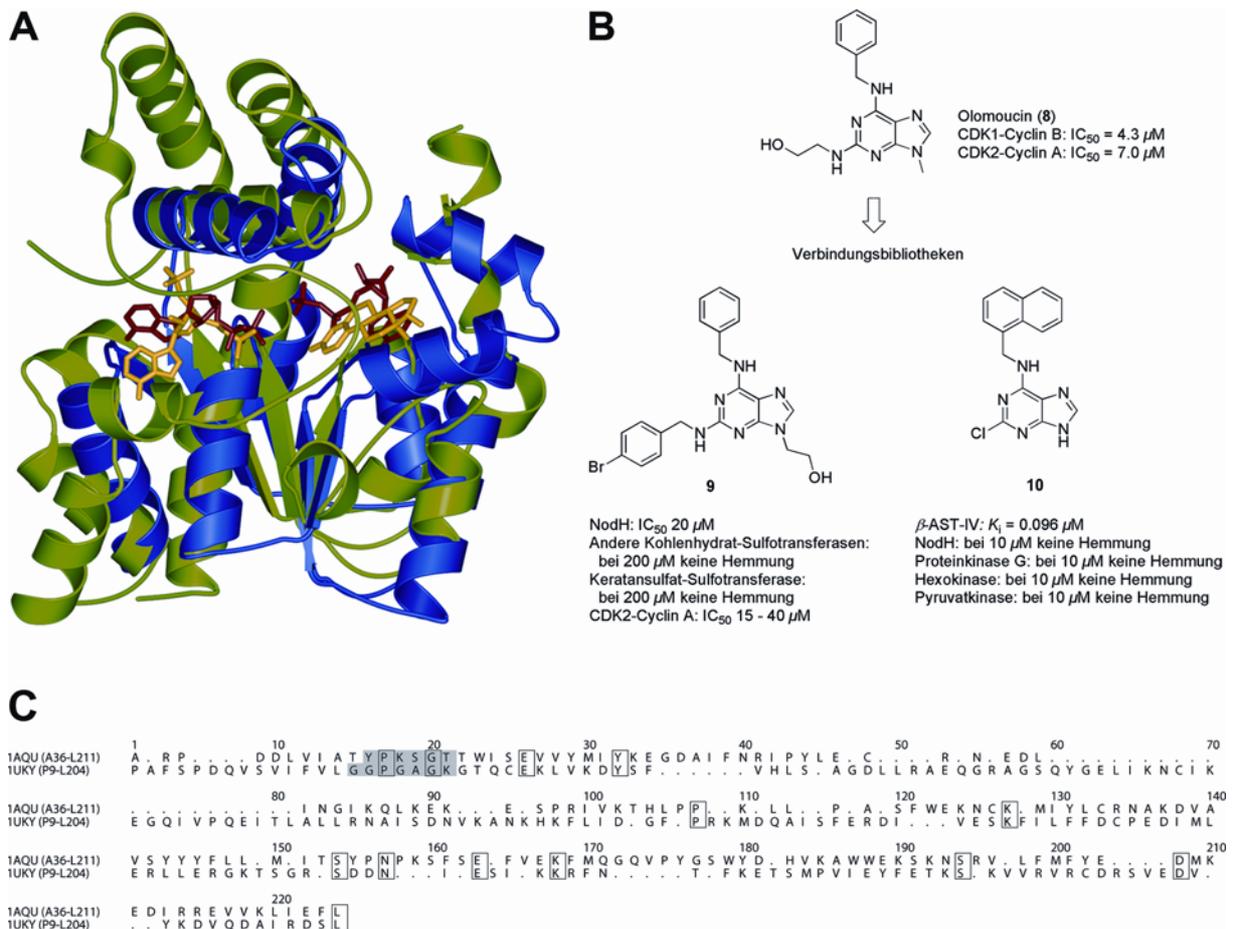


Abb. 5: Von Kinasen und Sulfotransferasen katalysierte Reaktionen.

Abb. 6: A) Überlagerung der Kristallstrukturen der katalytischen Domänen der Estrogensulfotransferase<sup>[61]</sup> (PDB-Code 1AQU) und der Uridylatkinase<sup>[65]</sup> (PDB-Code 1UKY), jeweils in Komplex mit Cofaktoren und Substraten. Der RMSD-Wert der gezeigten Überlagerung beträgt 3.27 Å (Alignmentlänge: 143 Aminosäuren).

EST (grün) mit verbrauchtem Cofaktor (PAP, gelb) und Substrat ( $17\beta$ -Estradiol, gelb), UK (blau) mit verbrauchtem Cofaktor (ADP, rot) und Substratanalogon (ADP, rot). B) Von Olomoucine abgeleitete Verbindungen, die Sulfotransferasen hemmen. C) Strukturabgeleitetes Sequenzalignment von EST und UK. Das 5'-Phosphatbindungsmotiv der EST ( $T_{45}YPKSGT$ )<sup>[61]</sup> und das ‚P-Loop‘-Motiv der UK ( $G_{23}XXGXGK$ )<sup>[61]</sup> sind jeweils grau unterlegt. Interessanterweise sind hier – trotz struktureller Ähnlichkeit – nur P47 und G50 der EST und die korrespondierenden Aminosäuren der UK, P25 und G28, konserviert.

Zunächst wurde die von Olomoucine abgeleitete Verbindungsbibliothek auf Inhibitoren der Kohlenhydrat-Sulfotransferase NodH aus *Rhizobium meliloti* hin untersucht. Es konnten PAPS-kompetitive NodH-Hemmer (siehe Verbindung **9**) mit schwacher inhibitorischer Aktivität (die  $IC_{50}$ -Werte lagen zwischen 20 und 40  $\mu$ M) identifiziert werden, die jedoch Selektivität gegenüber einigen getesteten Sulfotransferasen zeigten, wohingegen alle gefundenen Sulfotransferase-Hemmer auch verschiedene Kinasen im mikromolaren Bereich hemmten.<sup>[67]</sup> Ein Screening gegen die murine Estrogen-Sulfotransferase führte zur Auffindung von Hemmstoffen im nanomolaren Bereich, die noch schwache Aktivität bei verschiedenen cyclin-abhängigen Kinasen, jedoch Selektivität für die murine EST gegenüber Mitgliedern der Familie der Kohlenhydrat-Sulfotransferasen aufwiesen.<sup>[68]</sup> Ein Screening gegen  $\beta$ -Arylsulfotransferase-IV ( $\beta$ -AST-IV) lieferte schließlich einen potenten und selektiven  $\beta$ -AST-IV-Hemmer (**10**) mit einer Hemmkonstanten  $K_i$  von 96 nM. Eine Profilierung dieser Verbindung gegen verschiedene Nukleotid-bindende Proteine bestätigte ihre Selektivität.<sup>[69]</sup>

### ***Entwicklung von Modulatoren nukleärer Hormonrezeptoren***

Nukleäre Rezeptoren (NR) sind Liganden-gesteuerte Transkriptionsfaktoren, die aus einer Ligandenbindungsdomäne und einer DNA-Bindungsdomäne aufgebaut sind. Nukleäre Rezeptoren sind phylogenetische verwandte Proteine, die sich durch divergente Evolution entwickelt haben und daher nach SCOP einer gemeinsamen Familie zugeordnet werden. Ein Strukturvergleich der mäßig konservierten Ligandenbindungsdomänen zeigt, daß diese Domänen einen kanonischen Faltungstyp, der aus 12  $\alpha$ -Helices aufgebaut ist, aufweisen, der besser konserviert ist als ihre Aminosäureprimärsequenz. In dem hydrophoben Kernbereich der Ligandenbindungsdomänen werden die Liganden gebunden. Nukleäre Rezeptoren binden hydrophobe Moleküle wie Steroidhormone (Estrogene, Glucocorticoide, Progesteron, Mineralocorticoide, Androgene, Vitamin D, Ecdyson, Oxysterole, Gallensäuren usw.), Retinsäuren, Schilddrüsenhormone, Fettsäuren, Leukotriene und Prostaglandine.<sup>[70]</sup> Da nukleäre Rezeptoren durch kleine Hormonmoleküle kontrolliert werden, die ähnliche

physikochemische Eigenschaften wie therapeutisch verwendete chemische Substanzen aufweisen, stellen sie intrinsisch sehr attraktive Zielproteine für eine pharmakotherapeutische Intervention dar. Beispiele für klinisch relevante Modulatoren von nukleären Rezeptoren sind Estrogenrezeptor- $\alpha$  (ER $\alpha$ )-Antagonisten (z.B. Tamoxifen) bei der Behandlung von Brustkrebs oder die Strukturklasse der Thiazolidindione oder Glitazone als PPAR $\gamma$ - (peroxisome proliferator-activated receptor  $\gamma$ )-Agonisten, die als sogenannte „Insulinsensitizer“ bei der Behandlung von Diabetes Typ 2 eingesetzt werden.<sup>[71, 72]</sup>

Der Farnesoid-X-Rezeptor (FXR) wurde als Gallensäure-Rezeptor identifiziert. Er spielt eine regulatorische Rolle im Cholesteroll-Metabolismus. FXR kontrolliert die Synthese von Gallensäuren, ihre Konjugation (mit Taurin oder Glycin) und ihren Transport sowie den Lipidmetabolismus. Die derzeitige Hypothese ist, daß FXR als Sensor für den Gallensäurespiegel fungiert. FXR vermittelt die transkriptionelle Repression der Gene, die für die Konversion von überschüssigem Cholesteroll zu Gallensäuren verantwortlich sind, und die entsprechende Induktion von Genen, die für den Transport von Gallensäuren erforderlich sind. Nach diesen neuen Erkenntnissen über die Biologie von FXR handelt es sich bei FXR um ein wertvolles und pharmakologisch interessantes Zielprotein für neue Wirkstoffe zur Behandlung von Dyslipidämie und Cholestase (Gallestau).<sup>[73]</sup>

Um FXR jedoch als potentiellles Target zu validieren, ist ein genaues Verständnis seiner physiologischen Rolle unerläßlich. Ein selektiver, zellgängiger und hochaffiner Agonist könnte als chemisches Werkzeug in einem kombinierten chemisch-biologischen Ansatz bei der Aufklärung FXR-vermittelter Effekte überaus hilfreich sein. Um einen geeigneten Ausgangspunkt im chemischen Strukturraum für eine bibliotheksbasierte Entwicklung einer derartigen Verbindung zu finden, wäre eine Gruppierung der Ligandenbindungsdomänen verschiedener nukleärer Rezeptoren auf der Grundlage ihrer strukturellen Ähnlichkeit in der Retrospektive zielführend gewesen. So wäre ein Clustering der Ligandenbindungsdomänen der Kernrezeptoren ER $\beta$ , PPAR $\gamma$  und FXR nach struktureller Ähnlichkeit trotz ihrer vergleichsweise geringen Sequenzhomologie erfolgt. Die Sequenzidentität zwischen FXR und ER $\beta$  beträgt 18.2% bei einer Alignmentlänge von 159 Aminosäuren. Wenn FXR und PPAR $\gamma$  miteinander verglichen werden, ergibt sich eine Sequenzidentität von 27.8% bei einer Alignmentlänge von 176 Aminosäuren (siehe Abb. 7C). Die Überlagerung der Ligandenbindungsdomänen dieser Kernrezeptoren in Abb. 7A zeigt ihre strukturelle Verwandtschaft.

Der Naturstoff Genistein, ein Isoflavon-Phytoestrogen (**11**, Abb. 8), das in größeren Mengen in Soja-Bohnen und -Produkten vorkommt, bindet an beide Estrogenrezeptor-(ER)-Isoformen  $\alpha$  und  $\beta$  mit mäßiger Affinität und zeigt eine Präferenz für ER $\beta$ , wobei es als partieller Agonist wirkt.<sup>[74]</sup> Es wurde herausgefunden, daß Genistein auch ein PPAR $\gamma$ -Agonist ist.<sup>[75]</sup> Ein anderer bekannter synthetischer PPAR $\gamma$ -Agonist ist Troglitazon (**12**, Abb. 8), das als antidiabetischer Wirkstoff klinische Verwendung fand, aufgrund seiner Lebertoxizität jedoch wieder vom Markt genommen wurde.<sup>[76]</sup> Genistein und Troglitazon haben die Benzopyran-Partialstruktur gemeinsam. Das Benzopyran-Grundgerüst stellt ein privilegiertes Motiv dar und kommt in vielen Naturstoffen vor, die ein breites Spektrum biologischer Aktivitäten wie Antitumor- und antibakterielle Aktivität oder estrogene Wirkungen abdecken.<sup>[77]</sup> Es ist daher davon auszugehen, daß eine Verbindungsbibliothek, die durch einen Naturstoff-abgeleiteten Modulator eines Mitglieds der NR-Klasse inspiriert ist, im Zusammenhang mit der strukturellen Konservierung der Ligandenbindungsdomänen ebenfalls Hits für FXR liefern müßte.

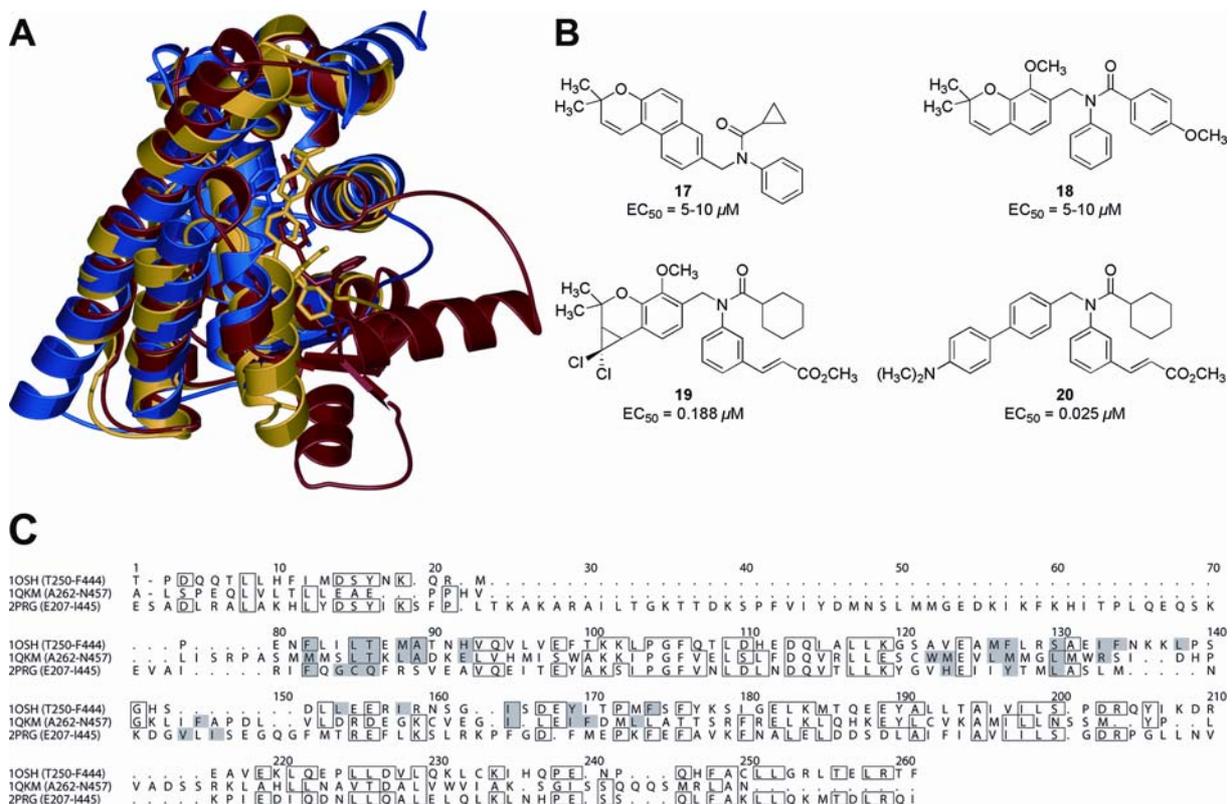


Abb. 7: A) Überlagerung der Kristallstrukturen der Ligandenbindungsdomänen des FXR<sup>[78]</sup> (PDB-Code 1OSH, mit gebundener Verbindung **20**, gelb), des ER $\beta$ <sup>[74]</sup> (PDB-Code 1QKM, mit gebundenem Genistein (**11**), blau) und des PPAR $\gamma$ <sup>[79]</sup> (PDB-Code 2PRG, mit gebundenem Rosiglitazon, rot). Die RMSD-Werte betragen 2.16 Å (Alignmentlänge: 159 Aminosäuren) bzw. 1.84 Å (Alignmentlänge: 176 Aminosäuren) für die gezeigte Überlagerung von FXR und ER $\beta$  bzw. von FXR und PPAR $\gamma$ . B) Benzopyran-inspirierte FXR-Agonisten. C) Strukturabgeleitetes Sequenzalignment der Ligandenbindungsdomänen des FXR, des ER $\beta$  und des PPAR $\gamma$ . Die für die Ligandenbindung relevanten Aminosäuren sind jeweils grau unterlegt. Einige der an der Bindung des

jeweiligen Liganden beteiligten Aminosäuren sind zwar nicht sequentiell aber räumlich in allen drei Ligandenbindungsdomänen konserviert.

Dies soll retrospektiv am Beispiel einer im weitesten Sinne von Genistein (**11**) abgeleiteten Bibliothek gezeigt werden. Ein initiales Screening einer kombinatorischen, naturstoffartigen und diversitätsorientierten Bibliothek von 10 000 Benzopyran-abgeleiteten Verbindungen, die von Nicolaou und Mitarbeitern generiert wurde,<sup>[77, 80]</sup> lieferte einige Leitstrukturen (**17** and **18**, Abb. 7B) mit niedrig mikromolarer Aktivität ( $EC_{50}$ -Werte zwischen  $5 \mu\text{M}$  und  $10 \mu\text{M}$ ) in einem Zell-basierten Assay auf FXR-Aktivierung. Die weitere Bearbeitung der identifizierten Leitstrukturen führte zu FXR-Liganden mit  $EC_{50}$ -Werten im niedrig nanomolaren Bereich, z.B. Verbindung **19** ( $EC_{50}$  188 nM) oder Verbindung **20** ( $EC_{50}$  25 nM), in der die Benzopyraneinheit zum privilegierten Biarylmotiv weiter abstrahiert wurde.<sup>[78, 81]</sup>

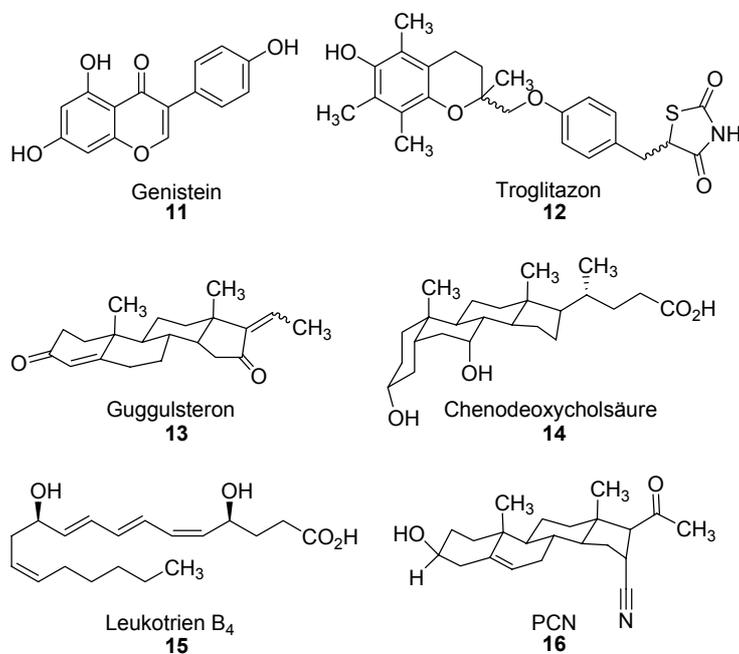


Abb. 8: Strukturell diverse natürliche und nicht-natürliche Liganden einiger nukleärer Rezeptoren.

Ein weiterer Hinweis, daß die Strategie, das Strukturmotiv eines bekannten Liganden eines Kernrezeptors für die Entwicklung von Liganden für andere Kernrezeptoren zu nutzen, zielführend ist, ist die Tatsache, daß die Ligandenbindungsdomänen der Kernrezeptoren eine gewisse Plastizität aufweisen, die es ihnen erlaubt, strukturell völlig unterschiedliche Liganden zu binden.<sup>[82]</sup> So zeigten Kristallstrukturen von Komplexen der Ligandenbindungsdomäne des Ecdyson-Rezeptors mit verschiedenen steroidal und nicht-steroidal Liganden deutlich unterschiedliche und auch nur teilweise überlappende Bindungsregionen, die nicht durch Molecular Modeling oder Docking-Studien vorhergesagt

werden konnten.<sup>[83]</sup> Eine Analyse von Kristallstrukturen der LXR $\beta$ -Ligandenbindungsdomäne mit unterschiedlichen synthetischen und nicht-steroidalen Liganden ergab, daß auch LXR $\beta$  eine flexible Bindungstasche aufweist.<sup>[84]</sup>

Darüber hinaus kann ein- und derselbe Ligand eine Kreuzreaktivität aufweisen und mit verschiedenen Kernrezeptoren interagieren, wie dies am Beispiel des pflanzlichen Sterols Guggulsteron (4,17(20)-Pregnadien-3,16-dion, **13**, Abb. 8) im folgenden verdeutlicht werden soll. Guggulsteron wird aus einem Extrakt des Gummiharzes des Guggul-Baums (*Commiphora mukul*) isoliert und senkt beim Menschen den LDL-(,Low-Density Lipoprotein')-Cholesterol-Spiegel. Guggulsteron (**13**) ist ein hocheffizienter Antagonist des FXR-Kernrezeptors, dessen natürliche Liganden Gallensäuren sind, wie z.B. Chenodeoxycholsäure (3 $\alpha$ ,7 $\alpha$ -Dihydroxy-5 $\beta$ -cholan-24-säure, **14**). Guggulsteron hemmt kompetitiv die Bindung von **14** an die Ligandenbindungsdomäne von FXR. Es bindet jedoch auch an den PPAR $\alpha$ -Kernrezeptor, dessen natürlicher Ligand Leukotrien B<sub>4</sub> (**15**) ist, und an den Pregnan-X-Rezeptor (PXR). Guggulsteron aktiviert PXR ungefähr halb so effektiv wie der spezifische, nicht-natürliche PXR-Agonist Pregnenolon-16 $\alpha$ -carbonitril (PCN, **16**).<sup>[85]</sup> Da Guggulsteron als Kongener von PCN aufgefaßt werden kann, überrascht es nicht, daß Guggulsteron nicht nur an FXR, sondern auch an PXR bindet. Viel überraschender ist seine Affinität zu PPAR $\alpha$ , dessen natürlicher Ligand Leukotrien B<sub>4</sub> (**15**) ist, wenn man die völlig verschiedenen chemischen Strukturen vergleicht.

In einem Vergleich der natürlichen Kernrezeptorliganden wurde herausgefunden, daß sie trotz ihrer chemischen Diversität ein konserviertes van der Waals-Volumen von  $318 \pm 53 \text{ \AA}^3$  aufweisen. Ihr mittleres Molekulargewicht von  $368 \pm 110 \text{ g mol}^{-1}$  ist weniger konserviert. Dies spricht für eine Co-Evolution von Rezeptoren und ihren Liganden, die zu einer Selektion von Liganden mit konservierten Volumina führte. Das Selektionskriterium war damit wohl die Fähigkeit, die Ligandenbindungskavität in der Ligandenbindungsdomäne räumlich auszufüllen. Die kanonische Faltung der Ligandenbindungsdomänen der Kernrezeptoren determiniert das Volumen der Bindungstasche, das der Ligand adressieren muß, und legt damit die tolerierte Bandbreite von Ligandenvolumina fest. Das molekulare Volumen könnte daher als Kriterium für die Beurteilung potentieller Kernrezeptor-Liganden herangezogen werden.<sup>[86]</sup>

### ***Zwischenergebnis***

Die der Literatur entnommenen und vor dem Hintergrund der eingangs formulierten Hypothesen reinterpretierten Beispiele belegen die Anwendbarkeit strukturbasierter Prinzipien, um Proteine im Kontext der Entwicklung von Liganden für die medizinisch-chemische und chemisch-biologische Forschung zu clustern. Bei allen Beispielen konnte jedoch allein schon aus sequenz- und funktionsbasierten Hinweisen eine Gruppierung der betrachteten Proteine vorgenommen werden. So konnte eine Verwandtschaft zwischen der LTA<sub>4</sub>H und den Metallopeptidasen über ein Signaturmotiv, im Falle der Nukleotidkinasen und Sulfotransferasen aufgrund der Ähnlichkeiten der gebundenen Cofaktoren und des Reaktionsmechanismus und im Falle der nukleären Rezeptoren aufgrund ihres – bereits bekannten – kanonischen Faltungstyps etabliert werden.

Die Frage, ob *ausschließlich* strukturelle Argumente (d.h. auch bei geringer Sequenzähnlichkeit und in Abwesenheit von charakteristischen Sequenzsignaturen) für ein Clustering herangezogen werden können, bleibt damit noch unbeantwortet. Die retrospektive Analyse der Modellfälle legt jedoch nahe, daß der Konservatismus der Natur in bezug auf die räumliche Struktur von Proteinen genutzt werden könnte, um Proteine analog zu dem Konzept der Target-Familien in Clustern nach dem Kriterium ihrer dreidimensionalen Ähnlichkeit zusammenzufassen. Ein Ligand eines Proteins einer derartigen Familie von strukturell ähnlichen Proteinen könnte als Leitprinzip für die Synthese einer fokussierten Verbindungsbibliothek herangezogen werden, mit der die anderen, strukturell ähnlichen Proteine adressiert werden könnten. Da die Ligandenbindungsstellen der Proteine aufgrund der geringen Sequenzhomologie und trotz prinzipieller struktureller Ähnlichkeit deutlich unterschiedlich sein können, ist ein kombinatorischer Ansatz für die Ligandenentwicklung unerlässlich. So entstehen durch ein strukturelles Leitprinzip inspirierte Verbindungsbibliotheken, die eine chemische Diversität aufweisen sollen, die mit der biologischen in Einklang zu bringen ist (siehe Abb. 9). Dies wird bei den retrospektiv analysierten Modellfällen sehr deutlich. Bei den Beispielen LTA<sub>4</sub>H-ACE-Thermolysin und Sulfotransferasen-Kinasen liegen die Sequenzidentitäten im Bereich 8-14% und damit insgesamt sehr niedrig, wobei hier konservierte Sequenzmotive im Bereich des aktiven Zentrums bzw. der Cofaktorbindungsstellen identifiziert wurden. Im Falle der Kernrezeptoren lagen die Sequenzidentitäten der Alignments höher, nämlich bei 18% für FXR und ER $\beta$  bzw. 28% für FXR und PPAR $\gamma$ . Die Analyse der jeweils für die Ligandenbindung verantwortlichen Aminosäuren ergab, daß bei der Betrachtung des Alignments von FXR und ER $\beta$  von den

insgesamt (für beide Sequenzen zusammengenommen) 22 bindungsrelevanten Aminosäuren 12 Aminosäuren align wurden (54.5%) und nur 4 identisch waren. Das Alignment von FXR und PPAR $\gamma$  zeigt, daß von insgesamt 21 bindungsrelevanten Aminosäuren lediglich 5 Aminosäuren align wurden (23.8%) und nur eine identisch war. Dies überrascht, da die Gesamtsequenzidentität für FXR/PPAR $\gamma$  mit 28% deutlich höher liegt als für FXR/ER $\beta$  mit 18%. Alle Beispiele belegen, daß die Ligandenbindungsstellen in ihrer Sequenz in einem Cluster strukturell ähnlicher Proteine deutlich verschieden sein können. Eine prinzipielle strukturelle Ähnlichkeit bleibt davon unberührt.

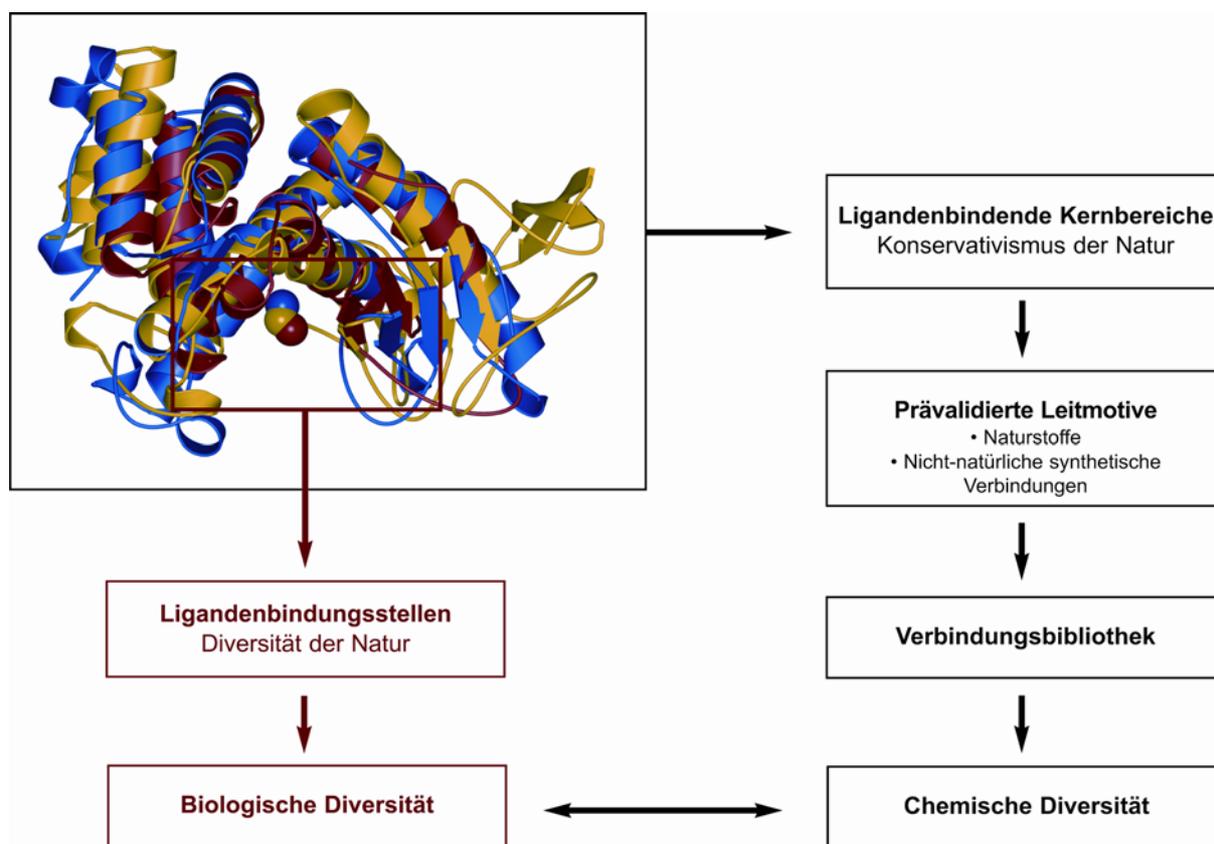


Abb. 9: Struktureller Konservativismus der Proteinfaltung führt zur Identifizierung von biologisch relevanten Gerüststrukturen als Leitprinzipien für die Konzipierung von Verbindungsbibliotheken. Der strukturellen Diversität der Bindungsstellen muß jedoch durch eine geeignete chemische Diversität Rechnung getragen werden.

Diese ersten Hinweise auf die grundsätzliche Anwendbarkeit einer Gruppierung von Ligandenbindungsdomänen bzw. ligandenbindenden Kernbereichen von Proteinen als abstrahierendes Leitprinzip für die Entwicklung von Proteinliganden mündeten in die Entwicklung einer Strategie, interessante Fälle von struktureller Ähnlichkeit bei niedriger

Sequenzidentität gezielt zu identifizieren, und in erste vorwärtsgerichtete und experimentell verifizierte Anwendungsbeispiele.

#### 4.1.2 Cdc25A-Phosphatase, Acetylcholinesterase und 11 $\beta$ -Hydroxysteroid-Dehydrogenasen – Ein Proteinstruktur-Ähnlichkeitscluster

##### *Suchstrategie – Definition des ligandenbindenden Proteinkernbereichs – Überlagerung von Proteinstrukturen*

Es wurde eine initiale Suchstrategie für die Identifizierung von Proteinstruktur-Ähnlichkeitsclustern entwickelt (siehe Abb. 10). Dabei wurden die folgenden, frei verfügbaren und Web-basierten Datenbanken genutzt:

- Zur prinzipiellen Klassifizierung und Einordnung einer Proteinstruktur in ihren strukturellen und evolutionären Kontext wurde die SCOP-(Structural Classification of Proteins)-Datenbank<sup>[29, 30, 87]</sup> (<http://scop.mrc-lmb.cam.ac.uk/scop/>) genutzt. In dieser Datenbank sind Proteindomänen, deren 3D-Struktur bekannt ist, aufgrund geometrischer und evolutionärer Verwandtschaft hierarchisch klassifiziert. Die höchste Hierarchieebene, die Klasse, beschreibt die wesentlichen Sekundärstrukturelemente. Die nächste Ebene, die Faltung, führt zu einer Gruppierung nach struktureller Ähnlichkeit, wohingegen die beiden Folgeebenen, Superfamilie und Familie, entfernte und nahe evolutionäre Verwandtschaftsverhältnisse beschreiben. Die SCOP-Datenbank wurde manuell aufgebaut und basiert auf visueller Prüfung und Vergleichen von Strukturen, Sequenzen und Funktionen. Alternativ dazu kann die CATH-(Class-Architecture-Topology-Homologous Superfamily)-Datenbank<sup>[88]</sup> zur Klassifizierung herangezogen werden. Proteindomänen werden hier semi-automatisch nach fünf hierarchischen Ebenen kategorisiert: Klasse (häufigste Sekundärstrukturelemente), Architektur (grobe Anordnung der Sekundärstrukturen, ohne die Konnektivität zu berücksichtigen), Topologie (Faltungsfamilien, entspricht im wesentlichen der SCOP-Faltungsebene), Homologe Superfamilie (sehr ähnliche Struktur und Funktion, gemeinsamer Vorfahre) und Sequenzfamilie (Sequenzidentität >35%). Die Einordnung auf der strukturellen Abstraktionsebene der Topologie erfolgt – wie auch bei SCOP – durch visuelle Prüfung.
- Die Dali/FSSP-(Fold Classification based on Structure-Structure alignment of Proteins)-Datenbank<sup>[89-94]</sup> (<http://www.ebi.ac.uk/dali/>) erlaubt, eine gegebene Proteinstruktur mit einer repräsentativen, nicht-redundanten Fraktion der ‚Protein Data Bank‘ (PDB, <http://www.rcsb.org/pdb/>), der PDB90<sup>[95]</sup>, in der keine zwei

Aminosäuresequenzen mit einer Sequenzidentität  $> 90\%$  vorkommen, rein strukturbasiert zu vergleichen. Die repräsentativen Proteinstrukturen in der PDB90 werden basierend auf der Auflösung, dem R-Faktor und ProCheck-Werten nach bestmöglicher struktureller Qualität ausgewählt. Der Dali-Algorithmus führt einen voll automatisierten paarweisen Strukturvergleich durch und berechnet ein strukturelles Ähnlichkeitsmaß zwischen Paaren von Proteinstrukturen, das als Z-Score dargestellt wird. Dies wird mit allen möglichen Paaren von Repräsentanten durchgeführt. Der Z-Score ist ein Maß für die statistische Signifikanz eines Ergebnisses bezogen auf ein Alignment von randomisierten Strukturen. Typischerweise weisen Proteine mit einer ähnlichen Faltung einen Z-Score von 3.5 oder höher auf. Die 3D-Koordinaten der Proteine werden zunächst in 2D-Distanzmatrizen überführt, die auf Abständen zwischen den durch ihre  $C^\alpha$ -Atome definierten Aminosäuren basieren. Diese Distanzmatrizen werden in elementare Kontaktmuster, z.B. Hexapeptid-Hexapeptid-Submatrizen, zerlegt. Es werden Paarungen von ähnlichen Kontaktmustern in den beiden Matrizen vorgenommen und diese dann zu größeren Einheiten zusammengefügt, wobei ein Monte-Carlo-Verfahren Anwendung findet.<sup>[96]</sup>

- Die CE-(Combinatorial Extension)-Datenbank<sup>[97]</sup> (<http://cl.sdsc.edu/ce.html>), erlaubt – wie Dali/FSSP – den Vergleich einer Proteinstruktur als Suchmotiv gegen die PDB. Es wird dazu jedoch ein anderer Algorithmus verwendet. Der CE-Algorithmus führt eine kombinatorische Erweiterung (CE) eines Alignmentpfads, der durch alignierte Fragmentpaare (AFPs) definiert wird, durch. AFPs stellen Paare von Fragmenten, jeweils eines für jedes Protein, dar, die eine strukturelle Ähnlichkeit aufweisen. AFPs sind kurze (8 Aminosäuren) Abschnitte einer Proteinstruktur und basieren mehr auf lokaler Geometrie als auf globalen Eigenschaften wie die Orientierung von Sekundärstrukturen oder die Gesamttopologie. Kombinationen von AFPs, die einen möglichen kontinuierlichen Alignmentpfad darstellen, werden selektiv erweitert oder verworfen. Dies führt letztlich zu einem einzigen optimalen Strukturalignment. Die Kriterien, die dem Algorithmus zugrundeliegen, sind im wesentlichen (i) die auf den Koordinaten der  $C^\alpha$ -Atome basierenden Abstände zwischen den Aminosäuren im Verlauf der Detektion des längsten Pfads von AFPs und (ii) im letzten Schritt die optimale strukturelle Überlagerung basierend auf dem RMSD-Wert (s.u.). Mittlerweile gibt es auch eine webserverbasierte CE-MC-Version dieses Algorithmus (<http://cemc.sdsc.edu/>), der zusätzlich eine Monte-Carlo-Optimierung durchführt.<sup>[98]</sup>

Darüber hinaus stehen u.a. die folgenden Webserver für Strukturalignments zur Verfügung:

- Der BioInfo3D-Server<sup>[99]</sup> (<http://bioinfo3d.cs.tau.ac.il/>), der für diese Dissertation nicht herangezogen wurde.
- Der FoldMiner-Webserver<sup>[100, 101]</sup> (<http://dlb4.stanford.edu/foldminer/>). Der in diesem Server implementierte LOCK2-Algorithmus wurde für Überlagerungen im Einzelfall verwendet.

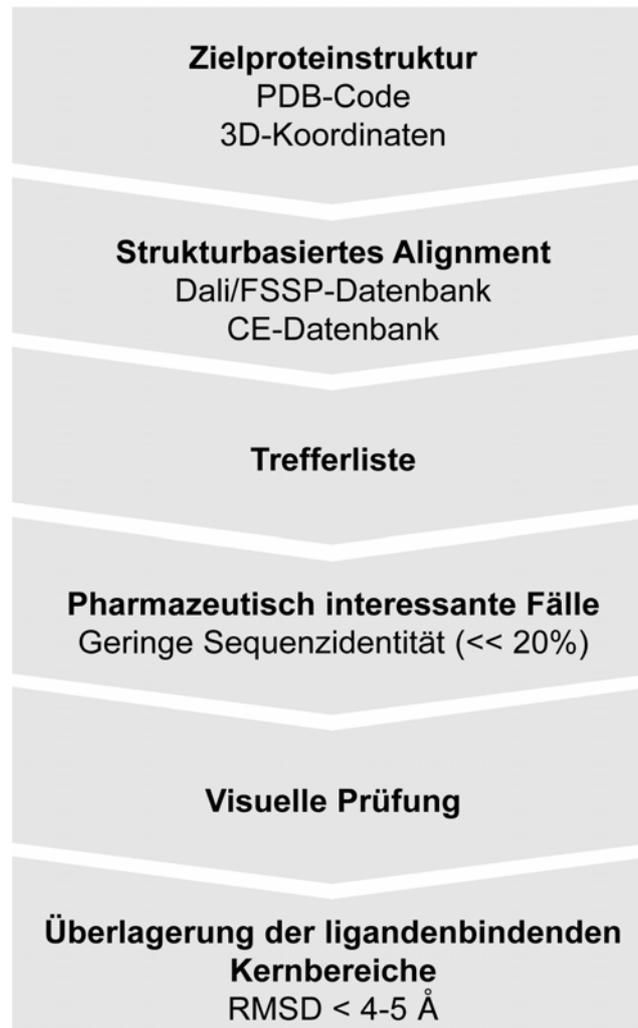


Abb. 10: Datenbanksuchstrategie zur Identifizierung von Proteinstruktur-Ähnlichkeitsclustern.

Mit den 3D-Koordinaten eines Proteins, das als Suchmotiv diente, wurden zunächst die Dali/FSSP- und die CE-Datenbanken durchforstet. Das Ergebnis einer derartigen Recherche war eine Trefferliste von vom jeweiligen Algorithmus als strukturell verwandt identifizierten Proteinen, in der die einzelnen Treffer mit sinkender Signifikanz, d.h. sinkendem Z-Score, gelistet waren. In einem ersten Selektionsschritt wurden pharmazeutisch relevante Treffer herausgefiltert, wenn diese eine geringe Sequenzidentität (Sequenzidentität << 20%) mit dem

Suchmotiv aufwiesen. Das letzte Kriterium war dabei im Hinblick auf eine erweiterte Anwendbarkeit auf ggf. auch nicht evolutionär verwandte jedoch strukturell ähnliche Proteine entscheidend. Die gewählten Fälle sollten nach Möglichkeit nicht auf der Grundlage von Sequenzhomologie detektierbar sein. Eine Recherche zur Klassifizierung der Treffer-Proteine nach SCOP sollte ggf. eine etwaige schon literaturbekannte Verwandtschaft des Suchproteins zum Trefferprotein ausschließen. Zur Absicherung der Treffer wurden Kreuzrecherchen sowohl mit dem Suchmotiv als auch mit den Treffern in der jeweils anderen Datenbank durchgeführt. Insgesamt jedoch waren die Ergebnisse einer derartigen Suchstrategie kritisch zu beurteilen. Das galt vor allem dann, wenn die Sequenzidentität um 10% und darunter lag. Hier lieferten die Recherchen zum Teil insignifikante Ergebnisse, d.h., es wurden für das Ligandendesign nicht-relevante (d.h., nicht die ligandenbindenden bzw. katalytischen Domänen) oder wenig aussagefähige, kleine Abschnitte der Proteine als ähnlich identifiziert. Diese Problematik trat umso häufiger auf, je größer der Unterschied in den räumlichen Dimensionen der verglichenen Proteine war. Daher war eine visuelle Prüfung des im Rahmen der Datenbankrecherche generierten Strukturalignments unumgänglich. Bei Proteinen, die aus mehreren Domänen bestehen, mußte sichergestellt werden, daß die für die Ligandenbindung relevanten Proteindomänen als ähnlich identifiziert und überlagert wurden. In den Fällen, wo sich die verglichenen Proteindomänen deutlich in ihrer Größe unterschieden, war es für ein optimales Überlagerungsergebnis erforderlich, zunächst die ligandenbindenden bzw. katalytischen Kernbereiche zu definieren, um so zu Domänenpartialstrukturen zu gelangen, die von ihrer Größe her auch vergleichbar waren. Zur Definition einer solchen ligandenbindenden oder katalytischen Domänenkernstruktur wurde ein Radius von ca. 10-15 Å um die Bindungstasche definiert, in dessen Bereich zunächst alle (auch teilweise) erfaßten Sekundärstrukturelemente zunächst zur Gänze berücksichtigt wurden. Der Größenbereich von 10-15 Å hat sich bewährt, um Bereiche mit konservierter Geometrie in Proteindomänen zu identifizieren.<sup>[102]</sup> Alternativ kann ggf. auch die Größe der kleineren Domäne für die Definition des Kernbereichs in der größeren Domäne herangezogen werden. Bei sehr kleinen Proteindomänen fallen in der Regel Domäne und Kernbereich zusammen (ggf. wurden in diesen Fällen ungeordnete Randbereiche entfernt). Es war zu prüfen, ob die wesentlichen Sekundärstrukturelemente der kleineren Domäne in der größeren Proteindomäne als Substrukturen zu finden waren. In diesem Fall und wenn die in der größeren Domäne identifizierte Substruktur den für die Ligandenbindung bzw. Katalyse relevanten Kernbereich beschrieb, konnte das kleinere Suchmotiv als Templat für den ligandenbindenden Kernbereich herangezogen werden. Sekundärstrukturelemente außerhalb des durch einen Radius um die

Bindungstasche oder das Suchmotiv als Templat definierten Kernbereichs wurden an weniger geordneten Schleifenbereichen abgetrennt, indem die entsprechenden Aminosäuren und Koordinaten aus der PDB-Datei durch Löschen entfernt wurden.

Die Ligandenbindungs- bzw. katalytischen Domänen oder die ligandenbindenden bzw. katalytischen Kernbereiche wurden unter Anwendung des CE-Algorithmus<sup>[97]</sup> oder des LOCK 2-Algorithmus über den FoldMiner-Webserver<sup>[100, 101]</sup> (<http://dlb4.stanford.edu/foldminer/>) überlagert. In einer aktuellen Evaluierung von Methoden des Strukturalignments von Proteinen wurden die folgenden, frei verfügbaren Programme getestet: SSAP<sup>[103]</sup>, STRUCTAL<sup>[104]</sup>, DALI<sup>[96]</sup> (wurde für die initiale Recherche genutzt), LSQMAN<sup>[105]</sup>, CE<sup>[97]</sup> (wurde für die Initialrecherche und für die Überlagerung der ligandenbindenden Kernbereiche genutzt) und SSM<sup>[106]</sup>. Die Evaluierung ergab, daß STRUCTAL und SSM, gefolgt von LSQMAN und CE, die besten Ergebnisse erbrachten.

Als Kriterium für die Beurteilung der Signifikanz des Überlagerungsergebnisses in bezug auf eine strukturelle Ähnlichkeit wurde die mittlere quadratische Abweichung (RMSD, Root Mean Square Deviation) der Abstände der einander als strukturell vergleichbar zugeordneten  $C^\alpha$ -Atome, stellvertretend für das jeweilige Proteinerückgrat, als quantitative Größe herangezogen. Der RMSD-Wert ist wie folgt definiert:

$$RMSD = \sqrt{\frac{\sum_{i=1}^N d_i^2}{N}}$$

$N$  = Anzahl der  $C^\alpha$ -Atome

$d_i$  = Abstand zwischen zwei korrespondierenden Atomen  $i$  in zwei Strukturen

RMSD-Werte sind abhängig von der Anzahl der korrespondierenden Paare von  $C^\alpha$ -Atomen, und es ist von entscheidender Bedeutung, daß das generierte Strukturalignment relevant ist, d.h., daß die ligandenbindenden Kernbereiche erfaßt sind.

Für die Beurteilung eines Strukturalignments im Sinne einer verwertbaren strukturellen Ähnlichkeit wurde vor allem die visuelle Prüfung des Überlagerungsergebnisses herangezogen. Der menschliche Verstand ist wie kein anderes Verfahren dazu in der Lage, gemeinsame Muster und somit eine gemeinsame Architektur zu erkennen.<sup>[107]</sup> Darüber hinaus wurde als quantitatives Kriterium eine RMSD-Obergrenze von 4 bis 5 Å als Richtwert angenommen. In einer Clusteranalyse von konservierten strukturellen Kernbereichen mit dem

Programm SCALI (Structural Core ALignment) wurden für eine gemeinsame Gruppierung die folgenden Ausschlußkriterien definiert:  $\text{RMSD} < 4.0 \text{ \AA}$  bei  $> 50$  alignen Aminosäuren.<sup>[108]</sup> Andere Arbeiten wiederum definieren strukturell verwandte Proteine durch einen  $\text{RMSD} < 5 \text{ \AA}$  (ohne Angabe einer Alignmentlänge).<sup>[109, 110]</sup> In einer statistischen Untersuchung, in der eine Methode zur Generierung von strukturabgeleiteten Sequenzalignments optimiert werden sollte, wurde ein Trainingsset von 412 Proteinpaaren verwendet, die zwar strukturell ähnlich (RMSD-Werte  $\sim 4 \text{ \AA}$ ) waren, jedoch keine detektierbare Sequenzverwandtschaft aufwiesen (durchschnittliche paarweise Sequenzidentität = 8%).<sup>[111]</sup>

### Die Cdc25A-Phosphatase als Suchmotiv

Das humane Genom codiert mindestens drei Cdc25-Enzyme (A, B und C), die Spezifität für unterschiedliche CDK-Cyclin-Komplexe aufweisen. Die Cdc25A-Phosphatase, kurz Cdc25A, ist eine dualspezifische Phosphatase, die die Progression durch den Zellzyklus am G1→S-Checkpoint reguliert, indem sie Tyr<sup>15</sup> und Thr<sup>14</sup> von CDK2-Cyclin A- und CDK2-Cyclin E-Komplexen dephosphoryliert und so den Eintritt in die S-Phase fördert.<sup>[112-114]</sup> Aufgrund der Kristallstruktur des Fragments der katalytischen Domäne der Cdc25A (M335-E495, Sequenzlänge: 161 Aminosäuren) wurde dieses Enzym nach SCOP dem Faltungstyp ‚rhodanese/cell cycle control phosphatase‘ zugeordnet. Die katalytische Domäne der Cdc25A ist ein kleine  $\alpha/\beta$ -Domäne mit einem zentralen fünfsträngigen  $\beta$ -Faltblatt in der Reihenfolge 15423. Die katalytisch entscheidende Aminosäure dieses Enzyms ist Cys<sup>430</sup>, die sich in einer Phosphat-bindenden CX<sub>5</sub>R-Schleife befindet. Cys<sup>430</sup> fungiert im Zuge der Phosphathydrolyse als Nucleophil und wird dabei von Glu<sup>431</sup> als einer allgemeinen Säure unterstützt (siehe Abb. 11).<sup>[112, 115]</sup> Abgesehen vom CX<sub>5</sub>R-Sequenzmotiv weist Cdc25A keine Sequenzhomologie mit den katalytischen Domänen anderer Protein-Tyrosin- und dualspezifischer Phosphatasen auf. Cdc25A wird als Target für die Entwicklung neuer Antitumor-Wirkstoffe diskutiert.<sup>[116]</sup>

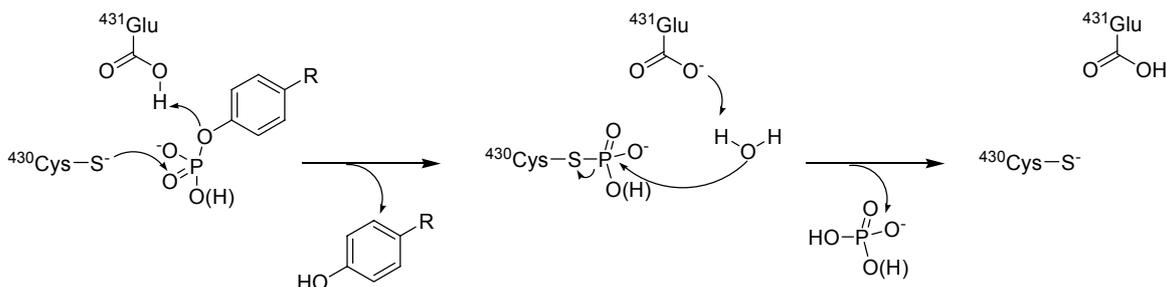


Abb. 11: Katalytischer Mechanismus der Cdc25A-Phosphatase.

Auf der Suche nach strukturell ähnlichen Proteinen wurde eine initiale Recherche in der Dali/FSSP- und der CE-Datenbank durchgeführt, wobei die 3D-Struktur der vollständigen katalytischen Domäne der Cdc25A, wie sie unter dem PDB-Code 1C25 abgelegt ist, als Suchmotiv diente.

Da die Dali/FSSP-Datenbank aufgrund der wachsenden Zahl von 3D-Proteinstrukturen in der PDB und damit auch in der PDB90 (s.o.) kontinuierlich modifiziert und aktualisiert wird, wurden die Ergebnisse der Suchstrategie vom März 2004, die von Koch *et al.* im November 2004 veröffentlicht wurden,<sup>[41]</sup> zum Zeitpunkt der Niederschrift dieser Dissertation noch einmal anhand der aktuellen, jedoch vorläufigen Version der Dali/FSSP-Datenbank vom März 2005 nachvollzogen. Die Analysen sind vom Ergebnis her vergleichbar:

- **Datenbankrecherche im März 2004**

Bei der Suche in der Dali/FSSP-Datenbank stellte sich die Hydroxynitril-Lyase aus *Hevea brasiliensis* (PDB-Code 3YAS, Sequenzlänge: 257 Aminosäuren), ein Eindomänenprotein, als erster Treffer heraus. Dieses Enzym teilt mit Cdc25A eine Sequenzidentität von 17% bei einem RMSD-Wert von 2.6 Å und einer Alignmentlänge von 76 Aminosäuren. Es gehört nach SCOP zur Superfamilie der  $\alpha/\beta$ -Hydrolasen. Eine weitere Suche nach strukturellen Nachbarn der Cdc25A in der CE-Datenbank führte zur Carboxylesterase EST2 aus *Alicyclobacillus acidocaldarius* (PDB-Code 1EVQ, Sequenzlänge: 310 Aminosäuren), die ebenfalls ein Mitglied der Enzym-Superfamilie der  $\alpha/\beta$ -Hydrolasen ist und mit Cdc25A eine Sequenzidentität von 10.9% bei einem RMSD von 3.5 Å und einer Alignmentlänge von 92 Aminosäuren teilt.

- **Datenbankrecherche im April 2005**

Die Recherche in Dali/FSSP unter Verwendung der Kristallstruktur der katalytischen Domäne der Cdc25A als Suchmotiv ergab zwei Treffer aus der Superfamilie der  $\alpha/\beta$ -Hydrolasen: die Lipase B aus *Candida antarctica* (PDB-Code 1TCC, Sequenzlänge: 317 Aminosäuren), die im Strukturalignment mit der Cdc25A eine Sequenzidentität von 6% bei einem RMSD von 4.6 Å und einer Alignmentlänge von 105 Aminosäuren aufweist, und die bereits oben beschriebene Carboxylesterase EST2 aus *Alicyclobacillus acidocaldarius* (PDB-Code 1QZ3, Sequenzlänge: 309 Aminosäuren), die gegenüber der Cdc25A eine Sequenzidentität von 7% bei einem RMSD von 3.3 Å

und einer Alignmentlänge von 94 Aminosäuren zeigt. Die Liste der Treffer und ihre Charakterisierung sind in Abb. 12 wiedergegeben.

### Dali database: select structural neighbours of 1cwsA

Please cite: L. Holm and C. Sander (1996) Science 273(5275):595-60.

structure alignment	structure+sequence alignment	3D superimposition	PDB format	Reset selection					
neighbour	Z	%ide	rmsd	lali	lseq2	PDB	compound		
<input type="checkbox"/>	0:	<a href="#">1cwsA</a>	35.3	100	0.0	178	178	<a href="#">PDB</a>	CDC25 B-TYPE TYROSINE PHOSPHATASE
<input checked="" type="checkbox"/>	1:	<a href="#">1c25</a>	26.1	66	1.3	157	161	<a href="#">PDB</a>	CDC25A
<input type="checkbox"/>	2:	<a href="#">1dp2A</a>	7.1	21	2.5	105	293	<a href="#">PDB</a>	RHDANASE
<input type="checkbox"/>	3:	<a href="#">1nvbB</a>	3.5	14	4.2	95	391	<a href="#">PDB</a>	3-DEHYDROQUINATE SYNTHASE
<input type="checkbox"/>	4:	<a href="#">1i9tA</a>	3.2	8	3.9	91	190	<a href="#">PDB</a>	MRNA CAPPING ENZYME
<input type="checkbox"/>	5:	<a href="#">1d5rA</a>	3.2	13	4.3	94	307	<a href="#">PDB</a>	PHOSPHOINOSITIDE PHOSPHATASE PTEN
<input type="checkbox"/>	6:	<a href="#">1j5uA</a>	3.2	12	3.3	83	430	<a href="#">PDB</a>	UDP-N-ACETYLMURAMATE--ALANINE LIGASE
<input type="checkbox"/>	7:	<a href="#">1hrkA</a>	2.9	4	3.3	83	359	<a href="#">PDB</a>	FERROCHELATASE
<input type="checkbox"/>	8:	<a href="#">1gqvB</a>	2.9	10	3.2	79	469	<a href="#">PDB</a>	UDP-N-ACETYLMURAMATE-L-ALANINE LIGASE
<input type="checkbox"/>	9:	<a href="#">1n13A</a>	2.8	7	4.3	91	839	<a href="#">PDB</a>	PREPROTEIN TRANSLOCASE SECA 1 SUBUNIT
<input type="checkbox"/>	10:	<a href="#">1q19B</a>	2.8	7	4.0	92	1020	<a href="#">PDB</a>	REVERSE GYRASE
<input type="checkbox"/>	11:	<a href="#">1i5eA</a>	2.8	7	4.3	96	208	<a href="#">PDB</a>	URACIL PHOSPHORIBOSYLTRANSFERASE
<input type="checkbox"/>	12:	<a href="#">1bgw</a>	2.8	8	4.0	86	679	<a href="#">PDB</a>	TOPOISOMERASE
<input type="checkbox"/>	13:	<a href="#">1emtA</a>	2.7	6	4.0	90	456	<a href="#">PDB</a>	METABOTROPIC GLUTAMATE RECEPTOR SUBTYPE 1
<input type="checkbox"/>	14:	<a href="#">1coulA</a>	2.7	6	3.7	96	645	<a href="#">PDB</a>	PROTEASE/HELICASE NS3
<input type="checkbox"/>	15:	<a href="#">1jz2A</a>	2.7	14	3.0	78	260	<a href="#">PDB</a>	UROPORPHYRINOGEN-III SYNTHASE
<input type="checkbox"/>	16:	<a href="#">1cvrA</a>	2.6	9	3.7	81	432	<a href="#">PDB</a>	GINGIPAIN R
<input type="checkbox"/>	17:	<a href="#">1f7uA</a>	2.5	6	2.8	73	606	<a href="#">PDB</a>	ARGINYL-TRNA SYNTHETASE
<input type="checkbox"/>	18:	<a href="#">1jxaA</a>	2.5	13	3.7	90	608	<a href="#">PDB</a>	GLUCOSAMINE 6-PHOSPHATE SYNTHASE
<input type="checkbox"/>	19:	<a href="#">1ngvA</a>	2.5	9	3.5	85	529	<a href="#">PDB</a>	AMYLOID BETA PRECURSOR PROTEIN-BINDING PROTEIN 1
<input type="checkbox"/>	20:	<a href="#">1cqxA</a>	2.5	5	3.8	91	403	<a href="#">PDB</a>	FLAVOHEMOPROTEIN
<input type="checkbox"/>	21:	<a href="#">1in5A</a>	2.5	5	3.9	87	301	<a href="#">PDB</a>	HOLLIDAY JUNCTION DNA HELICASE RUVB
<input type="checkbox"/>	22:	<a href="#">1d21A</a>	2.4	7	3.8	88	223	<a href="#">PDB</a>	DNA (5
<input type="checkbox"/>	23:	<a href="#">1irxB</a>	2.4	6	4.5	72	508	<a href="#">PDB</a>	LYSYL-TRNA SYNTHETASE
<input type="checkbox"/>	24:	<a href="#">1mgpA</a>	2.4	5	4.1	87	276	<a href="#">PDB</a>	HYPOTHETICAL PROTEIN TM841
<input type="checkbox"/>	25:	<a href="#">2shpA</a>	2.3	9	3.8	84	491	<a href="#">PDB</a>	SHP-2
<input checked="" type="checkbox"/>	26:	<a href="#">1tccA</a>	2.3	6	4.6	105	317	<a href="#">PDB</a>	LIPASE (E.C.3.1.1.3) (TRIACYLGLYCEROL HYDROLASE
<input checked="" type="checkbox"/>	27:	<a href="#">1civA</a>	2.2	7	2.9	67	374	<a href="#">PDB</a>	NADP-MALATE DEHYDROGENASE
<input checked="" type="checkbox"/>	28:	<a href="#">1qz3A</a>	2.2	7	3.3	94	309	<a href="#">PDB</a>	CARBOXYLESTERASE EST2
<input type="checkbox"/>	29:	<a href="#">1mioA</a>	2.2	2	3.6	76	525	<a href="#">PDB</a>	NITROGENASE MOLYBDENUM-IRON PROTEIN
<input type="checkbox"/>	30:	<a href="#">1larA</a>	2.2	6	4.8	86	566	<a href="#">PDB</a>	LAR
<input type="checkbox"/>	31:	<a href="#">1mqgA</a>	2.2	4	4.8	85	588	<a href="#">PDB</a>	SLY1 PROTEIN
<input type="checkbox"/>	32:	<a href="#">1prxA</a>	2.2	6	3.9	87	220	<a href="#">PDB</a>	HORF6
<input type="checkbox"/>	33:	<a href="#">1eudA</a>	2.1	4	3.1	71	306	<a href="#">PDB</a>	SUCCINYL-COA SYNTHETASE, ALPHA CHAIN
<input type="checkbox"/>	34:	<a href="#">1jftA</a>	2.1	10	3.2	73	340	<a href="#">PDB</a>	PURINE NUCLEOTIDE SYNTHESIS REPRESSOR
<input type="checkbox"/>	35:	<a href="#">1eqnA</a>	2.1	8	4.1	79	315	<a href="#">PDB</a>	DNA PRIMASE
<input type="checkbox"/>	36:	<a href="#">1regA</a>	2.1	2	3.4	83	727	<a href="#">PDB</a>	METHYLMALONYL-COA MUTASE
<input type="checkbox"/>	37:	<a href="#">1mnaB</a>	2.1	4	3.8	103	278	<a href="#">PDB</a>	POLYKETIDE SYNTHASE IV
<input type="checkbox"/>	38:	<a href="#">1ddzA</a>	2.1	3	5.2	94	481	<a href="#">PDB</a>	CARBONIC ANHYDRASE
<input type="checkbox"/>	39:	<a href="#">1lw7A</a>	2.1	7	4.2	84	344	<a href="#">PDB</a>	TRANSCRIPTIONAL REGULATOR NADR
<input type="checkbox"/>	40:	<a href="#">1ddgA</a>	2.1	3	3.7	89	374	<a href="#">PDB</a>	SULFITE REDUCTASE (NADPH) FLAVOPROTEIN ALPHA
<input type="checkbox"/>	41:	<a href="#">2pia</a>	2.0	10	4.6	69	321	<a href="#">PDB</a>	PTHALATE DIOXYGENASE REDUCTASE (E.C.1.18.1
<input type="checkbox"/>	42:	<a href="#">1lnoA</a>	2.0	12	3.5	77	301	<a href="#">PDB</a>	POTASSIUM CHANNEL RELATED PROTEIN

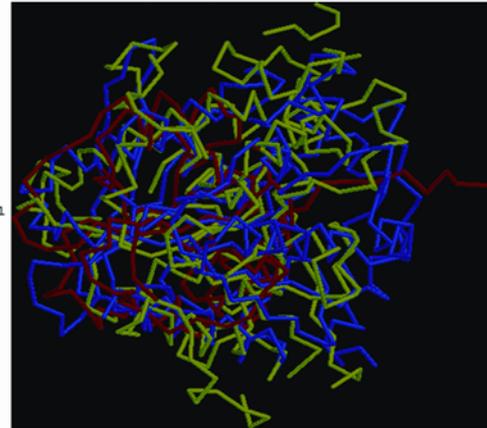


Abb. 12: Dali/FSSP-Recherche mit dem Suchmotiv der katalytischen Domäne der Cdc25A-Phosphatase (PDB-Code 1C25). Links ist die Trefferliste ( $Z$  = Z-Score, %ide = Sequenzidentität in %, lali = Alignmentlänge, lseq2 = Länge der zur Sequenz des Suchmotivs alignen zweiten Sequenz) und rechts das von Dali generierte Strukturalignment als Überlagerung der drei Proteine, 1C25 (blau), 1TCC (grün) und 1QZ3 (rot), dargestellt (zu sehen ist jeweils das  $C^\alpha$ -Rückgrat).

Interessanterweise gehört die Acetylcholinesterase als pharmazeutisch relevantes Mitglied zur Superfamilie der  $\alpha/\beta$ -Hydrolasen.

Die Acetylcholinesterase (AChE) hydrolysiert den Neurotransmitter Acetylcholin (**21**) und beendet dadurch die Impulsweiterleitung in cholinergen Synapsen (siehe Abb. 13).<sup>[117]</sup> Die AChE aus *Torpedo californica* (S4-T535, PDB-Code 2ACE) ist strukturell homolog zu verwandten Enzymen, die in den Nerven und Muskeln von Wirbeltieren vorkommen, und zeigt das typische  $\alpha/\beta$ -Hydrolase-Faltungsmotiv. Ser<sup>200</sup>, His<sup>440</sup> und Glu<sup>327</sup> bilden die katalytische Triade, die sich im esteratischen Zentrum der Enzymbindungsstasche befindet. AChE-Inhibitoren werden zur medikamentösen Therapie verschiedener Leiden eingesetzt. Dazu gehören *Myasthenia gravis*, das Glaucom und *Morbus Alzheimer*.<sup>[118]</sup>

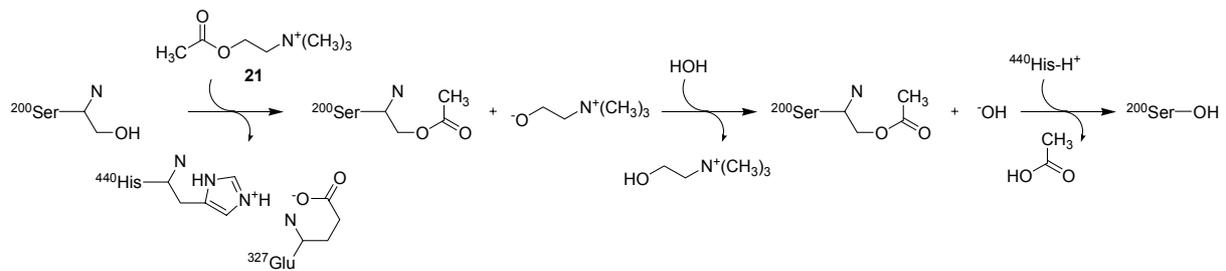


Abb. 13: Katalytischer Mechanismus der Acetylcholinesterase.

Da die AChE (PDB-Code 2ACE bzw. 2DFP, Sequenzlänge: 534 Aminosäuren) im Vergleich zur katalytischen Domäne der Cdc25A (Sequenzlänge: 161 Aminosäuren) ein sehr großes Eindomänenprotein ist, wurde eine sorgfältige vergleichende strukturelle Exploration des katalytischen Kernbereichs der AChE vorgenommen. Es wurde geprüft, ob das kleinere Strukturmotiv der Cdc25A im katalytischen Kernbereich der AChE zu finden ist. Hierzu wurde der katalytische Kernbereich der deutlich größeren AChE sukzessive herausgearbeitet, wobei die Cdc25A-Struktur, bei der zur Definition des Kernbereichs lediglich ungeordnete Bereiche im Randbereich der katalytischen Domäne entfernt wurden (M335-S360, S484-E495), als Templat verwendet wurde (siehe Abb. 14). Anschließend wurden die katalytischen Kernbereiche der AChE und der Cdc25A unter Verwendung des CE-Algorithmus überlagert. Das Überlagerungsergebnis wurde visuell geprüft. In Abbildung 15 sind die beiden überlagerten Kernstrukturen zu sehen. Sie weisen eine signifikante dreidimensionale Ähnlichkeit auf.

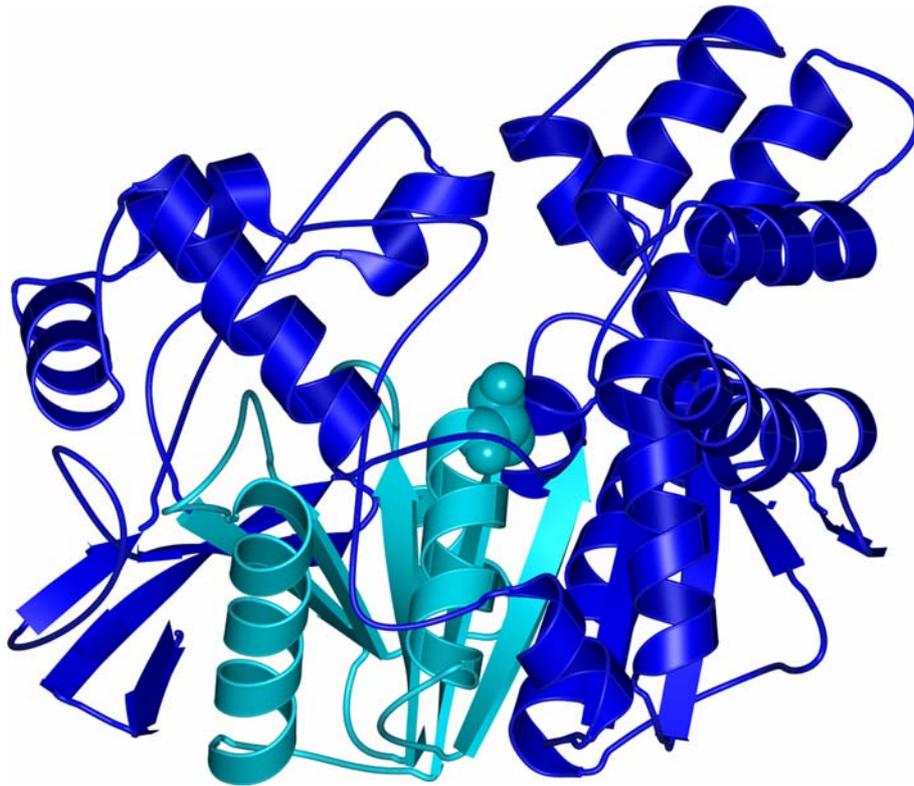
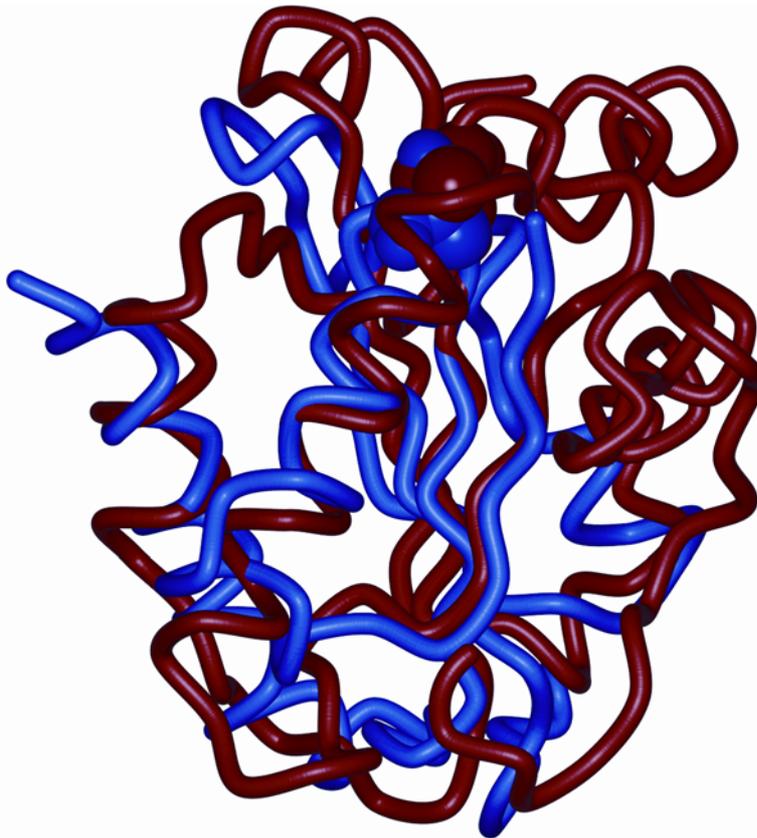


Abb. 14: Katalytischer Kernbereich (cyan, N98-Q225, Sequenzlänge: 128 Aminosäuren) der AChE (Strukturbereiche außerhalb des katalytischen Kerns sind blau dargestellt, PDB-Code 2ACE, Sequenzlänge: 527 Aminosäuren).

A



B

```

1C25 (E405-L465)  1  E E V E D F L L K K P . . . I V P T D . . G K R . . . V I V V F H C E F S S E R G P R M C R Y V R E
2ACE (N167-Q225)  1  N V G L L D Q R M A L Q W V H D N I Q F F G G D P K T V T I F G E S . . . . . A G G A S V G M H I L

1C25 (E405-L465)  60  R . . D R L G N E Y P K L H Y P E L Y V L
2ACE (N167-Q225)  60  S P G S R . . . . . D L F R R A I L Q

```

Abb. 15: A) Überlagerte katalytische Kernbereiche der Acetylcholinesterase (blau) und der Cdc25A (rot). Der RMSD beträgt 2.74 Å bei einer Alignmentlänge von 49 Aminosäuren, wobei die Sequenzidentität 8.2% beträgt. In CPK-Darstellung sind die katalytischen Aminosäuren, Ser<sup>200</sup> und Cys<sup>430</sup>, gezeigt. Sie sind gleich lokalisiert. B) Strukturabgeleitetes Sequenzalignment der überlagerten katalytischen Kernbereiche der Cdc25A (PDB-Code 1C25) und der AChE (PDB-Code 2ACE). Die katalytischen Schlüsselaminosäuren, C430 und S200, sind im Sequenzalignment korreliert und jeweils grau unterlegt.

Darüber hinaus zeigten sowohl das Strukturalignment als auch das daraus abgeleitete Sequenzalignment, daß die zentralen katalytischen Aminosäuren, Cys<sup>430</sup> (Cdc25A) und Ser<sup>200</sup> (AChE), jeweils korreliert waren (siehe Abb. 15B)

Bei der Dali/FSSP-Recherche mit dem Suchmotiv der Cdc25A-Struktur bzw. der ähnlichen AChE-Struktur ergaben sich noch weitere Hinweise auf strukturelle Verwandtschaften:

- **Datenbankrecherche im März 2004**

Die Methylen-tetrahydromethanopterin-Dehydrogenase aus *Methylobacterium extorquens* (PDB-Code 1LU9, Sequenzlänge: 287 Aminosäuren) wurde von

Dali/FSSP als im Vergleich zur Cdc25A ähnlich gefaltetes Protein als Treffer ausgegeben. Der RMSD-Wert des von Dali generierten Strukturalignments betrug 4.9 Å bei einer Alignmentlänge von 89 Aminosäuren und einer Sequenzidentität von 7%. Dieses Enzym wird nach SCOP dem NAD(P)-bindenden Rossmann-Faltungstyp zugeordnet. Eine weitere Dali/FSSP-Recherche mit der 3D-Struktur der AChE, die bereits als strukturell ähnlich zu Cdc25A identifiziert wurde, lieferte zusätzlich die Tropinonreduktase aus *Datura stramonium* (PDB-Code 1AE1, Sequenzlänge: 258 Aminosäuren) als Treffer mit einem RMSD von 3.9 Å bei einer Alignmentlänge von 155 Aminosäuren und einer Sequenzidentität von 6%. Auch dieses Protein zeigt die Rossmann-Faltung.

- **Datenbankrecherche im April 2005**

Die Suche in Dali/FSSP mit der 3D-Struktur der AChE (PDB-Code 2DFP bzw. 2ACE) lieferte – wie schon im März 2004 – ebenfalls die Tropinonreduktase als Treffer. Das Ergebnis dieser Dali/FSSP-Recherche und die Charakterisierung der Treffer sind in Abb. 16 wiedergegeben. Als Treffer wurden u.a. weitere Mitglieder der Superfamilie der  $\alpha/\beta$ -Hydrolasen identifiziert, wie z.B. verschiedene Lipasen, Carboxylesterasen oder die humane Acylprotein-Thioesterase 1 (*hAPT1*, PDB-Code 1FJ2). Implikationen dieses Ergebnisses für weitere Forschungsprojekte werden an anderer Stelle ausführlich besprochen (siehe Kapitel 4.1.4).

## Dali database: select structural neighbours of 2dfpA

Please cite: L. Holm and C. Sander (1996) Science 273(5275):595-60.

	structure alignment	structure+sequence alignment	3D superimposition	PDB format	Reset selection				
<input checked="" type="checkbox"/>									
neighbour	Z	%ide	rmsd	lali	lseq2	PDB	compound		
<input checked="" type="checkbox"/>	0:	<a href="#">2dfpA</a>	73.9	100	0.0	534	534	<a href="#">PDB</a>	ACETYLCHOLINESTERASE
<input type="checkbox"/>	1:	<a href="#">lg0iA</a>	62.9	53	1.3	523	523	<a href="#">PDB</a>	CHOLINESTERASE
<input type="checkbox"/>	2:	<a href="#">lf8uA</a>	58.7	58	1.1	529	531	<a href="#">PDB</a>	ACETYLCHOLINESTERASE
<input type="checkbox"/>	3:	<a href="#">lmaad</a>	58.5	57	1.1	533	541	<a href="#">PDB</a>	ACETYLCHOLINESTERASE
<input type="checkbox"/>	4:	<a href="#">lgonA</a>	55.8	38	1.7	524	540	<a href="#">PDB</a>	ACETYLCHOLINESTERASE
<input type="checkbox"/>	5:	<a href="#">lmx9D</a>	51.2	33	2.0	501	533	<a href="#">PDB</a>	LIVER CARBOXYLESTERASE I
<input type="checkbox"/>	6:	<a href="#">lk4yA</a>	50.8	34	2.0	475	501	<a href="#">PDB</a>	LIVER CARBOXYLESTERASE
<input type="checkbox"/>	7:	<a href="#">lf6wA</a>	48.8	31	2.4	493	533	<a href="#">PDB</a>	BILE SALT ACTIVATED LIPASE
<input type="checkbox"/>	8:	<a href="#">lakn</a>	44.9	32	2.2	493	547	<a href="#">PDB</a>	BILE-SALT ACTIVATED LIPASE
<input type="checkbox"/>	9:	<a href="#">lc7jA</a>	44.6	30	2.8	471	485	<a href="#">PDB</a>	PARA-NITROBENZYL ESTERASE
<input checked="" type="checkbox"/>	10:	<a href="#">lgz7A</a>	42.0	28	2.6	469	534	<a href="#">PDB</a>	LIPASE 2
<input checked="" type="checkbox"/>	11:	<a href="#">lllfA</a>	41.8	29	2.7	466	534	<a href="#">PDB</a>	LIPASE 3
<input checked="" type="checkbox"/>	12:	<a href="#">lthg</a>	36.8	27	2.9	472	544	<a href="#">PDB</a>	LIPASE (E.C.3.1.1.3) TRIACYLGLYCEROL HYDROLASE
<input type="checkbox"/>	13:	<a href="#">ljkmB</a>	19.6	15	2.8	232	361	<a href="#">PDB</a>	BREFELDIN A ESTERASE
<input type="checkbox"/>	14:	<a href="#">ljjiA</a>	18.3	15	2.8	226	311	<a href="#">PDB</a>	CARBOXYLESTERASE
<input type="checkbox"/>	15:	<a href="#">lorwA</a>	15.1	14	3.4	224	729	<a href="#">PDB</a>	DIPEPTIDYL PEPTIDASE IV
<input type="checkbox"/>	16:	<a href="#">lmpxA</a>	15.0	11	4.7	257	614	<a href="#">PDB</a>	ALPHA-AMINO ACID ESTER HYDROLASE
<input checked="" type="checkbox"/>	17:	<a href="#">lhlgA</a>	14.6	8	4.0	232	368	<a href="#">PDB</a>	LIPASE, GASTRIC
<input type="checkbox"/>	18:	<a href="#">ll7aA</a>	14.5	11	3.7	219	318	<a href="#">PDB</a>	CEPHALOSPORIN C DEACETYLASE
<input checked="" type="checkbox"/>	19:	<a href="#">ljfrA</a>	14.3	14	2.9	199	260	<a href="#">PDB</a>	LIPASE
<input type="checkbox"/>	20:	<a href="#">lmu0A</a>	14.2	12	2.7	198	294	<a href="#">PDB</a>	PROLINE IMINOPEPTIDASE
<input type="checkbox"/>	21:	<a href="#">lc4xA</a>	14.0	8	3.8	211	281	<a href="#">PDB</a>	2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENOATE
<input type="checkbox"/>	22:	<a href="#">ljjfA</a>	13.9	12	3.1	210	255	<a href="#">PDB</a>	ENDO-1,4-BETA-XYLANASE Z
<input type="checkbox"/>	23:	<a href="#">livyA</a>	13.1	12	3.7	220	452	<a href="#">PDB</a>	HUMAN PROTECTIVE PROTEIN
<input type="checkbox"/>	24:	<a href="#">lazwA</a>	13.0	9	4.4	215	313	<a href="#">PDB</a>	PROLINE IMINOPEPTIDASE
<input type="checkbox"/>	25:	<a href="#">liunB</a>	12.8	10	2.9	186	276	<a href="#">PDB</a>	META-CLEAVAGE PRODUCT HYDROLASE
<input type="checkbox"/>	26:	<a href="#">la8g</a>	12.8	14	2.8	192	274	<a href="#">PDB</a>	BROMOPEROXIDASE A1
<input checked="" type="checkbox"/>	27:	<a href="#">lthtA</a>	12.7	9	3.6	214	294	<a href="#">PDB</a>	THIOESTERASE
<input type="checkbox"/>	28:	<a href="#">limjA</a>	12.6	9	3.0	186	208	<a href="#">PDB</a>	CCG1-INTERACTING FACTOR B
<input type="checkbox"/>	29:	<a href="#">lqgvA</a>	12.6	11	0.0	194	232	<a href="#">PDB</a>	DIENELACTONE HYDROLASE
<input type="checkbox"/>	30:	<a href="#">laurA</a>	12.5	11	2.6	181	218	<a href="#">PDB</a>	CARBOXYLESTERASE
<input type="checkbox"/>	31:	<a href="#">lqtrA</a>	12.5	10	3.9	209	314	<a href="#">PDB</a>	PROLYL AMINOPEPTIDASE
<input checked="" type="checkbox"/>	32:	<a href="#">lfj2A</a>	12.4	11	2.8	185	229	<a href="#">PDB</a>	ACYL PROTEIN THIOESTERASE 1
<input type="checkbox"/>	33:	<a href="#">linsA</a>	12.3	8	3.7	254	763	<a href="#">PDB</a>	X-PROLYL DIPEPTIDYL AMINOPETIDASE
<input type="checkbox"/>	34:	<a href="#">lhdeA</a>	11.8	9	3.3	208	310	<a href="#">PDB</a>	HALOALKANE DEHALOGENASE
<input checked="" type="checkbox"/>	35:	<a href="#">ltccA</a>	11.3	9	3.4	200	317	<a href="#">PDB</a>	LIPASE (E.C.3.1.1.3) (TRIACYLGLYCEROL HYDROLASE
<input type="checkbox"/>	36:	<a href="#">lehyA</a>	11.2	10	3.0	185	282	<a href="#">PDB</a>	SOLUBLE EPOXIDE HYDROLASE
<input type="checkbox"/>	37:	<a href="#">lm33A</a>	10.9	12	4.3	190	255	<a href="#">PDB</a>	BIOH PROTEIN
<input type="checkbox"/>	38:	<a href="#">lcqwA</a>	10.7	9	4.2	203	295	<a href="#">PDB</a>	HALOALKANE DEHALOGENASE; 1-CHLOROHEXANE
<input type="checkbox"/>	39:	<a href="#">lgo7A</a>	10.7	11	4.3	227	385	<a href="#">PDB</a>	EPOXIDE HYDROLASE
<input type="checkbox"/>	40:	<a href="#">ll7rA</a>	10.6	10	3.9	248	573	<a href="#">PDB</a>	COCAINE ESTERASE
<input type="checkbox"/>	41:	<a href="#">lgpl</a>	10.2	12	3.7	199	432	<a href="#">PDB</a>	RP2 LIPASE
<input checked="" type="checkbox"/>	42:	<a href="#">li6wB</a>	10.1	13	0.0	158	180	<a href="#">PDB</a>	LIPASE A
<input checked="" type="checkbox"/>	43:	<a href="#">lethA</a>	10.0	11	3.6	198	448	<a href="#">PDB</a>	TRIACYLGLYCEROL ACYL-HYDROLASE
<input checked="" type="checkbox"/>	44:	<a href="#">loi1A</a>	9.8	8	4.4	193	320	<a href="#">PDB</a>	LIPASE
<input checked="" type="checkbox"/>	45:	<a href="#">lku0A</a>	9.7	10	4.2	202	388	<a href="#">PDB</a>	L1 LIPASE
<input type="checkbox"/>	46:	<a href="#">lbu8A</a>	9.5	13	3.9	192	446	<a href="#">PDB</a>	PANCREATIC LIPASE RELATED PROTEIN 2
<input checked="" type="checkbox"/>	47:	<a href="#">ltahA</a>	9.3	7	4.8	197	318	<a href="#">PDB</a>	LIPASE (E.C.3.1.1.3) (TRIACYLGLYCEROL HYDROLASE
<input checked="" type="checkbox"/>	48:	<a href="#">lhplA</a>	9.0	10	4.3	194	449	<a href="#">PDB</a>	LIPASE (E.C.3.1.1.3) (TRIACYLGLYCEROL HYDROLASE
<input type="checkbox"/>	49:	<a href="#">ljmkO</a>	8.8	13	3.6	178	230	<a href="#">PDB</a>	SURFACTIN SYNTHETASE
<input type="checkbox"/>	50:	<a href="#">2cut</a>	7.1	11	3.7	153	198	<a href="#">PDB</a>	CUTINASE (E.C.3.1.1.-) COMPLEXED WITH THE INHIBITOR DIETHYL
<input checked="" type="checkbox"/>	51:	<a href="#">lgxsA</a>	6.2	12	3.8	159	267	<a href="#">PDB</a>	HYDROXYNITRILE LYASE
<input type="checkbox"/>	52:	<a href="#">5gpb</a>	4.9	9	4.2	181	833	<a href="#">PDB</a>	GLYCOGEN PHOSPHORYLASE \$B (E.C.2.4.1.1) (T STATE) COMPLEX
<input checked="" type="checkbox"/>	53:	<a href="#">lae1B</a>	4.6	6	3.9	155	258	<a href="#">PDB</a>	TROPINONE REDUCTASE-I
<input type="checkbox"/>	54:	<a href="#">lnfrA</a>	4.5	5	3.7	141	244	<a href="#">PDB</a>	PUTATIVE OXIDOREDUCTASE RV2002
<input type="checkbox"/>	55:	<a href="#">lhdcA</a>	4.4	2	3.7	150	253	<a href="#">PDB</a>	3-ALPHA, 20-BETA-HYDROXYSTEROID DEHYDROGENASE (E.C.1.1.1.53
<input type="checkbox"/>	56:	<a href="#">lahiA</a>	4.2	10	3.6	55	255	<a href="#">PDB</a>	7 ALPHA-HYDROXYSTEROID DEHYDROGENASE
<input type="checkbox"/>	57:	<a href="#">llkjA</a>	3.8	8	5.0	93	146	<a href="#">PDB</a>	CALMODULIN
<input type="checkbox"/>	58:	<a href="#">lr4qA</a>	3.6	3	3.4	51	53	<a href="#">PDB</a>	RNA POLYMERASE ALPHA SUBUNIT
<input type="checkbox"/>	59:	<a href="#">lqmpA</a>	3.4	6	0.0	102	126	<a href="#">PDB</a>	SPOOA
<input type="checkbox"/>	60:	<a href="#">loi14A</a>	3.3	9	4.0	132	319	<a href="#">PDB</a>	SUGAR KINASE, PFKB FAMILY
<input type="checkbox"/>	61:	<a href="#">lcde</a>	3.2	9	4.1	138	210	<a href="#">PDB</a>	PHOSPHORIBOSYLGLYCINAMIDE FORMYLTRANSFERASE (E.C.2.1.2.2
<input type="checkbox"/>	62:	<a href="#">lionA</a>	3.2	8	4.5	133	243	<a href="#">PDB</a>	PROBABLE CELL DIVISION INHIBITOR MIND
<input type="checkbox"/>	63:	<a href="#">lh9cA</a>	2.9	4	2.8	92	106	<a href="#">PDB</a>	PTS SYSTEM, CHITOBIOSE-SPECIFIC IIB COMPONENT
<input type="checkbox"/>	64:	<a href="#">lewtA</a>	2.8	7	4.3	161	456	<a href="#">PDB</a>	METABOTROPIC GLUTAMATE RECEPTOR SUBTYPE 1
<input type="checkbox"/>	65:	<a href="#">li80A</a>	2.5	6	3.8	125	266	<a href="#">PDB</a>	PURINE NUCLEOSIDE PHOSPHORYLASE
<input type="checkbox"/>	66:	<a href="#">lotrA</a>	2.4	13	3.4	46	49	<a href="#">PDB</a>	PROTEIN CUE2
<input type="checkbox"/>	67:	<a href="#">lpuxA</a>	2.3	4	3.2	98	124	<a href="#">PDB</a>	SPORULATION INITIATION PHOSPHOTRANSFERASE F
<input type="checkbox"/>	68:	<a href="#">ljfjA</a>	2.3	6	3.6	64	134	<a href="#">PDB</a>	CALCIUM-BINDING PROTEIN
<input type="checkbox"/>	69:	<a href="#">ldelA</a>	2.2	8	3.0	57	241	<a href="#">PDB</a>	DEOXYNUCLEOSIDE MONOPHOSPHATE KINASE

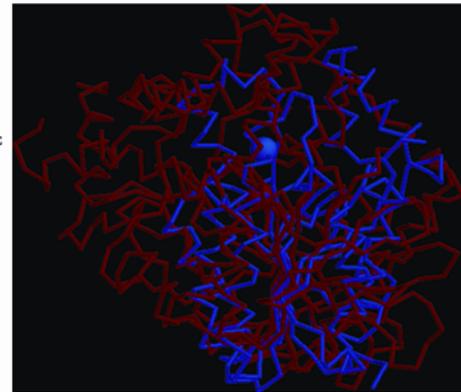


Abb. 16: Dali/FSSP-Recherche mit dem Suchmotiv der 3D-Struktur der Acetylcholinesterase (PDB-Code 2DFP). Links ist die Trefferliste ( $Z$  = Z-Score, %ide = Sequenzidentität in %, lali = Alignmentlänge, lseq2 = Länge der zur Sequenz des Suchmotivs alignierten zweiten Sequenz) und rechts das von Dali generierte Strukturalignment als Überlagerung von 2DFP (blau) und 1AE1 (rot) dargestellt. Das jeweilige  $C^\alpha$ -Rückgrat ist zu sehen, wobei die  $C^\alpha$ -Atome der katalytischen Schlüsselaminosäuren, S200 (2DFP) und Y171 (1AE1), in CPK-Darstellung zu sehen sind.

Diese Ergebnisse führten dazu, als drittes Mitglied des Proteinstruktur-Ähnlichkeitsclusters ein Enzym mit Rossmann-Faltung miteinzubeziehen. Interessanterweise gehört die Tropinonreduktase zur Familie der kurzkettigen Dehydrogenasen/Reduktasen (short-chain dehydrogenases, SDR-Familie), der auch die pharmazeutisch hochrelevanten Hydroxysteroid-Dehydrogenasen angehören. Aus dieser Familie wurden die  $11\beta$ -Hydroxysteroid-Dehydrogenasen ( $11\beta$ HSDs) Typ 1 und Typ 2 im Hinblick auf ihre strukturelle Ähnlichkeit mit Cdc25A bzw. AChE näher untersucht.

Die beiden  $11\beta$ HSD-Isoformen regulieren die wechselseitige Umwandlung der biologisch inaktiven 11-Ketosteroide (Cortison (**22**), 11-Dehydrocorticosteron) in die entsprechenden aktiven  $11\beta$ -Hydroxyglucocorticoide (Cortisol (**23**), Corticosteron) und umgekehrt (siehe Abb. 17).<sup>[119]</sup> Die Enzyme der SDR-Familie verfügen über ein katalytisches Tyrosin, das in ein konserviertes YXXXK-Motiv eingebettet ist, und haben eine Glycin-reiche N-terminale Cosubstrat-Bindungsstelle gemeinsam.<sup>[120]</sup>  $11\beta$ HSD1 ist für die lokale und gewebespezifische Aktivierung des Glucocorticoid-Rezeptors essentiell, da es die Oxoreduktion von Cortison zu Cortisol katalysiert.  $11\beta$ HSD1 könnte ein aussichtsreiches therapeutisches Target für die Antagonisierung Glucocorticoid-vermittelter Effekte sein.<sup>[119]</sup> Eine Hemmung der  $11\beta$ HSD1 wird als vielversprechender Ansatz bei der Behandlung der Fettsucht (Adipositas),<sup>[121]</sup> des metabolischen Syndroms,<sup>[122]</sup> des Typ-2-Diabetes<sup>[123, 124]</sup> und der kognitiven Dysfunktion<sup>[125]</sup> diskutiert. Das  $11\beta$ HSD2-Isoenzym katalysiert ausschließlich die Oxidation von Cortisol. Eine Hemmung der  $11\beta$ HSD2 führt zu Natriumretention und Bluthochdruck.<sup>[126]</sup> Aus diesem Grund wäre eine Isoenzym-spezifität für  $11\beta$ HSD1 eine wesentliche Voraussetzung für den klinischen Einsatz von  $11\beta$ HSD1-Inhibitoren.

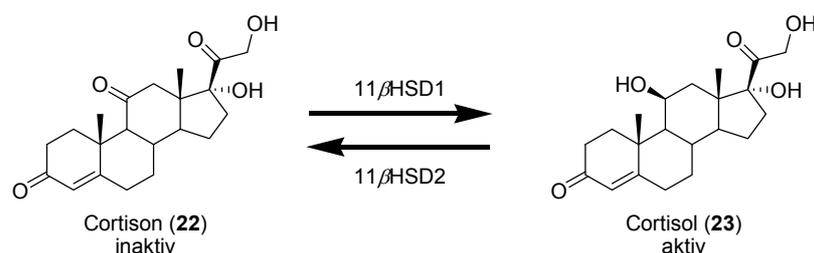


Abb. 17: Von  $11\beta$ -Hydroxysteroid-Dehydrogenasen katalysierte Reaktionen.

Da Kristallstrukturen der  $11\beta$ -Hydroxysteroid-Dehydrogenasen nicht verfügbar waren, wurden Homologiemodelle der beiden Isoenzyme auf der Grundlage ihrer Aminosäureprimärsequenz erstellt (s.u.).

Zur Untersuchung der Strukturähnlichkeit wurde der substratbindende Bereich der  $11\beta$ -Hydroxysteroid-Dehydrogenasen, bei denen es sich um Eindomänenproteine handelt, vom Cofactor-bindenden Teil der katalytischen Domäne jeweils abgetrennt (siehe Abb. 18) und einem Strukturvergleich durch Überlagerung mit dem CE-Algorithmus unterzogen. Auch hier konnte eine signifikante strukturelle Ähnlichkeit der katalytischen Kernbereiche der beiden  $11\beta$ HSD-Isoformen mit der katalytischen Domäne der Cdc25A festgestellt werden (siehe Abb. 19).

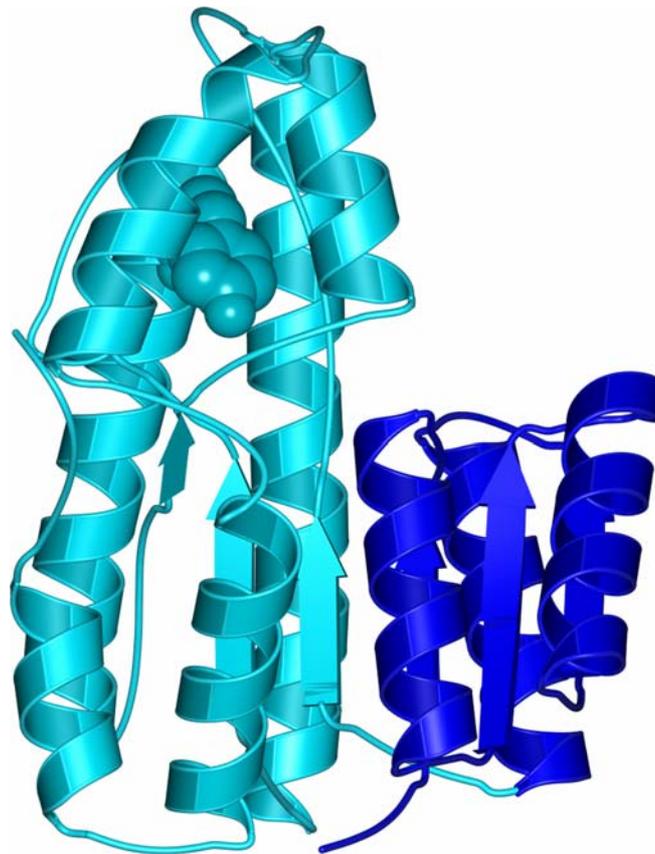
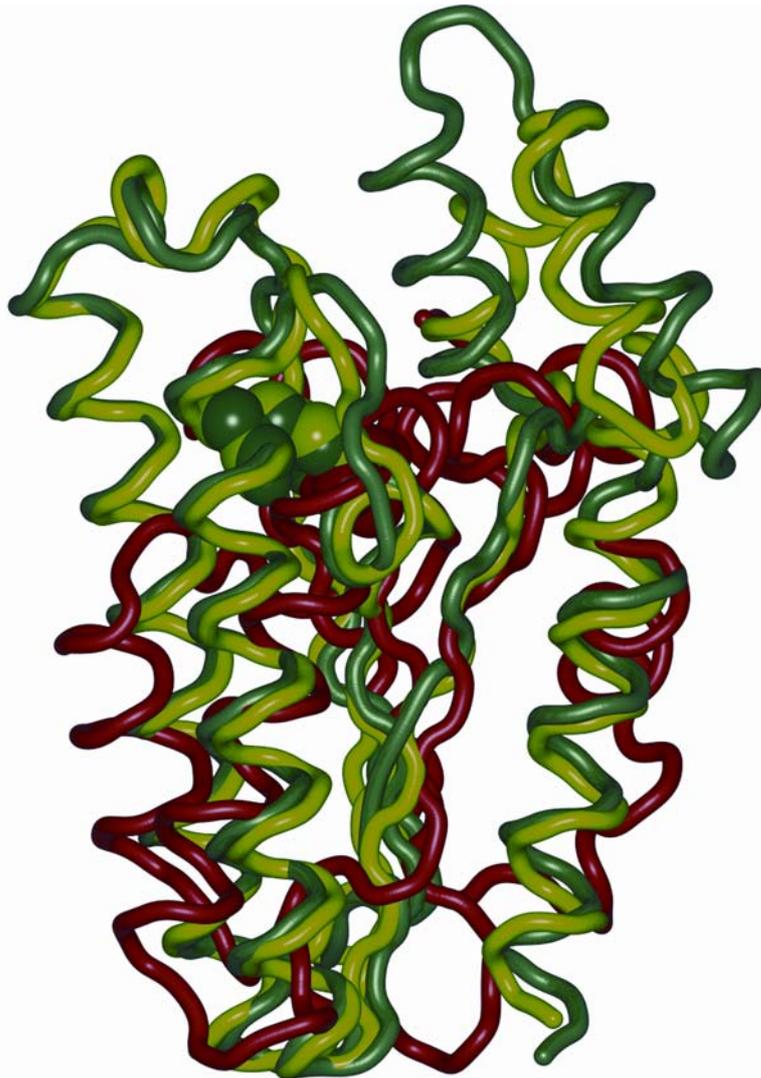


Abb. 18: Substratbindender Kernbereich (cyan, G111-N291, Sequenzlänge: 181 Aminosäuren) am Beispiel des Homologiemodells der  $11\beta$ HSD1 (Sequenzlänge: 261 Aminosäuren). Das katalytische Tyr<sup>183</sup> ist in CPK-Darstellung zu sehen. Der Cofactor-bindende Teil ist blau dargestellt.

A



B

	1	10	20	30	40	50
1C25 (K377-K475)	K E F V I L D C R	· · · · ·	· · · · ·	Y P Y E Y E G G H I K G A V N L H M E E E V E D		
HSD1 (D114-T222)	D M L L L N H I T N T S L · · · · ·	· · · · ·	· · · · ·	R K S M · · · · ·	· · · · ·	· · · · ·
HSD2 (W162-S269)	W G L V N N A G H N E V V A D A E L S P V A T F R S C M · · · · ·	· · · · ·	· · · · ·	· · · · ·	· · · · ·	· · · · ·
	60	70	80	90	100	
1C25 (K377-K475)	F L L K K P I V P T D G K R V I V V F H C E F S S E · · · · ·	· · · · ·	· · · · ·	· · · · ·	· · · · ·	· · · · ·
HSD1 (D114-T222)	L T V A A L P M L K Q S N G S I V V V S S · · · · ·	· · · · ·	· · · · ·	· · · · ·	· · · · ·	· · · · ·
HSD2 (W162-S269)	L T K G L L P L L R S S R G R I V T V G S · · · · ·	· · · · ·	· · · · ·	· · · · ·	· · · · ·	· · · · ·
	110	120				
1C25 (K377-K475)	D R L G N E Y P K L H Y P E L · Y V L K G G Y K E F F M K					
HSD1 (D114-T222)	R K · · E Y S V S R V N V S I T L C V L G L · · I D T E T					
HSD2 (W162-S269)	S C · · E L L P W G V K V S · · I I Q P G C · · F K T E S					

Abb. 19: A) Die überlagerten katalytischen Kernbereiche der Cdc25A (rot), der 11 $\beta$ HSD1 (dunkelgrün) und der 11 $\beta$ HSD2 (hellgrün). Cdc25A und das als Target relevante Isoenzym 11 $\beta$ HSD1 weisen einen RMSD von 4.13 Å bei einer Alignmentlänge von 80 Aminosäuren auf, wobei die Sequenzidentität 5.0% beträgt. In CPK-Darstellung sind die zentralen katalytischen Aminosäuren der Enzyme zu sehen: Cys<sup>430</sup> (Cdc25A), Tyr<sup>183</sup> (11 $\beta$ HSD1) und Tyr<sup>232</sup> (11 $\beta$ HSD2). Sie sind räumlich ähnlich lokalisiert. B) Strukturabgeleitetes Sequenzalignment der überlagerten katalytischen Kernbereiche der Cdc25A (PDB-Code 1C25), der 11 $\beta$ HSD1 (HSD1) und der 11 $\beta$ HSD2 (HSD2). Die katalytischen Schlüsselaminosäuren, C430, Y183 und Y232, sind im Sequenzalignment *nicht* korreliert und jeweils grau unterlegt.

Die katalytischen Aminosäuren Cys<sup>430</sup> (Cdc25A) und Tyr<sup>183</sup> (11 $\beta$ HSD1)/Tyr<sup>232</sup> (11 $\beta$ HSD2) waren im strukturabgeleiteten Sequenzalignment nicht miteinander korreliert (siehe Abb.

19B), obwohl sie – wie in der Überlagerung der Strukturen in Abb.13 zu sehen – durchaus eine ähnliche räumliche Positionierung innerhalb der Faltung aufweisen. Dies gilt besonders für die katalytisch entscheidenden funktionellen Gruppen (Sulfhydrylfunktionalität in Cys<sup>430</sup> und die phenolische Hydroxylfunktion in Tyr<sup>183</sup>).

Auf der Grundlage dieser Datenbank- und Strukturanalyse (siehe Abb. 20 für eine graphische Zusammenfassung des Ergebnisses der Suchstrategie vom März 2004) wurden die Enzyme Cdc25A, AChE und die 11 $\beta$ -Hydroxysteroid-Dehydrogenasen in einem Proteinstruktur-Ähnlichkeitscluster zusammengefaßt.<sup>[41]</sup> Weitere Evidenz für die 3D-Ähnlichkeit der katalytischen Kernbereiche dieser Enzyme lieferte eine Analyse der aktiven Zentren (siehe Abb. 21), die die überlappende Positionierung der katalytischen Schlüsselaminosäuren bei signifikanter struktureller Ähnlichkeit in ihrer unmittelbaren Umgebung bestätigt. Dieser Clustering-Ansatz basiert auf rein strukturellen Erwägungen. Die als strukturell ähnlich klassifizierten Enzyme weisen in ihren jeweiligen katalytisch relevanten Kernbereichen Sequenzidentitäten im Bereich 5% bis 8% auf, was nicht für eine evolutionäre Verwandtschaft spricht. Zusätzlich wird dies durch die Tatsache untermauert, daß sie sehr unterschiedliche Reaktionen katalysieren und auch nach SCOP unterschiedlichen Faltungstypen zugeordnet sind.

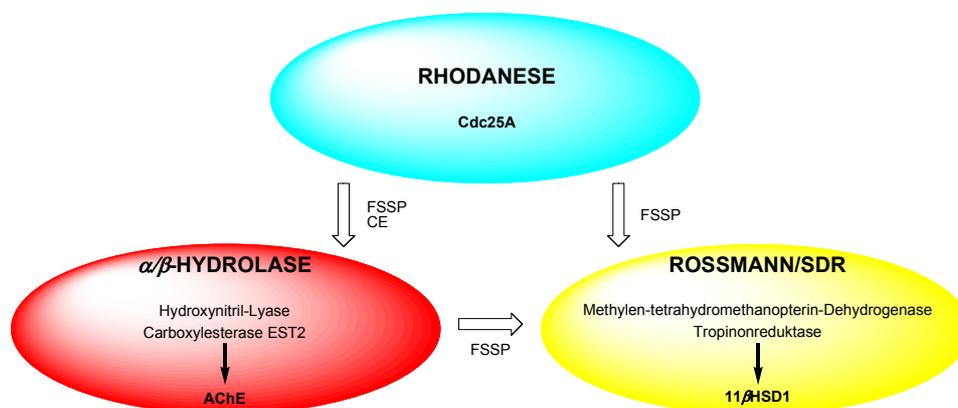


Abb. 20: Graphische Zusammenfassung der Suchstrategie vom März 2004, die zur Identifizierung des Proteinstruktur-Ähnlichkeitsclusters bestehend aus Cdc25A, AChE und 11 $\beta$ HSD1 führte.

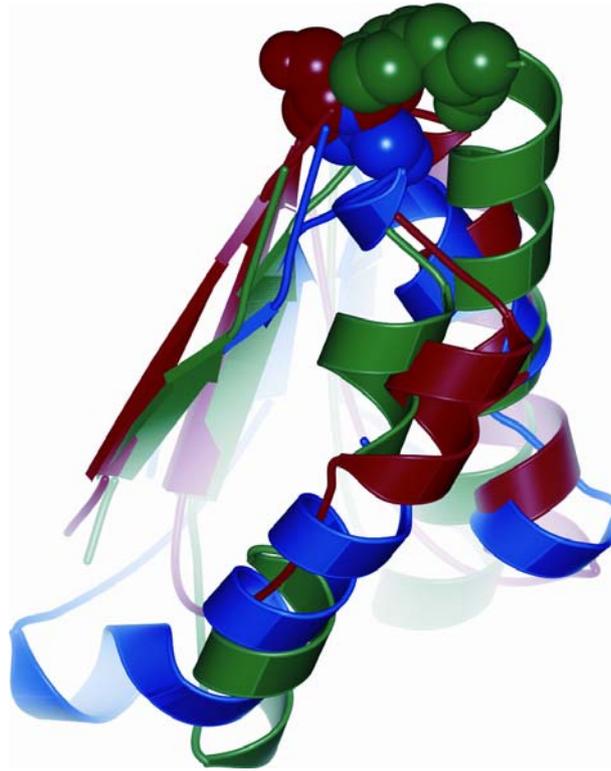


Abb. 21: Aufsicht auf die katalytischen Zentren von Cdc25A (rot), 11 $\beta$ HSD1 (grün) und AChE (blau). Die Schlüsselaminosäuren für die Katalyse, Cys<sup>430</sup> (Cdc25A), Tyr<sup>183</sup> (11 $\beta$ HSD1) und Ser<sup>200</sup> (AChE), die in CPK-Darstellung gezeigt sind, sind innerhalb der Faltung ähnlich lokalisiert.

### ***Homologiemodelle der 11 $\beta$ -Hydroxysteroid-Dehydrogenasen***

Grundsätzlich wäre ein grobes Modell der dreidimensionalen Struktur der 11 $\beta$ -Hydroxysteroid-Dehydrogenasen im Sinne des Konzepts des Proteinstruktur-Ähnlichkeitsclustering ausreichend, da eine Fokussierung auf prinzipielle strukturelle Ähnlichkeiten in den ligandenbindenden Kernbereichen von Proteinen erfolgt. Auch würde die Erkennung der Faltung aus der Aminosäureprimärsequenz schon ausreichen, um nach grundlegender struktureller Ähnlichkeit zu clustern. Dennoch wurde nach einer Methode gesucht, möglichst genaue und zuverlässige 3D-Darstellungen beider Isoenzyme mit vertretbarem Aufwand zu erhalten.

Die Aminosäureprimärsequenzen beider 11 $\beta$ HSD-Isoenzyme wurden der ‚SWISS-PROT Protein knowledgebase‘<sup>[127]</sup> (<http://au.expasy.org/sprot/>) entnommen. Zur Generierung des Konsensus wurden die Sequenzen dem ‚Structure Prediction Meta-Server‘<sup>[128]</sup> (<http://bioinfo.pl/meta/>) übermittelt. Dieser Meta-Server leitete die Sequenzen an mehrere autonome Server weiter, die dann jeweils den Konsensus berechneten. Das im SPMS implementierte „3D-Jury-System“ bewertet und rankt als „Meta-Prädiktor“ die Lösungen der

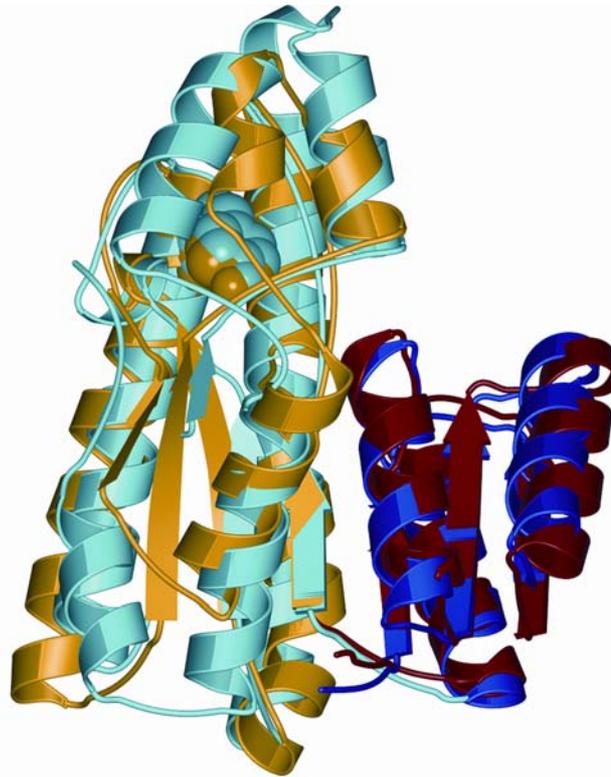
einzelnen autonomen Server nach einheitlichen Kriterien, wobei ein spezieller Score jedem Alignment zugeordnet wird. Ein Metaserver-Ansatz wurde aufgrund der höheren Qualität der Konsensusbildung gewählt. Die Konsensusbildung bei Enzymen der SDR-Familie ist zudem alles andere als trivial, da diese nur eine sehr geringe Sequenzähnlichkeit (Sequenzidentitäten im Bereich 15–30%) untereinander aufweisen. Das 3D-Jury-System schnitt im übrigen als bester Meta-Prädiktor im CASP5-Assessment von Methoden zur Proteinstrukturvorhersage im Bereich der Faltungserkennung und der Homologiemodellierung ab.<sup>[129]</sup> Es schnitt sogar genauso gut ab wie die besten menschlichen Prädiktoren. Auch im Zuge der automatisierten LiveBench-6-Evaluierung von Proteinstrukturvorhersage-Servern wurde das 3D-Jury-System als eines der besten Systeme bewertet.<sup>[130]</sup> Die Bildung des Konsensus, d.h. das Alignment zweier Proteinsequenzen, wobei von einer eine 3D-Struktur bekannt ist, ist für die Qualität des Homologiemodells entscheidend.

Mit diesem Verfahren wurden Alignments der Aminosäureprimärsequenzen von 11 $\beta$ HSD1 und 11 $\beta$ HSD2 zu den Sequenzen der Templatstrukturen der Levodionreduktase (PDB-Code 1IY8), das durch den FFAS03-Algorithmus<sup>[130]</sup> im Falle der 11 $\beta$ HSD1 generiert wurde, und der Glucosedehydrogenase (PDB-Code 1GCO), das durch den mGenTHREADER-Algorithmus im Falle der 11 $\beta$ HSD2 generiert wurde,<sup>[131]</sup> erstellt. Diese Alignments wurden mit einem 3D-Jury-Score von 175.53 bzw. 167.42 als höchstrangige Lösungen bewertet. Dieser Score spiegelt die Konsistenz der von den einzelnen autonomen Vorhersageservern generierten Modelle wieder. Eine im wesentlichen korrekte – d.h., wenn die modellierte Faltung der experimentell bestimmten Struktur ähnelt – Strukturvorhersage weist einen 3D-Jury-Score > 50 auf.<sup>[132]</sup> Da das Alignment die Qualität des Modells stark beeinflusst, wurde jeweils genau untersucht, ob z.B. Lückenpositionen sinnvoll gesetzt oder ob Sekundärstrukturelemente durch Lücken unterbrochen waren. Die Sequenzidentitäten in den Alignments betragen 17.4% für 11 $\beta$ HSD1 und 17.9% für 11 $\beta$ HSD2.

Die so generierten Alignments wurden zur Erstellung dreidimensionaler Modelle mit dem Programm MODELLER 6v2<sup>[133]</sup> (<http://salilab.org/modeller/>) herangezogen. Die Modelle wurden anschließend mit ProQ („Protein Quality Predictor“, <http://www.sbc.su.se/~bjorn/ProQ/>)<sup>[134]</sup>, einem neuronalen Netzwerk zur Identifizierung korrekter Proteinmodelle, evaluiert. Die ausgegebenen Zahlenwerte, LGscore and MaxSub, betragen 6.271 und 0.650 für 11 $\beta$ HSD1 und 5.088 und 0.427 für 11 $\beta$ HSD2, und bestätigten jeweils die gute Qualität der Modelle. Darüber hinaus wurden zusätzlich WhatCheck-<sup>[135]</sup> und ProCheck-<sup>[136]</sup> Analysen zur weiteren Qualitätsabsicherung durchgeführt.

Die Güte des Homologiemodells der 11 $\beta$ HSD1 konnte zudem anhand der zwischenzeitlich veröffentlichten Kristallstruktur<sup>[137]</sup> (PDB-Code 1XU7) dieses Enzyms überprüft werden. 85.3% der Aminosäuren konnten zwischen Modell und Kristallstruktur in einem strukturabgeleiteten Sequenzalignment unter Verwendung des CE-Algorithmus miteinander korreliert werden, wobei der RMSD-Wert 1.76 Å betrug (siehe Abb. 22).

**A**



**B**

11βHSD1 (MODEL)	1	10	20	30	40	50
1XU7	M L Q G K K V I V T G A S K G I G R E M A Y H L A K M G A H V V V T A R S K E T L Q K V V S H C L E					
11βHSD1 (MODEL)	60	70	80	90	100	
1XU7	L G A A S A H Y I A G T M E D M T F A E Q F V A Q A G K L M G G L D M L I L N H I T . N T S L N L F					
11βHSD1 (MODEL)	110	120	130	140	150	
1XU7	H D D I H H V R K S M E V N F L S Y V V L T V A A L P M L K Q S N G S I V V V S S L A G K V A Y P M					
11βHSD1 (MODEL)	160	170	180	190	200	
1XU7	V A A Y S A S K F A L D G F F S S I R K E Y S . . V S R V N V S I T L C V L G L I D T E T A M K A V					
11βHSD1 (MODEL)	210	220	230			
1XU7	S G I V H M Q A A P K E E C A L E I I K G G A L R Q E E V Y Y D S S L W T T L					
	. . . . . S G I V H M Q . . . . . A A P K E E C A L E I I K G G A L					

Abb. 22: Evaluierung des Homologiemodells der 11 $\beta$ HSD1. A) Überlagerung von Modell (blau – Cofactorbindender Bereich, cyan – Substratbindender Bereich) und Kristallstruktur (PDB-Code 1XU7, (rot – Cofactorbindender Bereich, orange – Substratbindender Bereich). Das katalytische Tyr<sup>183</sup> ist jeweils in CPK-Darstellung zu sehen. B) Strukturabgeleitetes Sequenzalignment der überlagerten 3D-Strukturen von Homologiemodell und Kristallstruktur. Die Signaturen für die Cofaktorbinding, GXXXGXG, und für das katalytische Zentrum, YXXXK, sind jeweils grau unterlegt.

### ***Dysidiolid als Ausgangspunkt***

Cdc25A, AChE und die 11 $\beta$ -Hydroxysteroid-Dehydrogenasen wurden wie oben beschrieben unter rein strukturellen Gesichtspunkten in einem Proteinstruktur-Ähnlichkeitscluster zusammengefaßt. Unter Berücksichtigung des strukturellen Konservativismus der Natur beim prinzipiellen Aufbau der katalytischen Kernbereiche dieser Enzyme wurde ein natürlich vorkommender Hemmstoff eines Enzyms dieses Clusters, das Sesterterpen Dysidiolid (**24**, siehe Abb. 23), als Ausgangspunkt für die Synthese einer fokussierten Verbindungskollektion ausgewählt.

### ***Synthese einer Dysidiolid-inspirierten Verbindungskollektion***

Aufgrund früherer Erkenntnisse<sup>[15, 16]</sup> und Literaturberichte über die Phosphatase-hemmende Aktivität von Dysidiolid und verwandten Naturstoffen<sup>[138]</sup> wurde die Hypothese aufgestellt, daß die  $\gamma$ -Hydroxybutenolidgruppe im Naturstoff seine Phosphatasehemmung wesentlich bestimmt. Auf dieser Grundlage synthetisierte Dr. Lars-Oliver Wittenberg eine Kollektion von 147  $\gamma$ -Hydroxybutenoliden und strukturell verwandten  $\alpha,\beta$ -ungesättigten Fünfring-Lactonen (siehe Abb. 23).<sup>[41, 139]</sup> Eine Auflistung sämtlicher Strukturen ist in Anhang I aufgeführt.

Die Synthese erfolgte unter Anwendung etablierter Verfahren.<sup>[16, 140-142]</sup> Furan oder 3-Bromfuran wurden selektiv in der 2- oder 3-Position lithiiert. Die Furyllithium-Zwischenprodukte wurden mit verschiedenen Elektrophilen abgefangen, um sekundäre und tertiäre Alkohole oder Furylketone als Intermediate zu generieren. Die Ketone und die sekundären Alkohole (diese nach ihrer Oxidation zum Keton) wurden mit Grignard- oder Aryllithium-Reagenzien weiter umgesetzt. Die Konversion der Furane zu den entsprechenden Hydroxybutenoliden erfolgte durch Cycloaddition von photochemisch generiertem Singulett-Sauerstoff und anschließender regioselektiver Öffnung der Cycloaddukte.<sup>[143, 144]</sup>  $\alpha,\beta$ -Ungesättigte Lactone wurden durch die Reaktion von verschiedenen Aldehyden mit 2-(Trimethylsilyloxy)furan in der Gegenwart von Lewis-Säure erhalten.<sup>[145]</sup> Ein paar wenige Thiophen-Derivate wurden in ähnlicher Weise synthetisiert.

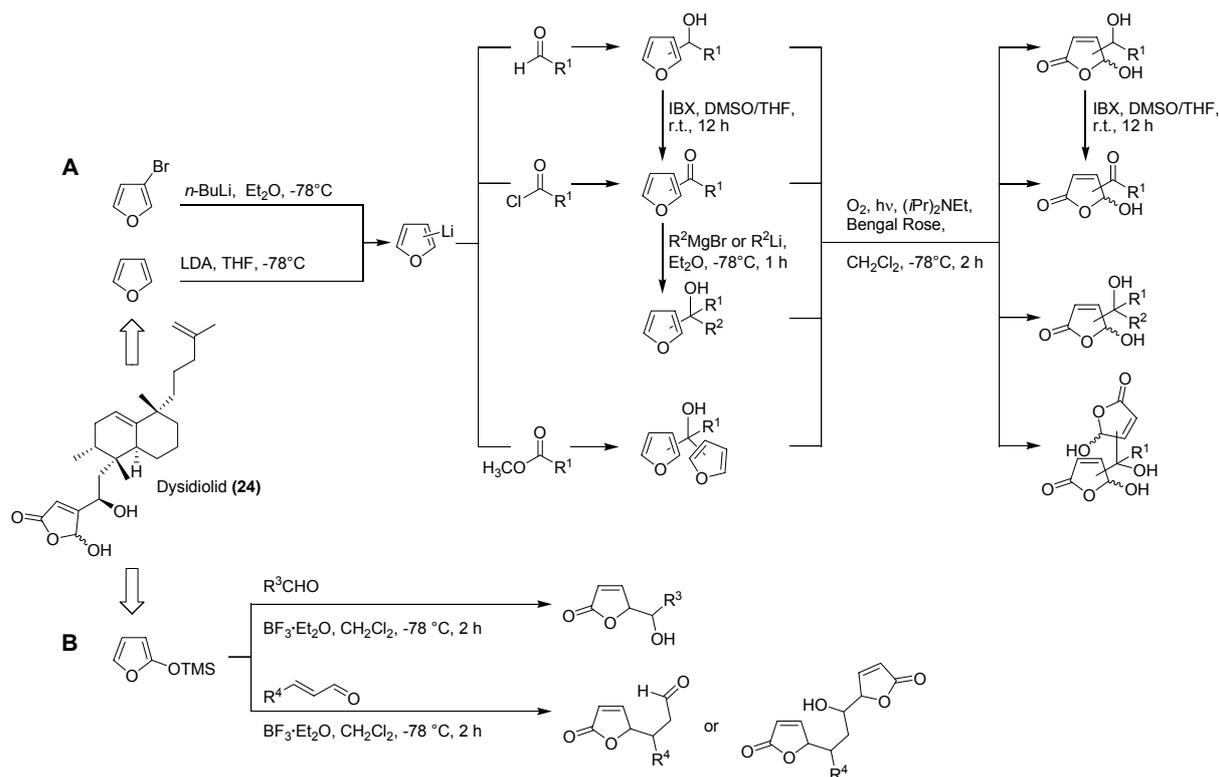


Abb. 23: A) Synthese von 2- und 3-substituierten Furanen und  $\gamma$ -Hydroxybutenoliden. B) Synthese von 5-substituierten Butenoliden und Bis-Butenoliden.

### Biochemische Evaluierung

Die gesamte vom Naturstoff Dysidiolide (**24**) abgeleitete Verbindungskollektion wurde in biochemischen Assay-Systemen im Hinblick auf eine mögliche Hemmung der Cdc25A, der AChE, der  $11\beta\text{HSD1}$  oder der  $11\beta\text{HSD2}$ , wie nachfolgend beschrieben, untersucht.

#### Hemmung der Cdc25A

Dieser Inhibitionsassay wurde in Zusammenarbeit mit Heike Rimpel und Walburga Hecker durchgeführt. Der Klon pET9d/His-Cdc25A wurde im *Escherichia coli*-Stamm BL21-DE3 exprimiert und in der Gegenwart von 8 M Harnstoff aufgereinigt. Das Testsystem wurde auf der Grundlage bekannter Verfahren entwickelt,<sup>[113, 146]</sup> wobei *p*-Nitrophenylphosphat als colorimetrisches Substrat verwendet wurde. Die generierten Daten, dargestellt als Mittelwert und Standardabweichung, wurden in drei unabhängigen Experimenten bestimmt.

#### Hemmung der $11\beta$ -Hydroxysteroid-Dehydrogenasen

Dieser Inhibitionsassay wurde in Kooperation mit PD Dr. Alex Odermatt, Universität Bern, durchgeführt. Die  $11\beta\text{HSD1}$ -abhängige Oxoreduktion von Cortison und die  $11\beta\text{HSD2}$ -abhängige Oxidation von Cortisol wurde in Lysaten von stabil transfizierten HEK-293-Zellen

nach etablierten Verfahren gemessen.<sup>[121]</sup> Die Geschwindigkeit der Umsetzung von Cortison zu Cortisol oder der umgekehrten Reaktion wurde bestimmt, indem [1,2,6,7-<sup>3</sup>H]-markiertes Substrat in einer Endkonzentration von 200 nM Cortison oder 25 nM Cortisol mit dem jeweiligen Inhibitor (0-200  $\mu$ M) vermessen wurde. Die Meßdaten stammten aus mindestens vier unabhängigen Experimenten. Berechnet wurden Mittelwert und Standardabweichung.

#### *Hemmung der Acetylcholinesterase*

Die inhibitorische Aktivität wurde mittels der auf Mikrotiterplattenformat adaptierten spektrophotometrischen Methode nach Ellman *et al.* bestimmt.<sup>[147, 148]</sup> Acetylthiocholiniodid wurde als Substrat der enzymatischen Reaktion verwendet und 5,5'-Dithiobis(2-nitrobenzoesäure) (DTNB) für die Messung der Cholinesterase-Aktivität. Die Meßdaten entstammen mindestens vier unabhängigen Experimenten, wobei Mittelwert und Standardabweichung berechnet wurden.

#### *Ausschluß der promiskuitiven Hemmung aufgrund von Aggregatbildung*

Um eine unspezifische Hemmung der untersuchten Enzyme durch aggregatbildende Verbindungen, die auf einer reversiblen physikalischen Assoziation zwischen Aggregaten und den Enzymen beruht, auszuschließen, wurden die identifizierten Hits für jedes Enzym jeweils nach Zugabe von 0.001–0.01% Triton-X-100 vermessen. Bei der 11 $\beta$ HSD2 war dies nicht möglich, da diese als membrangebundenes Enzym schon durch geringe Tensidkonzentrationen durch Herauslösen aus der Membran denaturiert wird.<sup>[149]</sup> Nicht-ionische Detergentien sind in der Lage, einen derartigen Inhibitionsmechanismus zu unterbinden. Dabei heben niedrige Konzentrationen von Triton X-100 die Aggregat-Enzym-Interaktion auf.<sup>[150]</sup>

### ***Ergebnis***

Verbindungen, die in den Inhibitionsassays IC<sub>50</sub>-Werte  $\leq 10 \mu$ M aufwiesen, wurden als Hit erachtet (siehe Tabelle 1). Nach diesem Ausschlußkriterium qualifizierten sich 42 der insgesamt 147 getesteten Verbindungen als Hits im Cdc25A-Inhibitionsassay. Mit einem IC<sub>50</sub> von 350 nM erwies sich Verbindung **25** als wirksamster Inhibitor. Dieser Wert ist signifikant niedriger als der für Dysidiolid in der Literatur berichtete IC<sub>50</sub>-Wert (9.4  $\mu$ M<sup>[151]</sup>). Drei Verbindungen (**26-28**) hemmten die AChE mit IC<sub>50</sub>-Werten zwischen 1.3  $\mu$ M und 4.5  $\mu$ M. Die Verbindungskollektion enthielt drei 11 $\beta$ HSD1-Inhibitoren (IC<sub>50</sub>-Werte 7.8-10  $\mu$ M) und

vier 11 $\beta$ HSD2-Hemmer (IC<sub>50</sub>-Werte 2.4-6.7  $\mu$ M). Somit lagen die Hitraten für die Enzyme, die als strukturell ähnlich zu Cdc25A klassifiziert wurden, im Bereich 2–3%. Dies ist ein mehr als akzeptabler Wert für ein auf die Identifizierung neuer Strukturklassen von Inhibitoren ausgerichtetes Screening, der in der Mehrzahl von Screens nicht erreicht wird. Die Ergebnisse der Profilierung für die wichtigsten Verbindungen sind in Tabelle 1 aufgelistet.

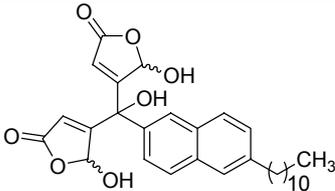
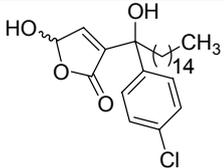
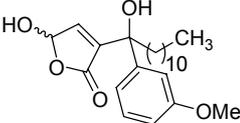
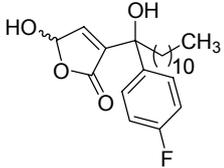
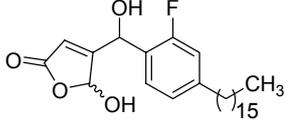
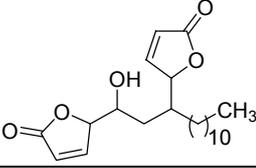
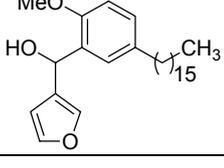
Vbdg	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
25		0.35±0.18	> 20	14±3	2.4±0.3
26		2.3±0.5	1.3±0.1	7.8±1.8	2.8±0.4
27		1.6±0.6	4.5±0.3	10±1	14±2
28		1.5±0.2	2.1±0.5	13±3	34±4
29		> 100	> 20	19±3	5.3±1.1
30		45	> 20	10±2	95±4
31		1.8±0.7	> 20	19±3	6.7±0.9

Tabelle 1: Synopse der Inhibitionsdaten für ausgewählte Verbindungen.

In dieser Substanzkollektion befanden sich trotz ihrer relativ geringen Größe einige Hits, die eine ausgesprochene Selektivität für die einzelnen Enzyme und auch für die Isoenzyme 11 $\beta$ HSD1 und 11 $\beta$ HSD2 zeigten.

Verbindung **25** stellte sich verglichen mit ihrer Inhibition der anderen Enzyme als ein deutlich stärkerer Hemmstoff der Cdc25A heraus. Verbindung **29** zeigte eine klare Präferenz für 11 $\beta$ HSD2. Besonders bemerkenswert ist das  $\alpha,\beta$ -ungesättigte Lacton (**30**), das lediglich das therapeutisch relevante Isoenzym 11 $\beta$ HSD1 und die anderen untersuchten Enzyme nicht oder nur sehr schwach hemmte. Auch ein Furan-Abkömmling (**31**) erwies sich als ein Inhibitor der Cdc25A und der 11 $\beta$ HSD2. Ein selektiver AChE-Hemmstoff wurde nicht entdeckt. Eine promiskuitive Hemmung der Enzyme aufgrund von Aggregatbildung wurde ausgeschlossen, indem Kontrollexperimente nach Shoichets etablierter Methode mit niedrigen Konzentrationen des Detergens Triton X-100 durchgeführt wurden.<sup>[150]</sup>

### *Diskussion*

Diese Ergebnisse bestätigen an einem ersten Beispiel die vorwärtsgerichtete Anwendbarkeit des Proteinstruktur-Ähnlichkeitsclustering (PSSC) als Filter für die initiale Identifizierung von (neuen) Verbindungsklassen als Ausgangspunkten, aus denen dann Liganden für multiple Mitglieder eines Ähnlichkeitsclusters hervorgehen können.

In diesem Zusammenhang ist zu betonen, daß letztlich die präzise strukturelle Organisation der Bindungsstellen der entscheidenden Faktor für die Bindung potentieller Liganden an Proteine ist. Da bei einem Clustering, wie es oben beschrieben wurde, die Bindungsstellen – obgleich topologisch in der Faltung ähnlich lokalisiert – sehr verschieden sein können, muß ein kombinatorischer Ansatz gewählt werden, um die biologische Diversität in den Bindungstaschen mit einer komplementären chemischen Diversität ansprechen zu können. PSSC sollte daher als initiales abstrahierendes Leitprinzip verstanden werden, das biologisch relevante Grundgerüste organischer Moleküle liefert, die in Folge jedoch kombinatorisch variiert werden müssen. Insgesamt liefert PSSC nicht Leitstrukturen, sondern Hits, die in medizinisch-chemischen Programmen optimiert werden müssen.

Die Ergebnisse zeigen auch, daß Potenz und Selektivität von Liganden für die einzelnen Mitglieder in einem Ähnlichkeitscluster auch in relativ kleinen Bibliotheken erreicht werden können.

Das für die Synthese der Verbindungskollektion durch PSSC bestimmte Startmolekül im chemischen Strukturraum war ein Naturstoff, was wiederum nahelegt, daß von Naturstoffen abgeleitete Verbindungsbibliotheken Proteinliganden mit hoher Zuverlässigkeit liefern

könnten. Naturstoffe sind biologisch prävalidierte Moleküle, die im Zuge ihrer Biosynthese und bei der Ausübung ihrer biologischen Funktion (z.B. chemische Verteidigung oder Kommunikation) an zahlreiche Proteine binden. Viele Naturstoffklassen weisen eine darüber hinausgehende biologische Aktivität auf, die vielfach in medizinische Anwendungen mündete, z.B. Morphin, Taxol, Colchicin, Pilocarpin, um nur einige wenige beispielhaft zu nennen. Aufgrund der Konservierung struktureller Elemente in Proteinen ist daher zu erwarten, daß den Naturstoffen zugrundeliegenden Strukturgerüsten die Fähigkeit, an Proteinrezeptoren zu binden, innewohnt, und daß diese Fähigkeit auch auf abgeleitete Strukturklassen übertragen werden kann. Verbindungskollektionen, die auf solchen privilegierten<sup>[14]</sup> und von der Natur im Laufe der Evolution selektierten Strukturen basieren, sollten daher hohe Hitraten biologisch prävalidierter Proteinliganden bei vergleichsweise kleiner Bibliotheksgröße liefern.

PSSC ermöglicht, vollständig neue und unerwartete Ligandentypen zu entdecken. Obwohl viele Inhibitor-Strukturklassen für die AChE<sup>[152, 153]</sup> und die 11 $\beta$ -Hydroxysteroid-Dehydrogenasen<sup>[119]</sup> bekannt sind, wurden  $\gamma$ -Hydroxybutenolide und  $\alpha,\beta$ -ungesättigte Fünfringlactone bislang noch nicht als Hemmstoffe der AChE oder der 11 $\beta$ -Hydroxysteroid-Dehydrogenasen beschrieben.

Die hier diskutierten Ergebnisse beschreiben die erste *De-novo*-Anwendung eines Clusterings nach struktureller Ähnlichkeit in den ligandenbindenden Kernbereichen von Proteindomänen. Dieser Clustering-Ansatz war rein strukturell ausgerichtet. Auf der Grundlage der Aminosäureprimärsequenz oder konservierter Sequenzmotive wäre eine gemeinsame Gruppierung dieser funktionell unterschiedlichen Enzyme nicht gelungen.

Ogleich die Enzyme des hier vorgestellten Proteinstruktur-Ähnlichkeitsclusters unterschiedlichen Faltungstypen (nach SCOP) zugeordnet sind, waren die Ligandenbindungsstellen innerhalb der Faltung jeweils topologisch ähnlich angeordnet. Dies spricht dafür, daß es sich hierbei um sogenannte ‚Supersites‘<sup>[154]</sup> handeln könnte. Derartige ‚Supersites‘ entstehen vermutlich aufgrund gewisser struktureller Prinzipien in Proteinen (wie z.B. der  $\alpha$ -Helixdipol oder Prinzipien der Bildung von Krümmungen und Oberflächen) oder chemischer Zwänge.<sup>[154]</sup> Die Prinzipien, die zur Existenz von ‚Supersites‘ führen, sind bislang wenig verstanden. Fest steht jedoch, daß eine ähnliche Lokalisierung von Bindungsstellen in einer Proteinfaltung im Sinne von ‚Supersites‘ selten auf eine funktionelle Ähnlichkeit hinweist, da die Substrate chemisch deutlich unterschiedlich sein können. Der hier

beschriebene Ansatz eines Clusterings der ligandenbindenden Kernbereiche von Proteindomänen nach struktureller Ähnlichkeit könnte dazu genutzt werden, derartige ‚Supersites‘ in Proteinen *in silico* zu identifizieren. Dieser Ansatz ergänzt bestehende *In-silico*-Verfahren<sup>[154]</sup> und eröffnet Wege, die Existenz von ‚Supersites‘ für die Identifizierung und Synthese von niedermolekularen Liganden dieser ‚Supersites‘ zu nutzen.

### 4.1.3 Ornithindecaboxylase und Dihydropteroatsynthetase – Ein Proteinstruktur-Ähnlichkeitscluster

#### *Die Dihydropteroatsynthetase als Suchmotiv*

Die Dihydropteroatsynthetase (DHPS) ist ein bakterielles Enzym, das bei der *De-novo*-Biosynthese von Folaten eine wichtige Rolle spielt. Prokaryoten und niedere Eukaryoten haben im Gegensatz zu höheren Eukaryoten kein Carrier-vermitteltes Transportsystem, das ihnen erlauben würde, Folate aufzunehmen. Sie sind daher auf eine eigene Folat-Synthese angewiesen.<sup>[155]</sup> Die DHPS ist in dem entsprechenden Biosyntheseweg ein entscheidendes Enzym, das die Kondensation von *p*-Aminobenzoesäure (*p*ABA) und 7,8-Dihydropterinpyrophosphat (DHPPP) unter Eliminierung von Pyrophosphat zu 7,8-Dihydropteroat katalysiert,<sup>[156, 157]</sup> das anschließend durch weitere Enzyme zu Tetrahydrofolat umgesetzt wird.

Die DHPS wird nach SCOP dem ‚TIM beta/alpha-barrel‘-Faltungstyp und einer eigenen Superfamilie der Dihydropteroatsynthetaseähnlichen zugeordnet. Das achtsträngige  $\alpha/\beta$ -Faltungsmotiv [ $(\beta/\alpha)_8$ - oder TIM-barrel] wurde zum ersten Mal bei der Triosephosphatisomerase (TIM) beobachtet.<sup>[158]</sup> Von diesem sehr vielseitigen Faltungstyp wird angenommen, daß sich die einzelnen Proteine durch divergente Evolution aus einem gemeinsamen Vorfahren entwickelt haben.<sup>[159]</sup> Die diesem Faltungstyp zugeordneten Enzyme katalysieren daher eine Vielzahl unterschiedlicher Reaktionen. Sie kommen aus der Gruppe der Oxidoreduktasen, der Transferasen, Hydrolasen, Lyasen oder Isomerasen. Viele  $(\beta/\alpha)_8$ -Faß-Enzyme enthalten über die kanonische Faltungstopologie hinaus Erweiterungen entweder an den N- oder C-Termini der Sequenz oder in Schlaufenbereichen.<sup>[159]</sup> Bei allen bekannten  $(\beta/\alpha)_8$ -Faß-Enzymen sind die katalytischen Zentren an der C-terminalen Seite in Schlaufenbereichen, die die  $\beta$ -Stränge mit den folgenden  $\alpha$ -Helices verbinden, gelegen. Es handelt sich um einen klassischen Fall von ‚Supersites‘.<sup>[154]</sup>

So war es nicht überraschend, daß die im vorhergehenden Kapitel dargestellte Suchstrategie in der Dali/FSSP-Datenbank mit der 3D-Struktur der Dihydropteroatsynthetase aus *Escherichia coli* (PDB-Code 1AJ0, M1-E282, Sequenzlänge: 282 Aminosäuren), einem Eindomänenprotein, als Suchmotiv die Ornithindecaboxylase aus *Trypanosoma brucei* (ODC, PDB-Code 1F3T, R14-S422, Sequenzlänge: 409 Aminosäuren) als Treffer lieferte (siehe Abb. 24). Die ODC besteht aus zwei Proteindomänen (siehe Abb. 25), einer C-

terminalen katalytischen Domäne, die eine  $(\beta/\alpha)_8$ -Fafstruktur zeigt, und einer N-terminalen Domäne. Die katalytische Domäne der ODC wird nach SCOP einer anderen Superfamilie (‚PLP-binding barrel‘) als die DHPS zugeordnet. Die ODC ist ein Pyridoxal-5'-phosphat-abhängiges homodimeres Enzym, das den ersten und geschwindigkeitsbestimmenden Schritt bei der Biosynthese von Polyaminen, nämlich die Decarboxylierung von Ornithin zu Putrescin<sup>[160]</sup>, katalysiert. Entsprechend dem Dali-Alignment teilt die ODC mit der DHPS eine Sequenzidentität von 10% bei einem RMSD-Wert von 3.2 Å und einer Alignmentlänge von 175 Aminosäuren.

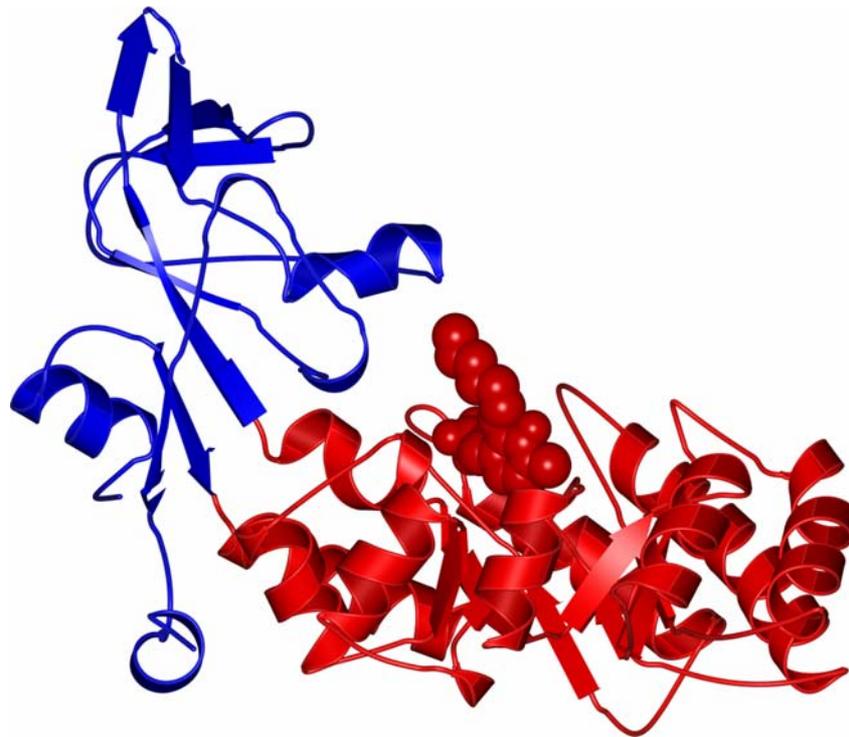


Abb. 25: Domänenarchitektur der ODC. Links, blau eingefärbt, die N-terminale Domäne (R14-A43, F284-S422), die nach SCOP dem ‚Greek key‘-(‚Domain of alpha and beta subunits of F1 ATP synthase-like‘)-Faltungstyp zugeordnet wird, und rechts, rot eingefärbt, die C-terminale katalytische Domäne der ODC (D44-A283) mit  $(\beta/\alpha)_8$ -Fafstruktur.

## Dali database: select structural neighbours of 1aj0

Please cite: L. Holm and C. Sander (1996) Science 273(5275):595-60.

structure alignment		structure+sequence alignment		3D superimposition		
PDB format	Reset selection					
neighbour	Z	wide	rmsd	lali	lseq2	compound
0: 1aj0	52.6	100	0.0	282	282	DIHYDROPTEROATE SYNTHASE
1: 1ad1A	33.1	37	2.3	256	264	DIHYDROPTEROATE SYNTHETASE
2: 1eyeA	30.4	40	1.7	237	256	DIHYDROPTEROATE SYNTHASE I
3: 1f6yA	21.3	16	2.7	222	258	5-METHYLTETRAHYDROFOLATE CORRINOID
4: 1g7vA	18.3	12	3.0	233	284	2-DEHYDRO-3-DEOXYPHOSPHOCTONATE A
5: 1fwvA	17.1	12	3.2	220	263	2-DEHYDRO-3-DEOXYPHOSPHOCTONATE A
6: 1cb7D	15.9	9	3.5	191	483	GLUTAMATE MUTASE
7: 1qfeA	15.7	9	2.9	208	252	3-DEHYDROQUINATE DEHYDRATASE
8: 1pk1G	15.6	8	3.2	203	498	PYRVUATE KINASE
9: 1b4eA	15.2	13	3.5	211	323	5-AMINOLEVULINIC ACID DEHYDRATASE
10: 112qA	14.9	9	3.5	175	457	MONOMETHYLAMINE METHYLTRANSFERASE
11: 1ho4A	14.9	13	3.2	192	242	PDB
12: 1gsgA	14.9	14	3.7	213	329	DELTA-AMINOLEVULINIC ACID DEHYDRAT
13: 1hfbC	14.8	7	3.4	209	348	TYROSINE-REGULATED 3-DEOXY-D-ARABI
14: 1rpxA	14.8	11	3.0	185	230	RIBULOSE-PHOSPHATE 3-EPIMERASE
15: 1h5yB	14.5	8	3.4	162	253	PDB
16: 1gppA	14.5	9	3.7	163	253	PDB
17: 1kflA	14.4	10	3.2	220	351	PDB
18: 1ka9F	14.4	9	3.5	161	251	PDB
19: 1eunA	14.3	10	3.9	165	213	PDB
20: 1h7rA	14.2	13	3.5	211	342	PDB
21: 1a3wA	14.2	12	3.1	203	492	PDB
22: 1k3uA	14.1	11	3.0	193	268	PDB
23: 1hlzA	14.1	13	3.3	184	219	PDB
24: 1kV8B	14.1	8	3.1	153	215	PDB
25: 1fs5A	14.0	14	3.2	197	293	PDB
26: 1hl2A	13.8	13	3.1	194	295	PDB
27: 1lt8A	13.7	8	3.3	209	348	PDB
28: 1mumA	13.7	9	4.2	152	289	PDB
29: 1g69B	13.7	10	3.3	189	228	PDB
30: 17req	13.6	8	3.0	206	623	PDB
31: 1llb1A	13.6	12	3.6	163	247	PDB
32: 1nsg	13.5	9	2.8	173	205	PDB
33: 1gqgB	13.5	12	3.5	198	245	PDB
34: 1agfa	13.4	9	3.2	203	519	PDB
35: 1mzhA	13.3	13	3.1	173	225	PDB
36: 1n8wA	13.3	9	3.1	211	718	PDB
37: 1jvnA	13.3	12	3.5	208	537	PDB
38: 1dvjA	13.1	13	3.4	184	239	PDB
39: 1j2wA	13.1	14	3.1	167	212	PDB
40: 1dhpA	12.9	13	3.0	188	292	PDB
41: 1i4nA	12.8	9	3.2	188	251	PDB
42: 1reqA	12.8	9	3.2	218	727	PDB
43: 1n7kA	12.8	12	3.2	170	234	PDB
44: 1qi0A	12.7	10	3.2	201	302	PDB
45: 1egzA	12.7	9	3.6	184	291	PDB
46: 1dxeA	12.7	11	2.9	177	253	PDB
47: 1rucA	12.6	6	3.5	207	467	PDB
48: 1pii	12.6	10	3.8	188	452	PDB
49: 1lt8A	12.5	13	3.4	198	577	PDB
50: 1b5cA	12.4	8	3.4	149	275	PDB
51: 1pvmA	12.4	10	3.4	183	291	PDB
52: 1b9bB	12.4	12	3.2	189	254	PDB
53: 1bqcA	12.3	9	3.3	197	302	PDB
54: 1o0yA	12.3	11	3.3	172	251	PDB
55: 1m6jA	12.3	8	3.0	166	260	PDB
56: 1go2A	12.3	12	3.3	185	241	PDB
57: 1gvfB	12.3	10	3.0	182	275	PDB
58: 1treA	12.3	6	3.7	164	255	PDB
59: 1gk8A	12.1	5	3.5	205	469	PDB
60: 1dquA	12.1	9	3.8	194	513	PDB
61: 1btmA	12.0	7	3.2	159	251	PDB
62: 1htiA	12.0	9	3.4	189	248	PDB
63: 1qnoA	12.0	9	3.9	211	344	PDB
64: 1hg3A	11.9	12	3.2	175	224	PDB
65: 1timA	11.8	6	3.9	173	247	PDB
66: 1hinA	11.8	10	3.4	193	304	PDB
67: 1a5cA	11.8	7	3.8	213	342	PDB
68: 1uahA	11.8	11	3.8	210	426	PDB
69: 1sbcaA	11.7	7	3.5	206	516	PDB
70: 1f2jA	11.7	7	3.8	213	357	PDB
71: 1ep3A	11.7	10	3.4	125	311	PDB
72: 2aldA	11.6	7	3.8	212	363	PDB
73: 1dquA	11.5	8	3.5	189	267	PDB
74: 1gonA	11.5	7	3.9	210	447	PDB
75: 1lzcA	11.5	12	3.1	178	299	PDB
76: 1olzA	11.5	12	3.3	170	226	PDB
77: 1d3hA	11.5	10	3.5	205	364	PDB
78: 1g0cA	11.5	11	3.3	200	358	PDB
79: 1od0B	11.4	5	3.8	204	446	PDB
80: 1dosA	11.4	7	3.3	204	358	PDB
81: 1eixC	11.3	12	3.4	180	232	PDB
82: 1ctdB	11.3	11	3.4	187	249	PDB
83: 1f07A	11.3	8	3.7	156	321	PDB
84: 1kc7A	11.3	9	3.4	186	872	PDB
85: 1iw6A	11.3	8	3.8	213	551	PDB
86: 1fbaA	11.3	5	3.9	216	360	PDB
87: 1f76A	11.2	8	3.2	185	336	PDB

88: 1mo0A	11.2	9	3.2	185	257	PDB	TRIOSEPHOSPHATE ISOMERASE
89: 1h62A	11.2	11	3.7	207	364	PDB	PENTAERYTHRITOL TETRANITRATE REDUC
90: 1m7oA	11.2	7	3.4	188	246	PDB	TRIOSEPHOSPHATE ISOMERASE
91: 1qdaA	11.2	11	3.3	185	250	PDB	TRIOSEPHOSPHATE ISOMERASE
92: 1eepA	11.2	9	3.2	191	314	PDB	
93: 1zjfA	11.2	9	3.2	194	476	PDB	INOSINE MONOPHOSPHATE DEHYDROGENAS
94: 1qvba	11.1	9	4.2	221	481	PDB	BETA-GLYCOSIDASE
95: 1jclA	11.1	8	3.3	171	252	PDB	DEOXYRIBOSE-PHOSPHATE ALDOLASE
96: 1kqgA	11.1	12	3.4	196	372	PDB	TRNA-GUANINE TRANSGLYCOSYLASE
97: 1edg	11.1	6	3.8	202	380	PDB	ENDOGLUCANASE A
98: 1kd0A	11.1	10	3.5	193	413	PDB	BETA-METHYLASPARTASE
99: 1jphA	11.1	8	3.8	196	357	PDB	UROPORPHYRINOGEN DECARBOXYLASE
100: 1gqxA	11.1	7	4.2	216	449	PDB	BETA-GLYCOSIDASE
101: 1eqcA	11.0	7	3.9	210	394	PDB	EXO- (B) - (1,3) -GLUCANASE
102: 1gthA	11.0	7	3.6	170	1019	PDB	DIHYDROPYRIMIDINE DEHYDROGENASE
103: 1gqwA	11.0	7	3.8	208	489	PDB	BETA-GLYCOSIDASE
104: 1fdjA	11.0	7	3.9	211	363	PDB	FRUCTOSE 1,6-BISPHOSPHATE ALDOLASE
105: 1exbA	11.0	12	4.0	169	326	PDB	KV BETA2 PROTEIN
106: 1j96A	10.9	9	4.1	197	323	PDB	3ALPHA-HYDROXYSTEROID DEHYDROGENAS
107: 1lbggA	10.9	5	4.2	192	447	PDB	BETA-GLUCOSIDASE A
108: 1ezwA	10.9	7	4.1	180	347	PDB	COENZYME F420-DEPENDENT N5,N10
109: 1jrlA	10.8	9	3.3	191	436	PDB	
110: 1eceA	10.8	6	3.6	204	358	PDB	ENDOCELLULASE E1
111: 1lbtA	10.8	7	3.1	135	237	PDB	
112: 1pbgA	10.7	7	4.0	208	468	PDB	6-PHOSPHO-BETA-D-GALACTOSIDASE
113: 1gwjA	10.7	11	4.1	177	374	PDB	MORPHINONE REDUCTASE
114: 1nouA	10.7	6	3.5	196	480	PDB	BETA-HEXOSAMINIDASE BETA CHAIN
115: 1c9wA	10.6	8	4.1	196	314	PDB	CHO REDUCTASE
116: 1qi3A	10.6	8	3.9	216	418	PDB	EXO-MALTOTETRAHYDROLASE
117: 1huvA	10.6	9	3.4	189	349	PDB	L(+)-MANDELATE DEHYDROGENASE
118: 1cen	10.6	9	3.6	195	334	PDB	CELLULOSE CELC
119: 1hqtA	10.6	9	4.1	193	324	PDB	ALDEHYDE REDUCTASE
120: 1eleB	10.6	8	4.0	207	495	PDB	BETA-GLUCOSIDASE
121: 1a0cA	10.5	6	3.8	209	437	PDB	XYLOSE ISOMERASE
122: 1el3A	10.5	7	4.2	198	316	PDB	ALDOSE REDUCTASE
123: 1me7A	10.4	6	4.0	125	365	PDB	
124: 1k8cB	10.3	9	4.0	162	319	PDB	XYLOSE REDUCTASE
125: 1m41A	10.3	8	4.4	181	328	PDB	FMNH2-DEPENDENT ALKANESULFONATE MO
126: 1lbgg	10.3	16	3.6	184	399	PDB	D-GLUCARATE DEHYDRATASE
127: 1j93A	10.3	9	3.9	186	343	PDB	UROPORPHYRINOGEN DECARBOXYLASE
128: 1d0rA	10.3	7	3.7	205	311	PDB	DIHYDROROTATE DEHYDROGENASE A
129: 1c7eA	10.3	7	3.3	205	858	PDB	BETA-N-ACETYLHEXOSAMINIDASE
130: 1l6wA	10.2	6	3.3	166	220	PDB	FRUCTOSE 6-PHOSPHATE ALDOLASE 1
131: 1f3tB	10.2	10	3.2	175	381	PDB	ORNITHINE DECARBOXYLASE
132: 1jneA	10.2	7	3.5	187	400	PDB	IMAGINAL DISC GROWTH FACTOR-2
133: 1lekaA	10.2	8	4.2	197	315	PDB	ALDOSE REDUCTASE
134: 1icpA	10.2	6	3.3	156	358	PDB	12-OXOPHYTODIENOATE REDUCTASE 1
135: 1f49A	10.2	10	4.3	192	1021	PDB	BETA-GALACTOSIDASE
136: 1afeA	10.2	10	4.1	196	319	PDB	3-ALPHA-HYDROXYSTEROID DEHYDROGENA
137: 1gvlA	10.2	12	3.8	196	352	PDB	GLYCOLATE OXIDASE (E.C.1.1.3.15) M
138: 1kbiA	10.1	9	3.7	201	504	PDB	CYTOCHROME B2
139: 1luoA	10.1	7	4.2	214	558	PDB	OLIGO-1,6-GLUCOSIDASE
140: 1d8cA	10.1	8	3.2	214	709	PDB	MALATE SYNTHASE G
141: 1gjjwA	10.1	8	3.7	205	636	PDB	MALTODEXTRIN GLYCOSYLTRANSFERASE
142: 1a80	10.1	10	3.9	184	277	PDB	2,5-DIKETO-D-GLUCONIC ACID REDUCTA
143: 1b54	10.1	6	3.6	164	230	PDB	YEAST HYPOTHETICAL PROTEIN
144: 1ljya	10.0	9	3.3	163	361	PDB	MGP-40
145: 1dtn	10.0	15	3.2	174	357	PDB	MANDELATE RACEMASE
146: 1kkoA	10.0	9	3.4	187	411	PDB	3-METHYLASPARTATE AMMONIA-LYASE
147: 1bw1A	9.9	11	4.1	220	399	PDB	NADPH DEHYDROGENASE 1
148: 1m01A	9.9	5	3.5	192	499	PDB	BETA-N-ACETYLHEXOSAMINIDASE
149: 1hjwA	9.9	8	4.1	193	362	PDB	CHITINASE-3 LIKE PROTEIN 1
150: 1liwA	9.9	9	3.9	191	436	PDB	ENDO-1,4-BETA-D-XYLANASE
151: 1k77A	9.8	8	4.1	184	260	PDB	HYPOTHETICAL PROTEIN YGBM
152: 1gveA	9.8	10	4.1	195	324	PDB	AFLATOXIN B1 ALDEHYDE REDUCTASE
153: 1b30A	9.7	9	3.9	189	302	PDB	XYLANASE
154: 1rubaA	9.7	8	4.1	173	459	PDB	RIBULOSE-1,5-BISPHOSPHATE CARBOXYL
155: 1e6eM	9.7	6	3.7	208	499	PDB	MYROSINASE
156: 1eh9A	9.7	7	3.8	199	557	PDB	GLYCOSYLTERALHOSE TREHALOHYDROLASE
157: 1eftA	9.6	17	3.7	170	382	PDB	ALANINE RACEMASE
158: 1fh9A	9.6	8	3.6	191	312	PDB	BETA-1,4-XYLANASE
159: 1ll7A	9.6	5	3.6	187	392	PDB	CHITINASE 1
160: 1e5nA	9.6	10	3.8	167	346	PDB	ENDO-1,4-BETA-XYLANASE A
161: 1m7xB	9.6	6	4.1	204	591	PDB	1,4-ALPHA-GLUCAN BRANCHING ENZYME
162: 1gtwA	9.5	8	3.8	195	285	PDB	ENDONUCLEASE IV
163: 1lglA	9.5	7	4.2	188	365	PDB	CHITOTRIOSIDASE
164: 1glvA	9.5	8	4.0	187	585	PDB	ALPHA-AMYLASE II
165: 1avaA	9.5	5	4.3	206	403	PDB	BARLEY ALPHA-AMYLASE 2 (CV MENNET
166: 1eswA	9.5	11	3.7				

186: <a href="#">1kfwA</a>	8.9	8	4.3	187	435	<a href="#">PDB</a>	CHITINASE B	235: <a href="#">1ef2A</a>	6.1	12	4.2	170	566	<a href="#">PDB</a>	UREASE ALPHA SUBUNIT
187: <a href="#">1jgmA</a>	8.9	10	4.7	210	333	<a href="#">PDB</a>	PHOSPHOTRIESTERASE	236: <a href="#">1fhvA</a>	5.7	11	3.6	159	322	<a href="#">PDB</a>	O-SUCCINYLBENZOATE SYNTHASE
188: <a href="#">1ea9D</a>	8.8	5	3.6	201	583	<a href="#">PDB</a>	CYCLOMALTODEXTRINASE	237: <a href="#">1m65A</a>	5.6	9	3.6	152	234	<a href="#">PDB</a>	HYPOTHETICAL PROTEIN YCDX
189: <a href="#">1bhgA</a>	8.8	9	3.6	193	611	<a href="#">PDB</a>	BETA-GLUCURONIDASE	238: <a href="#">1gxlA</a>	5.3	8	3.9	167	362	<a href="#">PDB</a>	CELLOBIHYDROLASE II
190: <a href="#">1lwjA</a>	8.8	6	3.8	192	441	<a href="#">PDB</a>	4-ALPHA-GLUCANOTRANSFERASE	239: <a href="#">1j5aA</a>	5.3	6	3.5	148	451	<a href="#">PDB</a>	URONATE ISOMERASE
191: <a href="#">1v1wA</a>	8.8	8	3.6	195	471	<a href="#">PDB</a>	ALPHA-AMYLASE	240: <a href="#">1ny1A</a>	5.2	8	3.4	135	235	<a href="#">PDB</a>	PROBABLE POLYSACCHARIDE DEACETYLASE
192: <a href="#">1k87A</a>	8.8	7	4.1	182	514	<a href="#">PDB</a>	PROLINE DEHYDROGENASE	241: <a href="#">1ubpc</a>	5.2	7	3.6	124	570	<a href="#">PDB</a>	UREASE GAMMA SUBUNIT
193: <a href="#">1kr0A</a>	8.8	4	3.8	180	273	<a href="#">PDB</a>	HEVAMINE A	242: <a href="#">1epxA</a>	5.1	7	4.1	151	357	<a href="#">PDB</a>	FRUCTOSE-1,6-BISPHOSPHATE ALDOLASE
194: <a href="#">1mw2A</a>	8.8	10	3.9	216	628	<a href="#">PDB</a>	AMYLOUCRASE	243: <a href="#">1dkrB</a>	5.0	9	3.9	110	306	<a href="#">PDB</a>	PHOSPHORIBOSYL PYROPHOSPHATE SYNTH
195: <a href="#">1gvvA</a>	8.8	7	3.7	194	376	<a href="#">PDB</a>	MANNAN ENDO-1,4-BETA-MANNOSIDASE	244: <a href="#">1j1lA</a>	4.9	2	4.4	166	637	<a href="#">PDB</a>	ALPHA-AMYLASE I
196: <a href="#">1i75A</a>	8.8	8	4.2	214	686	<a href="#">PDB</a>	CYCLODEXTRIN GLUCANOTRANSFERASE	245: <a href="#">1k3rB</a>	4.8	6	3.3	115	264	<a href="#">PDB</a>	CONSERVED PROTEIN MT0001
197: <a href="#">1cnv</a>	8.7	8	3.9	183	283	<a href="#">PDB</a>	CONCAVALIN B	246: <a href="#">1nfp</a>	4.5	4	4.7	144	228	<a href="#">PDB</a>	LUXF GENE PRODUCT (NONFLUORESCENT
198: <a href="#">1dhnA</a>	8.7	9	3.5	193	496	<a href="#">PDB</a>	PORCINE PANCREATIC ALPHA-AMYLASE	247: <a href="#">1fvpA</a>	4.5	6	4.5	143	231	<a href="#">PDB</a>	FLAVOPROTEIN 390
199: <a href="#">1chrA</a>	8.5	9	3.5	174	370	<a href="#">PDB</a>	CHLOROMUCONATE CYCLOISOMERASE (E.C	248: <a href="#">1jqnA</a>	4.5	5	3.4	160	874	<a href="#">PDB</a>	PHOSPHOENOLPYRUVATE CARBOXYLASE
200: <a href="#">1kbbA</a>	8.5	11	3.5	193	496	<a href="#">PDB</a>	ALPHA-AMYLASE, PANCREATIC	249: <a href="#">1tmyA</a>	4.3	14	3.3	83	118	<a href="#">PDB</a>	CHEY PROTEIN
201: <a href="#">1djgA</a>	8.4	6	3.7	189	729	<a href="#">PDB</a>	TRIMETHYLAMINE DEHYDROGENASE	250: <a href="#">1a04A</a>	4.1	9	4.9	92	205	<a href="#">PDB</a>	NITRATE/NITRITE RESPONSE REGULATOR
202: <a href="#">1pdz</a>	8.4	8	3.5	170	433	<a href="#">PDB</a>	ENOLASE	251: <a href="#">1a4mA</a>	4.1	10	4.5	167	349	<a href="#">PDB</a>	ADENOSINE DEAMINASE
203: <a href="#">2taaA</a>	8.3	8	4.2	197	478	<a href="#">PDB</a>	TAKA-*AMYLASE A (E.C.3.2.1.1	252: <a href="#">1o7dA</a>	4.0	9	4.2	161	291	<a href="#">PDB</a>	LYSOSOMAL ALPHA-MANNOSIDASE
204: <a href="#">1aq0A</a>	8.3	8	4.1	170	306	<a href="#">PDB</a>	1,3-1,4-BETA-GLUCANASE	253: <a href="#">1bf2</a>	3.7	5	4.2	172	750	<a href="#">PDB</a>	ISOAMYLASE
205: <a href="#">1iv8A</a>	8.2	8	3.9	144	720	<a href="#">PDB</a>	MALTOOLIGOSYL TREHALOSE SYNTHASE	254: <a href="#">6r1rA</a>	3.6	5	3.5	118	738	<a href="#">PDB</a>	RIBONUCLEOTIDE REDUCTASE R1 PROTEI
206: <a href="#">1izgA</a>	8.2	11	3.7	181	316	<a href="#">PDB</a>	TRANSALDOLASE B	255: <a href="#">1d1lA</a>	3.6	10	4.3	97	153	<a href="#">PDB</a>	PUTATIVE POTASSIUM CHANNEL PROTEIN
207: <a href="#">1ktbA</a>	8.2	8	4.3	177	388	<a href="#">PDB</a>	ALPHA-N-ACETYLGLACTOSAMINIDASE	256: <a href="#">1m5tA</a>	3.6	7	3.0	83	123	<a href="#">PDB</a>	CELL DIVISION RESPONSE REGULATOR D
208: <a href="#">1c92A</a>	8.1	9	3.6	169	265	<a href="#">PDB</a>	ENDO-BETA-N-ACETYLGLUCOSAMINIDASE	257: <a href="#">1f89A</a>	3.6	14	4.2	152	271	<a href="#">PDB</a>	HYPOTHETICAL 32.5 KDA PROTEIN YLR3
209: <a href="#">1oepA</a>	8.1	8	3.9	150	422	<a href="#">PDB</a>	MOL_ID: 1	258: <a href="#">1i3cA</a>	3.5	13	4.0	90	144	<a href="#">PDB</a>	RESPONSE REGULATOR RCP1
210: <a href="#">1hkvA</a>	8.1	7	3.2	130	447	<a href="#">PDB</a>	DIAMINOPIMELATE DECARBOXYLASE	259: <a href="#">1jhbA</a>	3.5	16	4.1	129	348	<a href="#">PDB</a>	NICOTINATE MONONUCLEOTIDE:5,6
211: <a href="#">1gkpA</a>	8.1	6	4.5	152	458	<a href="#">PDB</a>	HYDANTOINASE	260: <a href="#">1erzA</a>	3.3	12	4.6	124	303	<a href="#">PDB</a>	N-CARBAMYL-D-AMINO ACID AMIDOHYDRO
212: <a href="#">1j79A</a>	8.0	12	3.9	155	343	<a href="#">PDB</a>	DIHYDROOROTASE	261: <a href="#">2bnh</a>	3.1	9	6.7	126	456	<a href="#">PDB</a>	RIBONUCLEASE INHIBITOR
213: <a href="#">1k70A</a>	7.9	10	4.2	166	423	<a href="#">PDB</a>	CYTOSINE DEAMINASE	262: <a href="#">1cxzA</a>	3.0	4	4.4	107	182	<a href="#">PDB</a>	HIS-TAGGED TRANSFORMING PROTEIN RH
214: <a href="#">1itxA</a>	7.8	8	4.2	195	419	<a href="#">PDB</a>	GLYCOSYL HYDROLASE	263: <a href="#">1j8mF</a>	2.9	8	4.1	112	295	<a href="#">PDB</a>	SIGNAL RECOGNITION 54 KDA PROTEIN
215: <a href="#">1djwB</a>	7.7	7	3.8	171	561	<a href="#">PDB</a>	PHOSPHOINOSITIDE-SPECIFIC PHOSPHOL	264: <a href="#">1a4yA</a>	2.9	8	4.9	107	460	<a href="#">PDB</a>	RIBONUCLEASE INHIBITOR
216: <a href="#">1e9iD</a>	7.7	9	3.7	145	431	<a href="#">PDB</a>	ENOLASE	265: <a href="#">1m7bA</a>	2.9	7	3.8	93	179	<a href="#">PDB</a>	RND3/RHOE SMALL GTP-BINDING PROTEI
217: <a href="#">1gqpA</a>	7.6	9	4.4	102	289	<a href="#">PDB</a>	QUINOLINIC ACID PHOSPHORIBOSYLTRAN	266: <a href="#">1j85A</a>	2.7	9	3.1	88	156	<a href="#">PDB</a>	YIBK
218: <a href="#">1ehnA</a>	7.5	7	4.0	194	540	<a href="#">PDB</a>	CHITINASE A	267: <a href="#">1fqvA</a>	2.6	8	5.2	121	325	<a href="#">PDB</a>	SKP2
219: <a href="#">1btc</a>	7.4	7	3.9	182	491	<a href="#">PDB</a>	BETA-AMYLASE (E.C.3.2.1.2) COMPLEX	268: <a href="#">1m2eA</a>	2.4	12	3.7	83	135	<a href="#">PDB</a>	KAIA
220: <a href="#">1aod</a>	7.2	11	4.0	178	274	<a href="#">PDB</a>	PHOSPHATIDYLINOSITOL-SPECIFIC PHOS	269: <a href="#">1gcza</a>	2.4	10	6.9	98	162	<a href="#">PDB</a>	NS-CARBOXYAMINOIMIDAZOLE RIBONUCLE
221: <a href="#">1k1dA</a>	7.2	10	4.4	198	460	<a href="#">PDB</a>	D-HYDANTOINASE	270: <a href="#">1l1eA</a>	2.1	5	3.1	78	111	<a href="#">PDB</a>	HYPOTHETICAL PROTEIN MTH1491
222: <a href="#">1ebhA</a>	7.2	11	3.4	166	436	<a href="#">PDB</a>	ENOLASE (E.C.4.2.1.11) (2-PHOSPHO-	271: <a href="#">1b00A</a>	2.1	10	3.6	86	122	<a href="#">PDB</a>	PHOSPHATE REGULON TRANSCRIPTIONAL
223: <a href="#">7ptd</a>	7.2	7	4.1	180	296	<a href="#">PDB</a>	PHOSPHATIDYLINOSITOL-SPECIFIC PHOS	272: <a href="#">1uagA</a>	2.0	8	4.8	84	151	<a href="#">PDB</a>	CYTOSINE DEAMINASE
224: <a href="#">2ebn</a>	7.1	5	3.7	167	285	<a href="#">PDB</a>	ENDO-BETA-N-ACETYLGLUCOSAMINIDASE								
225: <a href="#">1ituA</a>	7.0	7	4.0	187	369	<a href="#">PDB</a>	RENAL DIPEPTIDASE								
226: <a href="#">1tml</a>	6.9	8	4.1	175	286	<a href="#">PDB</a>	ENDO-1,4-BETA-D-GLUCANASE (E.C.3.2								
227: <a href="#">1fcvA</a>	6.8	9	3.9	172	324	<a href="#">PDB</a>	HYALURONOGLUCOSAMINIDASE								
228: <a href="#">1j6oA</a>	6.7	10	3.7	160	260	<a href="#">PDB</a>	CONSERVED HYPOTHETICAL PROTEIN TMO								
229: <a href="#">1cb2A</a>	6.6	7	4.3	173	363	<a href="#">PDB</a>	CELLOBIHYDROLASE II								
230: <a href="#">1gprA</a>	6.6	9	4.4	103	284	<a href="#">PDB</a>	QUINOLINIC ACID PHOSPHORIBOSYLTRAN								
231: <a href="#">1eomA</a>	6.5	8	3.4	153	283	<a href="#">PDB</a>	ENDO-BETA-N-ACETYLGLUCOSAMINIDASE								
232: <a href="#">1jfxA</a>	6.3	10	3.7	144	217	<a href="#">PDB</a>	1,4-BETA-N-ACETYLMURAMIDASE M1								
233: <a href="#">1br1A</a>	6.2	7	4.3	154	340	<a href="#">PDB</a>	BACTERIAL LUCIFERASE								
234: <a href="#">1dysA</a>	6.2	8	4.4	172	345	<a href="#">PDB</a>	ENDOGLUCANASE								

Abb. 24: Dali/FSSP-Recherche mit dem Suchmotiv der 3D-Struktur der DHPS (PDB-Code 1AJ0). Die Trefferliste ( $Z = Z$ -Score, %ide = Sequenzidentität in %, lali = Alignmentlänge, lseq2 = Länge der zum Suchmotiv alignten zweiten Sequenz) enthält im wesentlichen Proteine mit TIM-Faßfaltung, darunter u.a. TIM (PDB-Code 1B9B u.a.),  $\alpha$ -Amylase (PDB-Code 1KBB) und die pharmazeutisch relevante Ornithindecaboxylase (Eintrag 131, PDB-Code 1F3T).

Die katalytischen Domänen beider Enzyme wurden mit dem LOCK 2-Algorithmus über den FoldMiner-Webserver<sup>[100, 101]</sup> (<http://dlb4.stanford.edu/foldminer/>) überlagert, wobei die Sequenzidentität 8.8% bei einem RMSD-Wert von 2.65 Å und einer Alignmentlänge von 160 Aminosäuren betrug. Beide katalytischen Domänen zeigen eine ausgesprochene strukturelle Ähnlichkeit (siehe Abb. 26) und auch die katalytischen Zentren sind topologisch gleichartig gelegen.

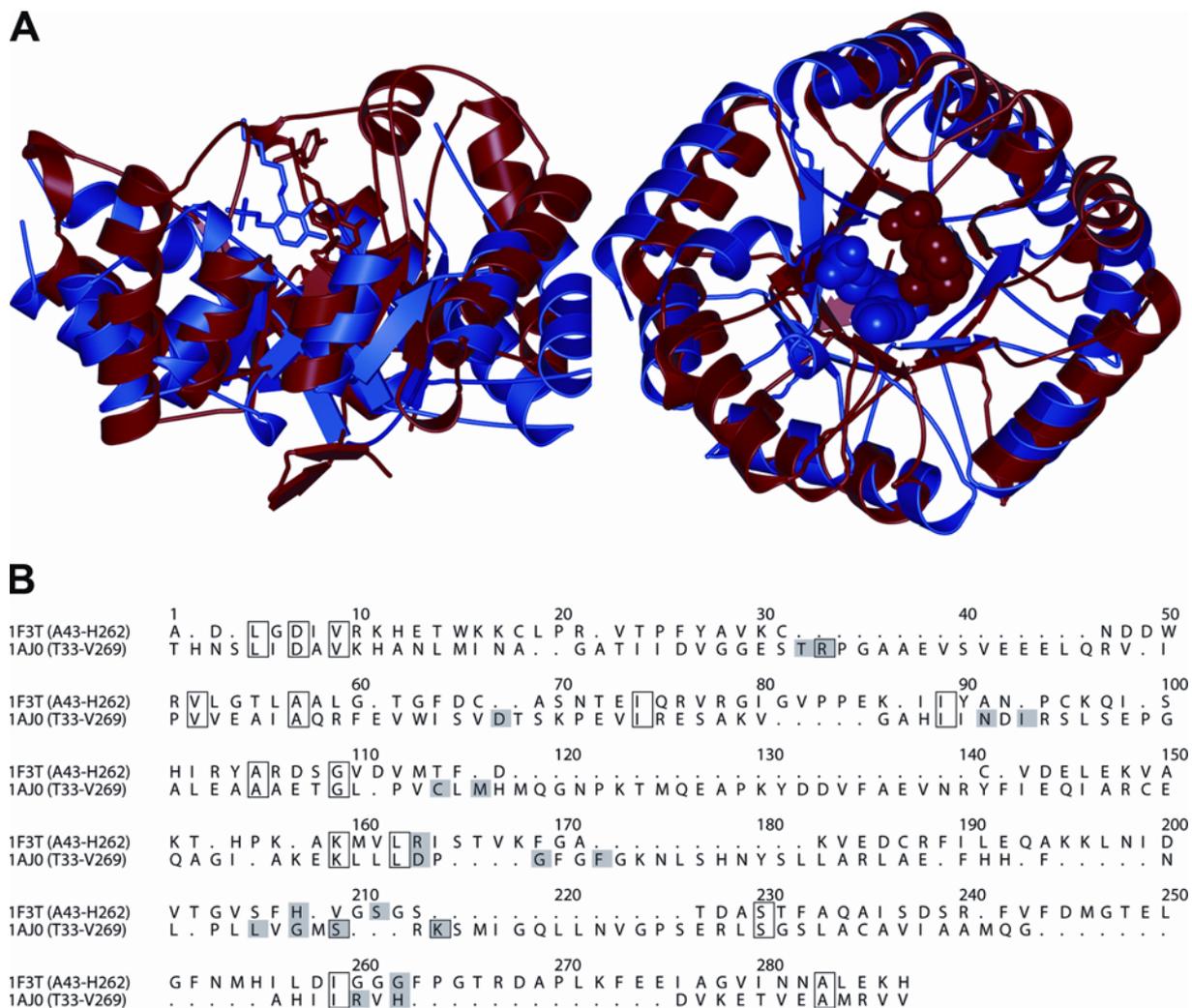


Abb. 26: A) Überlagerung (in Seitenansicht und Aufsicht) der katalytischen Domänen der ODC<sup>[161]</sup> (PDB-Code 1F3T, blau, mit gebundenem Imin aus dem Cofaktor Pyridoxal-5'-phosphat und Putrescin, dem Reaktionsprodukt der von der ODC katalysierten Reaktion) und der DHPS<sup>[155]</sup> (PDB-Code 1AJ0, rot, mit gebundenem Substrat 6-Hydroxymethyl-7,8-dihydropterin und Inhibitor Sulfanilamid). Der RMSD beträgt 2.65 Å bei einer Alignmentlänge von 160 Aminosäuren und einer Sequenzidentität von 8.8%. B) Strukturabgeleitetes Sequenzalignment der überlagerten katalytischen Domänen der ODC (1F3T) und der DHPS (1AJ0). Die zentralen in die Substrat- bzw. Cofaktorbinding involvierten Aminosäuren sind grau hinterlegt.<sup>[155, 160-163]</sup> Die an der Bindung von Sulfanilamid beteiligten Aminosäuren sind zusätzlich eingerahmt. Die Putrescin-bindenden Aminosäuren befinden sich in der N-terminalen Domäne der ODC.

### DHPS und ODC als Targets für Arzneistoffe

In der antimikrobiellen Therapie ist die DHPS ein schon seit langem bekanntes und genutztes Target. Die Sulfonamid-Chemotherapeutika wie z.B. ihr erster Vertreter, das Sulfanilamid (**32**, Abb. 27), hemmen die von der DHPS katalysierte Kondensation von *p*ABA (**33**) mit DHPPP zu Dihydrofolat, indem sie als *p*ABA-Analoga mit *p*ABA um die Bindung an die DHPS konkurrieren. Gerhard Domagk erkannte die antibakteriellen Eigenschaften des Diazofarbstoffs *Prontosil rubrum* (**34**).<sup>[164]</sup> Wenig später entdeckten die französischen Forscher um J. Tréfouel, daß nicht *Prontosil rubrum*, sondern sein schon lange bekanntes

Spaltprodukt Sulfanilamid, *Prontosil album* (**32**), das Bakterienwachstum hemmt.<sup>[165]</sup> Bei *Prontosil rubrum* handelt es sich folglich um ein „Prodrug“. Das aktive Prinzip wird durch Reduktion der Azobindung freigesetzt. Zahlreiche Variationen des Themas „*p*-Aminobenzensulfonamid“ folgten. Unterschiedliche Amidsubstituenten, v.a. 5- und 6-Ring-Heteroaromaten, bewirken unterschiedliche Wirkungsspektren und -stärken sowie unterschiedliche pharmakokinetische Eigenschaften. Außer bei der Behandlung von Harnwegsinfektionen oder einiger Protozoenerkrankungen in synergistischer Kombination mit Dihydrofolatreduktaseinhibitoren (z.B. Sulfamethoxazol/Trimethoprim) haben die Sulfonamid-Chemotherapeutika aufgrund zahlreicher Resistenzentwicklungen ihren therapeutischen Wert heute weitgehend verloren.

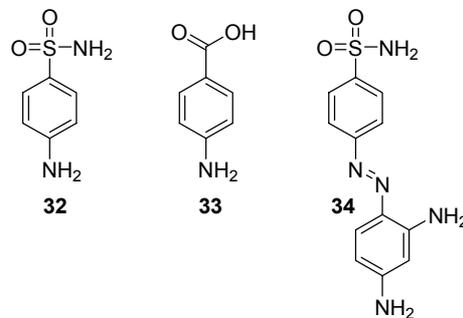


Abb. 27: Substrat-Analogie als Prinzip bei der Hemmung der DHPS.

Die ODC katalysiert den ersten Schritt der Polyamin-Biosynthese, indem sie Ornithin zu Putrescin decarboxyliert.<sup>[163]</sup> Polyamine kommen ubiquitär in allen Zellen vor und werden für Zellwachstum und -differenzierung benötigt. Eine Überexpression von ODC und folglich ein Konzentrationsanstieg intrazellulärer Polyamine wird mit der schnellen Proliferation verschiedener Zelltypen und somit mit Tumorentwicklung und -progression in Verbindung gebracht. Eine Hemmung der ODC scheint daher ein vielversprechender therapeutischer Ansatz bei der Behandlung von Krebs und Infektionskrankheiten zu sein. Die ODC wurde als Target in der Behandlung von Infektionen mit den protozoischen Parasiten *Trypanosoma brucei ssp.*, den Erregern der durch die Tsetse-Fliege übertragenen Afrikanischen Schlafkrankheit, etabliert. Der ODC-Hemmstoff *DL*- $\alpha$ -Difluormethylornithin (DFMO, **35**, Abb. 19) vermag *T. brucei*-verursachte Infektionen bei der Maus<sup>[166]</sup> und beim Menschen<sup>[167]</sup> zu heilen. DFMO bildet eine Schiff'sche Base mit Pyridoxalphosphat (analog dem Ornithin) und wirkt als Suizidinaktivator. Heute sind 50 Millionen Menschen dem Risiko ausgesetzt, sich mit *T. brucei* zu infizieren. Die derzeit verfügbaren antitrypanosomalen Arzneistoffe sind jedoch in ihrer Effektivität limitiert. Dem Erfolg von DFMO gegen *T. brucei gambiense*

stehen seine relativ schlechten pharmakokinetischen Eigenschaften und seine an klinischen Isolaten gezeigte Unwirksamkeit gegen den virulenteren *T. brucei rhodesiense*-Stamm gegenüber.<sup>[163]</sup> Es besteht daher ein großer Bedarf an neuartigen ODC-Hemmstoffen.

Eine Hemmung der Polyaminsynthese durch Hemmung der ODC hat sich als therapeutische Strategie gegen Krebs in klinischen Studien als allgemein ineffektiv erwiesen. Aus präklinischen Studien ging jedoch hervor, daß sich durch die Hemmung der ODC eine wirkungsvolle Chemoprävention gegen Krebs erzielen läßt.<sup>[168, 169]</sup>

Das aktive Zentrum der ODC ist zwischen Wirt und Parasit invariant. Eine selektive Toxizität von DFMO für den Parasiten wird demnach nicht durch eine differenzierende Bindung des Inhibitors an die jeweiligen Bindungsstellen erreicht.<sup>[163]</sup> Es wird stattdessen angenommen, daß diese aus metabolischen Unterschieden zwischen Wirt und Parasit resultiert. Dazu gehören die kürzere *in vivo*-Halbwertszeit der Wirt-ODC von 20 min verglichen mit der des Parasiten (> 1 Tag) und die Erfordernis des Polyamins Spermidin für die Bildung neuen Cofaktors Trypanothion, der für einen Erhalt reduzierter Thiol-Pools im Parasiten sorgt.<sup>[170, 171]</sup> Der  $K_i$ -Wert von DFMO gegenüber der Säugetier-ODC beträgt 40  $\mu\text{M}$ , wohingegen er für die *Trypanosoma*-ODC 130  $\mu\text{M}$  beträgt.<sup>[172]</sup>

Zahlreiche Substrat-(Ornithin)- und Produkt-(Putrescin)-analoge Inhibitoren der Ornithin-Decarboxylase wurden in der Literatur beschrieben.<sup>[173-177]</sup> Ein bekannter Naturstoff-Inhibitor der ODC ist das Phaseolotoxin (**36**, Abb. 28), ein Tripeptidphytoxin, das vom bohnenpathogenen Keim *Pseudomonas syringae pv. phaseolicola* produziert wird.<sup>[178]</sup> Es ist wahrscheinlich, daß das Toxin ein Prodrug darstellt. Es ist anzunehmen, daß das Hydrolyseprodukt  $N^\delta$ -( $N'$ -Sulfodiaminophosphinyl)-*L*-ornithin (**37**), das bereits *in planta* entsteht und das auch als ein überaus potenter Inhibitor der Ornithintranscarbamoylase bekannt ist,<sup>[179]</sup> das eigentlich wirksame Prinzip darstellt. Weiterhin sind u.a. Ifenprodil (**38**, ähnliche Potenz wie DFMO),<sup>[180]</sup> Ibuprofen (**39**)<sup>[181]</sup>, *p*-Phenylendiamin (**40**)<sup>[181]</sup> und Isoniazid (**41**)<sup>[182]</sup> (entnommen aus der BRENDA-Datenbank, <http://www.brenda.uni-koeln.de/>) als Hemmstoffe der ODC bekannt.

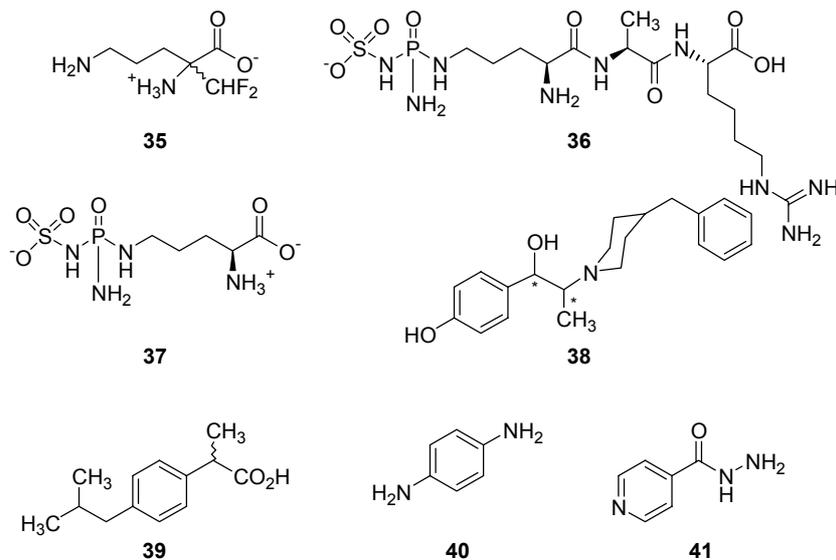


Abb. 28: Bekannte Inhibitoren der ODC.

### ***Das Sulfonamid-Strukturelement als privilegierte Partialstruktur***

Die Sulfonamid-Partialstruktur ist in zahlreichen Arznei- und Wirkstoffen vertreten. Über die bereits vorgestellten *p*-Aminobenzolsulfonamid-Chemotherapeutika hinaus zeigen nahe Strukturverwandte dieser Wirkstoffklasse, die sogenannten Sulfonyl-Harnstoffe, hypoglykämische Wirkung. Bereits 1947 wurde die hypoglykämische Wirkung eines Sulfonamids beobachtet und die potentielle Diabetes-Behandlung mit Sulfonyl-Harnstoffen erkannt. Zu einem therapeutischen Einsatz der Sulfonyl-Harnstoffe als perorale Antidiabetika kam es erst 1955 mit Carbutamid (**42**, siehe Abb. 29), das neben der chemotherapeutischen (aufgrund seiner *p*-Aminobenzolsulfonamid-Struktur) eine blutzuckersenkende Aktivität aufwies. Umfangreiche Bemühungen, die chemotherapeutische und blutzuckersenkende Wirkungskomponente synthetisch zu dissoziieren, führten zu einer Vielzahl verbesserter Wirkstoffe (geringere Dosis, längere Plasmahalbwertszeit, höhere Bioverfügbarkeit) vom Typ des Tolbutamids (**43**), das keine *p*-Aminofunktion mehr aufweist. Alle Sulfonyl-Harnstoffe blockieren den ATP-sensitiven Kaliumkanal ( $K^+_{ATP}$ -Kanal) der B-Zellen der Langerhans'schen Inseln des Pankreas und führen so zu einer vermehrten Freisetzung von Insulin.<sup>[183]</sup>

Die Sulfonamidpartialstruktur ist auch in Wirkstoffen mit saluretischer Wirkung zu finden. Zu nennen sind hier der Carboanhydrasehemmer Acetazolamid (**44**) und die Klasse der Thiazide, hier am Beispiel des Hydrochlorothiazids (**45**), die den  $Na^+/2Cl^-/K^+$ -Cotransporter hemmen.

Zonisamid (**46**) ist ein Sulfonamid mit antikonvulsiver Wirkung. Es blockiert spannungsempfindliche  $\text{Na}^+$ -Kanäle und spannungsabhängige T-Typ- $\text{Ca}^{2+}$ -Kanäle. Es unterdrückt die neuronale Hypersynchronisation. Zonisamid zeigt auch eine schwache Carboanhydrase-inhibitorische Aktivität (vgl. Acetazolamid (**44**)).

Auch in der Gruppe der COX-2-selektive Inhibitoren ist die Sulfonamidstruktur vertreten. Hier sei als Beispiel Celecoxib (**47**) genannt.

In der Stoffgruppe der Sulfonamid-Chemotherapeutika wurde das antibakteriell wirksame Sulfisoxazol (**48**) als Endothelin-Rezeptor-Antagonist identifiziert. Die  $\text{IC}_{50}$ -Werte betragen  $0.60 \mu\text{M}$  für  $\text{ET}_A$ - und  $22 \mu\text{M}$  für  $\text{ET}_B$ -Rezeptoren.<sup>[184]</sup> Eine weitere Bearbeitung dieser Leitstruktur führte zur Verbindung 5-(Dimethylamino)-*N*-(3,4-dimethyl-isoxazol-5-yl)naphthalin-1-sulfonamid (**49**), einem peroral wirksamen  $\text{ET}_A$ -Antagonisten mit einem  $\text{IC}_{50}$ -Wert von  $0.15 \mu\text{M}$ .<sup>[185]</sup>

Diese Analyse legt nahe, daß die Sulfonamid- und besonders die Arylsulfonamidpartialstruktur privilegierte Strukturelemente im chemischen Strukturraum darstellen. Arylsulfonamide könnten sich daher – ähnlich wie Naturstoffe – als biologisch relevante Ausgangspunkte für die Generierung von Proteinliganden eignen.

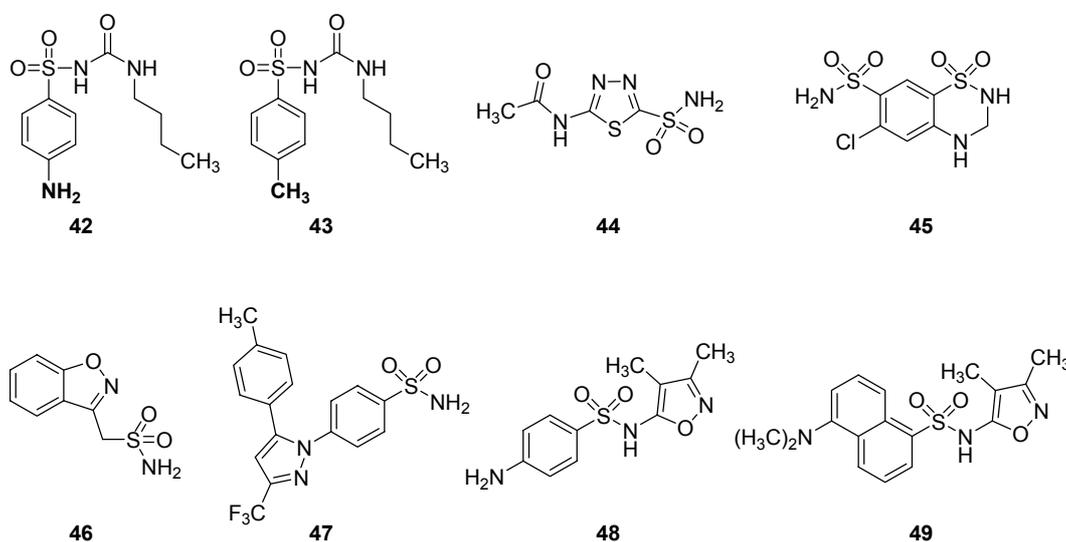


Abb. 29: Die Sulfonamidpartialstruktur als privilegiertes Strukturelement.

### Salazosulfapyridin

Salazosulfapyridin oder Sulfasalazin (**50**, Abb. 30) wird zur Behandlung entzündlicher Darmerkrankungen (*Colitis ulcerosa*, *Morbus Crohn*) eingesetzt. Sulfasalazin ist ein Prodrug,

das im Dickdarm durch Darmbakterien aktiviert wird und dabei 5-Aminosalicylsäure und Sulfapyridin freisetzt. Sulfapyridin wird praktisch vollständig im Dickdarm resorbiert. Sulfasalazin wird peroral verabfolgt. Die therapeutisch verwendete Dosis reicht von 2 g/Tag bei rheumatoider Arthritis und bis zu 4-6 g/Tag zur Behandlung der *Colitis ulcerosa*. Bei einer mehrfachen peroralen Verabfolgung von Sulfasalazin betragen die Steady-State-Plasmakonzentrationen von Sulfapyridin, abhängig vom Acetyliererstatus (schnell oder langsam), 44-128  $\mu\text{M}$  bei einer Gesamtdosis von 2 g/Tag und 132-385  $\mu\text{M}$  bei einer Gesamtdosis von 6 g/Tag.<sup>[186]</sup>

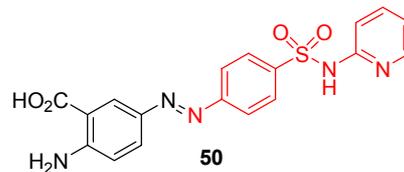


Abb. 30: Salazosulfapyridin, ein Prodrug. Rot dargestellt ist das Spaltprodukt Sulfapyridin.

Bei der Behandlung männlicher Patienten mit Salazosulfapyridin wurde festgestellt, daß diese im Zuge der Behandlung unter Oligospermie und Infertilität litten.<sup>[187]</sup> Das Sulfapyridin wird als das die Unfruchtbarkeit verursachende Prinzip angesehen.

Der antifertile Effekt verschiedener Sulfonamid-Chemotherapeutika, zu deren Substanzklasse auch das Sulfapyridin gehört, wurde im Tierversuch an männlichen Ratten untersucht.<sup>[188]</sup> Es wurde festgestellt, daß vor allem das Sulfapyridin einen signifikanten adversen Effekt auf die männliche Fruchtbarkeit ausübt. Wurde der Pyridinring durch Wasserstoff oder kurze aliphatische Ketten substituiert, blieb die antifertile Wirksamkeit erhalten oder wurde sogar verstärkt. Eine Substitution durch andere Heterocyclen führte meist zu einem Verlust der antifertilen Aktivität.

Das Target, über das Sulfapyridin seinen antifertilen Effekt ausübt, ist bis heute nicht bekannt.

### **Hypothesen**

1. Aufgrund der strukturellen Ähnlichkeit zwischen den katalytischen Kernen der ODC und der DHPS sollte es nach dem Konzept des Proteinstruktur-Ähnlichkeitsclustering möglich sein, ausgehend von Liganden der DHPS strukturell verwandte Liganden der humanen sowie der *Trypanosoma*-ODC abzuleiten. Als Ausgangspunkte im chemischen Strukturraum wurden *p*-Aminobenzolsulfonamide als wegen ihrer

Bindungseigenschaften an DHPS prävalidierte und damit biologisch relevante Startstrukturen verwendet. Das Benzolsulfonamid-Grundgerüst sollte dabei als privilegierte Partialstruktur konserviert bleiben, jedoch als Leitmotiv durchaus variiert werden können.

2. Ausgehend vom PSSC-Ansatz wurde die Hypothese aufgestellt, die antifertile Nebenwirkung des Sulfapyridins könnte auf einer Hemmung der ODC beruhen. Es wurden auch die Argumente erwogen, daß Sulfapyridin im Tierversuch mit männlichen Ratten antifertil wirkt<sup>[188]</sup> und daß eine Störung des ODC-Gens in *C. elegans* dazu führt, daß insbesondere männliche Würmer merklich unfruchtbar werden.<sup>[189]</sup> Außerdem führte die Behandlung von Puppen von *H. assulta* mit DFMO zu einer Verminderung der Konzentration testikulärer Polyamine und zu männlicher Sterilität.<sup>[190]</sup>

### ***Biochemische Evaluierung einer von p-Aminobenzolsulfonamid inspirierten Verbindungskollektion***

Eine kleine Kollektion von 18 kommerziell erhältlichen Sulfonamiden (eine Auflistung der Strukturen und Inhibitionsdaten ist Anhang III zu entnehmen) wurde gegen die humane ODC (*h*ODC) und die ODC aus *Trypanosoma brucei* (*t*ODC) profiliert. Beide Enzyme wurden von Professor Meg Phillips, Dallas, USA, zur Verfügung gestellt.

Zur Anwendung kam ein Assayverfahren, mit dem die durch ODC katalysierte Bildung von Kohlendioxid vermessen wird.<sup>[191]</sup> Mit dieser spektrophotometrischen Methode wird die Abnahme der Absorption durch NADH bei einer Wellenlänge  $\lambda_{\max}$  von 340 nm bestimmt. Der Assay wurde für die Durchführung in Mikrotiterplatten optimiert. Eine schematische Darstellung des Assayprinzips ist in Abb. 31 zu sehen.

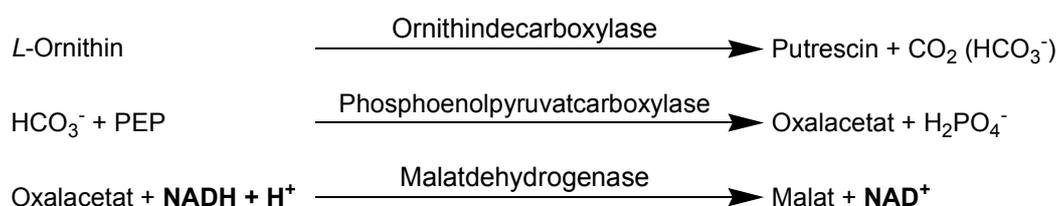


Abb. 31: Assay zur Bestimmung der ODC-Aktivität.

Die gewonnenen Daten beruhen auf mindestens drei unabhängigen Experimenten und sind als Mittelwert mit Standardabweichung angegeben.

### Ergebnis und Diskussion

Die Profilierung der Sulfonamidkollektion lieferte einen Hit, das 4-(Sulfamoylphenyl)hydrazin (**51**, Abb. 32), das als aza-homologes Sulfanilamid aufgefaßt werden kann. Alle anderen Aminobenzolsulfonamide erwiesen sich bei Konzentrationen unterhalb von 1000  $\mu\text{M}$  als inaktiv. Verbindung **51** zeigte für die *t*ODC einen  $\text{IC}_{50}$  von  $81 \pm 5 \mu\text{M}$  und für die *h*ODC einen  $\text{IC}_{50}$  von  $66 \pm 6 \mu\text{M}$ . Wenn diese Werte mit den  $K_i$ -Werten von DFMO (Säugetier-ODC:  $40 \mu\text{M}$ , *t*ODC:  $130 \mu\text{M}$ ) verglichen werden, so schneiden diese durchaus akzeptabel ab. Verbindung **51** ist nicht in der Lage, zwischen Wirts- und Parasit-ODC zu diskriminieren. Diese mangelnde Speziesselektivität von das aktive Zentrum der ODC adressierenden ODC-Inhibitoren ist ein bekanntes Problem.<sup>[170]</sup> Eine Recherche in der ChemBank<sup>[192]</sup> (<http://chembank.med.harvard.edu/>) nach Hydrazinderivaten lieferte die Verbindungen **52-54** (siehe Abb. 32) als Treffer. Diese Verbindungen hemmen allesamt die Aromatische-*L*-Aminosäuren-Decarboxylase, die wie die ODC ein Pyridoxal-5'-phosphat (PLP)-abhängiges Enzym ist. Die Verbindungen **53**, Benserazid (ein Prodrug, aus dem die Hydrazinspezies freigesetzt wird), und **54**, *L*-Carbidopa, sind Arzneistoffe, die in der Therapie des *Morbus Parkinson* eingesetzt werden. Die Hydrazinpartialstruktur ist ganz offenbar Voraussetzung, um eine ausreichende Wirksamkeit (aufgrund der Bildung eines Hydrazons mit dem Cofaktor) zu erreichen. Das entsprechende Hydrazon stellt ein Übergangszustandsanalogon der Aldiminspezies dar, die der Cofaktor (PLP) mit dem Substrat eingeht.<sup>[193]</sup>

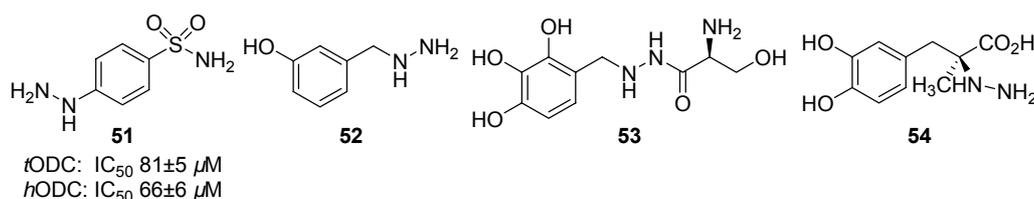


Abb. 32: Hemmstoffe der ODC und der Aromatischen-*L*-Aminosäuren-Decarboxylase mit Hydrazin-Partialstruktur.

Hypothese 1 konnte anhand einer sehr kleinen Verbindungskollektion verifiziert werden. Die Verwendung der Benzolsulfonamid-Partialstruktur als Leitprinzip für das Design einer fokussierten Verbindungskollektion lieferte **51** als Inhibitor. Verbindung **51** leitet sich strukturell von den Sulfonamid-Chemotherapeutika ab, deren Target die DHPS ist. Diese wurde mit der ODC in einem Proteinstruktur-Ähnlichkeitscluster zusammengefaßt. Verbindung **51** wurde im Zusammenhang mit Decarboxylase- oder ODC-Hemmung im speziellen noch nicht beschrieben. Der PSSC-Ansatz war demnach auch hier bei der

Identifizierung strukturell innovativer Inhibitoren ausgehend von einem biologisch relevanten Ausgangspunkt hilfreich. Es konnte an diesem Beispiel außerdem gezeigt werden, daß sich nicht nur Naturstoffstrukturen, sondern ganz allgemein auch andere als „privilegiert“ zu bezeichnende und bereits bekannten Wirkstoffen zugrundeliegende Strukturen als biologisch relevante Ausgangspunkte für die Konzipierung von Verbindungsbibliotheken eignen.

Hypothese 2 konnte nicht verifiziert werden. Sulfapyridin (und Sulfanilamid) hätten im therapeutisch relevanten Konzentrationsbereich 44-385  $\mu\text{M}$  ODC-inhibitorische Aktivität aufweisen müssen. Dies war nicht der Fall. Der adverse Effekt von Sulfapyridin auf die männliche Fruchtbarkeit konnte durch eine Inhibition der ODC *nicht* erklärt werden. Es bleibt noch zu prüfen, ob sich das PSSC-Konzept zur Vorhersage bzw. zur Erklärung unerwünschter pharmakologischer Wirkungen von Arzneistoffen bzw. Arzneistoffkandidaten als hilfreich erweist.

#### 4.1.4 Konzeptionelle Arbeiten zu PSSC

Die Anwendbarkeit der nachfolgend beschriebenen Proteinstruktur-Ähnlichkeitscluster für die Entwicklung von niedermolekularen Proteinliganden wurde experimentell noch nicht verifiziert. Diese Anwendungsbeispiele befinden sich noch im Konzeptstadium und können daher nur als Vorhersagen bzw. Extrapolationen betrachtet werden.

##### ***Zwei Proteinstruktur-Ähnlichkeitscluster strukturell unterschiedlicher ATP-Bindungsdomänen***

Die vorgehend beschriebenen Projekte können als valide Indizien für eine erfolgreiche Anwendung des PSSC-Konzepts im Kontext der Entwicklung von Proteinliganden erachtet werden, müssen jedoch mit weiteren Anwendungsbeispielen untermauert werden.

Im Rahmen dieser Arbeit wurden weitere Ähnlichkeitscluster identifiziert, die als Basis für eine experimentelle Validierung des PSSC-Konzepts herangezogen werden können.

Einer Analyse von Grishin zufolge, weisen die ATP-Bindungsdomänen von Proteinkinasen und *D-Ala-D-Ala-Ligase* (*DD-Ligase*) eine strukturelle Ähnlichkeit auf.<sup>[194]</sup> Die *DD-Ligase* aus *Escherichia coli*<sup>[195]</sup> (PDB-Code 2DLN, Sequenzlänge: 306 Aminosäuren) ist ein aus zwei Domänen aufgebautes Protein. Es besteht aus einer N-terminalen Domäne, die nach SCOP dem ‚PreATP-grasp domain‘-Faltungstyp zugeordnet wird, und einer C-terminalen Domäne, der ATP-bindenden Domäne, die eine ‚ATP-grasp‘-Faltung aufweist (siehe Abb. 33). Die *DD-Ligase* stellt ein attraktives enzymatisches Target dar. Mit Inhibitoren der *DD-Ligase* kann die Biosynthese des Peptidoglykans der bakteriellen Zellwand blockiert werden. In Abb. 34 wird dieses Enzym in einer vereinfachten Darstellung in den Gesamtprozeß der bakteriellen Zellwandsynthese eingeordnet.

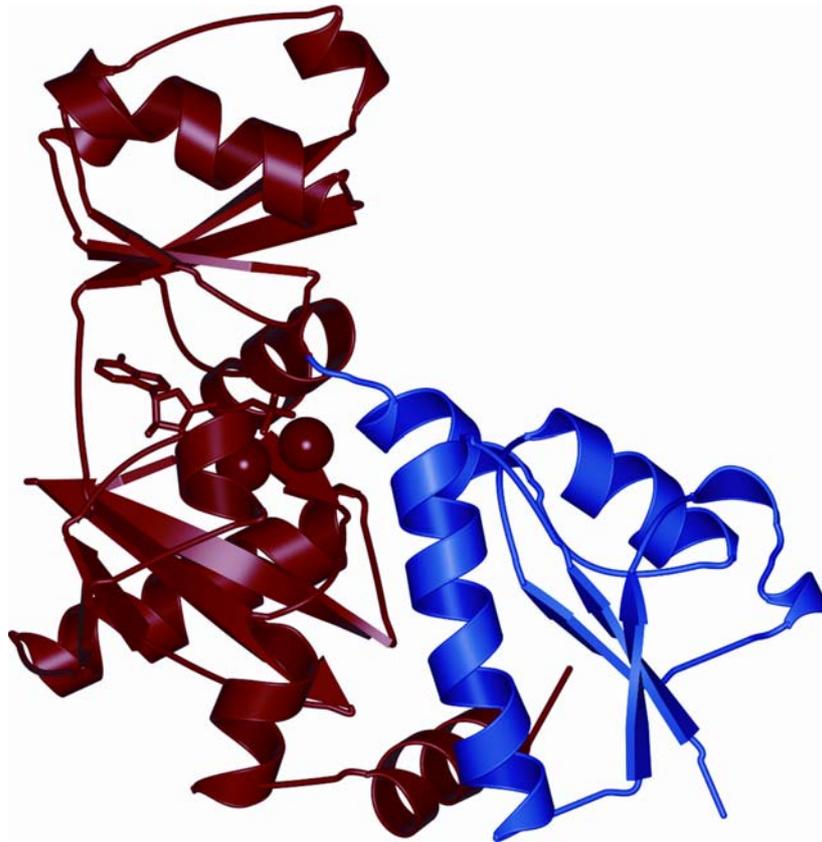


Abb. 33: Domänenarchitektur der *DD*-Ligase. Die C-terminale ATP-Bindungsdomäne mit gebundenen Magnesiumionen und ADP (K97-D306) ist rot eingefärbt (N-terminale Domäne, M1-D96, blau).

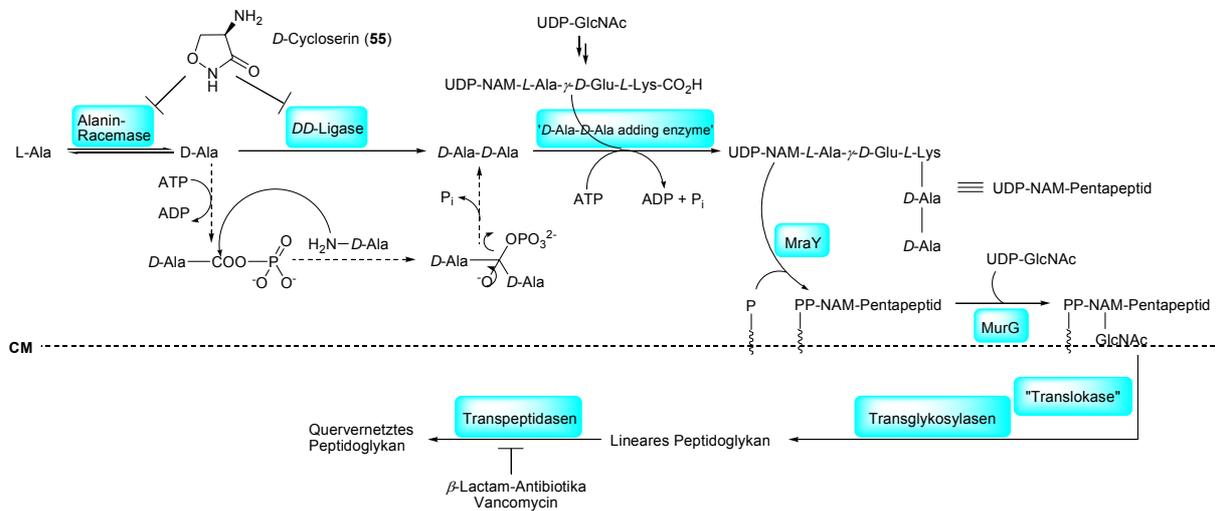


Abb. 34: Einordnung der von der *DD*-Ligase katalysierten Reaktion in den Gesamtprozess der bakteriellen Zellwandbiosynthese in Gram-positiven Bakterien (vereinfacht und ergänzt nach Ritter und Wong<sup>[196]</sup>). Die *DD*-Ligase wird durch *D*-Cycloserin (**55**) gehemmt. Das von der *DD*-Ligase generierte *D*-Ala-*D*-Ala-Dipeptid wird auf ein aus UDP-*N*-Acetylglucosamin (UDP-GlcNAc) in mehreren Schritten biosynthetisiertes UDP-*N*-Acetylmuraminsäure-(UDP-NAM)-verknüpftes Tripeptid übertragen. Das so generierte NAM-Pentapeptidphosphat wird von seinem UDP-Derivat auf Undecaprenylphosphat, ein in der cytoplasmischen Membran verankertes Trägermolekül, übertragen. Ein zweites GlcNAc wird an die 4-Position des NAM-Pentapeptids gebunden. Durch einen unbekanntem Mechanismus („Translokase“) wird das GlcNAc-NAM-Peptid durch die cytoplasmische Membran an die Zelloberfläche befördert (CM=Cytoplasmamembran). Dort werden die Disaccharideinheiten durch mehrere Transglykosylasen polymerisiert. Schließlich katalysieren Transseptidasen (Angriffspunkte von  $\beta$ -Lactamen und Vancomycin) die Vernetzung.

Die strukturelle Ähnlichkeit der ATP-Bindungsdomänen von *DD*-Ligase und Kinasen am Beispiel der lymphozytenspezifischen Kinase p56<sup>lck</sup> ist in einer Überlagerung der Kristallstrukturen dieser Proteindomänen zu sehen (siehe Abb. 35A). Diese Proteindomänen wurden deshalb, dem PSSC-Ansatz folgend, in einem Ähnlichkeitscluster zusammengefaßt, der als Ausgangspunkt für die Entwicklung neuer antibakteriell wirksamer Hemmstoffe der *DD*-Ligase ausgehend von bekannten Grundgerüsten von Kinaseinhibitoren dienen könnte. Zur Steigerung von Selektivität und Wirksamkeit könnten Bisubstrat-Inhibitoren synthetisiert werden, die aus einer ATP-kompetitiven und einer über einen Linker verknüpften *D*-Ala-(Substrat)-kompetitiven Komponente bestehen. Hier käme z.B. *D*-Cycloserin (**55**, Abb. 34 und 35B), ein in der Therapie der Tuberkulose eingesetzter Wirkstoff, in Frage.



Domänen des Hitzeschockproteins 90 (Hsp90), der DNA-Gyrase B (DNAGB), der DNA-Topoisomerase VI-B (eine Topoisomerase vom Typ IIB) und der Histidinkinase subsumiert (siehe Abb. 36). Aufgrund der Tatsache, daß diese Proteine analog zu den Kinasen ATP binden, könnte es auch hier aussichtsreich sein, Strukturen von natürlichen oder synthetischen ATP-kompetitiven Kinaseinhibitoren als Ausgangspunkte für die Synthese von Verbindungsbibliotheken zu nutzen, um diese Proteine zu adressieren. Umgekehrt könnten die Strukturen von Hemmstoffen dieser Proteine als Ausgangspunkte für die Entwicklung von Kinaseinhibitoren dienen. Hsp90 hat sich bereits in klinischen Studien als vielversprechendes Zielprotein in der Krebstherapie erwiesen. Derzeit befindet sich ein Derivat des Ansamycinantibiotikums Geldanamycin (ATP-kompetitiver Hsp90-Hemmstoff) in der klinischen Prüfung (17-Allylaminogeldanamycin, **57**, Abb. 36B).<sup>[198]</sup> Die Topoisomerase II ist das Zielprotein von einigen der aktivsten Antitumor-Wirkstoffe, die derzeit klinisch verwendet werden, wie z.B. Etoposid, Teniposid, Doxorubicin, Mitoxantron und Amsacrin.<sup>[199]</sup> Die bakterielle DNA-Gyrase ist ein in der antibakteriellen Chemotherapie genutztes Zielprotein und ist der Angriffspunkt der klinisch verwendeten Chinolone (z.B. Sparfloxacin). Abgesehen von den Chinolonen hemmen Cumarin-Naturstoffe, zu denen Novobiocin (**58**, Abb. 36B), Chlorobiocin (**59**) und Coumermycin A<sub>1</sub> (**60**) gehören, als ATP-kompetitive Inhibitoren die DNA-Gyrase.<sup>[200]</sup> Auch Histidinkinasen könnten potentielle Zielproteine für die antibakterielle Chemotherapie sein.<sup>[201]</sup>

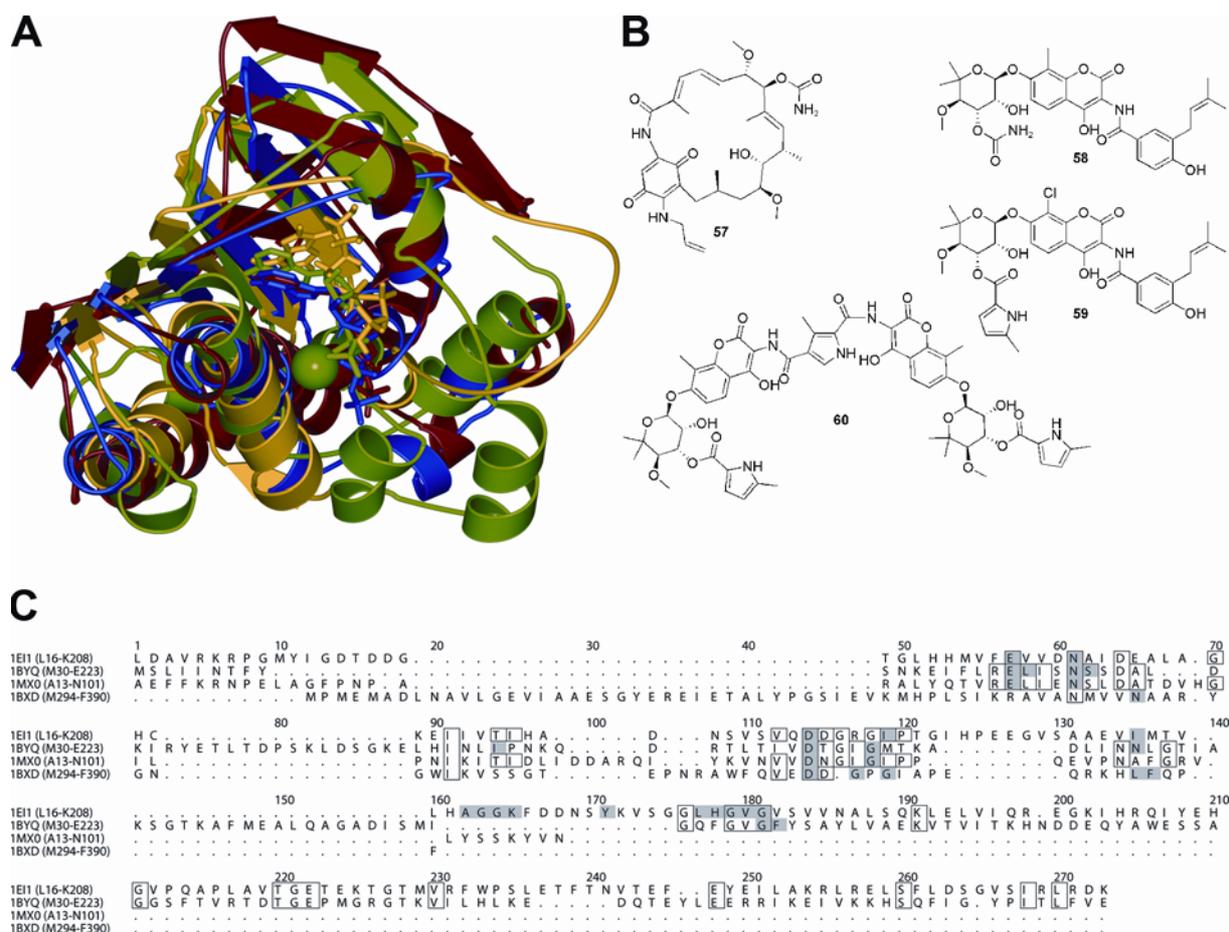


Abb. 36: A) Überlagerung der 3D-Strukturen der ATPase-Domänen der DNAGB aus *E. coli*<sup>[202]</sup> (PDB-Code 1E11, blau, mit gebundenem ANP, einem nicht-hydrolysierbaren ATP-Analogen), des humanen Hsp90<sup>[203]</sup> (PDB-Code 1BYQ, grün, mit gebundenem ADP und Mg<sup>2+</sup>), der Topoisomerase VI-B aus *Sulfolobus shibatae*<sup>[204]</sup> (PDB-Code 1MX0, rot, mit gebundenem ANP und Mg<sup>2+</sup>) und der Histidinkinase aus *E. coli*<sup>[205]</sup> (PDB-Code 1BXD, gelb, mit gebundenem ANP). Die RMSD-Werte betragen 3.37 Å (Alignmentlänge: 153 Aminosäuren, Sequenzidentität: 15.7%) für 1E11/1BYQ, 1.73 Å (Alignmentlänge: 82 Aminosäuren, Sequenzidentität: 22.0%) für 1E11/MX0 und 3.89 Å (Alignmentlänge: 105 Aminosäuren, Sequenzidentität: 12.5%) für 1E11/1BXD. B) ATP-kompetitive Naturstoff(derivat)inhibitoren, die das Hsp90 bzw. die DNA-Gyrase hemmen. C) Strukturabgeleitetes Sequenzalignment der ATPase-Domänen der DNAGB (1E11), des Hsp90 (1BYQ), der Topoisomerase VI-B (1MX0) und der Histidinkinase (1BXD). Die für die Liganden- bzw. Magnesiumbindung relevanten Aminosäuren sind jeweils grau unterlegt. Es können nur wenige sequentielle und räumliche Korrelationen festgestellt werden.

### Erweiterung des in Kapitel 4.1.2 beschriebenen Strukturähnlichkeitsclusters

Wie in Kapitel 4.1.2 bei der Diskussion der Trefferliste für die Dali/FSSP-Recherche mit dem Suchmotiv der 3D-Struktur der AChE bereits angedeutet, kann der Proteinstruktur-Ähnlichkeitscluster, der – wie bislang experimentell verifiziert – die Cdc25A, die AChE sowie die 11 $\beta$ -Hydroxysteroid-Dehydrogenasen umfaßt, um weitere Mitglieder der Superfamilie der  $\alpha/\beta$ -Hydrolasen erweitert werden. Hier kommen als pharmazeutisch relevante Enzyme insbesondere Lipasen (speziell die Pankreaslipase) und die humane Acylprotein-Thioesterase 1 (*hAPT1*) in Frage. Die Domänenarchitektur der Pankreaslipase ist

in Abb. 37A dargestellt. Die katalytischen Kernbereiche der Pankreaslipase (Abb. 37A) und der *h*APT1 (Abb. 37B) wurden unter Verwendung der Struktur der katalytischen Domäne der Cdc25A als Templat definiert und mit der katalytischen Domäne der Cdc25A überlagert (siehe Abb. 38A).

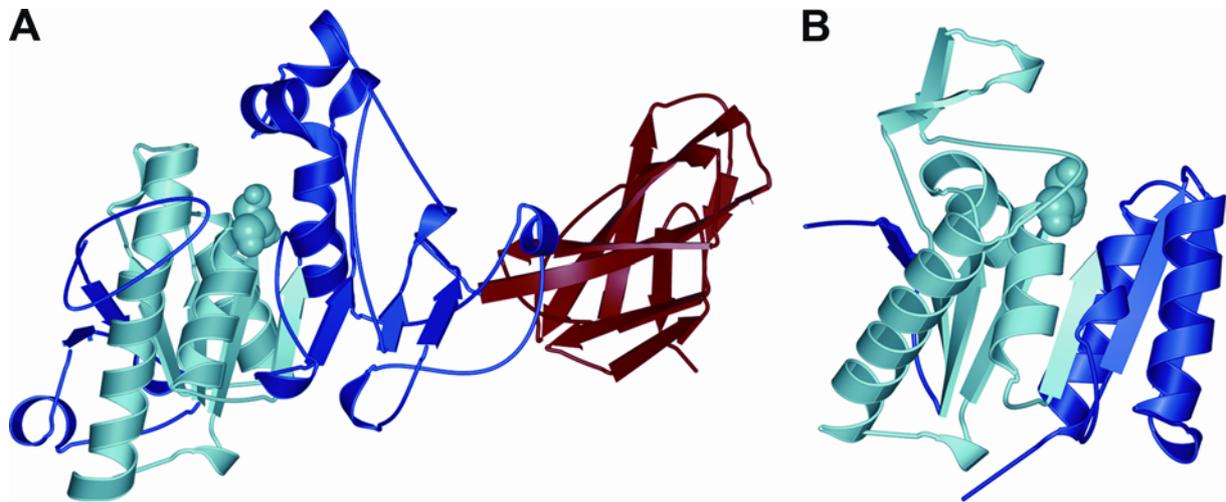


Abb. 37: A) Domänenarchitektur und katalytischer Kernbereich der Schweinepankreaslipase<sup>[206]</sup> (PDB-Code 1ETH). Sie besteht aus einer N-terminalen katalytischen (S1-F336, blau und cyan) und einer C-terminalen Domäne (A337-C448, rot), die an der Colipase-Bindung beteiligt ist. Innerhalb der katalytischen Domäne kann ein katalytischer Kernbereich (cyan, R69-A179) um das katalytische Serin (S153, in CPK-Darstellung) definiert werden. B) Die *h*APT1<sup>[207]</sup> (PDB-Code 1FJ2, Sequenzlänge: 229 Aminosäuren, blau und cyan) ist ein Eindomänenprotein. Der katalytische Kernbereich (T16-S138) mit S114 (in CPK-Darstellung) als katalytischer Schlüsselaminosäure ist cyan eingefärbt.



Krebstherapie genutzt werden. Die Kristallstruktur der FAS-Thioesterasedomäne zeigt, daß diese Domäne aus zwei Untereinheiten besteht, und die Untereinheit A  $\alpha/\beta$ -Hydrolase-Faltung zeigt (siehe Abb. 39). Eine Analyse des katalytischen Kernbereichs zeigt, daß auch die FAS dem oben beschriebenen Proteinstruktur-Ähnlichkeitscluster zugeordnet werden kann.

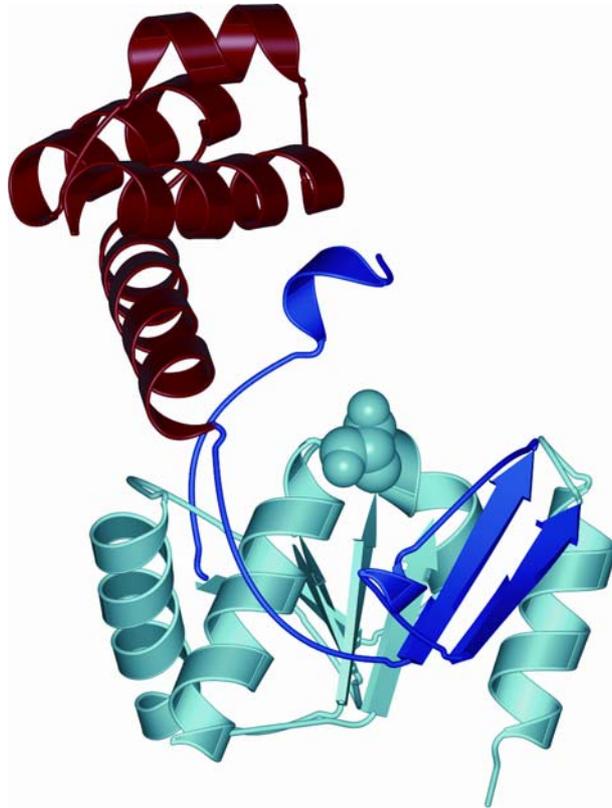


Abb. 39: Kristallstruktur der Thioesterasedomäne (Untereinheit A cyan und blau, Untereinheit B rot) der FAS<sup>[209]</sup> (PDB-Code 1XKT, Sequenzlänge: 260 Aminosäuren). Die größere Untereinheit A (cyan und blau) zeigt  $\alpha/\beta$ -Hydrolase-Faltung. Der katalytische Kernbereich (V2217-Y2343, G2489-A2502) mit der katalytischen Schlüsselaminosäure S2308 (in CPK-Darstellung) ist cyan eingefärbt.

Diese Befunde legen nahe, z.B. das Naturstoffderivat Orlistat (**61**) und strukturell verwandte Naturstoffe, die Ebelactone (**62** und **63**), die als Liganden von zwei Enzymen mit  $\alpha/\beta$ -Hydrolase-Faltung bekannt sind, als strukturelle Leitprinzipien für den Aufbau von Verbindungsbibliotheken zu nutzen. Diese Verbindungsbibliotheken könnten auch Liganden für die übrigen, strukturell verwandten Enzyme des Proteinstruktur-Ähnlichkeitsclusters liefern.

## 4.2 Strukturelle Klassifizierung der Naturstoffe (SCONP)

Historisch haben sich Naturstoffe als wertvolle Quelle für Arzneimittel erwiesen. Die bisherige industrielle Erforschung von Naturstoffen erlebte ihren Höhepunkt von 1970 bis 1980 und mündete in die Entwicklung zahlreicher Naturstoff- und Naturstoffabgeleiteter Wirkstoffe.<sup>[210]</sup>

Trotz dieser Erfolge nahm die Bedeutung der Naturstoffforschung für die Wirkstoffentwicklung innerhalb der beiden letzten Jahrzehnte ab. Die Gründe dafür sind vielfältig. Die Einführung von HTS gegen definierte pharmakologische Targets ist hier als wesentliche Ursache zu nennen. Diese Technologie favorisierte den Einsatz von definierten synthetisch-chemischen anstelle von Naturstoffextrakt-Bibliotheken. Die Synthese der hierfür benötigten großen Verbindungsbibliotheken wurde durch die rasante Entwicklung der Kombinatorischen Chemie im gleichen Zeitraum ermöglicht. Hinzu kam die wachsende Vielfalt bekannter molekularer Targets aufgrund von Fortschritten in der Molekular- und Zellbiologie und der Genomik. Diese Entwicklungen nährten die v.a. kommerziellen Erwartungen, daß durch ein schnelles Screening großer, gar randomisierter Verbindungskollektionen die Hitraten erhöht werden könnten, und die Hits schnell zu neuen Leitstrukturen für potentielle Wirkstoffkandidaten weiterentwickelt werden könnten. Das Ergebnis dieser Forschungsstrategie blieb jedoch deutlich hinter diesen hochgesteckten Erwartungen zurück.

Eine logische Konsequenz, die sich vor dem Hintergrund dieser Erfahrungen aufdrängt, könnte es sein, wieder auf die altbewährten Naturstoffe zurückzugreifen, nachdem es sich bei ihnen um biologisch relevante Ausgangspunkte im chemischen Strukturraum handelt, die als Inspirationsquelle für eine zielgerichtete Entwicklung von Verbindungskollektionen dienen könnten. Die Beobachtung, daß die HTS-Hitraten steigen, wenn Naturstoffe in ein Screening miteinbezogen werden, könnte hierfür als zusätzliches Indiz dienen.<sup>[11-13, 41]</sup>

Naturstoffe stellen einen Pool meist komplexer und strukturell diverser Verbindungen dar. Im Hinblick auf ein besseres Verständnis des von den Naturstoffen eingenommenen Strukturraums, wäre eine strukturbasierte Kartierung oder Kategorisierung von Naturstoffen von erheblichem Wert. Diese könnte als Navigator dienen, mit dessen Hilfe strukturelle Zusammenhänge zwischen den Naturstoffen hergestellt werden könnten.



### *Ulocladol – ein natürlich vorkommender Kinasehemmer mit atypischem Grundgerüst*

Der Naturstoff Ulocladol (**67**, Abb. 41, 4,10,11-Trihydroxy-2,9-dimethoxy-7*H*-dibenzo[*c,e*]oxepin-5-on) ist ein polyketider Sekundärmetabolit, der aus einem Extrakt des marinen Pilzes *Ulocladium botrytis* isoliert wurde. Ulocladol wurde als mikromolarer Hemmstoff der lymphozytenspezifischen Proteintyrosinkinase p56<sup>lck</sup> identifiziert, die zur Familie der *src*-Kinasen gehört. Von der Struktur seines Grundgerüsts ausgehend, ist Ulocladol als atypischer Proteinkinaseinhibitor zu klassifizieren, der jedoch das als „privilegiert“<sup>[14]</sup> geltende Biarylstrukturmotiv aufweist.

### *Synthese einer Ulocladol-inspirierten Verbindungskollektion*

Auf der Basis des Ulocladol-Grundgerüsts wurde eine kleine Verbindungskollektion (Abb. 41) synthetisiert. Alle Moleküle bis auf eines leiten sich synthetisch vom Naturstoff Ellagsäure (**68**) ab, der seinerseits ein potenter, jedoch nichtselektiver Inhibitor der tyrosinspezifischen Proteinkinase pp60<sup>src</sup> ist.<sup>[214]</sup> Ein fluoriertes Dibenzoxepinon (**69**) wurde auf anderem Wege dargestellt.

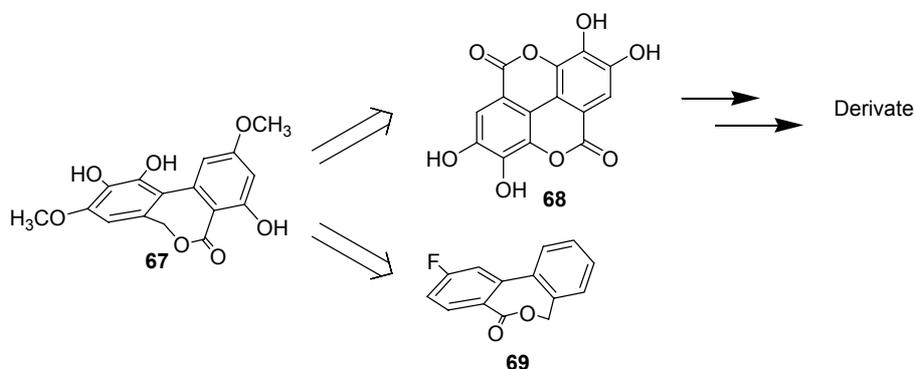


Abb. 41: Ulocladol-basierte Verbindungskollektion.

### *Synthese der Ellagsäure-Derivate*

Die Ellagsäure-Derivate wurden der Route von Kashiwada *et al.*<sup>[215]</sup> folgend synthetisiert (Abb. 42).

Ellagsäure (**68**) wurde perbenzyliert und das Reaktionsprodukt **70** mit LAH zu **71** reduziert. Daraufhin teilte sich die Syntheseroute. Über Route 1 wurden die phenolischen Hydroxylfunktionen der Verbindung **71** benzyliert und das Produkt **72** säurekatalysiert zum Dibenzoxepin **73** zyklisiert. Verbindung **73** wurde vollständig debenzyliert. Das

Reaktionsprodukt **74** wurde zu **75** permethyliert. Über Route 2 wurde **71** cyclisiert. Das Reaktionsprodukt **76** wurde zu **77** methyliert.

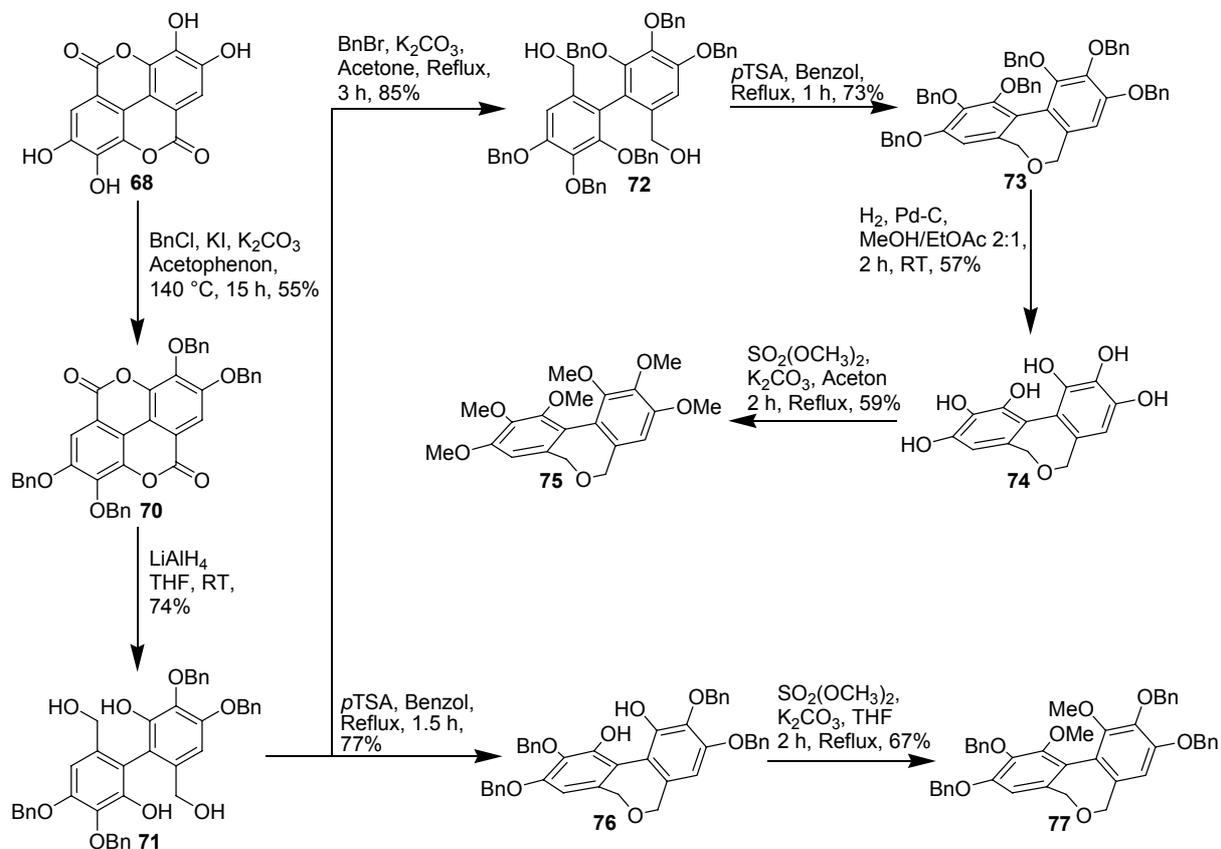


Abb. 42: Vom Naturstoff Ellagsäure ausgehend synthetisierte Ulocladol-Analoga.

#### Synthese des Fluor-substituierten Dibenzoxepinons

2-Brom-4-fluorbenzoesäure (**78**, Abb. 43) wurde zum Methylester **79** umgesetzt, der anschließend Pd-katalysiert zu **80** boryliert wurde. Das Boran **80** wurde anschließend mit der silylgeschützten Verbindung **82**, die aus dem entsprechenden Benzylalkohol **81** nach Standardverfahren synthetisiert wurde, Pd-katalysiert gekuppelt. Das Kupplungsprodukt **83** wurde danach zum Lacton **69** durch Abspaltung der Silyl-Schutzgruppe cyclisiert.

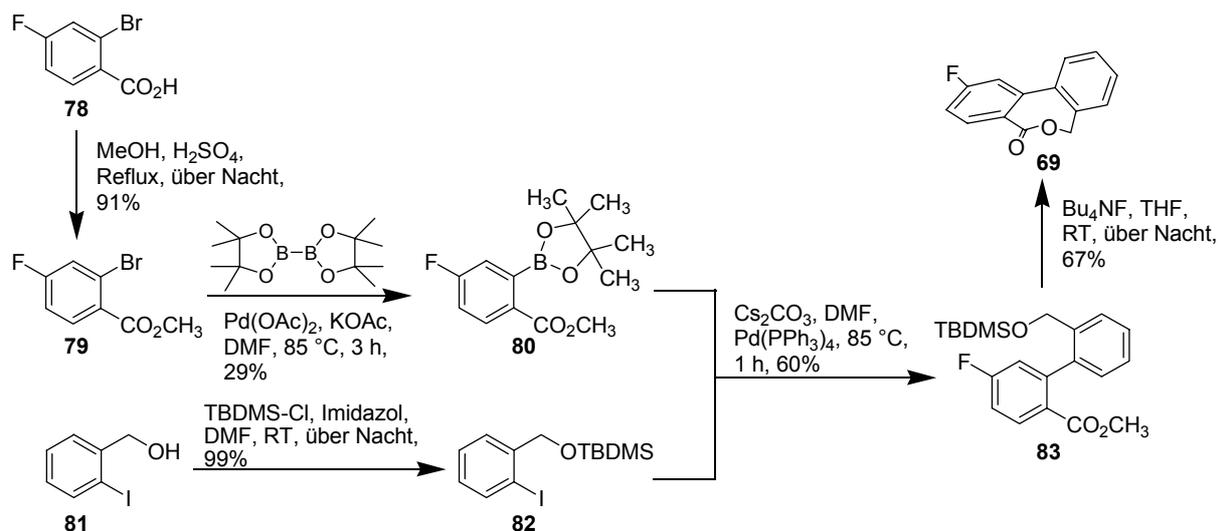


Abb. 43: Darstellung des fluorierten Dibenzoexepinons **69**.

### ***Profilierung der von Ulocladol abgeleiteten Verbindungskollektion gegen Kinasen***

Die Verbindungen wurden im Rahmen einer Kooperation mit der Firma Aventis, Frankfurt, gegen eine Gruppe von 14 verschiedenen Kinasen, die im folgenden kurz vorgestellt werden, gescreent.

Die Serin/Threonin-Kinase Akt (PKB) phosphoryliert Substrate, die das Überleben der Zelle und ihr Wachstum fördern und pro-apoptische Signale blockieren. Zahlreiche Studien haben gezeigt, daß eine Dysregulation von Akt wesentlich zur Entstehung von Tumoren beiträgt. Akt stellt daher ein potentielles Target in der Krebstherapie dar.<sup>[216]</sup>

STK15/Aurora2 ist eine Zentrosom-assoziierte Serin/Threonin-Kinase. Ihre Konzentration und Kinaseaktivität steigen zwischen der G2- und Mitose-Phase des Zellzyklus. Eine Überexpression von STK15 induziert Tumorgenese und kommt verstärkt in verschiedenen Krebsarten und Tumorzelllinien vor. STK15 stellt daher ein interessantes Target für Inhibitoren dar, die ihre Aktivität und damit die Zellproliferation hemmen.<sup>[217]</sup>

Die  $\beta$ -adrenerge-Rezeptorkinase Typ 1 (BARK1) hat eine regulatorische Funktion im Rahmen der Signaltransduktion über  $\beta$ -adrenerge Rezeptoren. Erhöhte Konzentration und Aktivität von BARK1 treten bei der Herzinsuffizienz auf.<sup>[218]</sup>

Die Familie der Cyclin-abhängigen Kinasen („cyclin-dependent kinases“, CDKs) spielt eine wichtige Rolle bei der Kontrolle des Zellzyklus. Sie bilden Assoziate mit spezifischen regulatorischen Untereinheiten (den Cyclinen). Von den insgesamt 9 CDKs wurden bereits 5

identifiziert, die im Zellzyklus von Säugern eine Rolle spielen: CDK4, CDK6 und CDK2 sind während der G1-, CDK2 während der S- und CDK1 während der G2- und M-Phase aktiv, wohingegen CDK7 in allen Phasen des Zellzyklus aktiv ist. Cycline assoziieren mit jeweils einer oder zwei CDKs und die meisten CDKs interagieren mit einem oder zwei Cyclinen. Im Zellzyklus ist Cyclin D in der G1-, die Cycline E und A in der S- und die Cycline B und A in der M-Phase an der Regulation beteiligt. Da Tumorzellen dysregulierte Zellzyklen aufweisen, sind Inhibitoren der CDKs als vielversprechende therapeutische Ansätze in der Behandlung von Krebs anzusehen.<sup>[219]</sup>

Die fokale Adhäsionskinase (FAK) ist eine Nicht-Rezeptor-Tyrosinkinase, die an den Kontaktpunkten von Zellen mit der extrazellulären Matrix (sogenannte fokale Adhäsionen) lokalisiert ist. FAK wird im Zusammenhang mit verschiedenen zellulären Prozessen wie Invasion, Motilität, Proliferation und Apoptose diskutiert. Außerdem wurde gezeigt, daß die FAK-Expression in verschiedenen menschlichen Krebsarten erhöht ist.<sup>[220]</sup>

Glycogensynthase-Kinase-3 (GSK3) spielt eine Rolle in der multifaktoriellen Ätiologie der muskulären Insulinresistenz bei Typ 2-Diabetikern. Die möglichen involvierten molekularen Mechanismen sind noch nicht ganz eruiert. GSK3-Inhibition führt zu verstärkter Insulinsensitivität.<sup>[221]</sup>

IGF1R (insulin-like growth factor receptor type 1) ist eine Rezeptor-Tyrosinkinase (RTK) und ist in vielen Tumoren überexprimiert. IGF1R vermittelt Proliferation, Motilität und Schutz vor Apoptose. Tumorwachstum und -metastasierung können durch IGF1R-Inhibitoren gehemmt werden. IGF1R ist daher ein vielversprechendes Target bei der Behandlung von Krebs.<sup>[222]</sup>

JNKs (c-Jun N-terminal kinases) sind wichtige Enzyme in Rahmen der Zellfunktion. JNK3 findet sich vor allem in ZNS-Neuronen, und wird mit verschiedenen neurodegenerativen Krankheiten wie *Morbus Alzheimer*, *Morbus Parkinson* und Schlaganfall in Verbindung gebracht. Vor allem spielt JNK3 eine übergeordnete Rolle bei der neuronalen ischämischen Apoptose. JNK3 ist stark exprimiert und aktiviert in *post-mortem*-Gehirnen von Alzheimer-Patienten.<sup>[223]</sup>

Die p38-MAP-(mitogen activated protein)-Kinasen stellen eine Familie von Serin/Threonin-Proteinkinasen dar, die eine wichtige Rolle bei zellulären Antworten auf externen Streß und Signale spielen. Sie sind in die Produktion einer Vielzahl von proinflammatorischen Cytokinen durch verschiedene Zelltypen involviert. Inhibitoren der p38-Familie zeigten eine

antientzündliche Wirkung in präklinischen Krankheitsmodellen, v.a. durch die Hemmung der Bildung von Entzündungsmediatoren.<sup>[224, 225]</sup>

Die p21-aktivierten Serin/Threoninkinasen (PAKs) nehmen eine wichtige Rolle in einer Vielzahl zellulärer Funktionen wie Zellmorphogenese, Motilität, Überleben, Angiogenese und Mitose ein. PAKs sind in viele zelluläre Signalwege eingebunden.<sup>[226, 227]</sup>

PLKs (Polo-like kinases) spielen eine entscheidende Rolle bei der Regulation des Zentrosomenzyklus und damit auch des Zellzyklus. PLK1 ist das am besten charakterisierte Mitglied der Familie der Säugetier-PLKs. PLK1 fördert stark die Progression durch den Zellzyklus und ist in einer Reihe menschlicher Tumore überexprimiert. Die PAK1-Expression korreliert mit zellulärer Proliferation und der Prognose von Tumorpatienten.<sup>[228, 229]</sup>

Proteinkinasen aus der *src*-Familie sind tyrosinspezifisch. *Src* selbst ist ein Protoonkogen retroviralen Ursprungs (*v-src*, virales *src*). *C-src* („cellular *src*“) stellt ein interessantes Target bei der Behandlung von Krebs und Leukämien dar.<sup>[230, 231]</sup> Zu dieser Kinasefamilie gehört auch die lymphozytenspezifischen Proteintyrosinkinase p56<sup>lck</sup>.

Tie2 ist eine endotheliale zell-spezifische RTK, die fast ausschließlich auf der Oberfläche des vaskulären Endothels exprimiert wird. Tie2 spielt eine wichtige Rolle bei der Tumorangiogenese. Tie2-Inhibitoren könnten daher therapeutisches Potential bei der Behandlung solider Tumore besitzen.<sup>[232]</sup>

### ***Ergebnis und Diskussion***

Bei der Profilierung erwiesen sich lediglich zwei Verbindungen als aktiv (eine komplette Auflistung der Strukturen und Profilierungsergebnisse findet sich in Anhang IV). Der Naturstoff Ellagsäure (**59**) und das von ihm partialsynthetisch abgeleitete sechsfach hydroxylierte Dibenzoxepin (**65**) zeigten eine selektive Hemmung bestimmter Kinasen unterhalb einer Ausschlußkonzentration von 30  $\mu\text{M}$ . Die übrigen Verbindungen zeigten keine inhibitorische Wirksamkeit bei einer Konzentration  $< 30 \mu\text{M}$ .

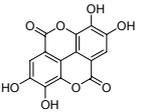
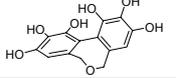
Vbdg	Struktur	AKT (ATP=80µM)	AURORA2 (ATP=2µM)	bARK1 (ATP=6µM)	CyclinD1/CDK4 (ATP=230µM)	CyclinE/CDK2 (ATP=40µM)	FAK (ATP=1µM)	GSK3b (ATP=16µM)	IGF1R (ATP=40µM)	JNK3 (ATP=6µM)	P38 (ATP=250µM)	PAK3 (ATP=1.7µM)	PLK1 (ATP=40µM)	SRC (ATP=40µM)	TIE2 (ATP=20µM)
		IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)	IC50 (µM)
68		1.93	8.1	3.08	>30	>30	>30	>30	1.82	>30	>30	0.36	1.49	>30	10.2
74		>30	1.07	n.b.	>30	>30	6.6	>30	1.38	>30	1.3	1.04	5.88	3.7	0.55

Tabelle 2: Aktive Verbindungen der von Ulocladol inspirierten Verbindungskollektion.

Ellagsäure (**68**) und Verbindung **74** zeichnen sich im Vergleich zu den übrigen Verbindungen der Kollektion durch mindestens zwei am aromatischen Ring benachbarte Hydroxylfunktionen aus. Blockiert man Hydroxylfunktionen durch Methyl- bzw. Benzylgruppen oder ersetzt man die Hydroxylfunktionen durch ein Fluoratom, so ist unterhalb von 30 µM keine inhibitorische Aktivität mehr festzustellen. Interessanterweise zeigte die planar gebaute Ellagsäure (**68**) bei einer Konzentration von 30 µM und einer ATP-Konzentration von 40 µM keine Hemmung der *src*-Kinase, zu deren Familie auch die von Ulocladol (**67**) inhibierte p56<sup>lck</sup> gehört, wohingegen die perpendicular aufgebaute, Ulocladol-ähnliche Verbindung **74** einen IC<sub>50</sub> von 3.7 µM aufwies (siehe Tabelle 2). Ein perpendicular aufgebautes, verbrücktes Biaryl-Grundgerüst scheint demnach eine wichtige Determinante für *src*-inhibitorische Wirksamkeit zu sein. Weiterhin fällt auf, daß **74** gegenüber **68** die Tie2-Rezeptortyrosinkinase um einen Faktor 19 stärker hemmt.

Im Rahmen von Molecular-Modeling-Untersuchungen unter Verwendung des Docking-Programms GOLD<sup>[233]</sup> ([http://www.ccdc.cam.ac.uk/products/life\\_sciences/gold/](http://www.ccdc.cam.ac.uk/products/life_sciences/gold/)) wurde ein wahrscheinlicher Bindungsmodus von Ulocladol (**67**) an die lymphozytenspezifische Kinase p56<sup>lck</sup> ermittelt. Die Docking-Methode wurde zunächst dadurch validiert, daß der mit der p56<sup>lck</sup> (Kristallstruktur mit dem PDB-Code 1QPE) kokristallisierte *src*-selektive und potente Inhibitor PP2 (1-*t*-Butyl-3-(4-chlorphenyl)-1*H*-pyrazolo[3,4-*d*]pyrimidin-4-ylamin, **84**, Abb. 44) gedockt wurde. Das Dockingergebnis war identisch zur nativen Position und Konformation des kokristallisierten Liganden PP2 (**84**). Ein flexibles Docking des Ulocladols (**67**) führte zur Hypothese eines wahrscheinlichen Bindungsmodus des Ulocladols an die p56<sup>lck</sup> (siehe Abb. 44). Die beiden *ortho*-ständigen Hydroxylgruppen in Position 10 und 11 im Ulocladol gehen ähnliche Wasserstoffbrückenbindungen wie N5 und die 4-Aminogruppe im PP2 ein. PP2 bildet Wasserstoffbrücken zum Rückgrat-NH des Met<sup>319</sup> respektive zum Rückgrat-Carbonyl des Glu<sup>317</sup> und zur Seitenketten-Hydroxylgruppe des Thr<sup>316</sup> aus. Darüber hinaus wären beim Ulocladol (**67**) Wasserstoffbrückenbindungen zwischen der 4-OH-Gruppe

und den Säurefunktionalitäten von Glu<sup>288</sup> und Asp<sup>382</sup> sowie zwischen der 5-Carbonylfunktion und dem Rückgrat-NH von Asp<sup>382</sup> denkbar. Eine prospektive Studie mit unterschiedlich substituierten Ulocladolderivaten ergab Substitutionsmöglichkeiten in den Positionen 2 (rot in Abb. 44, hier v.a. kleine hydrophobe Reste) und 9 (grün in Abb. 44, hier ist ein größerer Spielraum für mögliche Substituenten gegeben), mit denen die Affinität zur Bindungstasche ggf. gesteigert werden könnte.

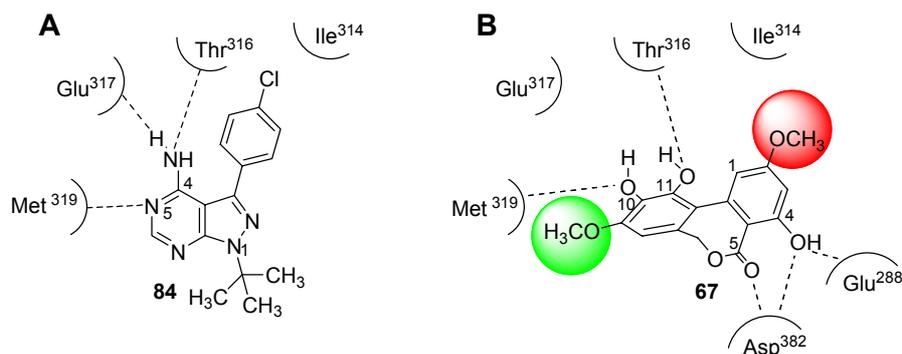


Abb. 44: Bindungsmodus von PP2 (**84**) und postulierter Bindungsmodus von Ulocladol (**67**) und Derivaten.

Insgesamt konnte aufgezeigt werden, daß auch in einer sehr kleinen, auf einem biologisch prävalidierten Naturstoff basierenden Verbindungskollektion Verbindungen, die sowohl das Ausgangsenzym als auch andere Proteine einer Target-Familie zu hemmen vermögen, identifiziert werden können.

## 4.2.2 SCONP – Ein Hypothesengenerator und Navigator durch den Strukturraum der Naturstoffe

Die effiziente Identifizierung kleiner organischer Moleküle, die die Eigenschaft besitzen, *in vitro* und *in vivo* die Funktion von Proteinen gezielt zu modulieren, ist von zentraler Bedeutung für die Forschung im Bereich der Chemischen Biologie und der Medizinischen Chemie und die Entwicklung neuer arzneistoffbasierter Therapien.

Als Schlüssel zu ihrer Entdeckung könnte eine Kartierung des biologisch relevanten Strukturraums dienen, d.h. der Bereiche des gesamten chemischen Strukturraums, die für die Biologie relevant sind.<sup>[11-13, 41, 234]</sup> Naturstoffe können als evolutionär selektierte Liganden für die ligandenbindenden Kernbereiche von Proteinen betrachtet werden. Sie werden durch Proteine biosynthetisiert und üben häufig verschiedene biologische Effekte durch Wechselwirkungen mit verschiedenen Proteinen aus.<sup>[41, 235]</sup> Die ihnen zugrundeliegenden Strukturen definieren strukturelle Bedingungen für die Bindung an Proteine und damit für biologische Aktivität.

Der gesamte biologisch relevante chemische Raum mag zwar größer sein als der von Naturstoffen abgedeckte Strukturraum, jedoch repräsentieren Naturstoffgrundgerüste die prävalidierten Bereiche des Strukturraums, die die Natur im Laufe der Evolution erkundet und genutzt hat. Folglich haben Verbindungsbibliotheken, die konzipiert wurden, um Strukturen und Eigenschaften von Naturstoffen nachzuahmen, eine größere biologische Relevanz als Verbindungsbibliotheken, die allein nach dem Gesichtspunkt der chemischen Machbarkeit synthetisiert wurden.<sup>[11, 13, 43, 234]</sup> Es ist daher zu erwarten, daß eine „Naturstoff-geleitete Entwicklung von Verbindungsbibliotheken“<sup>[11, 13, 41]</sup> eine wichtige Strategie für die Entwicklung biologisch wirksamer Moleküle bleiben wird.<sup>[235]</sup>

Für die Weiterentwicklung dieses Ansatzes wäre ein systematisches, strukturbasiertes Ordnungsprinzip der bekannten Naturstoffe mit Annotationen des biologischen Ursprungs und der biologischen Wirkung ein Werkzeug von unschätzbarem Wert. Es würde nicht nur die Regionen des von der Natur explorierten chemischen Raums kartieren und damit eine strukturelle Rationalisierung und Kategorisierung der Diversität der Naturstoffe erlauben. Ein derartiges Ordnungsprinzip würde auch als eine Art Navigator bei der Entwicklung naturstoffartiger Verbindungsbibliotheken im Hinblick auf strukturelle Vereinfachungen bei gleichzeitigem Erhalt der gewünschten pharmakologischen Aktivität fungieren.

Statistische Analysen verschiedener Naturstoff-Datenbanken wurden bereits in einigen wenigen Fällen vorgenommen.<sup>[10, 42, 43]</sup> Diese waren im wesentlichen darauf ausgerichtet, Unterschiede zwischen den Eigenschaften der Natur- und der Arzneistoffe zu analysieren. Obwohl wichtige Erkenntnisse im Rahmen dieser Studien gewonnen wurden, wurde von einer systematischen und annotierten strukturellen Kategorisierung von Naturstoffen noch nicht berichtet.

Im folgenden wird die Entwicklung einer strukturellen Klassifizierung von Naturstoffen (‘Structural Classification Of Natural Products’ – SCONP) als Ordnungsprinzip vorgestellt, das strukturelle Beziehungen zwischen verschiedenen Naturstoffklassen definiert. Die Anwendbarkeit von SCONP im Rahmen der Entwicklung von Verbindungsbibliotheken wurde experimentell verifiziert. Gleichzeitig wurde die synergistische Anwendung von SCONP zusammen mit Proteinstruktur-Ähnlichkeitsclustering (PSSC)<sup>[41]</sup> etabliert. Die nachfolgend geschilderten chemoinformatischen Analysen wurden in Zusammenarbeit mit Dr. Peter Ertl und Dr. Ansgar Schuffenhauer, beide Novartis Institutes of Biomedical Research, Basel, durchgeführt.

### ***Basisdatensatz***

Die wohl umfangreichste, derzeit verfügbare Naturstoff-Datenbank, das ‘Dictionary of Natural Products’ (DNP, Chapman & Hall/CRC, Taylor and Francis Books, v12.2.2004), die 161 278 Naturstoff-Datensätze enthält, diente als Grundlage für die chemoinformatische Analyse der Naturstoffstrukturen. Die nachfolgend geschilderten chemoinformatischen Analysen sind auch auf andere Datensätze übertragbar.

### ***Primärprozessierung des Basisdatensatzes***

Die Strukturdaten wurden zunächst vom MDL SDF-Molekülformat in Daylight SMILES-Strings (Simplified Molecular Input Line Entry Specification)<sup>[236]</sup> konvertiert, wobei fehlerhafte Einträge eliminiert wurden. SMILES ist eine flexible Linearnotation, die es ermöglicht, chemische Strukturen rechnerlesbar darzustellen. Es erfolgte eine weitere Standardisierung, wobei Ladungen normalisiert sowie Gegenionen und mit den Naturstoffen assoziierte kleine Moleküle, wie z.B. Wasser oder Salze, entfernt wurden. Die primäre chemoinformatische Bearbeitung des Basisdatensatzes wurde mit Programmen der Firma

Molinspiration (Molinspiration Cheminformatics, Bratislava, Slowakische Republik, <http://www.molinspiration.com>) durchgeführt.

Die Stereochemie wurde im Laufe dieser Analyse nicht berücksichtigt, zumal die Strukturdateien des DNP keine Stereoinformation enthalten. Die Fokussierung auf zweidimensionale Strukturen erlaubt jedoch eine Abstraktion der Strukturinformation, wie sie für ein nützliches strukturbasiertes Ordnungsprinzip für Naturstoffe, das von Medizinalchemikern intuitiv verwendet werden kann, erforderlich ist. Im übrigen wurde in früheren Studien gezeigt, daß chemoinformatische Analysen unter Verwendung von 3D-Moleküldekriptoren (noch) keine besseren Ergebnisse liefern als Analysen auf der Grundlage zweidimensionaler Strukturen.<sup>[237-239]</sup>

Duplikate wurden nicht aus dem Datensatz eliminiert, da es sich hierbei um Stereoisomere handeln könnte. Darüber hinaus kann ein- und dieselbe Struktur unterschiedlich annotiert sein (z.B. in bezug auf ihren biologischen Ursprung oder ihre Funktion).

Der nach der Primärprozessierung erhaltene Datensatz enthielt 161 110 Naturstoffstrukturen. 141 838 Naturstoffe enthielten Ringe (88%). Da die überwältigende Mehrzahl der in der medizinisch- und biologisch-chemischen Forschung genutzten Verbindungen Ringe enthalten,<sup>[240]</sup> konzentrierte sich die weitere Analyse auf die Naturstoffe mit Ringstrukturelementen. Die Analyse der acyclischen Naturstoffe ergab, daß diese hauptsächlich aus Aminosäuren, Lipiden, Glykosiden mit acyclischen Aglyka und kleinen Molekülen bis zur Größe der Glucose bestehen (siehe Abb. 45). Es handelt sich um Substanzen, die kaum für biologische Studien in Frage kommen, zumal auch ihre physikochemischen Eigenschaften meist problematisch sind (z.B. Lipide und Glykolipide).

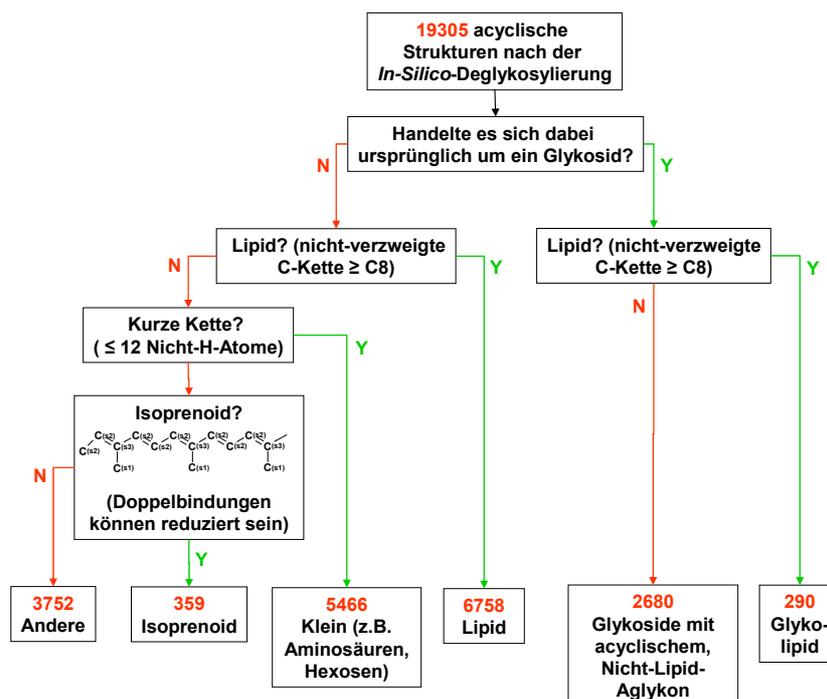


Abb. 45: Analyse und Kategorisierung der acyclischen Naturstoffe.

### Initiale chemoinformatische Analyse

Eine erste chemoinformatische Evaluierung der Naturstoffe mit Ringstrukturen begann mit der Entfernung acyclischer Substituenten. Die häufigsten Substituenten sind in Abb. 46 zu sehen.

Cyclische Substituenten wurden als Teil des Grundgerüsts erachtet. Es wurden 28 591 einzigartige Grundgerüste, d.h. Einheiten von Ringsystemen,<sup>[241]</sup> (einschließlich exocyclischer Doppelbindungen und verbindender Ketten zwischen Ringen) extrahiert. Eine sorgfältige Sichtanalyse dieses Ergebnisses ergab, daß Redundanzen aufgrund unterschiedlicher Glykosylierungsmuster ein- und desselben Aglykon-Grundgerüsts auftraten (siehe Abb. 47).

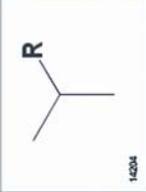
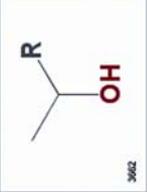
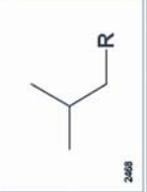
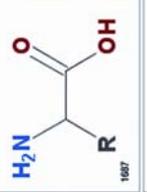
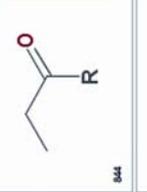
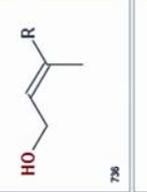
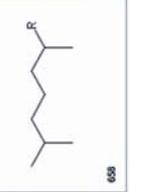
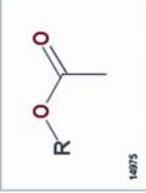
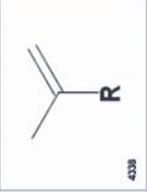
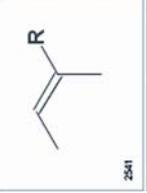
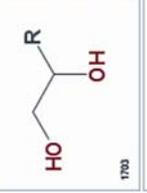
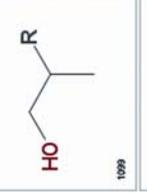
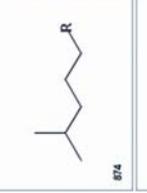
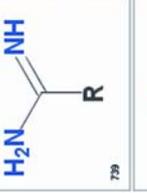
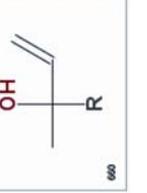
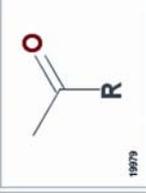
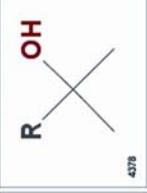
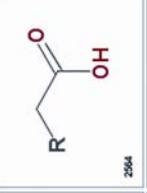
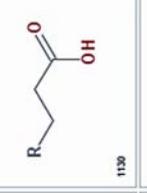
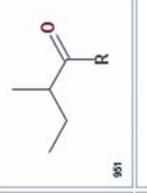
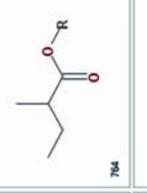
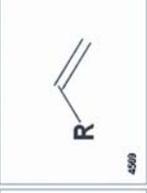
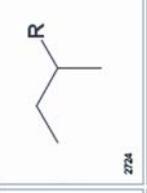
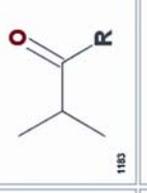
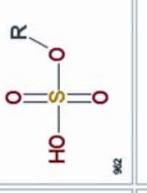
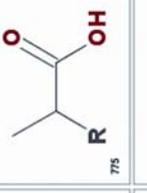
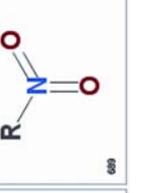
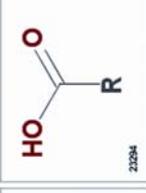
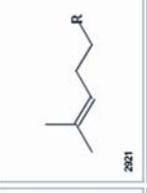
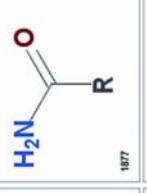
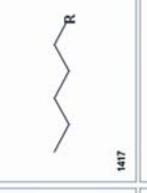
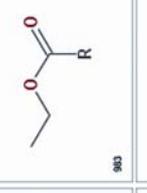
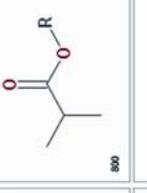
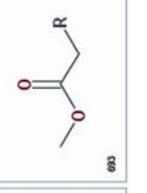
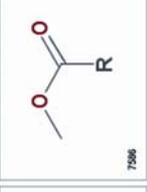
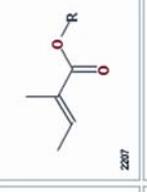
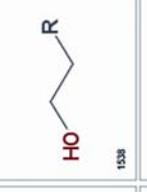
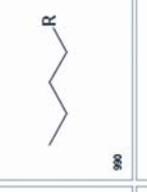
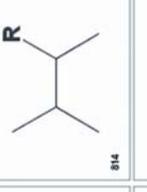
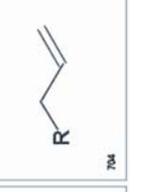
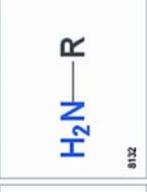
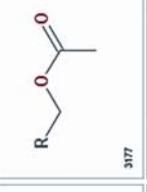
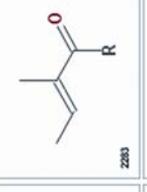
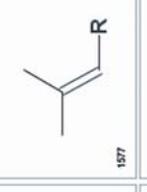
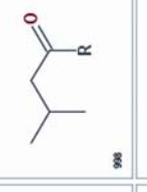
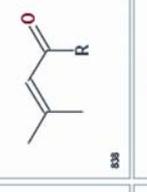
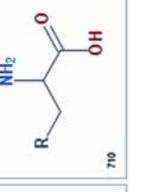
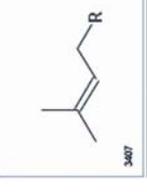
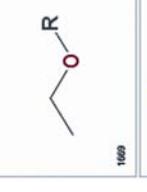
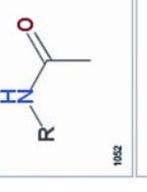
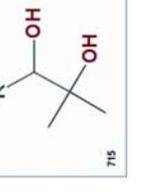
 128420	 10408	 3467	 2365	 1669	 1052	 840	 715
 121040	 8132	 3177	 2283	 1577	 998	 838	 710
 35070	 7596	 3116	 2207	 1538	 990	 814	 704
 23294	 8285	 2921	 1877	 1417	 893	 800	 693
 20000	 4569	 2724	 1838	 1183	 982	 775	 689
 19979	 4378	 2564	 1772	 1130	 951	 764	 670
 14975	 4338	 2541	 1703	 1099	 874	 739	 660
 14304	 3682	 2468	 1887	 1078	 844	 736	 658

Abb. 46: Die häufigsten Naturstoffsubstituenten (die Zahl in der jeweils linken unteren Ecke beschreibt die absolute Häufigkeit, d.h. die Zahl der Naturstoffe, in denen der Substituent mindestens einmal vorkommt). Sauerstoffhaltige Substituenten (z.B. Hydroxy, Methoxy, Carbonsäure, Acetyl u.a.) kommen zusammen mit aliphatischen (z.B. Methyl, Isopropyl, Ethyl, Vinyl, Dimethylallyl u.a.) am häufigsten vor. Diese sollten bei der Konzipierung von Naturstoff-abgeleiteten Bibliotheken berücksichtigt werden und ggf. in die Synthese – evtl. auch in entsprechender biososterer Form (z.B. Tetrazol statt Carbonsäure) – wieder einfließen (s.u.).

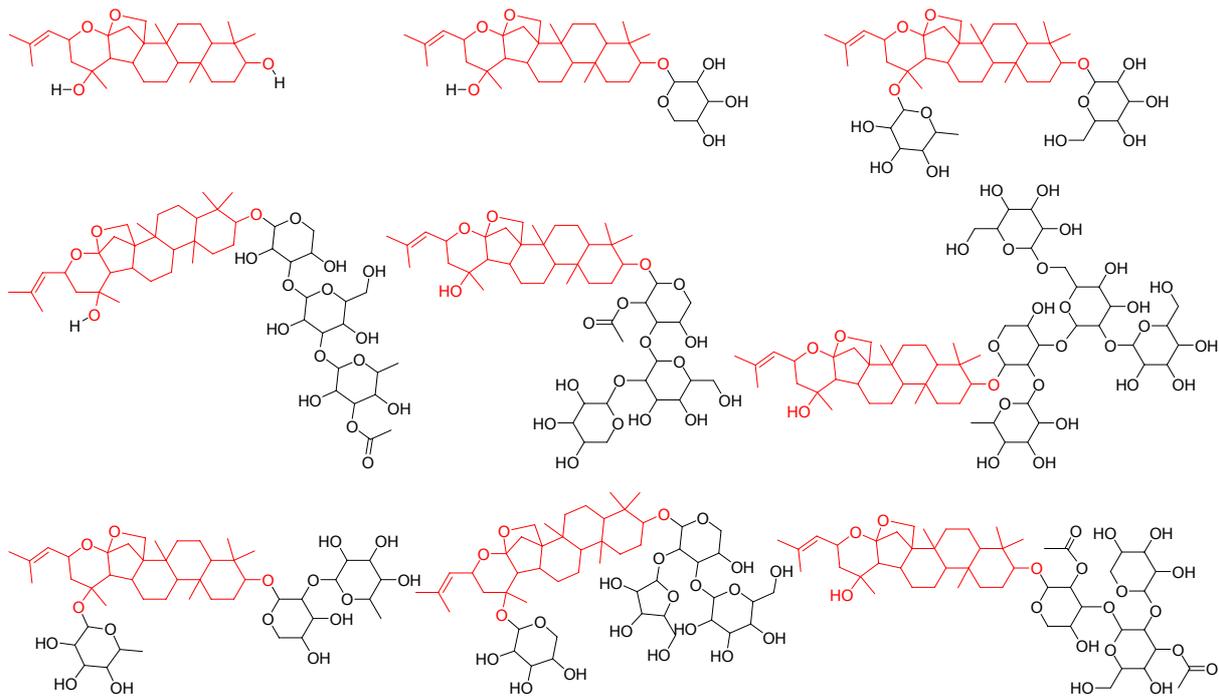


Abb. 47: Unterschiedliche Glykosylierungsmuster ein- und desselben Aglykon-Grundgerüsts (rot)

### ***Entwicklung eines Algorithmus zur In-Silico-Deglykosylierung***

Derzeit kann noch kein Muster der biologischen Aktivitäten der Glykoside im Vergleich zu den entsprechenden Aglyka abgeleitet werden.<sup>[242]</sup> Die glykosidischen Einheiten üben jedoch im allgemeinen einen modulierenden Effekt auf die biologische Aktivität des Aglykons aus, es sei denn, diese enthalten komplexe Zucker (z.B. glykosidische Antibiotika, wie die Aminoglykoside, Vancomycin u.a.).

Eine Glykosylierung, z.B. in herzwirksamen Glykosiden (Digitalis und verwandte), Saponinen oder dem Flavonolignan Silybin, beeinflusst im wesentlichen die Hydrophilie der Moleküle und damit ihre Bioverfügbarkeit und ihre pharmakokinetischen Eigenschaften.<sup>[242]</sup>

Einige Glykoside, wie die cyanogenen Glykoside und Glucosinolate, stellen inaktive Vorstufen von Abwehrstoffen dar. Sie treten zusammen mit Glykosidasen auf, die die jeweilige Zuckereinheit abspalten und so HCN respektive Isothiocyanate freisetzen. Um Autotoxizität zu verhindern, sind Glykoside und Glykosidasen räumlich entweder auf der subzellulären oder Gewebeebene voneinander getrennt. Eine Verletzung des Gewebes, wenn dieses z.B. durch Herbivoren gekaut wird, ist Voraussetzung dafür, daß die „chemische Waffe“ freigesetzt wird.<sup>[243, 244]</sup>

Um die potentiell irreführende strukturelle Variabilität aufgrund von Glykosylierung zu unterdrücken, wurde eine *In-Silico*-Deglykosylierung der Naturstoffe, die *O*-Glykoside enthalten, durchgeführt. Hierzu wurde ein spezieller rechnergestützter Algorithmus in der Programmiersprache Java entwickelt, der auf rekursivem Strukturabgleich basiert.

19 876 Naturstoffmoleküle der gesamten DNP-Datenbank (12.3%) enthielten einen oder mehrere unterschiedlich substituierte Zuckerringe. Die dem rekursiven Strukturabgleich zugrundeliegenden Pyranose- und Furanosesubstituenten sind Abb. 48 zu entnehmen.

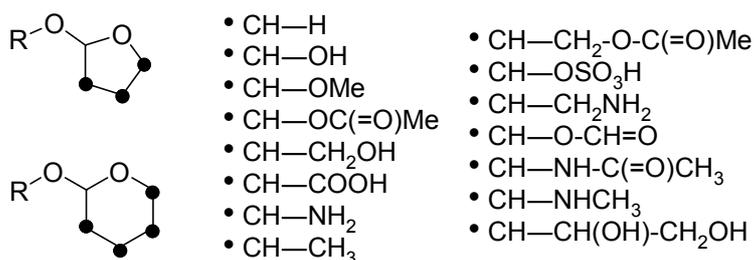


Abb. 48: Der Deglykosylierungsalgorithmus erfaßt Pyranosen und Furanosen. Die in den Grundgerüsten eingezeichneten Punkte können für einen der in der nebenstehenden Liste aufgeführten Substituenten stehen.

Die Naturstoffglykoside enthielten zwischen 1 und 11 Zuckereinheiten (siehe Tabelle 3 und Abb. 49). Aus der Annotation der Datenbank mit den jeweiligen Quellorganismen wurde ermittelt, daß eine hohe Variabilität der Zuckersubstituenten fast ausschließlich im Reich der Pflanzen zu finden ist. In Abb. 50 sind die Aglyka mit der höchsten Variabilität von Zuckersubstituenten gezeigt.

<b>N(Zuckereinheiten)</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>
<b>N(Glykoside)</b>	9836	4759	2811	1285	685	313	114	47	14	11	1

Tabelle 3: Verteilung der Anzahl der Zuckereinheiten in den Naturstoff-Glykosiden.

Die Anzahl der Grundgerüste reduzierte sich nach der Deglykosylierung des Primärdatensatzes um 16.0% von 28 591 auf 24 103 einzigartige Grundgerüststrukturen.

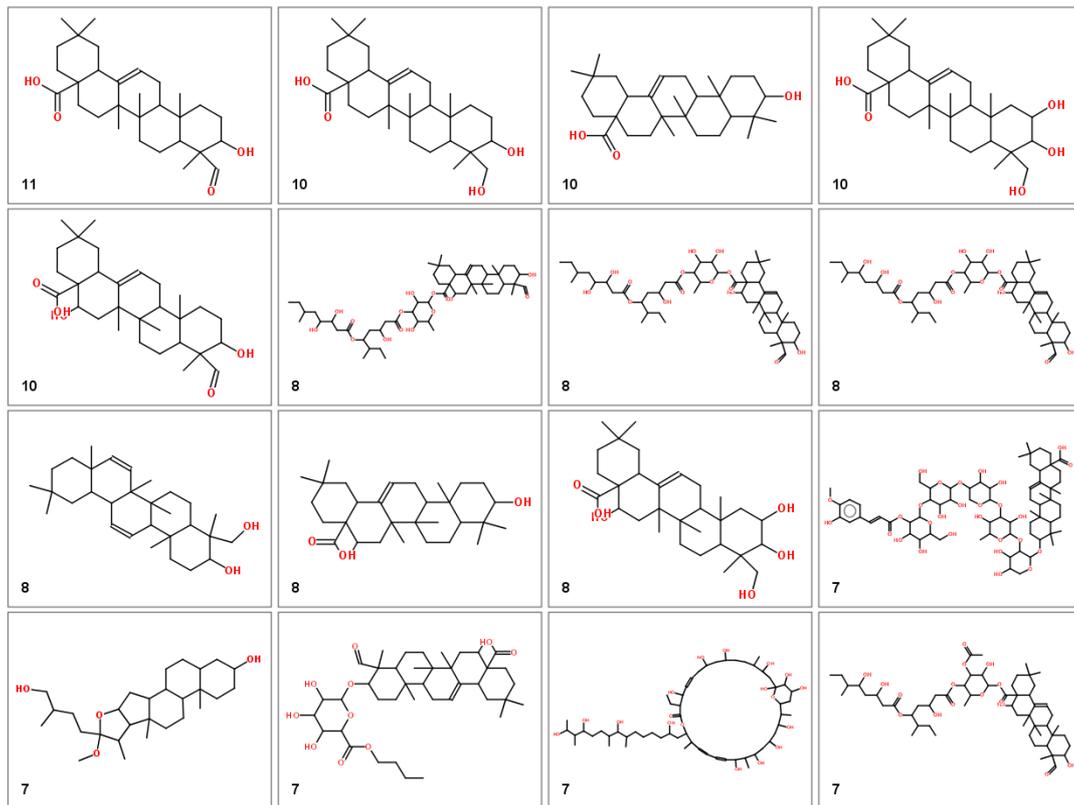


Abb. 49: Aglyka mit der höchsten Zahl verknüpfter Zuckereinheiten (jeweils in der linken unteren Ecke angegeben).

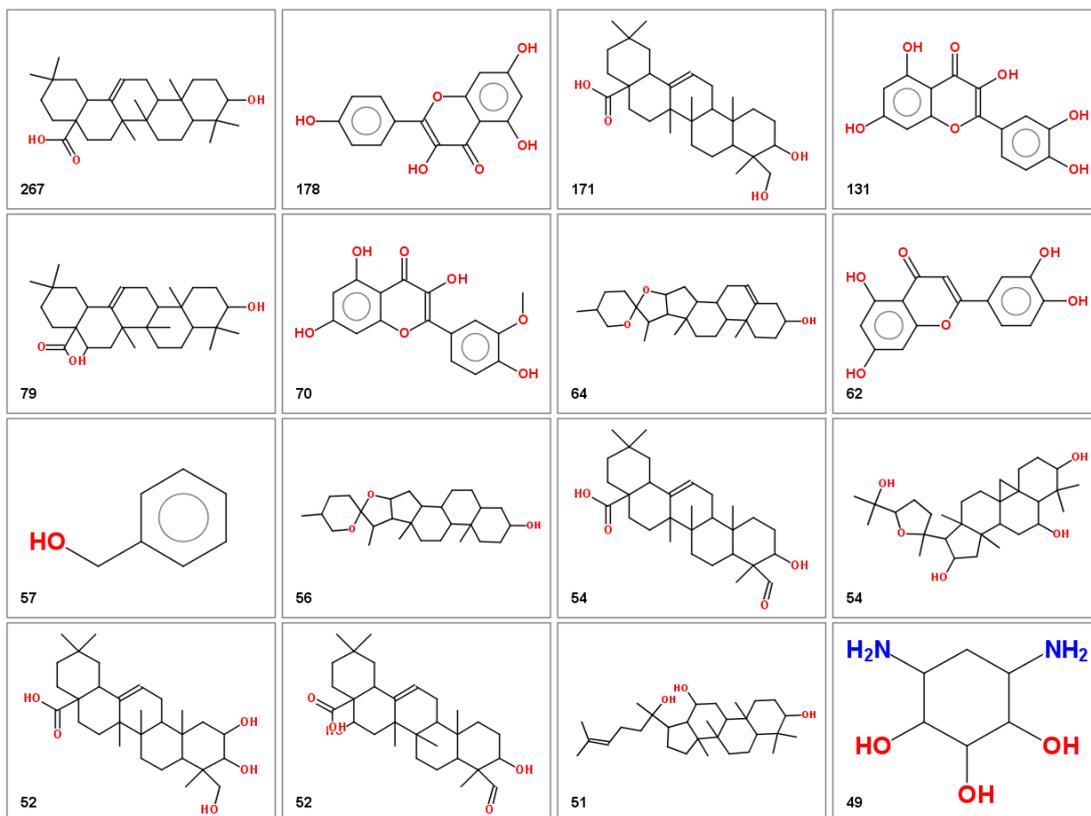


Abb. 50: Aglyka mit der höchsten Variabilität von Zuckersubstituenten. In der linken unteren Ecke ist jeweils die Zahl der entsprechenden Glykoside angegeben.

## Clustering

In einem ersten Anlauf, ein strukturbasiertes Ordnungsprinzip für Naturstoffe zu generieren, wurde zunächst eine Cluster-Analyse, wie im folgenden kurz beschrieben,<sup>[245]</sup> durchgeführt:

### 1. Generierung von ‚Fingerprints‘

Zunächst wurden ‚Fingerprints‘ (ASCII-Bit-Strings, Sequenzen aus den Ziffern „0“ und „1“) der Naturstoffmoleküle generiert. Fingerprints sind eine abstrakte Darstellung struktureller Gegebenheiten in einem Molekül. Es wird ein Muster für jedes Atom, ein Muster, das jedes Atom und seinen nächsten Nachbarn (Bindungen, die sie verbinden, eingeschlossen) repräsentiert, ein Muster, das jede Gruppe von Atomen und Bindungen, die über zwei Bindungen miteinander verknüpft sind, darstellt, usw. generiert.

### 2. Identifizierung potentieller Cluster-Zentroide

Ein Molekül innerhalb eines gegebenen Clusters, das die größte Anzahl von Nachbarn hat und daher dem Rest des Clusters am meisten ähnelt, wird als Cluster-Zentroid gewählt. Hierzu wird für jedes Molekül (dargestellt als Fingerprint) die Anzahl der Nachbarn basierend auf dem *Tanimoto*-Koeffizienten berechnet. Der *Tanimoto*-Koeffizient zwischen zwei Bitmaps (Fingerprints) ist definiert als die Zahl der Bits, die beide Bitmaps gemein haben, geteilt durch die Gesamtzahl der Bits.

$$Tanimoto = \frac{BG}{B1 + B2 - BG}$$

B1 = Bits in Fingerprint (F) 1, d.h. Anzahl der 1'en in F1

B2 = Bits in Fingerprint 2

BG = Anzahl der gemeinsamen 1'en in F1 und F2

$0 < Tanimoto < 1$ , wobei  $Tanimoto = 1$  identisches Molekül und  $Tanimoto = 0$  bedeutet, daß zwei Moleküle nichts gemeinsam haben.

Für das Clustering wurde  $Tanimoto \geq 0.9$  bestimmt. So wurden Listen generiert, bei denen die Verbindungen mit den meisten Nachbarn, d.h. die potentiellen Cluster-Zentroide, ganz oben aufgeführt sind. Die so ermittelten potentiellen Cluster-Zentroide werden nach absteigender Anzahl der Nachbarmoleküle, die sie beschreiben, sortiert.

### 3. Jarvis-Patrick-Clustering

Anschließend wurde ein Clustering nach der Jarvis-Patrick-Methode durchgeführt.<sup>[246]</sup> Dieser Clustering-Algorithmus wird in der pharmazeutischen Industrie weltweit am häufigsten verwendet. Für dieses Clustering wurde – wie unter Punkt 2 auch – das Ausschlußkriterium *Tanimoto*  $\geq 0.9$  gesetzt. Der Algorithmus startet zunächst mit der jeweils ersten Verbindung einer wie unter Punkt 2 beschriebenen generierten Liste von nächsten Nachbarn und berechnet den paarweisen *Tanimoto*-Ähnlichkeitsindex zu allen anderen Verbindungen. Alle Moleküle mit einem *Tanimoto*-Index  $\geq 0.9$  werden diesem Cluster zugeordnet und entsprechend von den Listen für die weiteren Vergleiche gestrichen, so daß sie weder ein Zentroid noch ein Mitglied eines anderen Clusters werden können.

Das Clustering des DNP ergab 42 große Cluster (bestehend aus  $\geq 100$  Naturstoffmolekülen), 1507 mittlere Cluster (bestehend aus  $\geq 10$  Naturstoffmolekülen) und 2851 kleine Cluster (bestehend aus  $\geq 5$  Naturstoffmolekülen). Weiterhin wurden 50 000 Naturstoffe als Singletons (einelementige Mengen) klassifiziert. Die Tendenz der Jarvis-Patrick-Methodologie, einen großen Anteil an Singletons und wenige sehr große Cluster zu generieren, ist als ihr entscheidender Nachteil bekannt.<sup>[237]</sup>

Diesem Ergebnis zufolge stellte sich ein Clustering-Ansatz als nicht geeignet heraus, um Naturstoffe sinnvoll strukturell zu klassifizieren.

### ***Baumanalyse als Basis für SCOMP***

Die Methode der Baumanalyse konnte aus dem logischen Grundprinzip der SCOP-(Structural Classification of Proteins)<sup>[29]</sup>-Datenbank, das sich bereits als erfolgreich für die Klassifizierung von Proteinen erwies, abgeleitet und entwickelt werden. In der SCOP-Datenbank werden 3D-Proteinstrukturen nach strukturellen und evolutionären Argumenten baumartig hierarchisch klassifiziert (siehe Abb. 51).

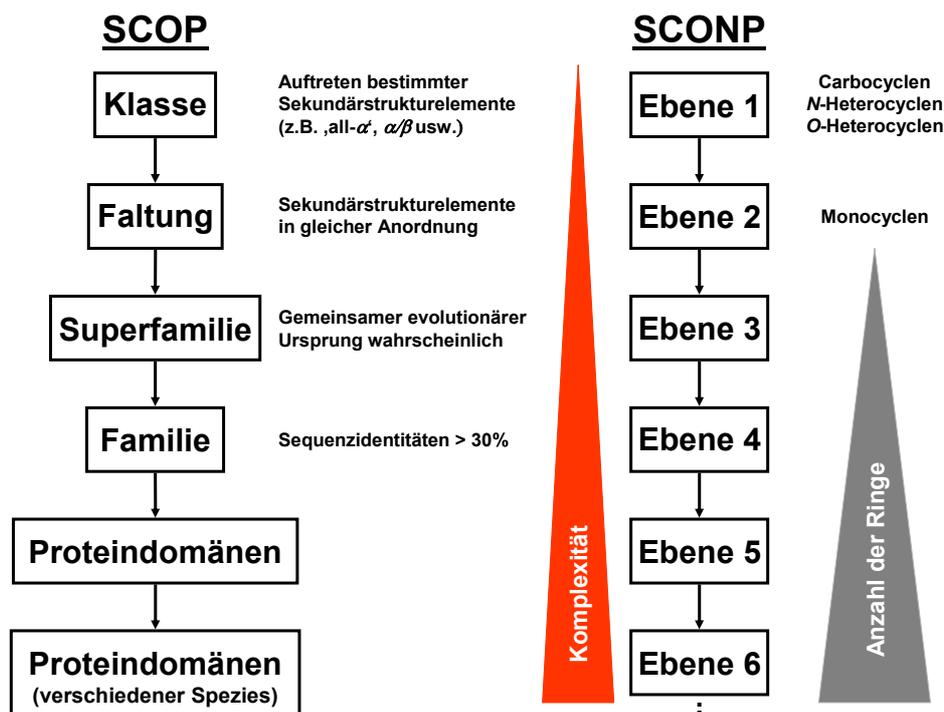


Abb. 51: Analogie zwischen der SCOP-Klassifizierung von Proteinen und der baumartig-hierarchischen Klassifizierung von Naturstoffgrundgerüsten nach SCOMP. Proteine können – wie in Kapitel 4.1.2 ausführlich beschrieben – hierarchisch klassifiziert werden. Ein solches Klassifikationssystem liegt der SCOP-Datenbank zugrunde. Die nach ihrer Herkunft aus verschiedenen Spezies differenzierten Proteindomänen werden hier Klassen, Faltungen, Superfamilien und Familien zugeordnet. Mit absteigender Hierarchieebene nimmt die Zahl der Taxa zu, d.h. es gibt wenige Klassen (7) und derzeit (Stand: Februar 2005) 887 Faltungstypen, 1 447 Superfamilien und 2 630 Familien, unter die 65 122 Proteindomänen subsumiert werden. Analog sollen Naturstoffe anhand der Anzahl der Ringe in ihren Grundgerüsten klassifiziert werden. Die in Naturstoffen vorkommenden prinzipiellen Strukturklassen von 1-Ring-Grundgerüsten (Carbocyclen, N- und O-Heterocyclen) bilden die Basis des Ordnungsprinzips. Die Naturstoffgrundgerüste werden nach der Anzahl der Ringe den einzelnen Hierarchieebenen zugeordnet und somit mit steigender struktureller Komplexität kategorisiert.

Um eine vergleichbare Analyse der Naturstoffgrundgerüste zu erreichen, wurde ein rechnergestützter Algorithmus entwickelt, der für jedes einzelne Grundgerüst das zugehörige Elter-Grundgerüst identifiziert. Der hier verwendete Begriff des Elters (beschreibt ein Elternteil geschlechtsneutral) wurde der klassischen Genetik entlehnt. Der für die Baumanalyse entwickelte Algorithmus sucht in jedem Grundgerüst nach Substrukturen, die ihrerseits wiederum Naturstoffgrundgerüste darstellen. Es wird im Prinzip eine strukturbasierte Genealogie der Naturstoffgrundgerüste erstellt, wobei diese dann hierarchisch mit steigender Anzahl von Ringen, d.h. Grundgerüstgröße, angeordnet werden. Das Elter-Grundgerüst stellt somit eine Substruktur seines Abkömmlings dar. Diese Analyse wurde mit jedem einzelnen Naturstoffgrundgerüst durchgeführt. Eine Korrelation der Analyseergebnisse lieferte einen Baum von Grundgerüsten mit verschiedenen hierarchischen Ebenen (siehe Abb. 53A), wobei die Einzelring-Grundgerüste als einfachste Grundgerüste die Wurzeln des Baums bilden. Die Einzelringe wurden in drei Klassen chemisch ähnlicher Cluster (Carbocyclen, N-

und *O*-Heterocyclen) nach Heteroatom und Ringgröße zusammengefaßt. Dabei wurden „dickere“ und „dünnere“ Hauptwurzeln entsprechend der Clustergröße gebildet. Jedes Grundgerüst einer hierarchischen Ebene stellt einen Knoten dar, von dem eine weitere Verzweigung zu Folgeknoten ausgehen kann, die komplexere Grundgerüste repräsentieren, d.h. Grundgerüste mit ( $N+1$ ) Ringen. Jedes Grundgerüst auf jeder hierarchischen Ebene repräsentiert mindestens eine Naturstoffstruktur.

Für die Festlegung des Elter-Grundgerüsts wurden einige Regeln aufgestellt, die den Denkmustern von Medizinal- und Synthesechemikern nahekommen:

1. Der Elter eines Grundgerüsts stellte eine Substruktur des Abkömmlings dar.
2. Ein Aromatizitätsmodell unterschied zwischen aromatischen und Doppelbindungen. Dies führte in manchen Fällen von Ringkondensationen zu unerwartetem Verhalten. So hat das Grundgerüst **c6c..1** (siehe Abb. 52A) kein bicyclisches Elter-Grundgerüst, da die rot markierte Bindung im 1,2-Dihydronaphtalin (ursprüngliche Kondensationsstelle) nicht aromatisch ist. Analog wurden anellierte und partiell hydrierte Benzo-Heterocyclen nicht auf den jeweiligen Heterocyclen, sondern auf Benzol (**c6c**) zurückgeführt (siehe Abb. 52B).
3. Das Aufbrechen von Ringbindungen in einem Abkömmling war nicht erlaubt, so daß Grundgerüst **c6c..1** nicht auf Biphenyl zurückgeführt wurde. Ein Elter-Grundgerüst durfte nicht mehr acyclische Bindungen als sein Abkömmling enthalten (siehe Abb. 52C).
4. Exocyclische Doppelbindungen wurden als Teil des Grundgerüsts beibehalten, um ein künstliches Aufbrechen „aromatischer“ Systeme zu vermeiden. In solchen Fällen tritt das Grundgerüst ohne exocyclische Doppelbindung meist nicht auf oder es ist sehr selten (siehe Abb. 52D). In manchen Fällen gibt es jedoch beide Grundgerüste (siehe Abb. 52E). Die Hierarchieebene des jeweiligen Grundgerüsts sollte der Anzahl der Ringe entsprechen. Im Falle des Methylencyclohexans wäre dies nicht der Fall. Es würde Hierarchieebene 2 zugeordnet werden, hätte aber nur einen Ring. In solchen Fällen wurde das Grundgerüst mit der exocyclischen Doppelbindung – sofern es nicht aromatisch war – mit dem „nackten“ Ring zusammengeführt. Dies geschah manuell, ließe sich aber prinzipiell auch automatisieren.

5. Wenn prinzipiell mehrere Kandidaten als Eltern in Frage kamen, wurde der Elter so gewählt, daß das Elter-Grundgerüst die größte Anzahl von Heteroatomen aufwies. Wenn dies nicht möglich war, wurde dasjenige Grundgerüst als Elter gewählt, das in Naturstoffen am häufigsten vorkommt.

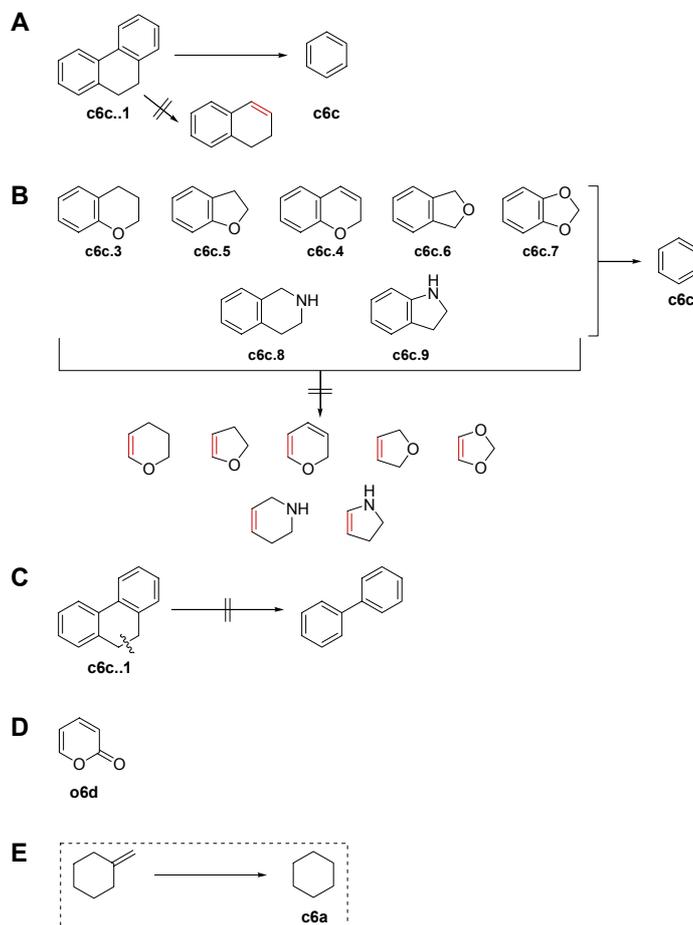


Abb. 52: Graphische Darstellung einiger Regeln der Baumanalyse. A) Das 3-Ring-Grundgerüst **c6c..1** wurde direkt auf Benzol zurückgeführt, da die rot markierte Bindung nicht aromatisch ist. B) Analog zu A wurden die Heterobicyclen **c6c.3-c6c.9** auf Benzol zurückgeführt. Die an der jeweiligen Kondensationsstelle bei der Trennung der Ringe resultierenden Doppelbindungen im Heterocyclen sind nicht aromatisch, da der resultierende Heteromonocyclen aufgrund partieller Hydrierung oder struktureller Gegebenheiten nicht aromatisch ist. C) Das Elter-Grundgerüst durfte nicht mehr acyclische Bindungen als sein Abkömmling enthalten. Daher wurde Grundgerüst **c6c..1** nicht durch das Aufbrechen einer Ringbindung auf Biphenyl zurückgeführt. D) Exocyclische Doppelbindungen wurden als Teil des Grundgerüsts beibehalten. Das entsprechende Grundgerüst ohne exocyclische Doppelbindung kommt nicht vor. E) Wenn das Grundgerüst sowohl ohne als auch mit exocyclischer Doppelbindung vorkommt, so wurden beide Grundgerüste der 1-Ring-Hierarchieebene zugeordnet, um das Ordnungsprinzip (Hierarchieebene entspricht der Anzahl der Ringe im Grundgerüst) beizubehalten.

Eine baumartige Anordnung von Naturstoffgrundgerüsten (Abb. 53A) erlaubt eine Einordnung auch sehr seltener Grundgerüste in ihren strukturellen Kontext. Sie werden als „Zweige“ mit Grundgerüsten in ihrer strukturellen Nachbarschaft klar korreliert. In einer

Clustering-Analyse würden derartige Grundgerüste als Singletons klassifiziert werden, wodurch strukturelle Ähnlichkeiten mit anderen Naturstoffen unentdeckt blieben.

Ein weiterer Vorteil einer Baumanalyse ist, daß sie eine modulare und dynamische Rationalisierung und Klassifizierung ermöglicht, d.h., daß das Baumdiagramm relativ einfach erweitert werden kann. Dies wird zusätzlich durch das vollautomatisierte *In-Silico*-Verfahren erheblich erleichtert. Im übrigen ist diese Art der Baumanalyse auch auf andere Datensätze übertragbar, z.B. Wirkstoff- und Leitstrukturdatenbanken.

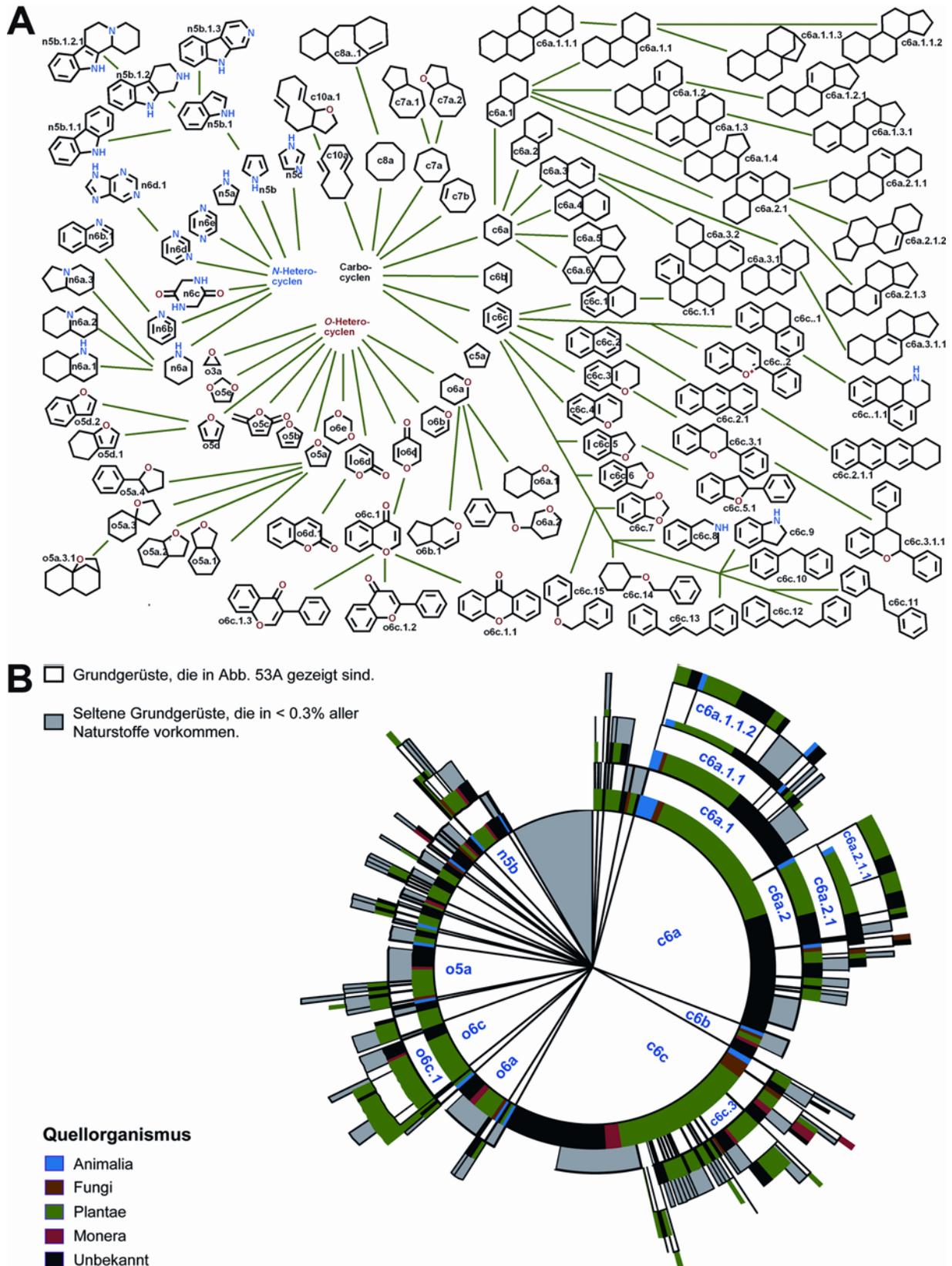


Abb. 53: A) Sternförmige graphische Darstellung des Baumdiagramms der Naturstoffgrundgerüste. Der Klarheit der graphischen Darstellung wegen sind nur die Grundgerüste gezeigt, die kumulativ wenigstens 0.3% der Naturstoffe im Datensatz beschreiben. Die Nomenklatur der Grundgerüste wurde so gewählt, daß der Name des Elter-Grundgerüsts im Namen seines Abkömmlings enthalten ist. Der Name des Elters kann aus dem des Abkömmlings durch Entfernen des letzten Suffixes abgeleitet werden.

B) Torten-Diagramm, das die gleiche Grundgerüstfraktion, wie in Abb. 53A dargestellt, repräsentiert. Die Segmentwinkel entsprechen den Häufigkeiten des Vorkommens der Naturstoffe und ihrer Erzeugerorganismen.

### *Quantitative Analysen und Annotation des SCONP-Baums der Naturstoffgrundgerüste*

Eine quantitative Analyse des Baumdiagramms der Naturstoffgrundgerüste ergab, daß Grundgerüste mit drei Ringen in Naturstoffen am häufigsten zu finden sind. Drei-Ring-Grundgerüste kommen in 20.8% aller Naturstoffe (nach Deglykosylierung) vor (siehe Abb. 54). Zusammen mit Zwei- und Vier-Ring-Grundgerüsten können 52.8% der deglykosylierten Naturstoffe des Datensatzes eingeordnet werden.

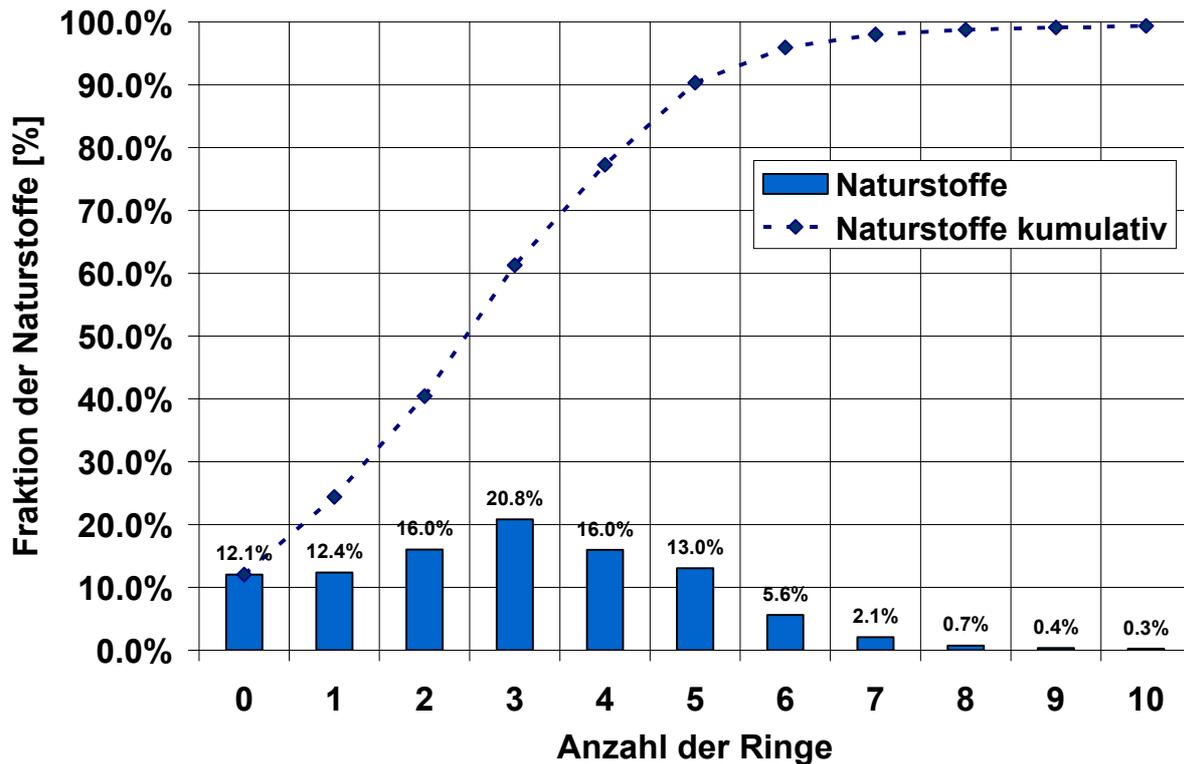


Abb. 54: Statistische Analyse des Grundgerüst-Baums. Die größte Fraktion der Naturstoffe kann durch Drei-Ring-Grundgerüste beschrieben werden, die 20.8% aller Naturstoffe (nach Deglykosylierung) repräsentieren. Unter Einbeziehung der Zwei- und Vier-Ring-Grundgerüste werden 52.8% aller Naturstoffe hinreichend in bezug auf ihre Grundgerüstarchitektur beschrieben.

Die hierarchische Ebene der Drei-Ring-Grundgerüste stellt demnach die im Zuge der Evolution der Sekundärmetabolite von der Natur am häufigsten selektierte Lösung für die Bindung an Proteine dar. Eine strukturell diverse Auswahl der häufigsten und der seltensten Drei-Ring-Grundgerüste ist jeweils Abb. 55 und 56 zu entnehmen.

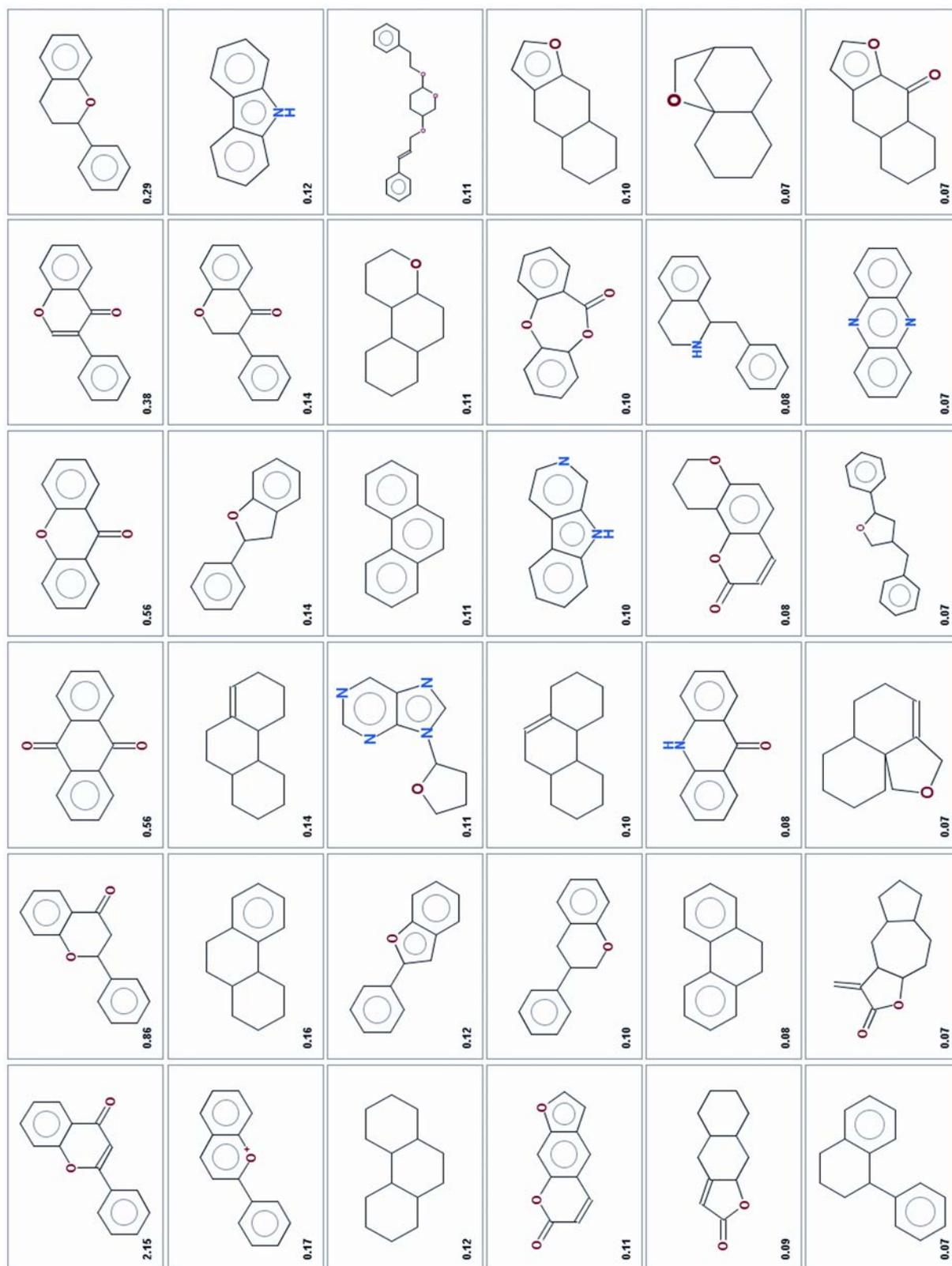


Abb. 55: Eine Auswahl der häufigsten Drei-Ring-Grundgerüste im deglykosylierten Naturstoffdatensatz. Die Zahl in der linken unteren Ecke stellt die vom jeweiligen Grundgerüst repräsentierte Fraktion der Naturstoffe in Prozent dar. Flavonoide und Anthranoide gehören zu den häufigsten Strukturklassen im untersuchten Datensatz.

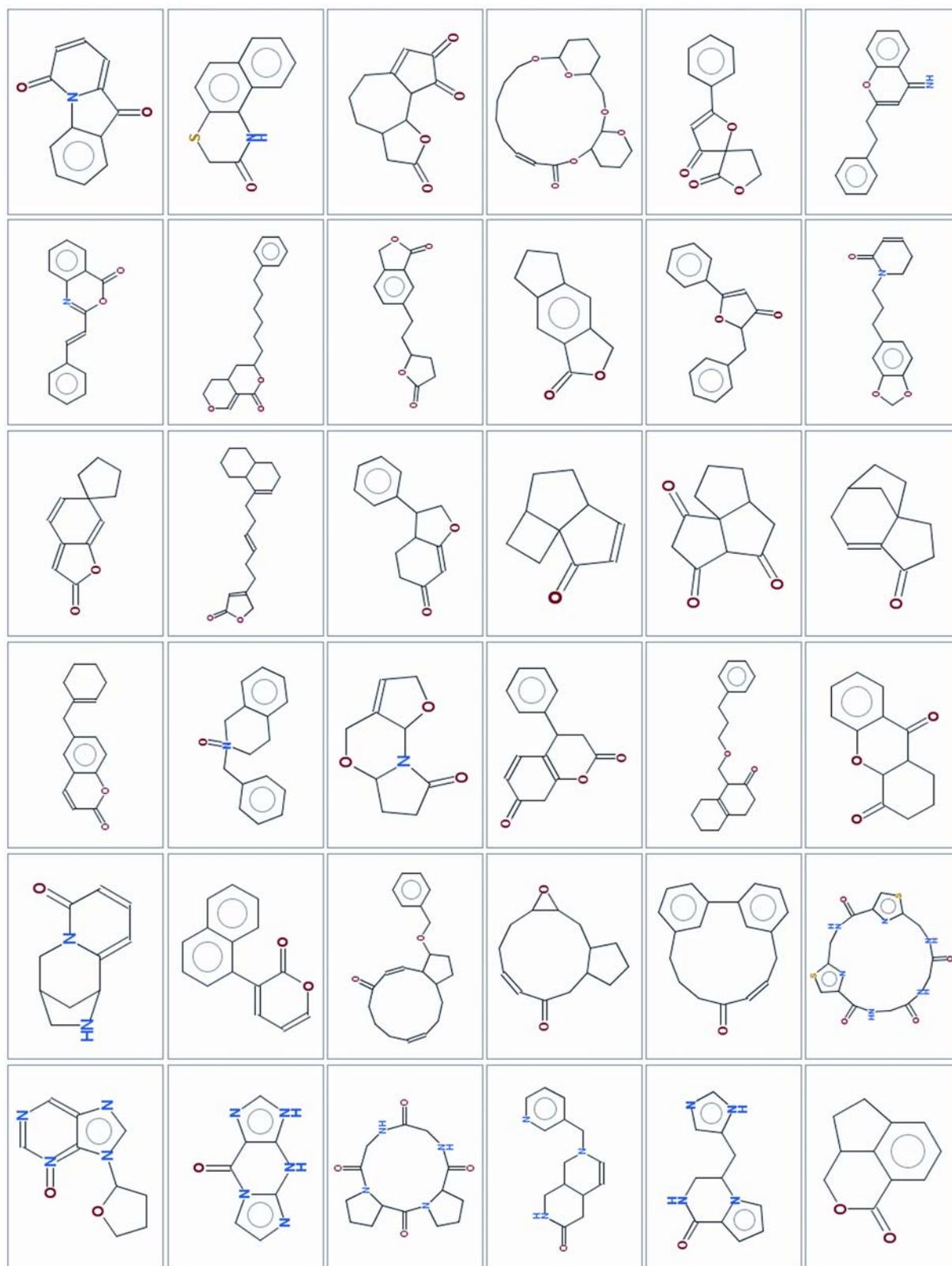


Abb. 56: Eine Auswahl singulärer Drei-Ring-Grundgerüste (Singletons).

Das statistische Ergebnis, daß mit Gerüststrukturen mit zwei bis vier Ringen über die Hälfte der deglykosylierten Naturstoffe bezüglich ihrer Architektur eingeordnet werden können, führte zur Hypothese, die van der Waals-Volumina der Naturstoffe der hierarchischen Ebenen

zwei bis vier könnten mit dem durchschnittlichen Volumen der Bindungstaschen in Proteinen korrelieren. Die van der Waals-Volumina der deglykosylierten Naturstoffe wurden mithilfe eines Java-Programms der Firma Novartis aus 3D-Strukturen berechnet, die von dem Programm CORINA 3.1 (Molecular Networks, Erlangen, Germany, <http://www.mol-net.de/>) aus den SMILES-Strings der Naturstoffmoleküle generiert wurden. Die Volumina der Mehrzahl der Naturstoffe mit zwei bis vier Ringen liegen im Bereich von  $100 \text{ \AA}^3$  bis  $500 \text{ \AA}^3$ , wobei das Maximum bei ca.  $250 \text{ \AA}^3$  liegt (siehe Abb. 57A). Eine statistische Auswertung der Volumina eines Datensatzes von 18 402 aus der PDB extrahierten Proteinkavitäten ergab, daß die meisten Kavitäten in einen Volumenbereich von  $300 \text{ \AA}^3$  bis  $800 \text{ \AA}^3$  fallen.<sup>[247]</sup> Aus dem Vergleich zwischen diesem Ergebnis und der Volumenanalyse der Naturstoffe kann geschlossen werden, daß die durchschnittlichen Volumina der Zwei- bis Vier-Ring-Naturstoffe auf die durchschnittlichen Dimensionen der Proteinkavitäten abgebildet werden können, wenn berücksichtigt wird, daß Proteinliganden oft nicht das gesamte Volumen einer Kavität ausfüllen. Eine ähnliche Analyse der Volumenverteilung von ca. 30 000 Wirkstoffen des World Drug Index (Thomson Derwent, <http://thomsonderwent.com>) ergab zudem, daß die Volumina der Zwei- bis Vier-Ring-Naturstoffe mit denen in Arzneistoffen vergleichbar sind (siehe Abb. 57B).

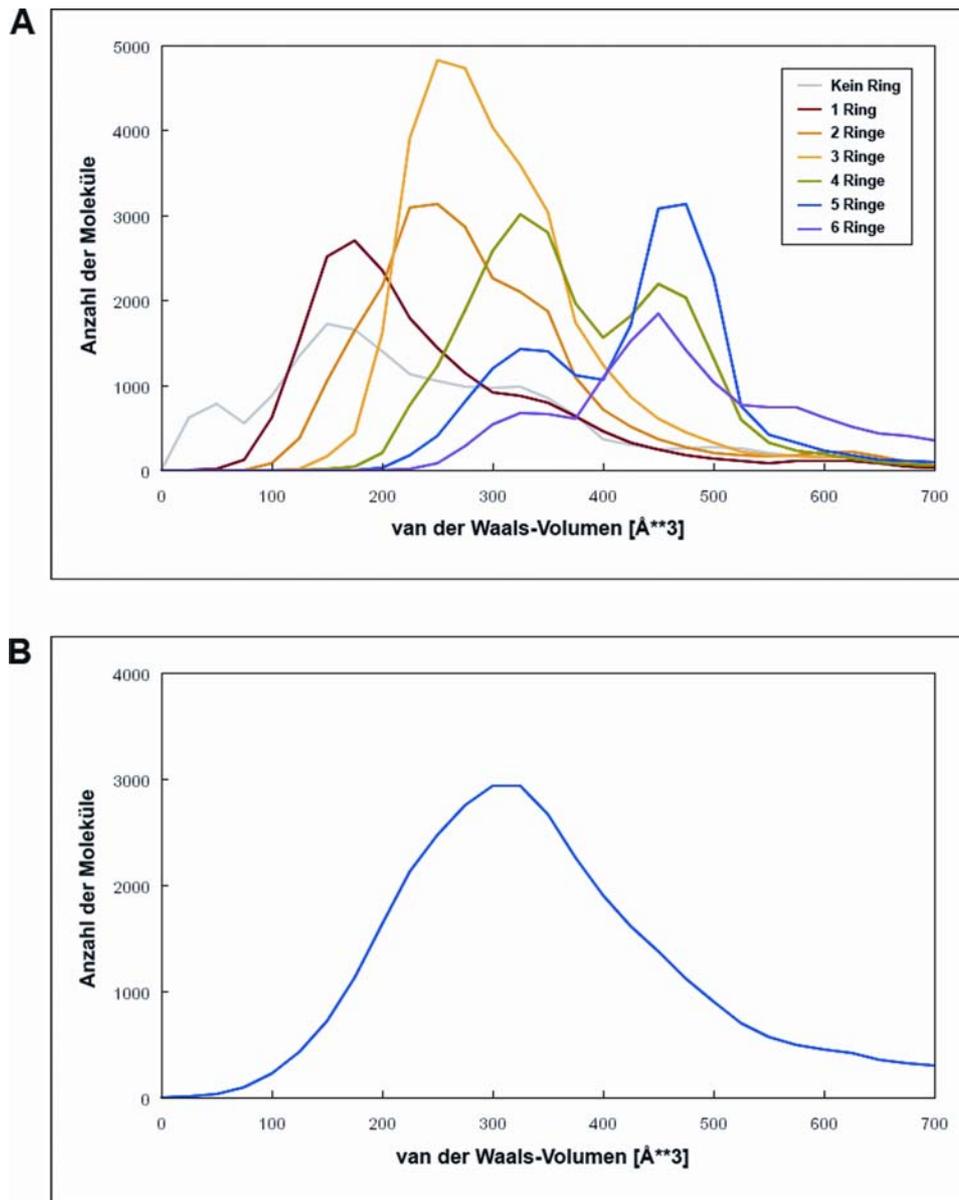


Abb. 57: A) Volumenverteilung der deglykosylierten und nach der Anzahl der Ringe in ihren Grundgerüsten sortierten Naturstoffe. B) Analoge Analyse von ca. 30 000 Wirkstoffen des World Drug Index.

Über diese statistischen Analysen hinaus, wurde mit der Annotation des Naturstoff-Grundstruktur-Baums mit biologischer Aktivität und Herkunft begonnen. Ein Vergleich der Naturstoffe des DNP mit den Strukturen der ‚MDL Drug Data Report‘-(MDDR)-Datenbank (<http://www.mdl.com>), in der die biologische Aktivität für 153 366 Verbindungen verzeichnet ist, ergab, daß nur 2 110 der Naturstoffe des Datensatzes eine biologische Aktivität aus der MDDR-Datenbank zugeordnet werden konnte. Darüber hinaus war die Beschreibung der biologischen Aktivität sehr uneinheitlich und reichte von der Angabe präziser Targets zu sehr allgemeinen Angaben wie z.B. „zytotoxisch“.

Die Annotation des Grundgerüst-Baums mit biologischer Herkunft ergab, daß die Naturstoffe im wesentlichen aus Pflanzen stammen. Für viele Naturstoffe war eine Annotation bzw. taxonomische Klassifizierung nicht möglich (siehe Abb. 58A). Im Reich der Pflanzen sind insbesondere die Divisionen der Angiospermen (Magnoliophyta) und der Gymnospermen (Coniferophyta) häufig vertreten (siehe Abb. 58B). Im Tierreich stammten die häufigsten Quellorganismen, Schwämme (Porifera) und Nesseltiere (Cnidaria, Quallen), aus dem Meer. Im Reich der Pilze herrschen Schlauchpilze (Ascomycota) als Erzeugerorganismen vor. 6407 Naturstoffe stammen aus Bakterien.

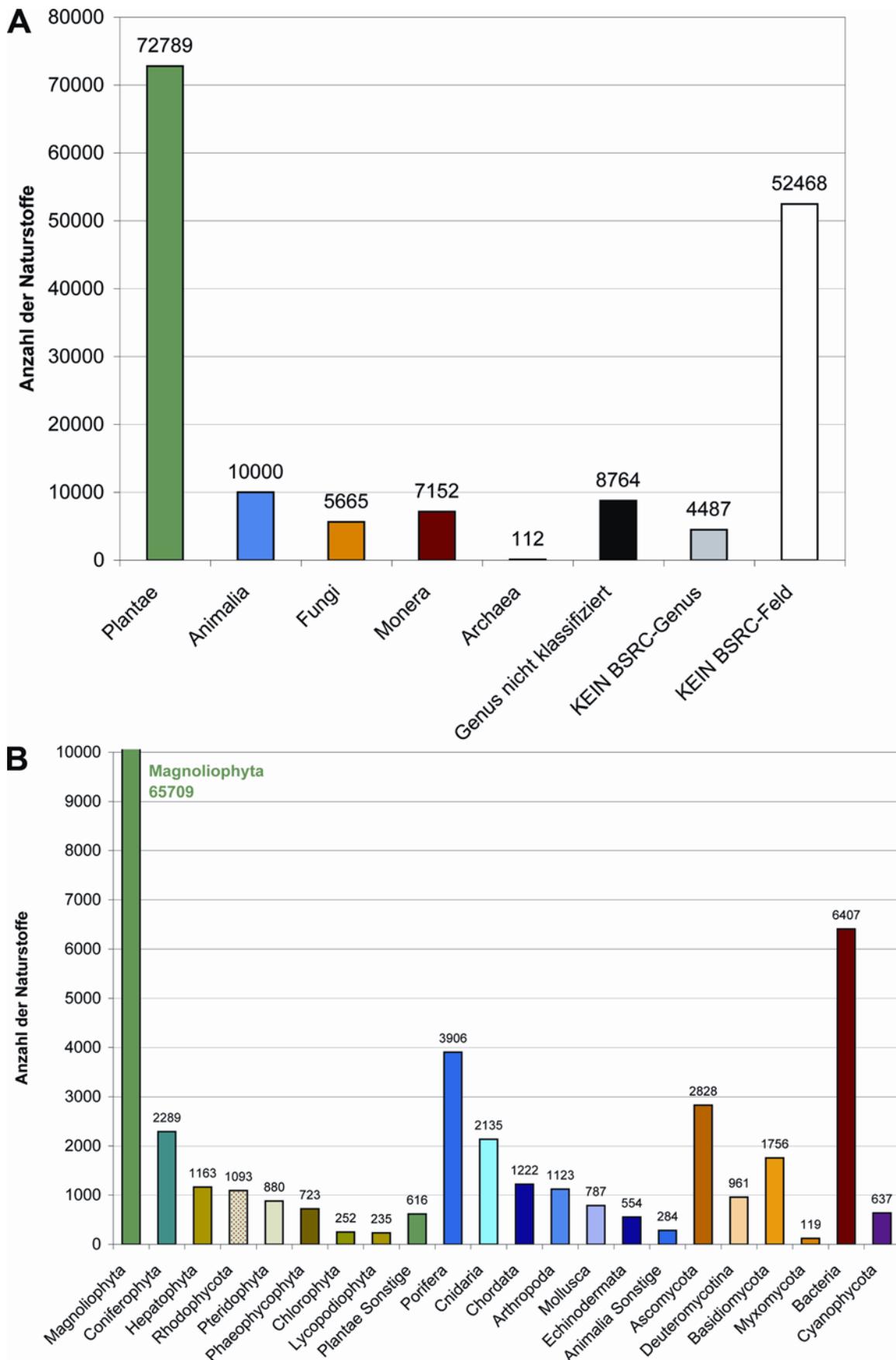


Abb. 58: A) Anzahl der Naturstoffe aufgetragen gegen die verschiedenen Reiche der Erzeugerorganismen. Die weit überwiegende Anzahl der Naturstoffe des Datensatzes (72 798 von 161 110 Naturstoffen, 45%) stammt aus Pflanzen. Insgesamt 65 719 Naturstoffe (41%) konnten nicht klassifiziert werden, u.a. weil im entsprechenden Datenfeld, dem BSRC-Feld, entweder kein Genus angegeben, oder dieses gar nicht vorhanden war. B) Anzahl

der Naturstoffe aufgetragen gegen die Phyla bzw. Divisionen der Erzeugerorganismen. Aus Magnoliophyta oder Angiospermen stammen 41% der Naturstoffe. Auf die Phyla Coniferophyta (Gymnospermen), Moose (Hepatophyta), Rotalgen (Rhodophycota), Farnpflanzen (Pteridophyta), Braunalgen (Phaeophycophyta), Grünalgen (Chlorophyta), Bärlappe (Lycopodiophyta) und sonstige verteilen sich 4% der Naturstoffe des Datensatzes. Bei den Divisionen des Reichs der Tiere herrschen Schwämme (Porifera) als Erzeugerorganismen vor. Sie sind die Quellorganismen von 2.4% der Naturstoffe des Datensatzes. Nesseltiere (Cnidaria, Quallen), Vertebraten (Chordata), Gliederfüßer (Arthropoda), Weichtiere (Mollusca), Stachelhäuter (Echinodermata, z.B. Seesterne, Seeigel) und sonstige Divisionen des Reichs der Animalia erzeugen 3.8% der Naturstoffe. Bei den Pilzen herrschen Ascomycota (Schlauchpilze) vor (1.8% der Naturstoffe des Datensatzes). Eine verhältnismäßig große Division von Quellorganismen stellen die Bakterien dar. Ihnen entstammen 4% der Naturstoffe.

## ***Opioide Analgetika – Ein pharmazeutisches Musterbeispiel für radikale***

### ***Grundgerüstvereinfachung***

Die konsequente medizinisch-chemische Bearbeitung des Strukturmotivs des opioiden Alkaloid-Naturstoffs Morphin (**85**, Abb. 59 und 60), der aus dem Schlafmohn (*Papaver somniferum*) stammt und als hochpotentes Analgetikum in der Schmerztherapie verwendet wird, ist ein Musterbeispiel dafür, daß radikale Grundgerüstvereinfachungen zu pharmakologisch wirksamen Analoga führen können. Die vom Morphin abgeleiteten Anlagetika können ähnliche, modifizierte oder sogar überlegene pharmakodynamische und pharmakokinetische Eigenschaften im Vergleich zum Ausgangsmolekül (analgetische Potenz, Nebenwirkungsprofil, Rezeptor(subtyp)selektivität, Bioverfügbarkeit) aufweisen. In Abb. 59 ist die strukturelle Einordnung von Morphin in den SCONP-Naturstoffbaum gezeigt.

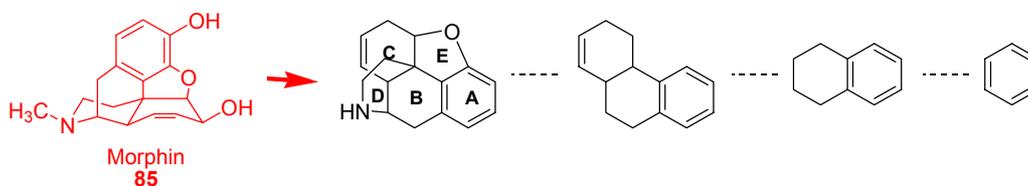


Abb. 59: Strukturelle Genealogie des Morphins nach dem SCONP-Baum der Naturstoffgrundgerüste. Offensichtlich kommen lediglich Grundgerüste mit ABC-, AB- und A-Ringsystem vor. Kombinationen von vier Ringen (ABCD, ABCE, ABDE, BCDE) und drei Ringen unter Einbeziehung von Ring D liegen keinem Naturstoff des Datensatzes als Gerüststruktur zugrunde.

Wenn eine SCONP-analoge Analyse mit den pharmakologisch aktiven Morphinabkömmlingen durchgeführt wird, resultiert eine strukturelle Klassifizierung, wie sie in Abb. 60 dargestellt ist. Jedem Ringsystem kann mindestens eine aktive Verbindung oder ein klinisch eingesetztes Pharmakon zugeordnet werden.

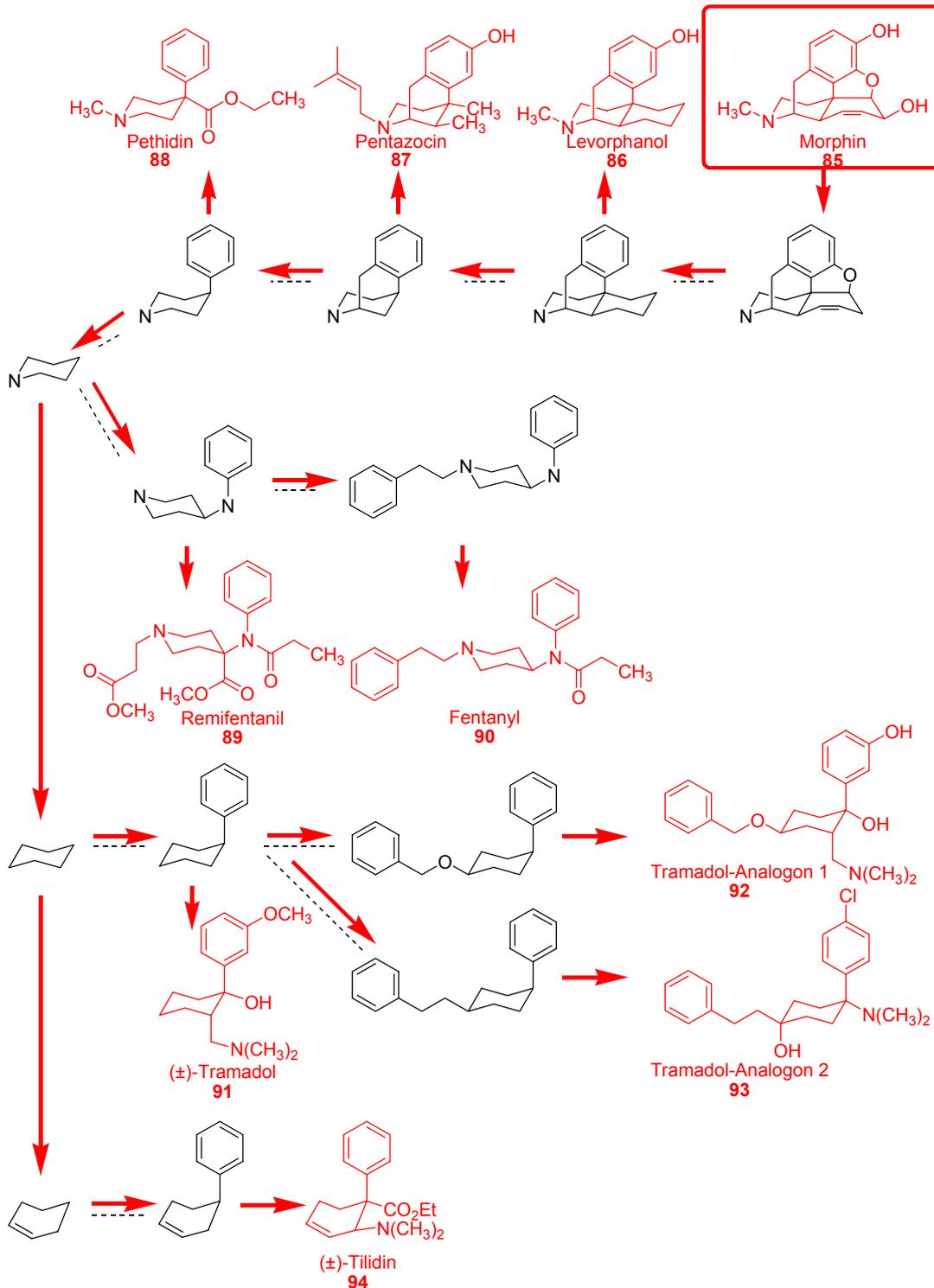


Abb. 60: SCONP-analyse des Naturstoffs Morphin und synthetischer Abkömmlinge. Es wird sehr deutlich, wie die komplexe Naturstoffstruktur sukzessive vereinfacht wurde. Die den jeweiligen Grundgerüsten zugeordneten Verbindungen sind entweder zur Schmerztherapie zugelassene Arzneistoffe oder patentierte Verbindungen (Tramadol-Analoga 1<sup>[248]</sup> und 2<sup>[249]</sup>) mit analgetischer Wirksamkeit.

Obwohl die von Morphin abgeleiteten Grundgerüste der in Abb. 60 gezeigten Wirkstoffe nicht von SCONP erfaßt werden – sie kommen in Naturstoffen nicht vor – zeigt dieses Anwendungsbeispiel der SCONP-Analyse die prinzipielle Machbarkeit auch von enormen

strukturellen Vereinfachungen unter Beibehaltung der Wirksamkeit. Interessanterweise ist in diesem Anwendungsbeispiel auch die Stereochemie meist vernachlässigbar, da einige gerade der analgetisch potentesten Morphinabkömmlinge achiral sind, wie z.B. Remifentanyl (**89**) und Fentanyl (**90**). Es wird insbesondere deutlich, daß eine SCONP-analoge Analyse auch erfolgreich auf Arzneistoff- und Leitstrukturen nicht-natürlichen Ursprungs angewendet werden kann.

### ***Erste Anwendung von SCONP als Navigator und Hypothesengenerator im Rahmen der Entwicklung von Verbindungsbibliotheken***

In der medizinisch- respektive biologisch-chemischen Forschung kann eine baumartig-hierarchische Klassifizierung von Naturstoffen basierend auf ihren Grundgerüsten als strategisches Werkzeug für die Auswahl von Naturstoffgrundgerüsten genutzt werden, die in der „Naturstoff-geleiteten Entwicklung von Verbindungsbibliotheken“<sup>[11-13, 41]</sup> als Leitprinzipien ihren Einsatz finden würden.

In einem solchen Ansatz wird ein als Suchmotiv dienender Naturstoff hinsichtlich seiner Ringstruktur analysiert. Aus dem Baumdiagramm können durch Bewegung entlang oder zwischen den Ästen (bildlich gesprochen: Brachiation = Schwinghängeln wie die Gibbons) mögliche strukturell vereinfachte Analoga unter Beibehaltung der biologischen Relevanz (nicht notwendigerweise der identischen Aktivität) identifiziert werden. Eine geleitete strukturelle Vereinfachung eines gegebenen Naturstoffgrundgerüsts resultiert aus dem Klassifikationsprinzip, das dem Baumdiagramm zugrundeliegt, d.h. inhärent daraus, daß Kernstrukturen weniger komplexer Naturstoffe Substrukturen des Suchmotivs darstellen können. Diesem Ansatz folgend, würden sich bei einer Brachiation wurzelwärts auf jeder hierarchischen Ebene vereinfachte Grundgerüste finden, die als Kern-Grundgerüste möglicher Verbindungsbibliotheken verwendet werden könnten. Da auf jeder Ebene geringerer Komplexität verschiedene Grundgerüste identifiziert werden können und weil *a priori* nicht offensichtlich ist, auf welcher Komplexitätsebene die Analyse angehalten werden soll, d.h. bis zu welcher Ebene die gewünschte biologische Aktivität erhalten bleibt, muß für die letzte Auswahl des Grundgerüsts der zu synthetisierenden Verbindungsbibliothek ein zweites Kriterium hinzugezogen werden.

Ein solches Kriterium liefert z.B. die biologische Aktivität von Verbindungsklassen, die durch die einzelnen Grundgerüste repräsentiert werden, oder der biologische Ursprung

(gleichartiges Taxon von Suchmotiv und vereinfachtem Grundgerüst). Die Information über Quellorganismen kann mit den Grundgerüsten korreliert und z.B. in einem Tortendiagramm, wie es in Abb. 53B zu sehen ist, dargestellt werden.

Obwohl die biologische Aktivität ein sehr relevantes zweites Kriterium sein mag, kann es bislang nur für eine kleine Fraktion der Naturstoffe in Betracht gezogen werden. Von Nachteil dabei ist, daß die Eigenschaft der biologischen Aktivität nicht generalisiert werden kann und zudem stark von der Art und Anzahl der biologischen Assay-Verfahren abhängt, denen die Naturstoffe unterzogen wurden und werden.

Die prinzipielle Fähigkeit, an Proteine zu binden, ist als hierarchisch höherrangig als Bioaktivität einzustufen, da sie auf der Tatsache gründet, daß Naturstoffe durch Proteine biosynthetisiert werden und damit prinzipiell an diese binden können. Daher ist eine funktionelle, mechanistische und/oder strukturelle Analogie zwischen möglichen Protein-Targets ein geeignetes zweites Kriterium für die Auswahl strukturell einfacherer Grundgerüste für die Entwicklung von Verbindungsbibliotheken.

Um die Möglichkeit zu prüfen, ob ein solcher Ansatz der strukturellen Einordnung in den SCONP-Baum und aus der SCONP-Klassifizierung abgeleiteter struktureller Vereinfachungen auf der Basis natürlich vorkommender Grundgerüste zielführend ist, wurde der Naturstoff Glycyrrhetinsäure (GS, **95**, Abb. 61) als Suchmotiv verwendet. GS ist ein natürlich vorkommender Hemmstoff der  $11\beta$ -Hydroxysteroid-Dehydrogenase Typ 1 und Typ 2 ( $11\beta$ HSD1 und  $11\beta$ HSD2). Diese Enzyme wandeln Cortison in Cortisol und umgekehrt um. Wie bereits in Kapitel 4.1.2 erwähnt, katalysiert  $11\beta$ HSD1 die Umsetzung inaktiver 11-Ketoglucocorticoide in die entsprechenden aktiven  $11\beta$ -Hydroxysteroiden.  $11\beta$ HSD1 ist ein vielversprechendes Target für die Entwicklung neuer Arzneistoffe für die Behandlung der Fettsucht,<sup>[121]</sup> des metabolischen Syndroms,<sup>[122]</sup> des Typ-2-Diabetes<sup>[123, 124]</sup> und der kognitiven Dysfunktion.<sup>[125]</sup>  $11\beta$ HSD2 stellt ein Anti-Target dar, weil dessen Hemmung zu Natriumretention und Bluthochdruck führt.<sup>[126]</sup> Isoenzymselektivität für  $11\beta$ HSD1 ist daher von herausragender Bedeutung.

Die Analyse der pentacyclischen Struktur des Naturstoffs GS (**95**) führte zu einer klaren Zuordnung des Naturstoffgrundgerüsts zu einem Ast des Naturstoffgrundgerüst-Baums (siehe Abb. 61). Brachiation in Richtung geringerer Komplexität, d.h. wurzelwärts, führte zu einer Gruppe von Zwei-, Drei- und Vier-Ringsystemen. Diese Grundgerüste wurden gewählt, da sie am häufigsten in Naturstoffen vertreten sind (s.o.).

Für die Auswahl des konkreten Grundgerüsts wurde als zweites entscheidendes Kriterium die strukturelle Ähnlichkeit zwischen möglichen Target-Proteinen herangezogen. In Kapitel 4.1.2 wurde das Konzept des Proteinstruktur-Ähnlichkeitsclustering (PSSC) als ein abstrahierendes Leitprinzip für die Entwicklung von Verbindungsbibliotheken vorgestellt. Auch der Ähnlichkeitscluster bestehend aus den  $11\beta$ -Hydroxysteroid-Dehydrogenasen, Acetylcholinesterase und Cdc25A-Phosphatase wurde in Kapitel 4.1.2 ausführlich beschrieben. Der Naturstoff Dysidiolid (**24**) ist als Inhibitor der Cdc25A bekannt.<sup>[151]</sup> Dieser Naturstoff enthält das 1,2,3,4,4a,5,6,7-Octahydronaphthalin-Grundgerüst, das auch als mögliches Grundgerüst für den Aufbau einer Verbindungsbibliothek durch Brachiation durch den Naturstoff-Grundstruktur-Baum ausgehend vom Naturstoff GS (**95**) identifiziert wurde (siehe Abb. 61). Hemmstoffe der  $11\beta$ HSD1, die dieses Grundgerüst enthalten, wurden bislang noch nicht beschrieben.

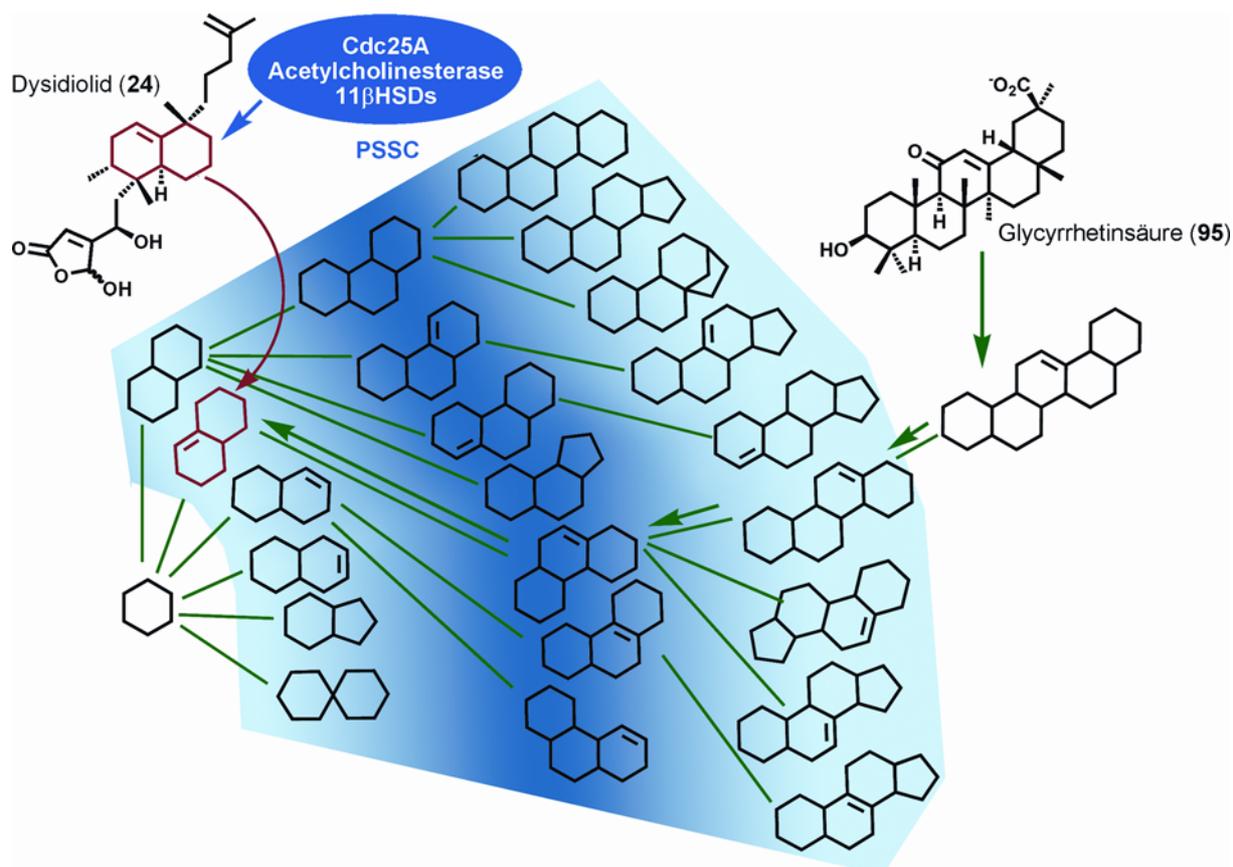


Abb. 61: Strategischer Gebrauch des SCONP-Baums der Naturstoffgrundgerüste. Die Häufigkeitsverteilung leitete die Brachiation vom komplexen pentacyclischen Start-Grundgerüst der GS (**95**) in Richtung geringerer Komplexität, nämlich zu Zwei-, Drei- und Vier-Ring-Naturstoff-Grundgerüsten, die bei Naturstoffen am häufigsten auftreten (die Intensität der blauen Schattierung in der Abbildung stellt die Häufigkeitsverteilung dar: Drei-Ringsysteme sind am häufigsten). Ein zweites, unabhängiges Kriterium (PSSC) führte letztlich zur Auswahl des 1,2,3,4,4a,5,6,7-Octahydronaphthalin-Grundgerüsts als biologisch relevantem Ausgangspunkt für die Generierung einer fokussierten Verbindungsbibliothek.

### ***Synthese einer Dysidiolid-inspirierten Verbindungsbibliothek von Octahydronaphthalinen und Decalinen***

Die strukturelle Exploration der Naturstoffgrundgerüste unter Verwendung von SCONP zusammen mit PSSC als zweitem unabhängigem Kriterium inspirierte die Synthese einer von Dysidiolid (**24**) und GS (**95**) abgeleiteten Bibliothek von 483 Verbindungen mit Octahydronaphthalin- und Decalin-Kernstrukturelement (siehe Abb. 62), die von Dr. Michael Scheck durchgeführt wurde.<sup>[250]</sup>

Verschieden funktionalisierte Octahydronaphthalin- und Decalin-Derivate wurden in Lösung synthetisiert und dann über eine Alkoholfunktion an die feste Phase gebunden. 3,4,8,8a-tetrahydronaphthalin-1,6(2*H*,7*H*)-dione (**96**) wurden durch enantioselektive Robinson-Anellierung dargestellt.<sup>[251, 252]</sup> Diese wurden reduktiv in die entsprechenden 4,4a,5,6,7,8-hexahydro-5-hydroxynaphthalin-2(3*H*)-one (**97**) überführt. Hydrierung der Verbindungen vom Typ **97** führte zu Hydroxydecalinonen (**98**). Die so generierten Grundgerüste, in Abb. 62 am Beispiel der Verbindung **99** gezeigt, wurden über einen Dihydropyranyl-Linker<sup>[253]</sup> an Merrifield-Harz immobilisiert. Die festphasengebundenen Ketone wurden Aldolkondensationsreaktionen mit verschiedenen Aldehyden unterworfen, wodurch exocyclische *E*-konfigurierte Olefine (**100**) entstanden.

Um die Diversität der Bibliothek zu erhöhen, wurden die immobilisierten Aldolkondensationsprodukte unterschiedlich derivatisiert. Wie aus Abb. 62 zu ersehen, umfaßten die Transformationen Sonogashira-, Suzuki- und Heck-Reaktionen, sowie Cu-katalysierte konjugierte Additionsreaktionen, Grignard-Reaktionen, Alkylierungsreaktionen in  $\alpha$ -Position zum Keton, Wittig-Reaktionen und reduktive Aminierungen.

Die Verbindungen wurden mittels Trifluoressigsäure von der festen Phase abgespalten und mittels präparativer HPLC aufgereinigt. Insgesamt wurden 483 Verbindungen im Multi-Milligramm-Maßstab synthetisiert.

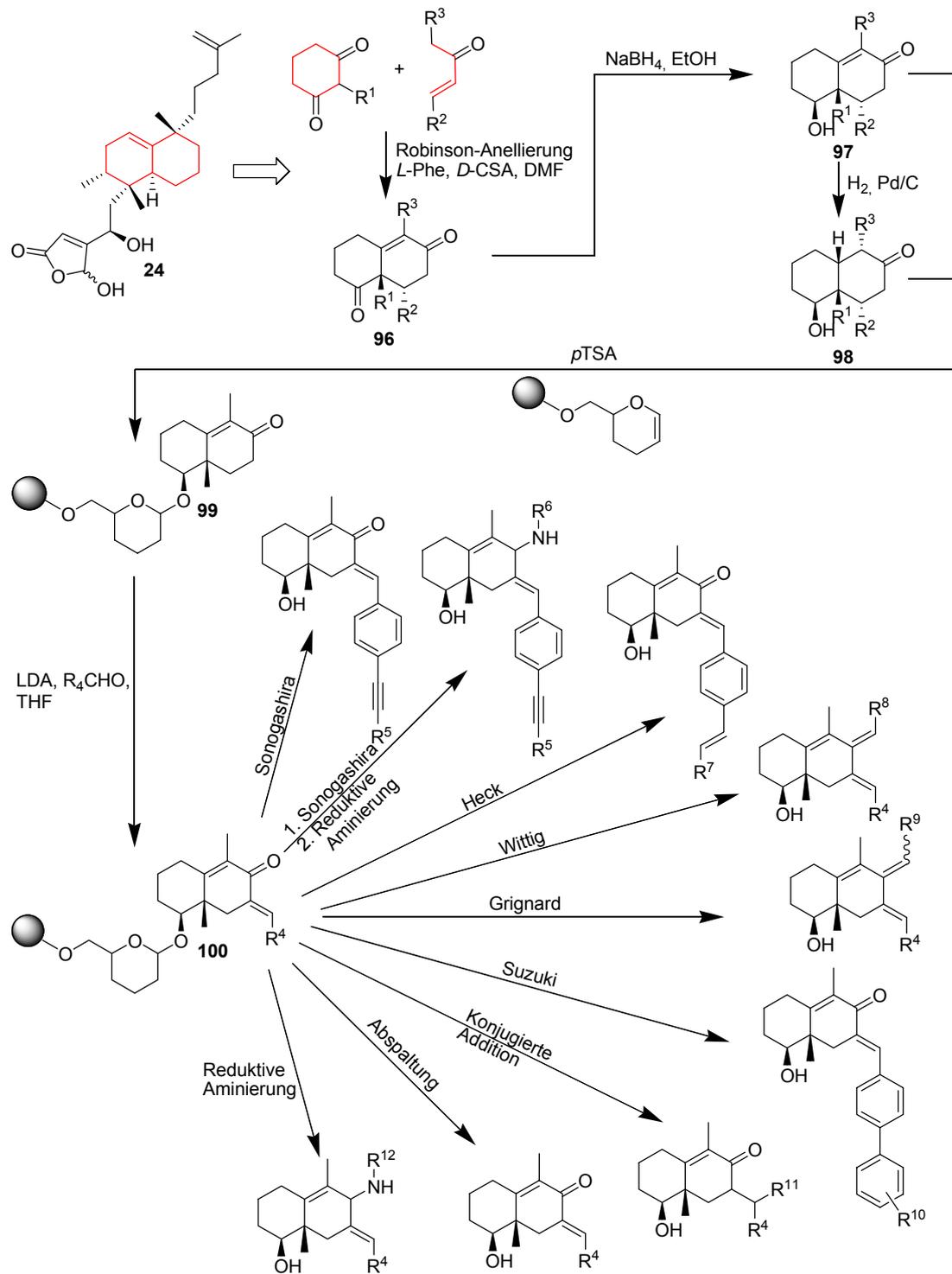


Abb. 62: Kombinierte Fest- und Flüssigphasensynthese einer Dysidiolid-inspirierten Verbindungsbibliothek.

### Biochemische Evaluierung der Verbindungsbibliothek

162 Verbindungen (eine Auflistung aller Strukturen und Inhibitionsdaten findet sich in Anhang II) der Octahydronaphthalin- und Decalin-basierten Bibliothek wurden auf inhibitorische Aktivität hinsichtlich der  $11\beta$ -Hydroxysteroid-Dehydrogenasen Typ 1 und Typ

2 getestet (in Zusammenarbeit mit PD Dr. Alex Odermatt, Universität Bern, Schweiz). Zudem wurden die Verbindungen gegen die übrigen Enzyme des Proteinstruktur-Ähnlichkeitsclusters, Cdc25A und AChE, profiliert. Die entsprechenden Assayverfahren wurden bereits in Kapitel 4.1.2 ausführlich diskutiert. Verbindungen mit  $IC_{50}$ -Werten  $\leq 10 \mu\text{M}$  wurden als Hits erachtet.

### ***Zellbiologische Charakterisierung***

Die neue Inhibitorklasse sollte hinsichtlich zellulärer Aktivität geprüft werden. Hierzu wurden anhand eines repräsentativen Beispiels Translokations- und Transaktivierungs-Assays von PD Dr. Alex Odermatt, Universität Bern, Schweiz, durchgeführt.

#### *Translokations-Assay*

HEK-293-Zellen, die keine  $11\beta$ -Hydroxysteroid-Dehydrogenasen und nur in sehr geringem Maße den Glucocorticoidrezeptor (GR) endogen exprimieren,<sup>[254]</sup> wurden mit einem Expressionsplasmid für das Fusionskonstrukt aus GFP (grün-fluoreszierendes Protein) und GR (Glucocorticoidrezeptor) und entweder humanem  $11\beta$ HSD1- oder leerem pcDNA3-Vektor transfiziert.<sup>[255, 256]</sup> Die Zellen wurden mit Inhibitor inkubiert, gefolgt von Zugabe von Cortison und weiterer Inkubation. Die Lokalisierung von GFP-GR in der Zelle (überwiegend zytoplasmatisch oder nukleär) wurde mittels Fluoreszenzmikroskopie bestimmt. Drei unabhängige Transfektionsexperimente wurden durchgeführt, wobei je Probe 300 fluoreszierende Zellen durch einen gegenüber der Zellvorbehandlung verblindeten Beobachter ausgewertet wurden.

#### *Transaktivierungs-Assay (GR-abhängig)*

HEK-293-Zellen wurden mit pMMTV-LacZ-Reporterplasmid, pCMV-LUC-Kontrollplasmid, GR-Expressionsvektor und entweder  $11\beta$ HSD1- oder leerem pcDNA3-Vektor transfiziert. Die Zellen wurden sowohl mit und ohne Steroidhormon und mit oder ohne  $11\beta$ HSD1-Inhibitor oder GR-Antagonist inkubiert, anschließend lysiert und mit dem Luciferase-Reporter-Assay-System (Promega) und dem Galacto-Light Plus Chemolumineszenz-Reporter-Assay-Kit (Tropix) analysiert. Die  $\beta$ -Galactosidaseaktivität wurde auf die interne Luciferasekontrolle normalisiert. Die Daten (Mittelwert  $\pm$  Standardabweichung) wurden als

Prozentwert in bezug auf die Kontrolle in Gegenwart von Steroid und Abwesenheit von Inhibitor ausgedrückt und ergaben sich aus drei unabhängigen Experimenten.

### ***Ergebnisse und Diskussion***

Die Verbindungskollektion enthielt 30 11 $\beta$ HSD1-Inhibitoren, die IC<sub>50</sub>-Werte zwischen 0.31  $\mu$ M und 9.1  $\mu$ M aufwiesen. Vier der Inhibitoren (**101-104**) hemmten die 11 $\beta$ HSD1 im nanomolaren Bereich (IC<sub>50</sub>-Werte 0.31-0.74  $\mu$ M). Drei Verbindungen inhibierten die 11 $\beta$ HSD2 mit IC<sub>50</sub>-Werten von 2.0-6.6  $\mu$ M. Die wichtigsten Ergebnisse sind in Tabelle 4 zusammengefaßt. Besonders hervorzuheben ist, daß die Hits insgesamt einen hohen Grad an Selektivität aufwiesen. 28 von 30 11 $\beta$ HSD1-Hits hemmten selektiv die 11 $\beta$ HSD1, sowohl im Hinblick auf das Isoenzym 11 $\beta$ HSD2 als auch auf die übrigen PSSC-Proteine, Cdc25A und AChE. Die Verbindungskollektion lieferte darüber hinaus 12 Hits im Cdc25A-Assay mit IC<sub>50</sub>-Werten von 1.2-9.2  $\mu$ M. Zwei Verbindungen hemmten die AChE mit IC<sub>50</sub>-Werten von 3.7  $\mu$ M und 5.3  $\mu$ M. Dieses Ergebnis steht im Einklang mit der gemeinsamen Klassifizierung dieser Proteine in einem Proteinstruktur-Ähnlichkeitscluster (siehe Kapitel 4.1.2).

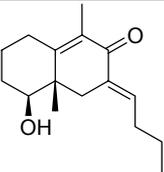
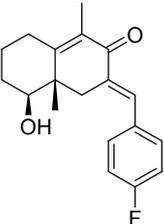
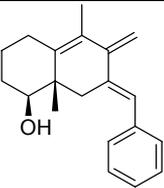
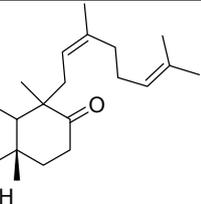
Vbdg	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
101		15 $\pm$ 6	5.3 $\pm$ 1.1	0.31 $\pm$ 0.03	6.6 $\pm$ 0.9
102		> 100	> 20	0.74 $\pm$ 0.11	> 30
103		> 100	> 20	0.35 $\pm$ 0.04	> 30
104		> 100	> 20	0.63 $\pm$ 0.12	> 30

Tabelle 4: Synopse der Inhibitionsdaten für ausgewählte Verbindungen.

Die zelluläre Aktivität dieser neuen Strukturklasse von Inhibitoren wurde anhand von Verbindung **103** als einem der wirksamsten und selektivsten 11 $\beta$ HSD1-Inhibitoren nachgewiesen. Im Translokations-Assay ist der GR in Abwesenheit von Cortisol zytosolisch lokalisiert. Wenn Cortison, das eine sehr geringe Bindungsaffinität zum GR hat,<sup>[257]</sup> durch Reduktion über die 11 $\beta$ HSD1 in den potenten GR-Agonisten Cortisol überführt wird, bindet das so generierte Cortisol an den GR, induziert dadurch die Translokation des Rezeptors in den Zellkern und führt auf diese Weise zu einer Stimulation der Transaktivierung.

Wie die Abbildungen 63 und 64A zeigen, wurde nach Zugabe von Cortison zu Zellen, die 11 $\beta$ HSD1 exprimieren, eine dosisabhängige Induktion der Translokation des GR in den Zellkern und Stimulierung der Transaktivierung beobachtet. Die Glucocorticoid-abhängige nukleäre Translokation des GR und Transaktivierung wurden durch den unspezifischen 11 $\beta$ HSD-Hemmstoff Glycyrrhetinsäure (GS, **95**, Abb. 61) blockiert. Sowohl nukleäre Translokation als auch GR-abhängige Transaktivierung fanden nach gemeinsamer Inkubation der Zellen mit Cortison und dem selektiven 11 $\beta$ HSD1-Inhibitor **103** nicht statt. Damit wurde nachgewiesen, daß **103** die Konversion von Cortison zu Cortisol in intakten Zellen zu

hemmen vermochte. In der Gegenwart von 3  $\mu\text{M}$  der Verbindung **103** waren ca. 50% der GR-Moleküle im Kern lokalisiert (siehe Abb. 63B), und die Transaktivierung durch GR auf ungefähr 40% vermindert. Die nukleäre Translokation und die Transaktivierung waren bei einer Konzentration der Verbindung **103** von 10  $\mu\text{M}$  blockiert.

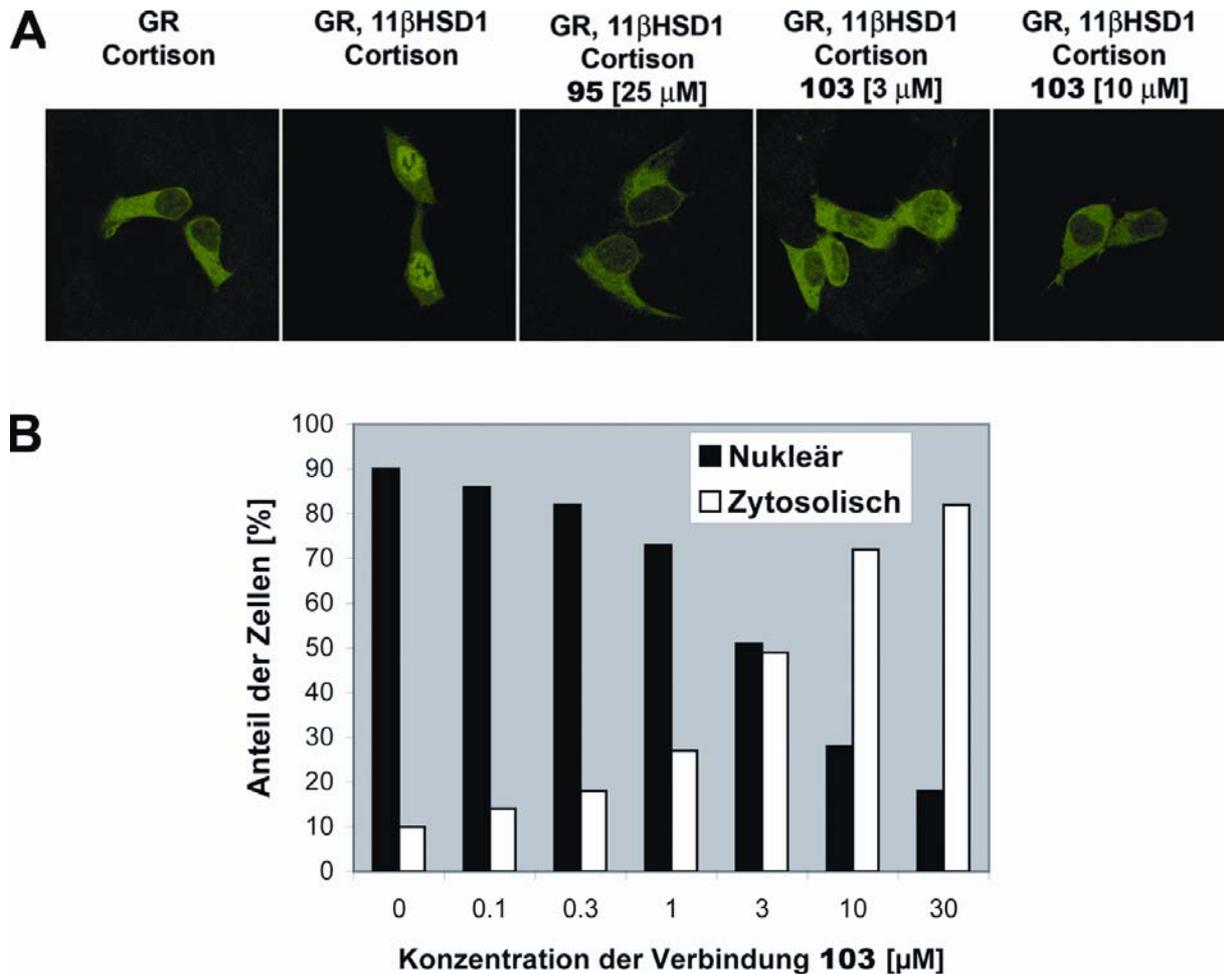


Abb. 63

A) Inhibition der 11 $\beta$ HSD1-abhängigen Glucocorticoidaktivierung und nukleären Translokation des GR. HEK-293-Zellen wurden mit GFP-GR allein oder zusammen mit 11 $\beta$ HSD1 transfiziert. Die Zellen wurden 6 h nach der Transfektion in Steroid-freiem Medium 18 h inkubiert. Die Zellen wurden 30 min mit 25  $\mu$ M Glycyrrhetinsäure (GS, **95**) bzw. 3  $\mu$ M oder 10  $\mu$ M des Inhibitors **103** präinkubiert. Daraufhin erfolgte eine weitere Inkubation für 40 min mit 500 nM Cortison. GFP-GR wurde mittels Fluoreszenzmikroskopie detektiert. Die 11 $\beta$ HSD1 vermittelte die nukleäre Translokation des GR, indem sie inaktives Cortison in aktives Cortisol überführte. Eine Hemmung der 11 $\beta$ HSD1 verhinderte die Bildung von Cortisol und die nukleäre Translokation des GR.

B) Mit Expressionsplasmiden für GFP-GR und 11 $\beta$ HSD1 transfizierte HEK-293-Zellen wurden wie unter A beschrieben mit verschiedenen Konzentrationen der Verbindung **103** behandelt. Die Zellen wurden fixiert und die Lokalisierung von GFP-GR mittels Fluoreszenzmikroskopie detektiert. Die Zellen wurden zwei Phänotypen zugeordnet: überwiegend nukleäre versus überwiegend zytosolische Lokalisierung von GFP-GR. Die gezeigten Ergebnisse wurden aus unabhängigen Transfektionsexperimenten erhalten und sind als Prozentsatz bezogen auf die Zahl fluoreszierender Zellen dargestellt. Wenigstens 300 fluoreszierende Zellen wurden pro Konzentration und Versuch ausgewertet.

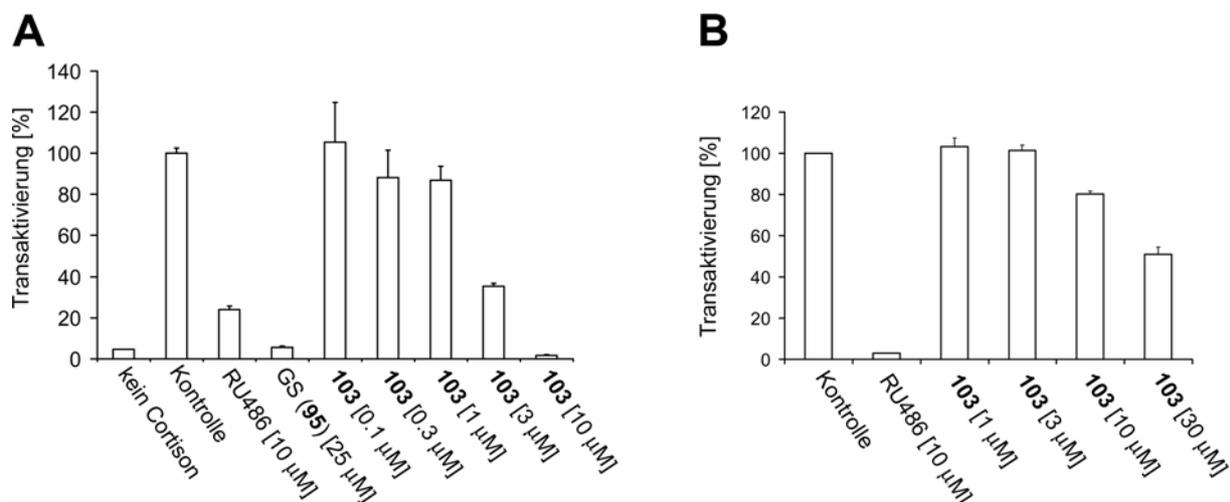


Abb. 64

Effekt des  $11\beta$ HSD1-Hemmers **103** auf die GR-abhängige Transaktivierung. HEK-293-Zellen wurden mit pMMTV-LacZ-Reporterplasmid, pCMV-LUC-Kontrollplasmid, GR-Expressionsvektor und entweder  $11\beta$ HSD1-(A)- oder leerem pcDNA3-(B)-Vektor transfiziert. Die Zellen wurden 6 h nach der Transfektion in Steroid-freiem Medium, das den jeweils angegebenen Inhibitor oder Antagonist enthielt, inkubiert, gefolgt von der Zugabe von entweder keinem Steroid (kein Cortisol), 500 nM Cortisol (A) oder 100 nM Cortisol (B) und weiterer Inkubation für 24 h. Die Galactosidase-Reporteraktivität wurde auf die interne Luciferasekontrolle normalisiert. Die Daten stellen Prozentwerte in bezug auf die Kontrolle in Gegenwart von Steroid und Abwesenheit von Inhibitor dar. Sie wurden aus drei unabhängigen Experimenten ermittelt und sind als Mittelwert mit Standardabweichung dargestellt.

Das Experiment B diente als Kontrolle zur Prüfung, ob der Effekt von **103** primär auf eine Hemmung der  $11\beta$ HSD1 und nicht auf GR-Antagonismus zurückzuführen war. Die Zellen wurden daher in Abwesenheit von  $11\beta$ HSD1 mit 100 nM Cortisol und verschiedenen Konzentrationen von **103** inkubiert. Als Kontrolle wurde der bekannte GR-Antagonist RU486 verwendet, der einen vollständigen Verlust der Transaktivierung bei einer Konzentration von 10  $\mu$ M bewirkte. Ein schwacher antagonistischer Effekt von **103** wurde beobachtet, wobei die halbmaximale Hemmung der Transaktivierung bei einer Konzentration von 30  $\mu$ M beobachtet wurde. Die antagonistischen Effekte traten aber erst bei etwa zehnfach höheren Konzentrationen bezogen auf die Hemmung der  $11\beta$ HSD1 auf.

Mit diesen Ergebnissen konnte der Nachweis erbracht werden, daß die in dieser Arbeit vorgestellte strukturelle Klassifizierung der Naturstoffe (SCONP) ein nützliches Leitprinzip für die Kartierung des chemischen Strukturraums, wie er für die Natur relevant ist und von dieser genutzt wird, darstellt. Die Anwendung von SCONP als Hypothesengenerator und Navigator für das Design von Naturstoff-abgeleiteten Verbindungskollektionen – vor allem synergistisch mit Proteinstruktur-Ähnlichkeitsclustering (PSSC) als zweitem Leitprinzip – kann zu Verbindungskollektionen führen, die als biologisch prävalidiert erachtet werden können. Diese liefern qualitativ hochwertige Proteinliganden bei vergleichsweise kleiner Bibliotheksgröße.

## 5 Zusammenfassung und Ausblick

In dieser Arbeit wurden strukturbasierte Prinzipien der „Proteinwelt“ und der „Naturstoffwelt“ genutzt, um Leitlinien für eine gerichtete Konzipierung von Verbindungsbibliotheken zu entwickeln.

Das PSSC-Konzept, das im ersten Teil dieser Arbeit beschrieben wird, beruht auf einer rein strukturorientierten Sicht auf Proteine. Es wurde eine Vorgehensweise entwickelt, mit der strukturelle Ähnlichkeiten in Proteinen identifiziert wurden. Dabei wurde der Domänenarchitektur der Proteine Rechnung getragen. Gegenstand für die vergleichende Betrachtung im Hinblick auf die Entwicklung niedermolekularer Proteinliganden waren die katalytisch bzw. für die Ligandenbindung relevanten Proteindomänen. Wurden Proteindomänen deutlich unterschiedlicher Größe miteinander verglichen, wurde ermittelt, ob wesentliche Strukturelemente der kleineren Domäne in der größeren zu finden waren. War dies der Fall, wurde untersucht, ob das Strukturmotiv der kleineren Proteindomäne auch den katalytischen bzw. ligandenbindenden Kernbereich der größeren Domäne beschreibt. Die vergleichende Beurteilung der Ähnlichkeit wurde auf der Basis von Überlagerungen der durch die Position der  $C^\alpha$ -Atome beschriebenen Proteinrückgrate vorgenommen. Dabei wurden visuelle Mustererkennung und die mittlere quadratische Abweichung der  $C^\alpha$ -Positionen als quantitatives Maß zur Beurteilung der Ähnlichkeit herangezogen. Zeigten Proteindomänen bzw. ihre Kernbereiche hinreichende strukturelle Ähnlichkeit, wurden sie unabhängig von Sequenzähnlichkeit und damit evolutionärer oder funktioneller Verwandtschaft, also rein strukturbasiert, in einen Ähnlichkeitscluster gruppiert. Die Grundgerüste bekannter Liganden eines Vertreters der einem derartigen Ähnlichkeitscluster zugeordneten Proteine wurden als Leitprinzipien für den Aufbau fokussierter Verbindungsbibliotheken gewählt, mit der selektiv die verschiedenen Proteine des Ähnlichkeitsclusters angesprochen werden sollten. Der Ligand eines Proteins stellte eine biologisch prävalidierte Startstruktur im chemischen Raum dar, von der aus Liganden für die übrigen, strukturell verwandten Proteine entwickelt werden konnten.

Dies wurde am Beispiel des Proteinstruktur-Ähnlichkeitsclusters der Cdc25A-Phosphatase (Cdc25A), der Acetylcholinesterase (AChE) und der  $11\beta$ -Hydroxysteroid-Dehydrogenasen ( $11\beta$ HSDs) gezeigt. Hier wurde ein bekannter Naturstoff-Ligand der Cdc25A, das Dysidiolid (**24**, Abb. 65), als Leitmotiv für die Synthese einer Verbindungskollektion genutzt. Diese wurde gegen die Mitglieder des Clusters gescreent. Die Bibliothek lieferte zahlreiche Inhibitoren für die Cdc25A, die mehrheitlich deutlich aktiver waren als Dysidiolid

(die potenteste Verbindung **25** ist in Abb. 65 zu sehen). Dieses Ergebnis lag im Rahmen der Erwartungen, nachdem ein bekannter Cdc25A-Hemmstoff als Ausgangspunkt verwendet wurde. Interessanterweise konnten in der Verbindungskollektion aber auch Inhibitoren der AChE und der pharmazeutisch relevanten 11 $\beta$ HSD1-Isoform mit einer Trefferrate von 2-3% identifiziert werden. Diese Enzyme sind zwar strukturell in ihren Kernbereichen, jedoch nicht sequentiell mit der Cdc25A verwandt. Die identifizierten Hemmstoffe zeigten teilweise eine ausgeprägte Selektivität für einzelne Mitglieder des Ähnlichkeitsclusters. Besonders bemerkenswert ist hier das  $\alpha,\beta$ -ungesättigte Lacton **30**, das nur die therapeutisch relevante 11 $\beta$ HSD1, aber nicht oder nur schwach die übrigen Enzyme des Ähnlichkeitsclusters hemmt. Selektive AChE-Inhibitoren wurden nicht identifiziert.

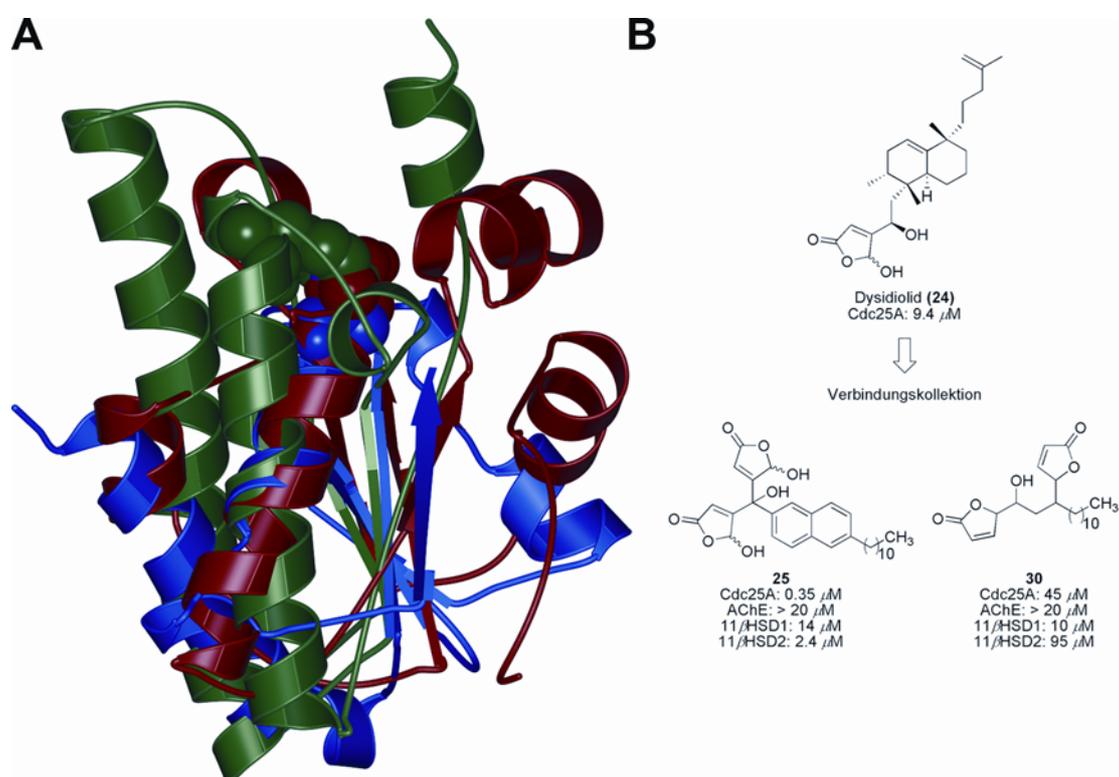


Abb. 65: A) Proteinstruktur-Ähnlichkeitscluster bestehend aus Cdc25A, AChE und 11 $\beta$ HSD1. B) Vertreter der vom Naturstoff Dysidiolid (**24**) abgeleiteten Verbindungskollektion.

Mit diesem Beispiel konnte zum ersten Mal gezeigt werden, daß ein Clustering von Target-Proteinen nach ausschließlich strukturellen Gesichtspunkten in Verbindung mit Naturstoff-inspirierten Verbindungskollektionen ein wertvolles Leitkonzept bei der gezielten Entwicklung von Proteinliganden darstellt. Dies gilt umso mehr, als die in diesem Projekt identifizierten Strukturtypen von Liganden bislang noch nicht in der Literatur als Hemmstoffe der jeweiligen Enzyme beschrieben wurden.

Darüber hinaus wurde ein weiterer Ähnlichkeitscluster bestehend aus Ornithindecaboxylase (ODC) und Dihydropteroatsynthase (DHPS) identifiziert (siehe Abb. 66A). Ausgehend von einer kleinen Kollektion von *p*-Aminobenzolsulfonamiden, die eine bekannte Klasse von antibakteriellen Chemotherapeutika darstellen und deren molekulares Target die DHPS ist, wurde ein mikromolarer Inhibitor der humanen und der *Trypanosoma*-ODC (**51**, Abb. 66B) gefunden, der bislang noch nicht beschrieben worden war. Anhand dieses Beispiels wurde gezeigt, daß nicht nur Naturstoffe, sondern auch bekannte Arzneistoffe oder sogenannte „privilegierte Strukturen“ als Ausgangspunkte für die Konzipierung biologisch relevanter Verbindungskollektionen im Rahmen des PSSC-Konzepts herangezogen werden können.

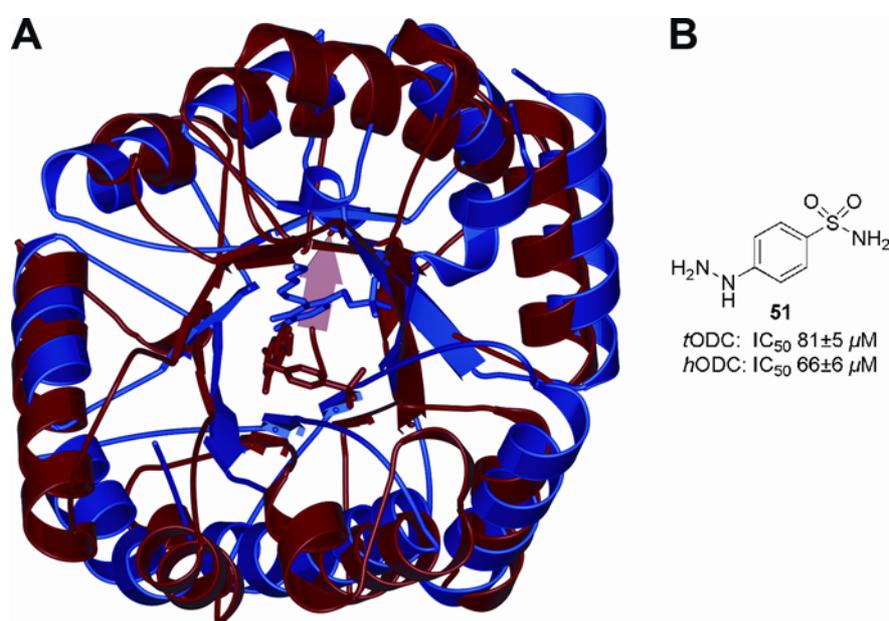


Abb. 66: A) Proteinstruktur-Ähnlichkeitscluster bestehend aus ODC und DHPS. B) Identifizierter ODC-Inhibitor.

Im Rahmen konzeptioneller Arbeiten wurde der Proteinstruktur-Ähnlichkeitscluster bestehend aus Cdc25A, AChE und den 11βHSDs um die Pankreaslipase und die hAPT1 erweitert sowie weitere Ähnlichkeitscluster identifiziert. Diese müssen jedoch noch in Ligandenentwicklungsprojekten experimentell verifiziert werden.

Über die weitere experimentelle Evidenz hinaus ist zu prüfen, wie groß der Anwendungsbereich des PSSC-Konzepts ist. Hierzu sollte eine Analyse der Proteinstrukturen des gesamten bekannten Proteoms erfolgen. Dies würde eine umfassende Auswertung der Daten und die Modellierung von Proteinstrukturen aus der Sequenz erfordern, sofern keine Strukturinformationen vorliegen. Dieser Vorgang ist nur automatisiert sinnvoll durchführbar. Für eine derartige Weiterentwicklung des PSSC-Konzepts ist die Verarbeitung großer

Datenmengen und damit eine verstärkte Integration von Methoden der (Bio-)Informatik unabdingbar.

Der zweite Teil der Arbeit beinhaltet eine komplementäre Betrachtung des chemischen Strukturraums der Naturstoffe. Naturstoffe sind Produkte von Enzymen und wurden im Laufe der Evolution in der Regel im Hinblick auf eine bestimmte Funktion, die meist ebenfalls über Interaktionen mit Proteinen vermittelt werden, selektioniert. Naturstoffstrukturen stellen daher inhärent „privilegierte“<sup>[14]</sup> Motive dar. Sie stecken den von der Natur genutzten Teil des chemischen Raums kleiner organischer Moleküle ab. Im Rahmen dieser Arbeit wurden die Gerüststrukturen der Naturstoffe untersucht. Es entstand ein strukturbasiertes Ordnungsprinzip der Naturstoffe, das als Navigator durch den Naturstoffstrukturraum fungieren kann. Die Naturstoffgrundgerüste wurden hierarchisch mit zunehmender Gerüstgröße und damit Komplexität in einem Baumdiagramm klassifiziert. Dazu wurden strukturbasierte Genealogien der Naturstoffgrundgerüste erstellt und miteinander korreliert. Jedes Naturstoffgrundgerüst wurde so auf einen Monocyclus als Grundtaxon zurückgeführt. So entstand eine Art phylogenetischer Baum, in dem die Naturstoffgrundgerüste entsprechend ihrer strukturellen Verwandtschaft und mit zunehmender Anzahl der Ringe im Grundgerüst geordnet wurden. Jeder Knoten im Baumdiagramm stellt eine bestimmte Grundgerüststruktur dar, von der aus eine weitere Verzweigung zu abgeleiteten, komplexeren Gerüststrukturen führen kann. Diese grundgerüstbasierte strukturelle Klassifizierung der Naturstoffe (‘Structural Classification Of Natural Products’, SCONP) wurde an jedem Knoten mit dem biologischen Quellorganismus und dem biologischen Effekt, soweit bekannt, annotiert. Die Anwendbarkeit von SCONP als Hypothesen- und Ideengenerator bei der Konzipierung von Verbindungsbibliotheken wurde experimentell verifiziert.

Unter Anwendung von SCONP in Verbindung mit PSSC als zweitem Kriterium konnten aus dem komplexen Naturstoff und unselektiven  $11\beta$ HSD-Hemmstoff Glycyrrhetinsäure (**95**, Abb. 67) strukturell vereinfachte, potente und hochselektive  $11\beta$ HSD1-Inhibitoren entwickelt werden. Das Grundgerüst der GS (**95**) wurde zunächst strukturell in den SCONP-Baum eingeordnet. Anschließend führte wurzelwärts gerichtete Brachiation im SCONP-Baum zu einer Gruppe von Zwei-, Drei-, und Vierring-Grundgerüsten, die in Naturstoffen am häufigsten vertreten sind. Aus dieser Untermenge der Naturstoffgrundgerüste wurde unter Einbeziehung eines zweiten, unabhängigen Kriteriums, PSSC, das Octahydronaphthalin-Grundgerüst ausgewählt, das auch im Dysidiolid zu finden ist. Eine Kollektion von 162 vom Dysidiolid (**24**) abgeleiteten Verbindungen wurde hinsichtlich einer Hemmung der  $11\beta$ HSDs

biochemisch evaluiert. In dieser Verbindungsbibliothek wurden insgesamt 30  $11\beta$ HSD1-Hemmstoffe identifiziert. Vier Inhibitoren hemmten die  $11\beta$ HSD1 im nanomolaren Bereich. Am Beispiel einer der aktivsten und selektivsten Verbindungen (**103**, Abb. 67) wurde die zelluläre Aktivität dieser neuen  $11\beta$ HSD1-Strukturklasse nachgewiesen. Es ist insbesondere hervorzuheben, daß im Zuge dieser Arbeiten die komplementären Welten der Naturstoffe und der Proteine durch den synergistischen Gebrauch von SCONP und PSSC bei der Konzipierung von Verbindungsbibliotheken zusammengeführt wurden.

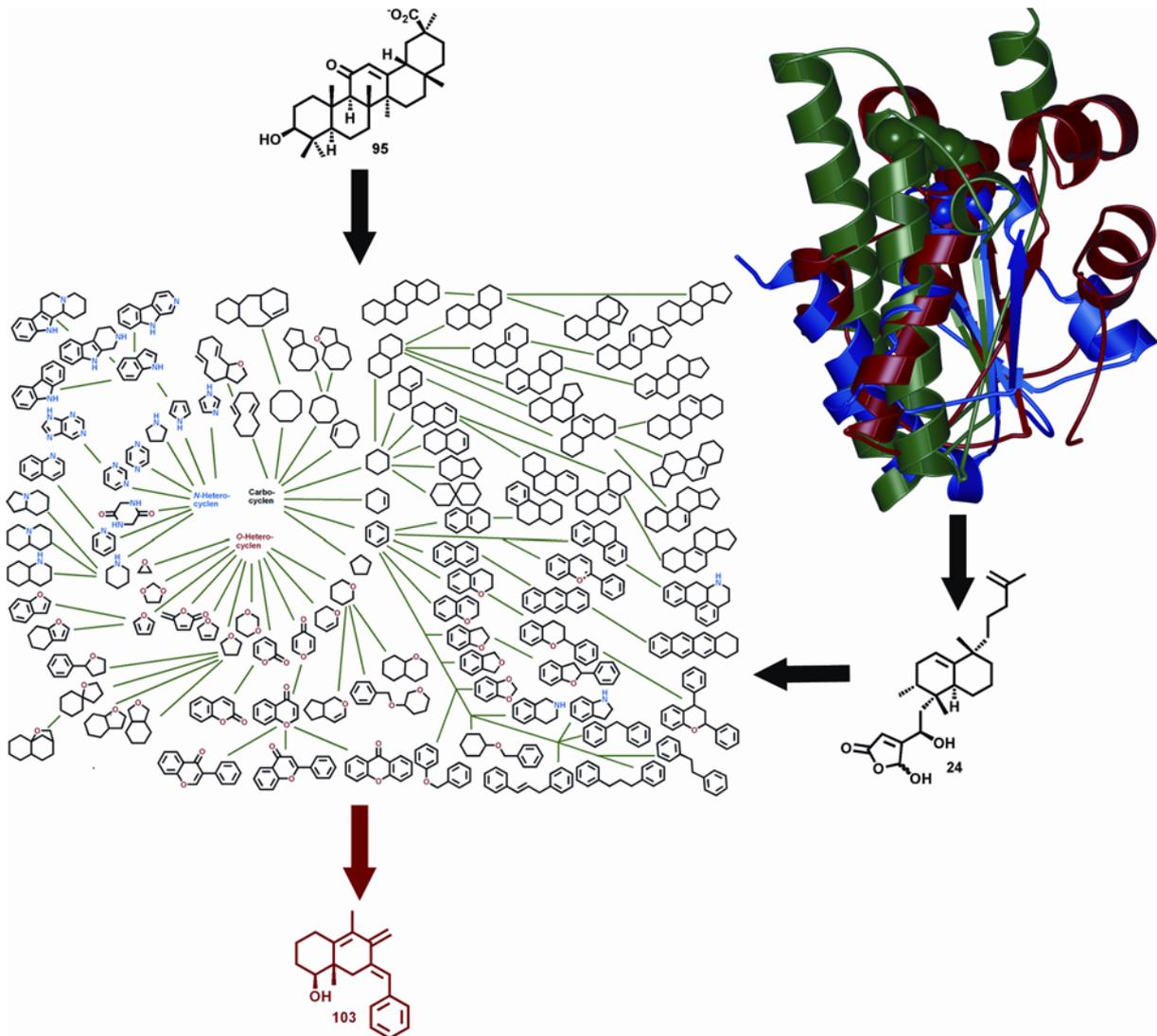


Abb. 67: Synergistische Anwendung von SCONP und PSSC im Rahmen der Entwicklung von  $11\beta$ HSD1-Inhibitoren. Ausgehend vom komplexen Naturstoff GS (**95**) wurde auf der Grundlage einer SCONP-Analyse und von PSSC als zweitem, unabhängigem Kriterium das OctahydronaphthalinGrundgerüst, das auch im Naturstoff Dysidiolid (**24**) vorkommt, als biologisch relevanter Startpunkt für die Entwicklung einer fokussierten Verbindungsbibliothek ausgewählt. In dieser Bibliothek wurden 30  $11\beta$ HSD1-Inhibitoren identifiziert. Verbindung **103** war einer der potentesten und selektivsten Hemmstoffe und zeigte zudem zelluläre Aktivität.

Die in dieser Arbeit entwickelten Konzepte könnten vor allem in der medizinisch-chemischen Forschung hilfreich sein, insbesondere in dem noch jungen und sich sehr schnell

entwickelnden Feld der Chemogenomik, in dem chemische Verbindungen als Werkzeuge dazu genutzt werden, biologische Systeme zu verstehen. Eine enger gefaßte Definition des Begriffs Chemogenomik ist die Identifizierung von Leitstrukturen, die es als Modulatoren ermöglichen, die Funktion eines Genprodukts oder von Produkten einer Genfamilie zu erforschen. Bei der Verfolgung dieses Ansatzes herrscht derzeit die Klassifizierung von Produkten einer Genfamilie auf der Basis von Sequenz und Funktion (z.B. Bildung von Target-Familien und -Subfamilien im Bereich der Kinasen, Phosphatasen, Proteasen usw.) vor (siehe Abb. 68). Ein rein strukturbasierter Ansatz der Gruppierung potentieller Target-Proteine wie PSSC könnte in diesem Zusammenhang ergänzend oder als Alternative einen wertvollen Beitrag leisten, da prinzipiell eine Kenntnis der Funktion oder des Target-Kontexts nicht vorausgesetzt wird. Die Betrachtung der Proteinstrukturen erweitert den Erkenntnisraum, zumal diese im Laufe der Evolution stärker konserviert sind als die sie determinierenden Sequenzen. Diese Sequenzen können so verschieden sein, daß durch sequenzbasierte Ansätze in den Fällen niedriger oder keiner Sequenzhomologie strukturelle Verwandtschaftsverhältnisse nicht detektierbar sind. PSSC stellt somit gleichsam eine Erweiterung des Spektrums der sequenzgetriebenen Methoden dar (siehe Abb. 41). Der Nachteil im Falle von PSSC, auf Strukturinformation angewiesen zu sein, wird zumindest teilweise dadurch relativiert, daß für eine prinzipielle Einordnung strukturell ähnlicher Proteine in einen Cluster im allgemeinen die nach dem Stand der Technik erstellten Strukturmodelle ausreichen.

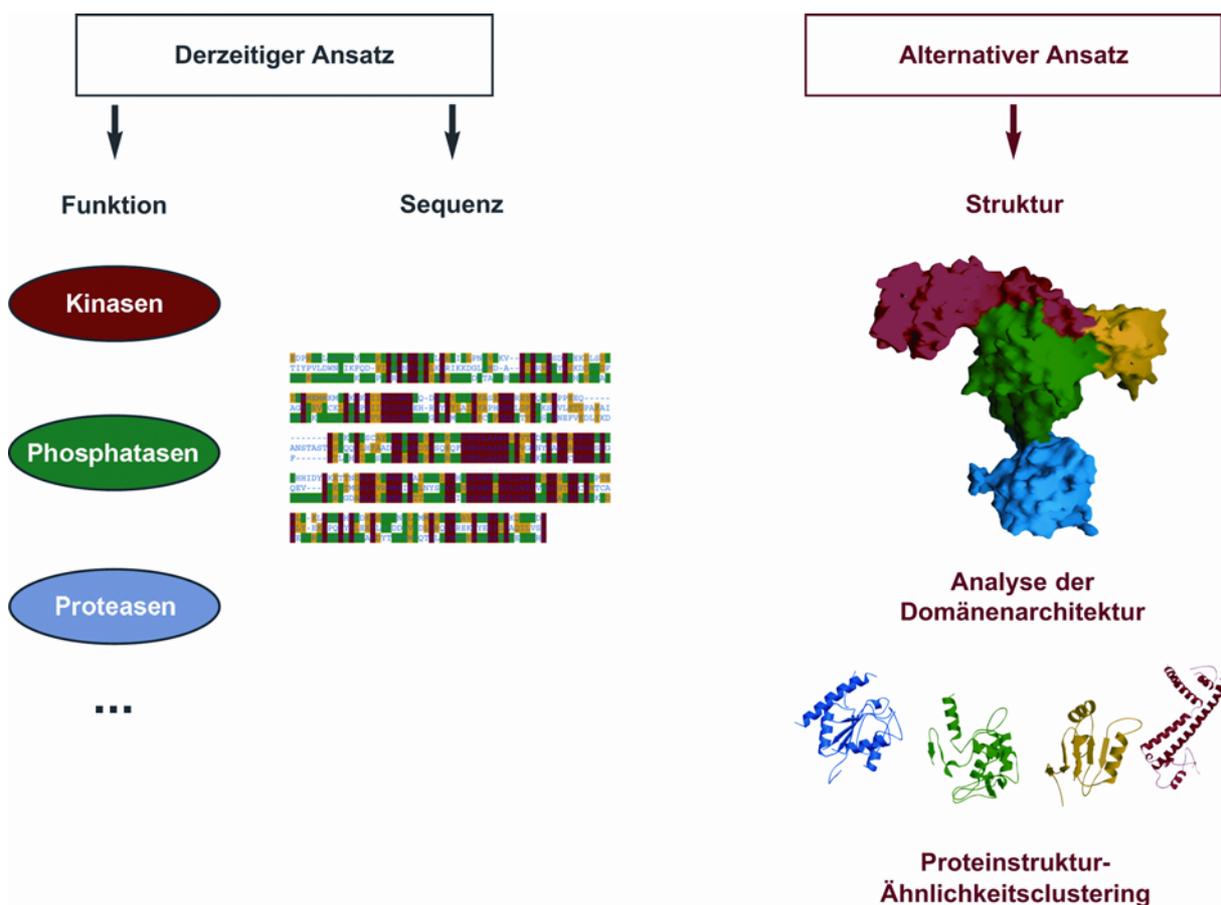


Abb. 68: Vorgehensweisen bei der Kategorisierung von Proteinen. Der derzeit im Rahmen der Chemogenomik vorherrschende Ansatz, Proteine nach ihrer Funktion und/oder Sequenz zu sogenannten Target-Familien zusammenzufassen, auf der einen Seite könnte durch eine komplementäre Betrachtung der Domänenarchitektur und eine Analyse struktureller Ähnlichkeit von ligandenbindenden Kernbereichen im Rahmen des Proteinstruktur-Ähnlichkeitsclustering-(PSSC)-Konzepts auf der anderen Seite erweitert werden.

Die komplementäre abstrahierende Kartierung des Strukturraums der Naturstoffe durch SCONP erlaubt die Navigation durch diese biologisch relevanten, von der Natur eingenommenen Gefilde des chemischen Raums. Auf hohem Abstraktionsniveau können ausgehend von komplizierten Naturstoffgrundgerüsten einfachere und strukturell innovative Lösungen unter Retention der biologischen Aktivität abgeleitet werden.

Beide Konzepte, PSSC und SCONP, beinhalten durch das hohe Abstraktionsniveau zwar eine gewisse Unschärfe, die durch die kombinatorische Bearbeitung der als vielversprechend identifizierten Grundgerüste jedoch aufgefangen wird. Diese muß erfolgen, um die biologische Diversität der Bindungstaschen mit einer geeigneten chemischen Diversität als Voraussetzung für Selektivität und Wirksamkeit ansprechen zu können.

PSSC oder SCONP, einzeln oder synergistisch angewandt, können vor allem in frühen Entwicklungsphasen, wenn nur wenig über die Funktion oder den biologischen Kontext eines potentiellen Protein-Targets bekannt ist, als rationale Ansätze hilfreich sein. Es können dabei

neue Verbindungsklassen gezielter identifiziert werden, mit denen die Biologie des potentiellen Targets erforscht werden kann. Es bleibt festzuhalten, daß die so identifizierten Modulatoren im Rahmen von medizinalchemischen Programmen in der Regel optimiert werden müssen, um die Selektivität zu erhöhen oder unerwünschte Wirkungen zu minimieren. Dennoch stellt bereits die Auffindung einer geeigneten Startstruktur für diese Optimierung einen großen Schritt nach vorne dar.

## 6 Experimenteller Teil

### 6.1 Bioinformatik und Molecular Modeling

#### 6.1.1 Allgemeine Arbeitsvorschriften (AAV)

Die verwendeten Skript- und Pfadbezeichnungen beziehen sich auf das Netzwerk des Max-Planck-Instituts für Molekulare Physiologie in Dortmund.

##### *AAV 1: Datenbanksuchstrategie*

Um Einblicke in die strukturelle Verwandtschaft von Proteinen zu erhalten, wurde typischerweise das folgende Zwei-Stufen-Schema angewandt:

- Proteindomänen, deren dreidimensionale Struktur (Röntgen- oder NMR-Struktur) bekannt ist, werden in der SCOP-(,Structural Classification of Proteins')[<sup>29, 30, 87</sup>]-Datenbank (<http://scop.mrc-lmb.cam.ac.uk/scop/>) hierarchisch entsprechend den Taxa Klasse, Faltung, Superfamilie und Familie hierarchisch klassifiziert. Die Klasse beschreibt die wesentlichen Sekundärstrukturelemente und der Faltungstyp geometrische Verwandtschaft. Die letzten beiden Hierarchieebenen, Superfamilie und Familie, beschreiben entfernte und nahe Verwandtschaftsverhältnisse. Im entsprechenden Suchfeld wird der vierstellige PDB-(,Protein Data Bank')-Code (<http://www.rcsb.org/pdb/>) eingegeben, worauf die Klassifizierung des Proteins nach SCOP baumartig ausgegeben wird. Jede Hierarchieebene ist individuell ansteuerbar, so daß man sich ein abgerundetes Bild der benachbarten Familien, Superfamilien und Faltungstypen machen kann. Dies erlaubt eine erste Einordnung der Proteindomäne und bildet einen Ausgangspunkt für die weitere, rein strukturorientierte Betrachtung.
- Die Suche in der Dali/FSSP-(,Fold Classification based on Structure-Structure alignment of Proteins')[<sup>90-94</sup>]-Datenbank (<http://www.ebi.ac.uk/dali/>) erlaubt es, die 3D-Struktur eines Proteins mit allen Proteinstrukturen, die sich derzeit in der PDB befinden, rein strukturbasiert zu vergleichen. Eine weitere Datenbank, die in diesem Kontext genutzt werden kann, ist die CE-(Combinatorial Extension)[<sup>97</sup>]-Datenbank (<http://cl.sdsc.edu/ce.html>). Auch hier können als Suchmotiv entweder ein PDB-Code oder die Raumkoordinaten eines Proteins (entweder experimentellen Ursprungs oder eines Modells) verwendet werden. Als Ergebnis erhält man eine Trefferliste von

strukturell verwandten Proteinen mit sinkender statistischer Signifikanz (dargestellt durch den Z-Score). Diese Liste wird hinsichtlich pharmazeutisch interessanter Fälle ausgewertet. Darüber hinaus wird mit einem Suchergebnis der Recherche in einer Datenbank eine Kreuzrecherche in der jeweils anderen Datenbank durchgeführt. Konvergieren beide Suchstrategien, kann die Relevanz der Suchergebnisse möglicherweise höher eingestuft werden. Insgesamt jedoch müssen die Ergebnisse einer derartigen Suchstrategie kritisch bewertet werden. Dies gilt vor allem dann, wenn die Sequenzidentität unter 10% liegt. Diese Fälle, die sich durch eine geringe Sequenzidentität auszeichnen, sind die interessantesten, da diese lediglich durch eine rein strukturbasierte Herangehensweise und nicht durch sequenzbasierte Algorithmen identifiziert werden können. Dabei muß insbesondere visuell geprüft werden, ob wirklich die für Ligandenbindung relevanten Teile einer Proteindomäne als ähnlich identifiziert und überlagert wurden. Besonders problematisch ist es, wenn Proteindomänen mit signifikant unterschiedlicher Größe miteinander verglichen und als „ähnlich“ klassifiziert werden. Häufig ist es erforderlich, zunächst die ligandenbindende Domäne (bei Proteinen, die aus mehreren Domänen bestehen) bzw. die ligandenbindende Kernstruktur einer Proteindomäne zu definieren (siehe AAV 2).

### ***AAV 2: Editieren einer PDB-Datei – Extraktion der ligandenbindenden (katalytischen) Domäne bzw. Kernstruktur***

Proteine, die aus mehreren Domänen bestehen, müssen hinsichtlich ihrer Domänenarchitektur untersucht werden. Die Domänenarchitektur kann der SCOP- bzw. der CATH<sup>[88]</sup>-Datenbank entnommen werden. Die für die Bindung niedermolekularer Liganden entscheidende ligandenbindende (LD) bzw. katalytische Domäne (KD) muß definiert und extrahiert werden. Werden Proteindomänen von deutlich unterschiedlicher Größe miteinander verglichen, so ist dahingehend zu prüfen, ob die wesentlichen Struktur motive der kleineren Domäne als Substruktur im ligandenbindenden (LK) bzw. katalytischen Kernbereich (KK) der größeren Domäne vorkommen. Für ein optimales Überlagerungsergebnis (siehe AAV 3) ist der LK bzw. KK zu definieren und zu extrahieren.

Zunächst wird die PDB-Datei des zu untersuchenden Proteins z.B. aus der lokalen Kopie der PDB in das Arbeitsverzeichnis kopiert. Dazu wird folgender Befehl auf der UNIX-Kommandozeile eingegeben:

```
getpdb <PDB code>
```

Diese PDB-Datei wird mit einem Editor (z.B. nedit) zur weiteren Bearbeitung geöffnet. Parallel dazu wird die Datei mit dem graphischen Visualisierungsprogramm RASMOL<sup>[258]</sup> geöffnet. Das Protein ist in der „Wireframe“-Darstellung zu sehen. Im RASMOL-Menü kann die Darstellung wie folgt geändert werden:

*Display>Backbone* (Anzeige der Spur der C<sup>α</sup>-Atome) oder

*Display>Ribbons* (Anzeige der Sekundärstrukturelemente)

Sind Liganden in der Proteinstruktur enthalten, werden diese von der Kommandozeile von RASMOL wie folgt selektiert:

```
select ligand
```

Die Darstellung wird wiederum über das *Display*-Menü ausgewählt (entweder *Sticks* (Stabdarstellung) oder *Spacefill* (CPK-Darstellung)).

Im Kopf der PDB-Datei befinden sich wichtige Informationen, allen voran die Primärliteratur in der JRNL-Sektion, die das Experiment zur Proteinstruktur beschreibt. Diese sollte konsultiert werden, um weitergehende Informationen zum Protein zu erhalten. Ebenso von Interesse ist die „Heterogen“-Sektion (HET, HETNAM, HETSYN, FORMUL), in der Moleküle aufgeführt sind, die keine proteinogenen Aminosäuren sind, wie z.B. prosthetische Gruppen, Inhibitoren, Lösungsmittelmoleküle oder Ionen, für die ebenfalls 3D-Koordinaten zur Verfügung stehen.

Zusätzlich sollten für die weitere Bearbeitung die Schlüsselaminosäuren für Katalyse und/oder Bindung lokalisiert werden, um den Kernbereich der Bindungstasche eingrenzen zu können. Diese werden – ähnlich wie Liganden – von der Kommandozeile von RASMOL selektiert (hier am Beispiel des katalytischen Serins (Ser<sup>200</sup>) der Acetylcholinesterase (PDB-Code 2ACE):

```
select ser200
```

Im *Display*-Menü wird die gewünschte Darstellungsweise gewählt. Zur Definition des LK bzw. KK wird zunächst ein Radius von ca. 10-15 Å um die Bindungstasche herum definiert. Dies kann abschätzungsweise in RASMOL und exakt unter Zuhilfenahme des Programms WIT!P („wit-not-pee“, entwickelt von Novartis Pharma AG, Kontakt: Armin Widmer, [armin.widmer@pharma.novartis.com](mailto:armin.widmer@pharma.novartis.com)) erfolgen. Hierzu wird WIT!P von der UNIX-

Kommandozeile des Arbeitsverzeichnisses aus gestartet, und die PDB-Datei des Proteins eingelesen. Im WIT!P-Menü wird ein Bereich in einem Protein mit

```
define>set
```

definiert und benannt (z.B. `core`). Das Protein, aus dem dieser Bereich stammt, wird ausgewählt und dann die Art des Sets festgelegt. Anschließend wird ein Bereich mit dem Kommando `rzone` und der Angabe eines Radius (10-15 Å) festgelegt. Anschließend wird ein Atom des Liganden bzw. der katalytischen Schlüsselaminosäure selektiert, um das der Radius festgelegt werden soll. Mit `done` wird das Set entgültig definiert. Von der Kommandozeile von WIT!P aus wird mit dem Befehl

```
move -atoms set core done core
```

aus dem selektierten Bereich ein eigenständiges Molekül generiert, das mit dem Befehl

```
write xpdb core.pdb core done
```

unter dem Dateinamen `core.pdb` abgespeichert wird.

Alle Sekundärstrukturelemente, die von diesem Bereich (auch teilweise) erfaßt werden, sind zunächst im Ganzen zu berücksichtigen. D.h., im Falle der Definition des LK bzw. KK über RASMOL werden Domänenelemente außerhalb des Kernbereichs an weniger geordneten Schleifenbereichen abgetrennt und entfernt. Dies wird durch Löschen der entsprechenden Aminosäuren und Koordinaten in der PDB-Datei bewerkstelligt. Im Falle der Definition des LK bzw. KK mit WIT!P müssen angeschnittene Sekundärstrukturelemente durch Hinzufügen der entsprechenden Aminosäuren und Koordinaten aus der Original-PDB-Datei des Gesamtproteins vervollständigt werden.

### ***AAV 3: Superposition von Proteindomänen bzw. Domänenkernbereichen***

Die Superposition kann in der Regel unter Verwendung der zwei nachfolgend beschriebenen Methoden erfolgen. Die entsprechenden Skripten werden von der UNIX-Kommandozeile im Arbeitsverzeichnis aufgerufen.

- Die CE-(Combinatorial Extension)<sup>[97]</sup>-Methode wird mit folgendem Befehl aufgerufen:

```
structalign pdbfile1 - pdbfile2 -
```

Dabei wird die Struktur in PDB-Datei 2 über die in PDB-Datei 1 überlagert. Die entsprechend veränderten Koordinaten von Struktur 2 werden als Datei `pdbfile2_mov` ausgegeben.

- Die Methode nach Taylor und Orengo<sup>[103]</sup> wird mit folgendem Befehl aufgerufen:

```
sap pdbfile1 pdbfile2
```

Die beim Überlagerungsprozeß veränderten Koordinaten von Struktur 2 werden als Datei `pdbfile2.fit` ausgegeben.

Templat-Datei (`pdbfile1`) und überlagerte Struktur (`pdbfile2_mov` bzw. `pdbfile2.fit`) werden anschließend mit folgendem Befehl von der UNIX-Kommandozeile aus zu einer Superpositionsdatei vereinigt:

```
cat pdbfile1 pdbfile2_mov > both.pdb1.pdb2.pdb   oder  
cat pdbfile1 pdbfile2.fit > both.pdb1.pdb2.pdb
```

Diese Superpositionsdatei kann mit RASMOL visualisiert werden.

Als alternative Methode für den Proteinstrukturvergleich kann der FoldMiner-Webserver<sup>[100, 101]</sup> (<http://dlb4.stanford.edu/foldminer/>) genutzt werden. Die PDB-Dateien der auf strukturelle Ähnlichkeit zu untersuchenden Proteinkernstrukturen bzw. Proteindomänen werden auf den Server hochgeladen bzw. über ihren PDB-Code definiert. Von den verschiedenen Optionen des LOCK 2-Algorithmus, der ein paarweises strukturelles Alignment durchführt, ist vor allem der Schwellenwert für den maximal zugelassenen Abstand zwischen  $C^\alpha$ -Atomen hervorzuheben. Dieser ist zwar standardmäßig auf 3 Å eingestellt, kann jedoch in schwierigen Fällen durchaus auf 4-5 Å hochgesetzt werden. Ausgegeben wird eine PDB-Datei mit den beiden überlagerten Strukturen. Sollen Liganden beim Überlagerungsprozeß mitberücksichtigt werden, so muß die Bezeichnung HETATM in der Koordinatensektion der PDB-Datei in ATOM geändert werden. Ein strukturbasiertes Sequenzalignment wird ebenso ausgegeben. Der RMSD-Wert für den überlagerten Kern ist angegeben, ebenso die Anzahl der strukturell einander zuordenbaren Aminosäuren. Die Sequenzidentität hingegen muß manuell berechnet werden. Sie errechnet sich aus dem Quotienten aus der Anzahl der alignen Aminosäuren geteilt durch die Länge des Alignments abzüglich der Lückenpositionen.

#### ***AAV 4: Strukturabgeleitete Sequenzalignments***

Von den Superpositionsprogrammen werden strukturabgeleitete Sequenzalignments generiert. Das Vorgehen zur weiteren Bearbeitung eines solchen Alignments soll am Beispiel der Ausgabedatei „structalign.log“ der CE-Methode illustriert werden:

```
*****
Structure Alignment Calculator, version 1.02, last modified: Jun 15, 2001.

CE Algorithm, version 1.00, 1998.

Ch1: 1c25.core.pdb:_ (Size=148)
Ch2: 2ace.core.pdb:B:1 (Size=60)

Alignment length = 49 Rmsd = 2.74A Z-Score = 3.5 Gaps = 22(44.9%) CPU = 0s
Sequence identities = 8.2%

Ch1: 70 EEVEDFLLKKP---IVPTD--GKR---VIVVFHCEFSSERGPRMCRYVRER--DRLGNEYPKLHYPELYV
Ch2:  2 NVGLLDQRMALQVHDNIQFFGGDPKTVTIFGES-----AGGASVGMHILSPGSR-----DLFRRAIL

Ch1:130 L
Ch2: 60 Q
*****
```

Dieses „structuralign“-Alignment muß editiert und in Blast-Format überführt werden:

```
*****
Query:  70  EEVEDFLLKKP---IVPTD--GKR---VIVVFHCEFSSERGPRMCRYVRER--DRLGNEYPKLHYPELYV 129
                G      V      G      R
Sbjct:   2  NVGLLDQRMALQWVHDNIQFFGGDPKTVTIFGES-----AGGASVGMHILSPGSR-----DLFRRAIL 59
```

Query:130 L

Sbjct: 60 Q

```
*****
```

Mit dem Skript „blast2msf“ wird Blast- in MSF-Format konvertiert, das von dem Programm SHADYBOX (Camson Huynh, Australian National Genomic Information Service) gelesen werden kann. Dieses Programm erlaubt die Bearbeitung des Alignments zur weiteren Auswertung. Für Details wird auf das Handbuch verwiesen (<http://www.bimcore.emory.edu/home/SeqAnal/manuals/shadybox.txt>).

### ***AAV 5: Erstellung von Strukturbildern***

Die PDB-Datei wird mit RASMOL visualisiert:

```
rasmol pdbfile
```

Das Protein wird durch Auswahl im RASMOL-Menü als Spur der C<sup>α</sup>-Atome dargestellt:

```
Display>Backbone
```

Darzustellende Aminosäuren oder Liganden werden mit dem select-Befehl von der RASMOL-Kommandozeile aus selektiert und durch Auswahl im RASMOL-Menü z.B. als Stab- oder CPK- oder Kalottenmodell dargestellt:

```
select ligand oder
select XaaYYY (z.B. select ser200)
```

```
Display>Sticks oder
```

```
Display>Spacefill
```

Von der Kommandozeile von RASMOL aus wird eine MolScript-Datei generiert:

```
write molscript <filename>.in
```

Diese Datei muß editiert werden, um Auflösung und Farbgebung festzulegen:

```
nedit <filename>.in
```

Unter den Koordinatenabschnitt wird folgender Absatz eingefügt, um eine hohe Bildqualität zu erhalten:

```
set segments 45;  
set smoothsteps 30;
```

Die Farbdefinition eines Proteins in Schlaufendarstellung erfolgt durch Voranstellung des folgenden Abschnitts vor die jeweilige Proteinbeschreibung, z.B.:

```
set planecolour blue;  
set plane2colour blue;
```

Liganden und ausgewählte Aminosäuren in CPK-Darstellung werden durch Voranstellung der folgenden Zeile eingefärbt:

```
set atomcolour in residue XXX blue;
```

Für weitere Darstellungsmöglichkeiten wird auf die MolScript-Homepage (<http://www.avatar.se/molscript/doc/molscript.html>) verwiesen.

Die Generierung des Bildes kann mit POV-RAY (<http://www.povray.org/>) erfolgen. Hierzu wird die MolScript-Datei in GL\_RENDER eingelesen. Die gewünschte Orientierung der Proteinstruktur wird eingestellt und eine POV-RAY-Rechnung aufgesetzt. Zur detaillierten Vorgehensweise wird auf die GL\_RENDER-Homepage verwiesen ([http://www.hhmi.swmed.edu/external/Doc/Gl\\_render/Html/gl\\_render.html](http://www.hhmi.swmed.edu/external/Doc/Gl_render/Html/gl_render.html)).

### ***AAV 6: Erstellung von Homologiemodellen***

Proteinsequenzen können der „SWISS-PROT Protein knowledgebase“<sup>[127]</sup> (<http://au.expasy.org/sprot/>) entnommen werden. Sequenzen im Fasta-Format werden dem „Structure Prediction Meta-Server“<sup>[128]</sup> (SPMS, <http://bioinfo.pl/meta/>) übermittelt. Dieser Meta-Server leitet eine Sequenz an mehrere autonome Server weiter, die dann den Konsensus berechnen. Das im SPMS implementierte „3D-Jury-System“ bewertet und rankt als „Meta-Prädiktor“ die Lösungen der einzelnen autonomen Server nach einheitlichen Kriterien, wobei ein spezieller Score jedem Alignment zugeordnet wird. Die Konsensus-Bildung, d.h. das Alignment zweier Proteinsequenzen, wobei von einer eine 3D-Struktur bekannt ist, ist für die

Qualität des Homologiemodells entscheidend. Die von dem 3D-Jury-System als am besten bewerteten Lösungen werden analysiert, wobei sicherzustellen ist, daß Lückenpositionen im Alignment jeweils sinnvoll gesetzt wurden. Sekundärstrukturelemente sollten nicht durch Lücken unterbrochen sein. Das Modell wird dann mit dem Programm MODELLER 6v2<sup>[133]</sup> konstruiert und die Modellqualität mit ProQ<sup>[134]</sup> („Protein Quality Predictor“, <http://www.sbc.su.se/~bjorn/ProQ/>), einer auf einem neuronalen Netz beruhenden Evaluierungsmethode, geprüft. Ausgegeben werden die Qualitätsmaßzahlen „LGscore“ und „MaxSub“. Sehr gute Modelle zeigen: LGscore > 5, MaxSub > 0.8 (max. 1.0). Zusätzlich werden zur Qualitätsprüfung der Modelle WhatCheck<sup>[135]</sup>- und ProCheck<sup>[136]</sup>-Analysen durchgeführt.

### ***AAV 7: Flexibles Liganden-Docking mit GOLD***

Wie unter AAV 2 beschrieben wird eine PDB-Datei einer Proteinstruktur mit einem gebundenen Liganden in das Arbeitverzeichnis kopiert. Aus dieser Datei werden Wassermoleküle, Ionen und Liganden entfernt. Bei der Entfernung von Wassermolekülen und insbesondere von Metallionen ist Vorsicht geboten, da über diese die Ligandenbindung in der Bindungstasche vermittelt werden kann. Um dies feststellen bzw. ausschließen zu können, ist ein intensives Studium der Primärliteratur zwingend. Für das Docking mit GOLD<sup>[233]</sup> ([http://www.ccdc.cam.ac.uk/products/life\\_sciences/gold/](http://www.ccdc.cam.ac.uk/products/life_sciences/gold/)) muß die Proteinstruktur im MOL2-Format vorliegen. Die für das Docking vorbereitete PDB-Datei muß entsprechend konvertiert werden. Dies gelingt mit dem Programm BABEL (<http://inka.mssm.edu/docs/molmod/babel/>), wobei die Koordinaten der PDB-Ausgangsdatei erhalten bleiben. Von der UNIX-Kommandozeile aus wird folgender Befehl ausgeführt:

```
babel -h -ipdb <filename>.pdb -omol2 <filename>.mol2
```

Der Zusatz „-h“ bewirkt, daß Protonen in korrekter Geometrie zu der Proteinstruktur hinzugefügt werden. Der Protonierungszustand v.a. in der Bindungstasche sollte überprüft und ggf. manuell korrigiert werden. Dies kann z.B. mithilfe des Programms WIT!P erfolgen.

3D-Modellstrukturen der zu dockenden Liganden werden mit dem Programm SYBYL ([www.tripos.com](http://www.tripos.com)) erstellt. Wie die Zeichnung und Kraftfeldminimierung genau durchgeführt wird, ist im SYBYL-Handbuch ausführlich beschrieben. Es ist auf korrekte Bezeichnung der Atomtypen (Parametrisierung) zu achten (siehe SYBYL-Handbuch). Für die

Kraftfeldminimierung kleiner Moleküle wird das Kraftfeld „MMFF94“ empfohlen, wobei als Minimierungsmethode „Conjugate Gradient“ angewendet werden sollte. Initial kann zur Optimierung die „Simplex“-Methode verwendet werden, wenn die Moleküle schlecht gezeichnet sind. Als Beendigungskriterium sollte „Gradient=0.05“ gewählt werden. Für das flexible Liganden-Docking ist eine maximale Anzahl von 1000 Iterationen völlig ausreichend. Die Strukturen sollten ebenfalls als MOL2-Dateien abgespeichert werden.

GOLD wird von der UNIX-Kommandozeile aus gestartet. Das GOLD-Eingabefenster erscheint. Die Standardeinstellungen werden gewählt. Rezeptor- und Ligandenstrukturen werden eingelesen und die Bindungstasche definiert. Hierzu wird ein Atom des in der Kristallstruktur enthaltenen Ausgangsliganden (drei Raumkoordinaten) als Punkt gewählt und die Bindungstasche durch einen Radius (z.B. 7.0 Å) um diesen Punkt herum definiert. Ebenso sollte der Detektionsalgorithmus für Kavitäten aktiviert sein.<sup>[259]</sup> Nach einem Docking-Lauf wird eine Liste (`bestranking.lst`) mit den bestgerankten Lösungen und den entsprechenden, die 3D Koordinaten der gedockten Moleküle enthaltenden MOL2-Dateien ausgegeben. Diese können mit dem Programm BABEL in PDB-Format konvertiert und mit der Proteinstruktur vereinigt werden:

```
cat pdbfile ligand.gold.pdb > protein.ligand.pdb
```

Diese kombinierte Datei kann zur weiteren Evaluierung und Abschätzung der Plausibilität des Docking-Ergebnisses herangezogen werden. Besondere Berücksichtigung sollte der Vergleich zwischen der gefundenen Position und Interaktionen des gedockten und des kokristallisierten Liganden finden. Dieser ist umso hilfreicher, je höher die strukturelle Verwandtschaft zwischen den Liganden ist. Im übrigen können in WIT!P Wasserstoffbrückenbindungen und van der Waals-Kontakte visualisiert werden. Eine zusätzliche Möglichkeit zur Überprüfung des Docking-Ergebnisses bietet das Programm LIGPLOT V.4.4.2<sup>[260]</sup> (<http://www.biochem.ucl.ac.uk/bsm/ligplot/ligplot.html>), mit dem Protein-Ligand-Wechselwirkungen (Wasserstoffbrücken und van der Waals-Kontakte) zweidimensional visualisiert werden können.

Einen weiteren Anhaltspunkt für das Docking-Ergebnis bietet der von der GOLD-Fitneß-Funktion ermittelte Fitneß-Wert. Dieser ist eine hybride Größe und errechnet sich aus dem Betrag mehrerer Energie-Termen (für Details siehe GOLD-Handbuch). Der Fitneß-Wert ist äußerst kritisch im Zusammenhang mit einer vergleichenden Beurteilung der unterschiedlichen Liganden hinsichtlich ihrer etwaigen Bindungsaffinität zu bewerten. Die

Fitneß-Funktion ist auf das Auffinden der Ligandenposition hin optimiert und nicht so sehr im Hinblick auf die Vorhersage von Bindungsaffinitäten. Vorsicht ist im übrigen geboten, wenn hydrophobe Liganden gedockt werden. Der Docking-Algorithmus von GOLD ist auf die Untersuchung des Bindungsmodus von hydrophilen Liganden hin optimiert!<sup>[233]</sup>

## 6.1.2 Retrospektive Analyse von Literaturbeispielen

### *Leukotrien-A<sub>4</sub>-Hydrolase und Aminopeptidasen*

Ausgehend von einer literaturbekannten Analyse<sup>[11]</sup> und weiteren Literaturrecherchen wurden Leukotrien-A<sub>4</sub>-Hydrolase (LTA<sub>4</sub>H, PDB-Code 1HS6) und die Aminopeptidasen Thermolysin (PDB-Code 5TMN) und Angiotensin-Conversionsenzym (ACE, PDB-Code 1O86) als strukturell ähnlich identifiziert. Die Proteine wurden entsprechend AAV 1 nach SCOP klassifiziert und ihre Domänenarchitektur analysiert. AAV 2 folgend wurde der KK des ACE identifiziert. Die Überlagerung der katalytischen Domänen bzw. Kernbereiche erfolgte gemäß AAV 3 nach der CE-Methode. AAV 4 und AAV 5 folgend wurden Sequenzalignments sowie Abbildungen der überlagerten Strukturen erstellt.

### **SCOP-Klassifizierung**

1HS6: Drei Domänen

- N-terminalen Domäne (P1-A208): *Faltung*: LTA<sub>4</sub>H N-terminal domain, *Superfamilie*: LTA<sub>4</sub>H N-terminal domain, *Familie*: LTA<sub>4</sub>H N-terminal domain.
- C-terminalen Domäne (D461-D610): *Faltung*:  $\alpha$ - $\alpha$ -Superhelix, *Superfamilie*: ARM repeat, *Familie*: LTA<sub>4</sub>H C-terminal domain.
- Katalytischen Domäne (L209-Y460): *Faltung*: Zinkin-artig, *Superfamilie*: Metalloproteases ("zincins"), catalytic domain, *Familie*: LTA<sub>4</sub>H catalytic domain.

5TMN: Eindomänenprotein

- *Faltung*: Zinkin-artig, *Superfamilie*: Metalloproteases ("zincins"), catalytic domain, *Familie*: Thermolysin-like.

1O86: Eindomänenprotein

- *Faltung*: Zinkin-artig, *Superfamilie*: Metalloproteases ("zincins"), catalytic domain, *Familie*: Neurolysin-like.

**KK**: 1O86 (P312-P424, I521-G583)



**RMSD: 3.27 Å**

### Strukturabgeleitetes Sequenzalignment:

```

1AQU (A36-L211) 1 10 20 30 40 50 60 70
1UKY (P9-L204) P A F S P D Q V S V I F V L G G P G A G K G T Q C E K L V K D Y S F . . . . V H L S . A G D L L R A E Q G R A G S Q Y G E L I K N C I K

1AQU (A36-L211) 80 90 100 110 120 130 140
1UKY (P9-L204) E G Q I V P Q E I T L A L L R N A I S D N V K A N K H K F L I D . G F . P R K M D Q A I S F E R D I . . . V E S K F I L F F D C P E D I M L

1AQU (A36-L211) 150 160 170 180 190 200 210
1UKY (P9-L204) V S Y Y Y F L L . M . I T S Y P N P K S F S E . F V E K F M Q G V P Y G S W Y D . H V K A W W E K S K N S R V . L F M F Y E . . . . D I M K
E R L L E R G K T S G R . S D D N . . . I . E S I . K K R F N . . . . T . F K E T S M P V I E Y F E T K S . K V V R V R C D R S V E D V .

1AQU (A36-L211) 220
1UKY (P9-L204) E D I R R E V V K L I E F L
. . Y K D V Q D A I R D S L

```

### Nukleäre Rezeptoren

Literaturrecherchen und der gemeinsamen Klassifizierung in SCOP zufolge wurden einige Kernrezeptoren untersucht: FXR (PDB-Code 1OSH), ER $\beta$  (PDB-Code 1QKM) und PPAR $\gamma$  (PDB-Code 2PRG). Die Proteine wurden entsprechend AAV 1 nach SCOP klassifiziert sowie ihre Domänenarchitektur jeweils analysiert. Von diesen Kernrezeptoren wurde jeweils die Ligandenbindungsdomäne (LD) kristallisiert. Die Überlagerung der LDs erfolgte gemäß AAV 3 mithilfe des FoldMiner-Webservers. Nach AAV 4 und AAV 5 wurden Sequenzalignments sowie Abbildungen der überlagerten Strukturen erstellt.

### SCOP-Klassifizierung

1OSH/1QKM/2PRG: LD

- *Faltung*: Nuclear receptor ligand-binding domain, *Superfamilie*: Nuclear receptor ligand-binding domain, *Familie*: Nuclear receptor ligand-binding domain.

**Alignmentlängen:** 176 Aminosäuren (1OSH/2PRG), 159 Aminosäuren (1OSH/1QKM)

**Sequenzidentitäten:** 27.8% (1OSH/2PRG), 18.2% (1OSH/1QKM)

**RMSD:** 2.16 Å (1OSH/2PRG), 1.84 Å (1OSH/1QKM)

### Strukturabgeleitetes Sequenzalignment:

```

1OSH (T250-F444) 1 10 20 30 40 50 60 70
1QKM (A262-N457) T P D Q O T L L H F L M D S Y N K Q R M . . . . .
2PRG (E207-4445) E S A D L R A L A K H L Y D S Y I K S F P L T K A K A R A I L T G K T T D K S P F V I Y D M N S L M M G E D K I K F K H I T P L Q E Q S K

1OSH (T250-F444) 80 90 100 110 120 130 140
1QKM (A262-N457) . . . P . . . . E N F L I L T E M A T N H V Q V L V E F T K K L P G F I Q T L D H E D Q L A L L K G S A V E A M F L R S A E I F N K K I P S
2PRG (E207-4445) E V A I . . . . R I E Q G C Q F R S V E A V Q E I T E Y A K S I P G F E V N I D L N D Q V T L L K Y G V H E I I Y T M L A S L M . . . . N

1OSH (T250-F444) 150 160 170 180 190 200 210
1QKM (A262-N457) G H S . . . . . D L L E E R I R N S G . . I S D E V I T P M E S E Y K S I G E L K M T Q E E Y A L L T A I V I L L S . P D R I Q V I K D R
2PRG (E207-4445) K D G M L I S E G Q G F M T R E F L K S L R K P F G D . F M E P K E E A V A K F N A L E L D D S D L A I F I A V I L L S I G D R P G L L N V

1OSH (T250-F444) 220 230 240 250 260
1QKM (A262-N457) . . . . E A V E K L Q E P L L D V L Q K L C K I H Q P E I N P . . . . Q H E A C L L G R L T E L R T F
2PRG (E207-4445) V A D S S R K L A H L L N A V T D A L V W V I A K S G T S S I Q Q S M R L A I N . . . . Q L F A K L L Q K M I D L R I Q I
. . . . K P I E D I Q D N I L L Q A L E L Q L K I L N H P E I S S . . . . Q L F A K L L Q K M I D L R I Q I

```

### 6.1.3 Arbeiten zum Proteinstruktur-Ähnlichkeitsclustering (PSSC)

*Katalytische Kernbereiche der Cdc25A-Phosphatase, der Acetylcholinesterase, der 11 $\beta$ -Hydroxysteroid-Dehydrogenasen Typ 1 und Typ 2, der Pankreaslipase (prospektiv) und der Acylprotein-Thioesterase 1 (prospektiv)*

Entsprechend AAV 1 wurden die Dali/FSSP- sowie die CE-Datenbank nach strukturell ähnlichen Proteinen durchforstet. Die Treffer der Datenbankrecherche wurden gemäß AAV 1 nach SCOP klassifiziert und ihre Domänenarchitektur analysiert. Auf der Basis dieser Klassifizierung wurden pharmazeutisch relevante Enzyme selektiert. AAV 2 folgend wurden die katalytischen Kernbereiche dieser Enzyme identifiziert und gemäß AAV 3 nach der CE-Methode überlagert. AAV 4 und AAV 5 folgend wurden Sequenzalignments sowie Abbildungen der überlagerten Strukturen erstellt.

#### **Dali/FSSP:**

*Suchmotiv:* Cdc25A-Phosphatase (PDB-Code 1C25)

*Ergebnis:* Hydroxynitril-Lyase aus *Hevea brasiliensis* (PDB-Code 3YAS): Sequenzidentität 17%, RMSD 2.6 Å, Alignmentlänge: 76 Aminosäuren; Methylen-tetrahydromethanopterin-Dehydrogenase aus *Methylobacterium extorquens* (PDB-Code 1LU9): Sequenzidentität 7%, RMSD 4.9 Å, Alignmentlänge: 89 Aminosäuren.

Aus der ersten Dali-Recherche ging das Suchmotiv der AChE hervor, die wie die Hydroxynitril-Lyase  $\alpha/\beta$ -Hydrolase-Faltung zeigt und ein pharmazeutisch relevantes Mitglied der Superfamilie der  $\alpha/\beta$ -Hydrolasen ist (s.u.).

*Suchmotiv:* AChE (PDB-Code 2ACE)

*Ergebnis:* Tropinon-Reduktase aus *Datura stramonium* (PDB-Code 1AE1): Sequenzidentität 6%, RMSD 3.9 Å, Alignmentlänge: 155 Aminosäuren.

#### **CE:**

*Suchmotiv:* Cdc25A-Phosphatase (PDB-Code 1C25)

*Ergebnis:* Carboxylesterase EST2 aus *Alicyclobacillus acidocaldarius* (PDB-Code 1EVQ): Sequenzidentität 10.9%, RMSD 3.5 Å, Alignmentlänge: 92 Aminosäuren.

#### **SCOP-Klassifizierung und Auswahl pharmazeutisch relevanter Enzyme:**

1C25:

- *Faltung:* Rhodanese/Cell cycle control phosphatase, *Superfamilie:* Rhodanese/Cell cycle control phosphatase, *Familie:* Cell cycle control phosphatase, catalytic domain.

## 3YAS/1EVQ:

- *Faltung*: alpha/beta-Hydrolases, *Superfamilie*: alpha/beta-Hydrolases, *Familie*: Hydroxynitrile lyase/Carboxylesterase.

→ Pharmazeutisch relevante Enzyme mit  $\alpha/\beta$ -Hydrolase-Faltung:

Acetylcholinesterase (PDB-Code 2ACE)

Pankreaslipase (PDB-Code 1ETH)

Humane Acylprotein-Thioesterase 1 (PDB-Code 1FJ2)

## 1LU9/1AE1:

- *Faltung*: NAD(P)-binding Rossmann-fold domains, *Superfamilie*: NAD(P)-binding Rossmann-fold domains, *Familie*: Aminoacid dehydrogenase-like, C-terminal domain/Tyrosine-dependent oxidoreductases (SDR family)

→ Pharmazeutisch relevante Vertreter der SDR-Familie:

11 $\beta$ -Hydroxysteroid-Dehydrogenasen Typ 1 und Typ 2 (Homologiemodelle, s.u.)

**KK:** 1C25 (P361-P483), 2ACE (N98-Q225), 1ETH (R69-A179), 1FJ2 (T16-S138), 11 $\beta$ HSD1 (G111-N291)

**Alignmentlängen:** 49 Aminosäuren (1C25/2ACE), 80 Aminosäuren (1C25/1ETH), 67 Aminosäuren (1C25/1FJ2), 80 Aminosäuren (1C25/11 $\beta$ HSD1)

**Sequenzidentitäten:** 8.2% (1C25/2ACE), 16.3% (1C25/1ETH), 11.9% (1C25/1FJ2), 5.0% (1C25/11 $\beta$ HSD1)

**RMSD:** 2.74 Å (1C25/2ACE), 2.55 Å (1C25/1ETH), 2.51 Å (1C25/1FJ2), 4.13 Å (1C25/11 $\beta$ HSD1)

**Strukturabgeleitete Sequenzalignments:**

```

1C25 (E405-L465)   1  E E V E D F L L K K P . . . I V P T D . . G K R . . . V I V V F H C E F S S E R G P R M C R Y V R E
2ACE (N167-Q225)  1  N V G L L D Q R M A L Q W V H D N I Q F F G G D P K T V T I F G E S . . . . . A G G A S V G M H I L

1C25 (E405-L465)   60  R . . D R L G N E Y P K L H Y P E L Y V L
2ACE (N167-Q225)  60  S P G S R . . . . . D L F R R A I L Q

1C25 (K377-K475)   1  K E F V I I D C R . . . . . Y P Y E Y E G G H I K G A V N L H M E E E V E D
HSD1 (D114-T222)  1  D M L L L N H I T N T S L . N L F H D D I H H V R K S M . . . . . E V N F L S Y V V
HSD2 (W162-S269)  1  W G L V N N A G H N E V V A D A E L S P V A T F R S C M . . . . . E V N F F G A L E

1C25 (K377-K475)   60  F L L K K P I V P T D G K R V I V V F H C E F S S E . . . . . R G P R M C R Y V R E
HSD1 (D114-T222)  60  L T V A A L P M L K Q S N G S I V V V S S . . L A G K V A Y P M V A A Y S A S K F A L D G F F S S I
HSD2 (W162-S269)  60  L T K G L L P L L R S S R G R L V T V G S . . P A G D M P Y P C L G A Y G T S K A A V A L L M D T F

1C25 (K377-K475)   110  D R L G N E Y P K L H Y P E L . Y V L K G G Y K E F F M K
HSD1 (D114-T222)  110  R K . . E Y S V S R V N V S I T L C V L G L . . I D T E T
HSD2 (W162-S269)  110  S C . . E L L P W G V K V S . . I I Q P G C . . F K T E S

```

```

1C25 (A365-C476) 1  A S V L N G K F A N L K E F V I I D . . . . . C R Y P Y E Y E G G H . I K . . . . . G A V I N . . . . . L H . . . . . 70
1ETH (K70-S274) . . . . . K T R F L I H . . . . . G F I D K . . . . . G E E D W L S N L C K N L F K V E S V N C I C V D W K G G S R T
1FJ2 (G38-G166) G . I . . . R S S H I . . . . . K Y I C P H J A P V R P V T L N M N V . . . . . A M . . . . . . . . . . . P . . . . . S W F

1C25 (A365-C476) 80  . . . . . M E E E V E D F L L K K P I V . . . . . P T D G K R V I V V F H C E F S S E R G P R M C R Y V R R E R D 140
1ETH (K70-S274) . . . . . N I R I V G A E V A Y F V E V L K S S L G Y S P S N V H V I G H S . . . . . L G S H A A G E A G R R T
1FJ2 (G38-G166) D I I G L S P D S Q E D E S G I K Q A A E N I K A L I D Q E V K N G I P S N R I T L L G G F . . . . . S Q G G A L S L Y T A L T I

1C25 (A365-C476) 150  R L G N E Y P K L H Y P E L Y V L 160 170 180 190 200 210
1ETH (K70-S274) N G . . . . . T I E R I T G L D P A E P C F Q G T P E L V R L D P S D A K F V D V I H T D A A P I I P N L G F G M S Q T V G H L D
1FJ2 (G38-G166) . . . . . Q Q K L A G V T J A L S C L P L R A S F P Q G P I G G A N R D I S I L Q C H G . . . . . . . . . . . . . . . . . . . . .

1C25 (A365-C476) 220 230 240 250 260
1ETH (K70-S274) F F P N G G K Q M P G C Q K N I L S Q I V D I D G I W E G T R D F V A C N H L R S Y K E F F M K C
1FJ2 (G38-G166) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

```

### ***Homologiemodell der 11 $\beta$ -Hydroxysteroid-Dehydrogenase Typ 1***

Gemäß AAV 6 wurde ein Homologiemodell der 11 $\beta$ HSD1 erstellt. Die 292 Aminosäuren lange Primärsequenz der 11 $\beta$ HSD1 wurde der „SWISS-PROT protein knowledgebase“ entnommen. Consensus-Berechnung erfolgte über den „Structure Prediction Meta-Server“. Die 3D-JuryA-all (3JAA)-Version des Meta-Prädiktors ordnete das unter Anwendung des FFAS03-Algorithmus<sup>[130]</sup> generierte Alignment der 11 $\beta$ HSD1-Sequenz zur Templatstruktur der Levodion-Reduktase (PDB-Code 1IY8, siehe Sequenzalignment 1) als höchstrangige Lösung ein. Die Konstruktion des Homologiemodells erfolgte mit dem Programm MODELLER 6v2. Die Qualität des Modells wurde mit ProQ, WhatCheck und ProCheck überprüft. Zusätzlich konnte die Modellqualität anhand der am 2. November 2004 veröffentlichten Röntgenstruktur der 11 $\beta$ HSD1 (PDB-Code 1XU7) überprüft werden (Sequenzalignment 2, Sequenzidentität 2, RMSD s.u.).

**Sequenzalignment 1:**

```

11BHSD1 (MODEL) 1 10 20 30 40 50
11Y8 . . . . . T A T S S P T T R F T D R V V L I T G A S K G I G R E M

11BHSD1 (MODEL) 60 70 80 90 100
11Y8 A Y H L A K M G A H V V V T A R S K E T L Q K V V S H C L E L G . A A S A H Y I A G T M E D M T F A
A V R L A A E G A K L S L V D V S S E G L E A S K A A V L E T A X D A E V L T T V A D V S D E A Q V

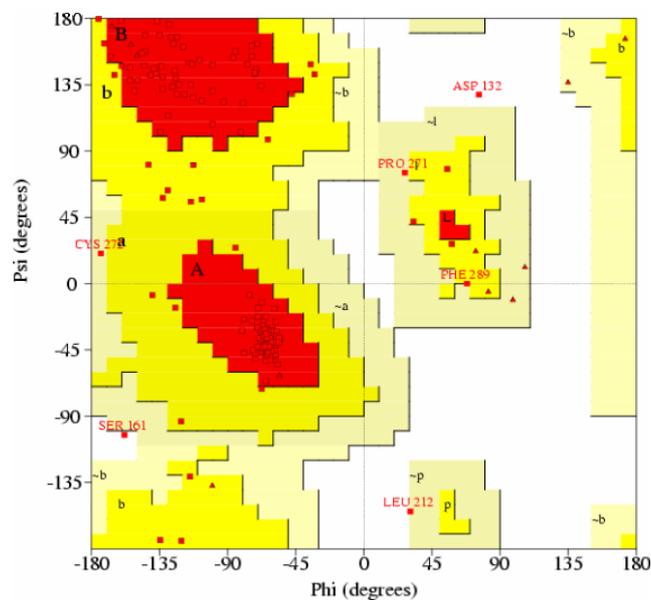
11BHSD1 (MODEL) 110 120 130 140 150
11Y8 E Q F V A Q A G K L M G G L D M L I L N H I T N . T S L N L F H D D I H H V R K S M E V N F L S Y V
E A Y V T A T T E R F G R I D G F F N N A G I E X K Q N P T E S F T A A E F D K V V S I N L R G V F

11BHSD1 (MODEL) 160 170 180 190 200
11Y8 V L T V A A L P M L . K Q S N G S I V V V S S L A G K V A Y P M V A A Y S A S K F A L D G F F S S I
L G L E K V L K I M X E Q G S G M V V N T A S V G I R G I G N Q S G Y A A A K H G V V G L T R N S

11BHSD1 (MODEL) 210 220 230 240 250
11Y8 R K E Y S V S R V N V S I T L C V L G L I D T E T A M K A V S G I V H M Q A A P . K E E C A L E I I
A V E Y G R Y G I R I N . . A I A P G A L W T P M V E N S M K Q L D P E N P R K X A E E F I Q V N P

11BHSD1 (MODEL) 260 270 280 290
11Y8 K G G A L R Q E E V Y Y D S S L W T T L L I R N P C R K I L E F L Y S T S Y N M D R F I N K
S K R Y G E A P E . I A A V V A F L L S . . . . . D D A S Y V N A T V V P I D G G Q S .

```

**Sequenzidentität 1:** 17.4%**3D-Jury:** 175.53**LGscore:** 6.271**MaxSub:** 0.650**Ramachandran-Diagramm:** 99.2% der Aminosäuren in erlaubten Bereichen (A, B, L, a, b, l, p, ~a, ~b, ~l, ~p)

**Sequenzalignment 2:**

```

11BHSD1 (MODEL) 1      10      20      30      40      50
1XU7             MLQGKKVIVT GASKGI GRE MAYHLAKMGAHV VVTARSKETLQKV VSHCLE
11BHSD1 (MODEL) 60      70      80      90      100
1XU7             LGAASAHYI AGT MEDMTFAEQFVAQAGKLMGGLDMLILNHIT. NTS L NLF
11BHSD1 (MODEL) 110     120     130     140     150
1XU7             HDDIHHVVRKSM EVNFLSYV VLTVAALPMLKQSN GSI VVVSSLAGKVAYPM
11BHSD1 (MODEL) 160     170     180     190     200
1XU7             VAAYSASKFALD GFFSSIRKEYS. VSRV NVSIT LCVLGLI DTETAMKAV
11BHSD1 (MODEL) 210     220     230
1XU7             SGI VHMQAAPKEECALEI I KGG ALRQEEVYYDSSLWTTLL
                . . . . . SGI VHMQ. . . . . A A P K E E C A L E I I K G G A L

```

**Sequenzidentität 2:** 85.3%**Alignmentlänge:** 218 Aminosäuren**RMSD:** 1.76 Å**Homologiemodell der 11 $\beta$ -Hydroxysteroid-Dehydrogenase Typ 2**

AAV 6 folgend wurde ein Homologiemodell der 11 $\beta$ HSD2 erstellt. Das unter Anwendung des mGenTHREADER-Algorithmus<sup>[131]</sup> generierte Alignment der 11 $\beta$ HSD2-Sequenz (Arg<sup>171</sup>–Arg<sup>405</sup>) zur Templatstruktur der Glucose-Dehydrogenase (PDB-Code 1GCO) wurde als höchstrangige Lösung bewertet. Das Homologiemodell wurde unter Anwendung des Programms MODELLER 6v2 erstellt und durch ProQ, WhatCheck und ProCheck evaluiert.

**Sequenzalignment:**

```

11BHSD2 (MODEL) 1      10      20      30      40      50
1GCO             RLARPQRLPVATRAVLITGCD S G F G K E T A K K L D S M G F T V L A T V L E L N - - S
                - M - - - Y K D L E G K V V I T G S S T G L G K S M A I R F A T E K A K V V V N Y R S K E D E A
11BHSD2 (MODEL) 60      70      80      90      100
1GCO             P G A I E L R T C C S P R L R L L Q M D L T K P G D I S R V L E F T K A H T T S T G L W G L V N N A
                N S V L E E I K K V G G E A I A V K G D V T V E S D V I N L V Q S A I K E F G K - - L D V M I N N A
11BHSD2 (MODEL) 110     120     130     140     150
1GCO             G H N E V V A D A E L S P V A T F R S C M E V N F F G A L E L T K G L L P L L R S S - - R G R I V T
                G L E N P V - S S H E M S L S D W N K V I D T N L T G A F L G S R E A I K Y F V E N D I K G T V I N
11BHSD2 (MODEL) 160     170     180     190     200
1GCO             V G S P A G D M P Y P C L G A Y G T S K A A V A L L M D T F S C E L L P W G V K V S I I Q P G C F K
                M S S V H E K I P W P L F V H Y A A S K I G G M K L M T E T L A L E Y A P K G I R V N N I L G P G A I N
11BHSD2 (MODEL) 210     220     230     240     250
1GCO             T E S V R N V G Q W E K R K Q L L L A N L P Q E L L Q A Y G K D Y I E H L H G Q F L H S L R L A M S
                T P I N A E K F A D P E Q R A D V E S M I P M G Y I G E P - - E E I A A V A A W L A S S E A S Y V T
11BHSD2 (MODEL) 260     270     280     290     300
1GCO             D L T P V V D A I T D A L L A A R P R R R Y Y P G Q G L G L M Y F I H Y Y L P E G L R R R F L Q A F
                G I T L F A D G G M T - - - - - Q Y P S F Q A G R G - - - - -
11BHSD2 (MODEL) 310     320     330
1GCO             F I S H C L P R A L Q P G Q P G T T P Q D A A Q D P N L S P G P S P A V A R
                - - - - -

```

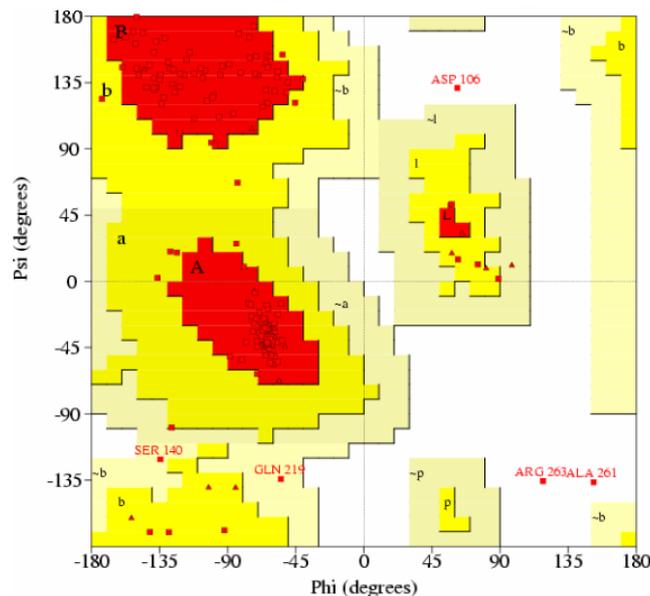
**Sequenzidentität:** 17.9%

**3D-Jury:** 167.42

**LGscore:** 5.088

**MaxSub:** 0.427

**Ramachandran-Diagramm:** 98.7% der Aminosäuren in erlaubten Bereichen (A, B, L, a, b, l, p, ~a, ~b, ~l, ~p)



### ***Katalytische Domänen der Ornithindecaboxylase und der Dihydropteroatsynthetase***

Entsprechend AAV 1 wurden die Dali/FSSP-Datenbank nach strukturell ähnlichen Proteinen durchforstet. Die Treffer der Datenbankrecherche wurden gemäß AAV 1 nach SCOP klassifiziert und die Domänenarchitektur analysiert. Nach AAV 3 wurden die katalytischen Domänen mithilfe des FoldMiner-Webservers überlagert. AAV 4 und AAV 5 folgend wurden Sequenzalignments sowie Abbildungen der überlagerten Strukturen erstellt.

#### **Dali/FSSP:**

*Suchmotiv:* Dihydropteroatsynthase aus *E. coli* (PDB-Code 1AJ0)

*Ergebnis:* Ornithindecaboxylase aus *Trypanosoma brucei* (PDB-Code 1F3T): Sequenzidentität 10%, RMSD 3.2 Å, Alignmentlänge: 175 Aminosäuren.

#### **SCOP-Klassifizierung:**

1AJ0: Eindomänenprotein

- *Faltung*: TIM beta/alpha-barrel, *Superfamilie*: Dihydropteroate synthetase-like, *Familie*: Dihydropteroate synthetase.

1F3T: Zwei Domänen

- N-terminale Domäne (R14-A43, F284-S422): *Faltung*: Domain of alpha and beta subunits of F1 ATP synthase-like (greek-key), *Superfamilie*: Alanine racemase C-terminal domain-like, *Familie*: Eukaryotic ODC-like
- C-terminale katalytische Domäne (D44-A283): *Faltung*: TIM beta/alpha-barrel, *Superfamilie*: PLP-binding barrel, *Familie*: Alanine racemase-like, N-terminal domain.

**Alignmentlänge:** 160 Aminosäuren

**Sequenzidentität:** 8.8%

**RMSD:** 2.65 Å

**Sequenzalignment:**

```

1F3T (A43-H262) 1   A . D . L I G D I V R K H E T W K K C L P R . V T P F Y A V K C . . . . . N D D W R V L G T L A A L G . T G F D C . . A S
1AJ0 (T33-V269) 1   T H N S L I D A V J K H A N L M I N A . . G A T I I D V G G E S T I R P G A A E V S V E E E L Q R V . I P V V E A I A Q R F E V W I S V D I T S K

1F3T (A43-H262) 80   N T E I Q R V R G I G V P P E K . I I Y A N . P C K Q I . S H I R Y A R D S G V D V M T F . D . . . . . 120   130   140
1AJ0 (T33-V269) 80   P E V I R E S A K V . . . . . G A H I I N I D I R S L S E P G A L E A A A E T G L . P V C L M H M Q G N P K T M Q E A P K Y D D V F A E V N R

1F3T (A43-H262) 150   C . V D E L E K V A K T . H P K . A K I M V L R I S T V K E G A . . . . . 170   180   190   200   210
1AJ0 (T33-V269) 150   Y F I E Q I A R C E Q A G I . A K E K L L L D P . . . . . G F G F G K N L S H N Y S L L A R L A E . F H H . F . . . . . N L . P L I V G M S .

1F3T (A43-H262) 220   S I G S . . . . . T D A S T F A Q A I S D S R . F V F D M G T E L G F N M H I L D I G G G F P G T R D A P L K F E E I A G V I
1AJ0 (T33-V269) 220   . . R K S M I G Q L L N V G P S E R L S G S L A C A V I A A M Q G . . . . . A H I R V H . . . . . D V K E T

1F3T (A43-H262)  N N A L E K H
1AJ0 (T33-V269)  V E A M R V V

```

## 6.1.4 Konzeptionelle Arbeiten

### *ATP-Bindungsdomänen von p56<sup>lck</sup> und D-Ala-D-Ala-Ligase*

Angeregt durch die Feststellung der strukturellen Ähnlichkeit zwischen Proteinkinasen und *DD*-Ligase<sup>[194]</sup> wurden die ATP-Bindungsdomänen der p56<sup>lck</sup> (PDB-Code 1QPE, stellvertretend für Proteinkinasen) und der *DD*-Ligase (PDB-Code 2DLN) nach AAV 3 mithilfe des FoldMiner-Webservers überlagert. Ebenso erfolgte eine Klassifizierung nach SCOP gemäß AAV 1. Gemäß AAV 4 wurde ein strukturabgeleitetes Sequenzalignment und gemäß AAV 5 eine Abbildung der überlagerten Proteinkernstrukturen generiert.

**SCOP-Klassifizierung:**

1QPE: KD

- *Faltung*: Protein kinase-like (PK-like), *Superfamilie*: Protein kinase-like (PK-like), *Familie*: Protein kinases, catalytic subunit.

2DLN: Zwei Domänen

- N-terminale Domäne (M1-D96): *Faltung*: PreATP-grasp domain, *Superfamilie*: PreATP-grasp domain, *Familie*: D-Alanine ligase N-terminal domain.
- C-terminale ATP-Bindungsdomäne (KD, K97-D306): *Faltung*: ATP-grasp, *Superfamilie*: Glutathione synthetase ATP-binding domain-like, *Familie*: ATP-binding domain of peptide synthetases.

**Alignmentlänge:** 99 Aminosäuren

**Sequenzidentität:** 8.1%

**RMSD:** 3.55 Å

**Sequenzalignment:**

```

2DLN (L96-M292)  1  L R S K L L W Q G A G L P . V A P W . V A L T . R A E F E K G L S D K Q L A E I S A L G L P V I V K
1QPE (L286-M444)  L A E A N L M K Q L Q H Q R L V R L Y A V V T Q E . . . . .

2DLN (L96-M292)  60  P S R E G S S V G M S K V V A E N A L Q D A L R L A F Q H D E E V L I E K W L S G P E F T V A I L G
1QPE (L286-M444)  . . . . . P . . I . Y I I T E Y M E . N G . . . . .

2DLN (L96-M292)  110  E E I L P S I R I Q P S G T F Y D Y E A K Y L S D E T Q Y F C P A . . . . . G L E . . . . A S Q E
1QPE (L286-M444)  . . . . . S L V . . . . . D F L K T P S G I K L T I N K L L D M A

2DLN (L96-M292)  160  A N L Q A L V L K A W T . . . T L G C K G . W G R I D V M L D S D G Q F Y L L E A N . . . . .
1QPE (L286-M444)  A Q I A E G M A F I . E E R N Y . I H R D L R A A N . I L V S D T L S C K I A D F G L A R L I E D N

2DLN (L96-M292)  210  . . . . . T S P G M T S . . . . . 220  230  240  250
1QPE (L286-M444)  E Z T A R E G A K F P I K W T A P E A I N Y G T F T I K S D V W S F G I L L T E I V T H G R I P Y P

2DLN (L96-M292)  Q A G M
1QPE (L286-M444)  . . G M

```

### *ATPase-Domänen der DNA-Gyrase B, der DNA-Topoisomerase VI-B und des Hitzeschockproteins 90 und Kinase-Domäne der Histidinkinase*

Gemäß AAV 1 wurde eine Dali/FSSP-Datenbankrecherche durchgeführt. Gemäß AAV 1 erfolgte die Klassifizierung nach SCOP. Nach AAV 3 wurden die ATPase-Domänen dieser Proteine nach der CE-Methode überlagert. AAV 4 und AAV 5 folgend wurden Sequenzalignments und eine Abbildung der überlagerten Strukturen generiert.

**Dali/FSSP:**

*Suchmotiv:* DNA-Gyrase B (PDB-Code 1EI1)

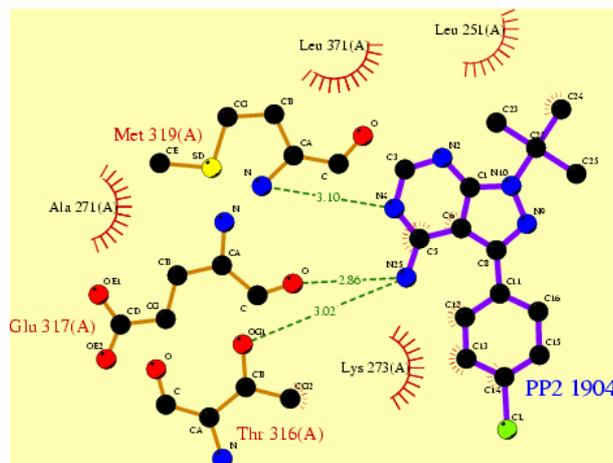


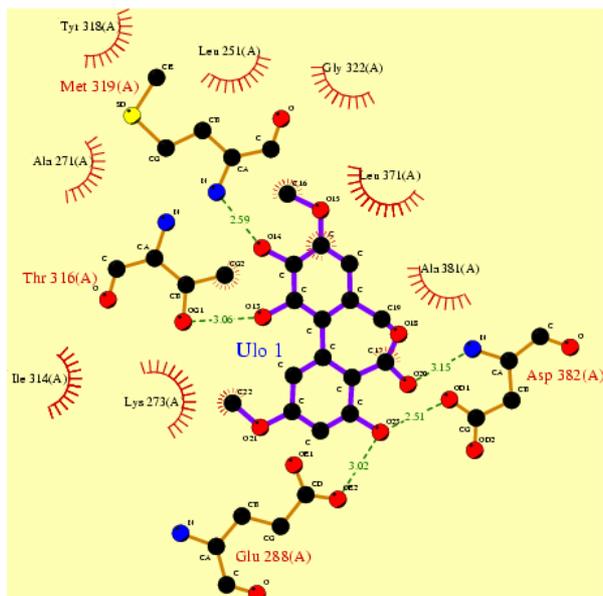
### 6.1.5 Docking

#### *Wahrscheinlicher Bindungsmodus von Ulocladol and die p56<sup>lck</sup> und virtuelles Screening von Ulocladol-Analoga*

Gemäß AAV 2 wurde die PDB-Datei 1QPE (p56<sup>lck</sup> in Komplex mit dem *src*-selektiven Inhibitor PP2) in das Arbeitsverzeichnis kopiert. Die Datei wurde dann gemäß AAV 7 für das Docking-Experiment vorbereitet. Eine 3D-Struktur von Ulocladol und Analoga wurde – wie unter AAV 7 beschrieben – in SYBYL generiert, die 3D-Struktur von PP2 wurde der PDB-Datei entnommen und mittels des Programms BABEL in MOL2-Format konvertiert.

Die GOLD-Rechnung wurde mit den Standardeinstellungen durchgeführt, wobei als Nebenbedingung für das Docking Templatähnlichkeit in bezug auf das Wasserstoffbrückenbindungsmuster von PP2 in der nativen Struktur gewählt wurde. Die Ligandenbindungsstelle wurde durch den Punkt  $P = (x, y, z) = (24.739, 35.127, 81.788)$  (N4 von PP2) und den Radius 5 Å definiert. Sowohl der Ausgangsligand PP2 als auch Ulocladol und Analoga wurden in die Bindungstasche eingepaßt. Als Fitneß-Werte wurden für PP2 54.07 und für Ulocladol 45.41 ermittelt. Der RMSD-Wert, berechnet mit dem GOLD-Utility „smart\_rms“, zwischen dem kokristallisierten und dem gedockten PP2 betrug 0.534 Å bezogen auf die Abweichungen der Atome mit Ausnahme von Wasserstoff. Ein RMSD-Wert von  $< 1.0$  Å wird in Docking-Studien als identisch zur nativen Ligandenposition und -konformation angesehen<sup>[261]</sup> und auch ein RMSD-Ausschlußkriterium von  $< 2.0$  Å kann in manchen Fällen sogar noch akzeptabel sein.<sup>[262]</sup> Diese Untersuchung diente als Validierung für die gewählten Docking-Bedingungen. Die LIGPLOT-Analyse der Bindungsmodi von PP2 und Ulocladol ergab eine weitgehende Übereinstimmung der für die Bindungsaffinität relevanten Wasserstoffbrückenbindungen





## 6.2 Biochemie und Zellbiologie

### 6.2.1 Inhibition der Cdc25A-Phosphatase

Der Cdc25A-Assay wurde von Heike Rimpel und Walburga Hecker, beide Max-Planck-Institut für Molekulare Physiologie, Dortmund, durchgeführt.

*Material:* Tris (Carl Roth GmbH & Co., Karlsruhe), NaCl (RdH Laborchemikalien GmbH & Co. KG, Seelze), Di-Natrium-EDTA-Dihydrat (GERBU Biochemicals GmbH, Gaiberg), DTE (GERBU Biochemicals GmbH, Gaiberg), *p*-NPP (Calbiochem, La Jolla, CA), DMSO (Serva Electrophoresis GmbH, Heidelberg), Triton X-100 (Serva Electrophoresis GmbH, Heidelberg), pET9d/His-Cdc25A-Klon (Dr. Ingrid Hoffmann, Heidelberg).

*Assay:* Der Klon pET9d/His-Cdc25A wurde im *E. coli* Stamm BL21-DE3 exprimiert. Die Aufreinigung erfolgte in der Gegenwart von 8 M Harnstoff.<sup>[113]</sup> Eine 30  $\mu\text{g}$ -Probe des gereinigten Enzyms wurde mit den Inhibitoren (gelöst in DMSO, Endkonzentrationen 0-100  $\mu\text{M}$ ) in 50  $\mu\text{L}$  eines Puffers mit 50 mM Tris, 50 mM NaCl, und 2 mM DTE, pH 8.0, 15 min bei Raumtemperatur inkubiert.<sup>[146]</sup> Anschließend wurden 50  $\mu\text{L}$  eines Puffers mit 50 mM Tris, 50 mM NaCl, 1 mM EDTA und 100 mM *p*-NPP, pH 8.0, hinzugefügt. Das Assay-Endvolumen betrug 100  $\mu\text{L}$ , die Endkonzentration von *p*-NPP 50 mM. Analog wurden Kontrollexperimente in Gegenwart von 0.001% (v/v) Triton X-100 durchgeführt.<sup>[150]</sup> Die Veränderung der Absorption bei einer Wellenlänge von 405 nm über die Zeit wurde auf einem Multiskan Ascent Mikrotiterplattenlesegerät der Firma ThermoLabsystems bei  $37.0 \pm 0.1$  °C 80 min lang kontinuierlich in Intervallen von 1 min aufgezeichnet.

*Auswertung:* Die Reaktionsgeschwindigkeit wurde aus der Absorptionsdifferenz zwischen 30 min and 60 min Reaktionszeit bestimmt. IC<sub>50</sub>-Werte wurden aus den Reaktionsgeschwindigkeiten bei unterschiedlichen Konzentrationen der Testverbindungen bestimmt und graphisch aus Log-Konzentration-Inhibitionskurven ermittelt.

*Ergebnisse:* Siehe Anhänge I und II.

### 6.2.2 Inhibition der 11 $\beta$ -Hydroxysteroid-Dehydrogenasen Typ 1 und Typ 2

Die 11 $\beta$ HSD-Assays wurden in Zusammenarbeit mit PD Dr. Alex Odermatt, Universität Bern, Schweiz, durchgeführt.

*Material:* Zellkulturreagenzien (Invitrogen AG, Basel), [1,2,6,7-<sup>3</sup>H]-Cortison (American Radiolabeled Chemicals, St. Louis, USA), [1,2,6,7-<sup>3</sup>H]-Cortisol (Amersham Biosciences, General Electrics Health Care, Piscataway, USA), Triton X-100 (Serva Electrophoresis GmbH, Heidelberg), Dünnschichtchromatographieplatten (SIL G-25 UV254, Macherey-Nagel, Oensingen, Schweiz), PCR-Reaktionsplatten (Applied Biosystems, Foster City, USA).

*Assay:* Die 11 $\beta$ HSD1-abhängige Oxoreduktion von Cortison und die 11 $\beta$ HSD2-abhängige Oxidation von Cortisol wurden in Lysaten von stabil transfizierten HEK-293-Zellen vermessen.<sup>[121]</sup> Die Zellen wurden 16 h in steroidfreiem Medium inkubiert, einmal mit PBS (phosphate-buffered saline) gewaschen, abgelöst und 3 min bei 150 $\times$ g zentrifugiert. Der Überstand wurde entfernt und das Zellpellet in einem Trockeneis-Ethanol-Bad schockgefroren. Am Tag des Experiments wurden die Zellpellets in Puffer TG1 (20 mM Tris-HCl, pH 7.4, 1 mM EGTA, 1 mM EDTA, 1 mM MgCl<sub>2</sub>, 100 mM NaCl, 20% Glycerol) resuspendiert, mit Ultraschall behandelt und die Aktivität unmittelbar bestimmt. Die Konversion von Cortisol in Cortison oder die umgekehrte Reaktion wurde in 96-well PCR-Reaktionsplatten in einem Endvolumen von 22  $\mu$ L durchgeführt. Die Reaktionen wurden gestartet, indem 10  $\mu$ L Zell-Lysat mit 12  $\mu$ L Reaktionslösung (TG1-Puffer, so angereichert, daß sich die folgenden Endkonzentrationen im Assay ergeben: 400  $\mu$ M NADP<sup>+</sup> oder NADPH, 30 nCi [<sup>3</sup>H]-markiertes Substrat, 200 nM unmarkiertes Cortison oder 25 nM unmarkiertes Cortisol, und 0-200  $\mu$ M Inhibitor), gemischt wurden. Anschließend wurde 10 min bei 37 °C inkubiert. Analog wurden Kontrollexperimente in der Gegenwart von 0.01% (v/v) Triton X-100 durchgeführt.<sup>[150]</sup> 11 $\beta$ HSD2 wird jedoch durch Detergentien inaktiviert,<sup>[149]</sup> so daß hier diese Kontrollexperimente nicht möglich waren. Die Reaktionen wurden durch Zugabe von 10  $\mu$ L einer 2 mM methanolischen Lösung von unmarkiertem Steroid gestoppt. Anschließend wurden die Steroide dünnschichtchromatographisch (Laufmittel: CHCl<sub>3</sub>/MeOH 9:1 (v/v)) aufgetrennt.<sup>[121]</sup> Die Substanzbanden, die den einzelnen Steroiden entsprachen, wurden unter UV-Licht detektiert, ausgeschnitten, in ein Szintillationsgefäß gefüllt und in Szintillationsflüssigkeit in einem Szintillationsdetektor vermessen.<sup>[263]</sup>

*Auswertung:* Die Enzymkinetik wurde durch nicht-lineare Regression unter Verwendung der DATA ANALYSIS TOOLBOX (Elsevier MDL, San Leandro, USA) und unter der Annahme einer Kinetik erster Ordnung analysiert.

*Ergebnisse:* Siehe Anhänge I und II.

### 6.2.3 Inhibition der Acetylcholinesterase

*Material:* Acetylcholinesterase (AChE, Typ III, Elektrischer Aal, Lösung, 200 Units, Sigma-Aldrich Chemie GmbH, Taufkirchen), Acetylthiocholiniodid (ATC, Sigma-Aldrich Chemie GmbH, Taufkirchen), 5,5'-Dithiobis-(2-nitrobenzoesäure) (DTNB, Sigma-Aldrich Chemie GmbH, Taufkirchen), Na<sub>2</sub>HPO<sub>4</sub>·2H<sub>2</sub>O (Mallinckrodt Baker, Inc., Deventer, Holland), NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O (Mallinckrodt Baker, Inc., Deventer, Holland), DMSO (Serva Electrophoresis GmbH, Heidelberg), Triton X-100 (Serva Electrophoresis GmbH, Heidelberg), Acetonitril (HPLC-Grad, LGC Promochem GmbH, Wesel), 96-Well Mikrotiterplatten (Falcon, Becton Dickinson GmbH, Heidelberg).

*Assay:* Die inhibitorische Aktivität wurde nach der spektrophotometrischen Methode von Ellman *et al.* gemessen,<sup>[147, 148]</sup> die zur Anwendung im Mikrotiterplattenformat adaptiert und optimiert wurde. ATC wurde als Substrat der enzymatischen Reaktion und DTNB für die Messung der Cholinesteraseaktivität verwendet. Die Mikrotiterplatten wurden vor jedem Assay mit 100 mM Natriumphosphatpuffer pH 7.2 gespült. In diesem Verfahren wurden 20  $\mu$ L einer verdünnten Enzymlösung (0.25 Units/ml) in 100 mM Natriumphosphatpuffer pH 7.2 mit jeweils 1  $\mu$ L Inhibitorlösung versetzt und 15 min bei Raumtemperatur inkubiert. Die Inhibitoren wurden in Acetonitril/DMSO-Mischungen gelöst, wobei die finale DMSO-Konzentration im Assay < 0.1% gehalten wurde (DMSO selbst ist ein kompetitiver Inhibitor der AChE<sup>[264]</sup>). Die Inhibition wurde auf Vergleichswerte mit entsprechenden DMSO/ACN-Mischungen bezogen. Nach der Inkubation wurden 79  $\mu$ L einer Lösung von 0.253 mM ATC und 0.380 mM DTNB in 100 mM Natriumphosphatpuffer pH 8.0 zupipettiert so daß sich ein Assay-Endvolumen von 100  $\mu$ L und Endkonzentrationen von 200  $\mu$ M ATC<sup>[265]</sup> und 300  $\mu$ M DTNB<sup>[266]</sup> ergaben. Analog wurden Kontrollexperimente in der Gegenwart von 0.01% (v/v) Triton X-100 durchgeführt.<sup>[150]</sup> Die Veränderung der Absorption bei einer Wellenlänge von 405 nm über die Zeit wurde auf einem Multiskan Ascent Mikrotiterplattenlesegerät der Firma ThermoLabsystems bei 25.0±0.1 °C 5 min lang in Intervallen von 30 s lang kontinuierlich aufgezeichnet. Nach einem initialen Screening der Bibliotheken bei 20  $\mu$ M wurden IC<sub>50</sub>-Werte bestimmt.

*Auswertung:* Die Reaktionsgeschwindigkeit wurde aus der Absorptionsdifferenz zwischen 30 min und 60 min Reaktionszeit bestimmt. Ohne Inhibitor betrug diese 0.059±0.003  $\Delta$ A/min.<sup>[266]</sup> IC<sub>50</sub>-Werte wurden aus den Reaktionsgeschwindigkeiten bei unterschiedlichen Konzentrationen der Testverbindungen bestimmt und aus Log-Konzentration-

Inhibitionskurven unter Anwendung der im Programm GRAFIT Version 5.0.4 (Erithacus, Surrey, Großbritannien) implementierten IC<sub>50</sub>-Fitfunktion ermittelt.

*Ergebnisse:* Siehe Anhänge I und II.

#### 6.2.4 Inhibition der Ornithindecaboxylase

*Material:* *Trypanosoma brucei*-ODC (*t*ODC)-Lösung (gepuffert, Proteinkonzentration 41 g/L, Prof. Meg Phillips, Dallas, USA), humane ODC (*h*ODC)-Lösung (gepuffert, Proteinkonzentration 2 g/L, Prof. Meg Phillips, Dallas, USA), DTT (GERBU Biochemicals GmbH, Gaiberg), Hepes (GERBU Biochemicals GmbH, Gaiberg), Hepes-Na (Serva Electrophoresis GmbH, Heidelberg), *L*-Ornithin-hydrochlorid (*L*-Orn-HCl, Sigma-Aldrich Chemie GmbH, Taufkirchen), Pyridoxal-5'-phosphat-Monohydrat (Sigma-Aldrich Chemie GmbH, Taufkirchen), Sulfonamide (Sigma-Aldrich Chemie GmbH, Taufkirchen; FLUKA, Taufkirchen; ABCR, Karlsruhe), DMSO (Serva Electrophoresis GmbH, Heidelberg), Triton X-100 (Serva Electrophoresis GmbH, Heidelberg), Thermo CO<sub>2</sub> Liquid Stable Reagent (Clindia Benelux BV, Leusden, Niederlande), 96-Well Mikrotiterplatten (Falcon, Becton Dickinson GmbH, Heidelberg).

*Assay:* Alle im Inhibitionstest verwendeten Lösungen wurden mit CO<sub>2</sub>-freiem Wasser hergestellt. Hierzu wurde Wasser 30 min lang zum Sieden erhitzt und anschließend unter Argonatmosphäre abgekühlt. Der Assay wurde für die Anwendung in Mikrotiterplatten entwickelt. Die Assay-Bedingungen von Osterman *et al.*<sup>[191]</sup> mußten entsprechend adaptiert und optimiert werden. Sämtliche Lösungen wurden mit 100 mM Hepes-Puffer pH 7.5, der zusätzlich 2 mM DTT enthielt, hergestellt. In einem Well wurden 50 µL einer Lösung von 0.02 g/L *h*ODC bzw. 0.164 g/L *t*ODC, 4 mM DTT, 0.04 mM PLP in Hepes-Puffer pH 7.5 mit den Sulfonamiden (in DMSO, Lösungsmittelleffekt wurde überprüft, 0-1000 µM) 15 min unter Argonatmosphäre bei Raumtemperatur inkubiert. Anschließend wurde eine Mischung aus 10 µL 3 mM *L*-Orn-HCl-Lösung in Hepes-Puffer pH 7.5 und 40 µL Thermo CO<sub>2</sub> Liquid Stable Reagent zugegeben. Analog wurden Blindwerte ohne *L*-Orn-HCl bestimmt, so daß die Hintergrundreaktion von aus der Umgebungsluft stammendem CO<sub>2</sub> mit dem CO<sub>2</sub>-Detektionsreagenz erfaßt werden konnte. Die Veränderung der Absorption bei einer Wellenlänge von 340 nm (Abnahme der Konzentration von NADPH) über die Zeit wurde auf einem Multiskan Ascent Mikrotiterplattenlesegerät der Firma ThermoLabsystems bei 37.0±0.1 °C 10 min lang in Intervallen von 30 s kontinuierlich aufgezeichnet.

*Auswertung:* Die Reaktionsgeschwindigkeit wurde aus der Absorptionsdifferenz zwischen 8 min und 9 min Reaktionszeit bestimmt. Abgezogen wurde jeweils die Hintergrundreaktion von aus der Umgebungsluft stammendem CO<sub>2</sub>. IC<sub>50</sub>-Werte wurden aus den Reaktionsgeschwindigkeiten bei unterschiedlichen Konzentrationen der Testverbindungen bestimmt und aus Log-Konzentration-Inhibitionskurven mit der im Programm GRAFIT Version 5.0.4 (Erithacus, Surrey, Großbritannien) implementierten IC<sub>50</sub>-Fitfunktion ermittelt.

*Ergebnisse:* Siehe Anhang III.

### **6.2.5 Translokation des Glucocorticoidrezeptors (GR)**

Dieser Versuch wurde in Kooperation mit PD Dr. Alex Odermatt (Universität Bern, Schweiz) durch Marco Casaulta durchgeführt.

HEK-293 Zellen (300 000 Zellen pro Well) wurden auf mit Poly-*L*-Lysin beschichteten Glasplättchen in 6-Well-Platten, die 2 mL Dulbecco's modifiziertes Eagle Medium (DMEM, mit 10% fötalem Kälberserum (FCS) angereichert) enthielten, gezüchtet. Die Zellen wurden nach der Calciumphosphat-Präzipitationsmethode mit 1 µg Expressionsplasmid für das Fusionskonstrukt aus GFP (grün-fluoreszierendes Protein) und GR (Glucocorticoidrezeptor) [255] und entweder 0.5 µg humanem 11βHSD1-Plasmid oder leerem pcDNA3-Vektor (Invitrogen) pro Well transfiziert. Nach 6 h wurden die Zellen zweimal mit Medium gewaschen, das zuvor zweimal über Aktivkohle filtriert wurde, um Steroide zu entfernen. Anschließend erfolgte eine Inkubation in diesem Medium für weitere 18 h. Die Zellen wurden anschließend 30 min lang mit einem Inhibitor oder einem GR-Antagonisten präinkubiert. Dann erfolgte die Zugabe von 500 nM Cortison gefolgt von einer weiteren 40-minütigen Inkubation. Die Zellen wurden gewaschen und mit 4% Paraformaldehyd 10 min lang fixiert und die Lokalisierung von GFP-GR mittels Fluoreszenzmikroskopie bestimmt. Drei unabhängige Transfektionsexperimente wurden durchgeführt, wobei je Probe 300 fluoreszierende Zellen durch einen Beobachter, der gegenüber der Vorbehandlung der Zellen verblindet war, ausgewertet wurden.

### **6.2.6 GR-abhängige Transaktivierung**

Dieser Versuch wurde in Kooperation mit PD Dr. Alex Odermatt (Universität Bern, Schweiz) durch Marco Casaulta durchgeführt.

HEK-293 Zellen wurden in mit Poly-*L*-Lysin 12-Well-Platten, die 1 mL Dulbecco's modifiziertes Eagle Medium (DMEM, mit 10% fötalem Kälberserum (FCS) angereichert) enthielten, gezüchtet. Die subkonfluenten Zellen (200 000 Zells pro Well) wurden mit 300 ng pMMTV-LacZ-Reporterplasmid, 100 ng pCMV-LUC-Kontrollplasmid, 300 ng GR-Expressionsvektor und entweder 300 ng 11 $\beta$ HSD1- oder leerem pcDNA3-Vektor transfiziert. Nach 6 h wurden die Zellen zweimal sorgfältig mit steroidfreiem Medium gewaschen. 11 $\beta$ HSD1-Inhibitor, GR-Antagonist und Steroidhormon wurden zugefügt und die Zellen für zusätzliche 24 h inkubiert. Die Zellen wurden in 50  $\mu$ L Lysis-Puffer lysiert. Die Lysate wurden mit dem Luciferase-Assaysystem (Promega) und dem  $\beta$ -Galactosidase Galacto-Light Plus Kit (Tropix) analysiert. Die Galactosidaseaktivität wurde auf die interne Luciferase-Kontrolle normalisiert. Die Daten (Mittelwert  $\pm$  Standardabweichung, drei unabhängige Experimente) wurden als Prozentwert relativ zur Kontrolle in Gegenwart von Steroid aber in Abwesenheit von Inhibitor dargestellt.

### 6.2.7 Inhibition von Kinasen

Die Kinase-Profilierung der Ulocladol-abgeleiteten Verbindungskollektion erfolgte bei der Firma Aventis Pharma Deutschland GmbH, Frankfurt.

*Ergebnisse:* Siehe Anhang IV.

## 6.3 Chemie

### 6.3.1 Meßgeräte und Hilfsmittel

$^1\text{H}$ - und  $^{13}\text{C}$ -NMR-Spektren wurden auf einem Varian Mercury 400 Spektrometer (400 MHz  $^1\text{H}$ -NMR, 100.6 MHz  $^{13}\text{C}$ -NMR) aufgenommen.

Die chemischen Verschiebungen sind in ppm angegeben und beziehen sich auf Tetramethylsilan (TMS,  $\delta = 0$  ppm) als internen Standard. Die Kopplungskonstanten  $J$  sind in Hertz (Hz) angegeben. Die Signalmultiplizitäten wurden wie folgt abgekürzt: s = Singulett, d = Dublett, dd = Doppeldublett, t = Triplett, m = Multipllett, br = breites Signal,  $m_{\text{verd}}$  = Signal teilweise vom Lösungsmittel verdeckt.

Schmelzpunkte wurden auf einem Schmelzpunktmeßgerät Büchi 530 bestimmt und sind unkorrigiert angegeben.

FAB-Massenspektren wurden mit einem Jeol SX 102A aufgenommen. Die Messung der MALDI-TOF Massenspektren erfolgte auf einer Voyager-DE Pro BioSpectrometry<sup>TM</sup> Workstation der Firma PerSeptive Biosystems mit 2,5 Dihydroxybenzoesäure (DHB) als Matrix.

### *Chromatographie*

Zur Dünnschichtchromatographie wurden Aluminiumplatten, beschichtet mit Kieselgel 60<sub>F254</sub> der Firma Merck, verwendet. Die entsprechenden Laufmittel und  $R_f$ -Werte sind bei den jeweiligen Substanzen angegeben. Zur Detektion wurde UV-Licht der Wellenlänge 254 nm und „Seebach-Reagenz“ (2.5 g Molybdato-phosphorsäure, 1 g Cer-(IV)-sulfat, 6 ml konz. Schwefelsäure und 94 ml Wasser) als Anfärbereagenz verwendet.

Säulenchromatographischen Trennungen erfolgten an Flash-Kieselgel der Firma J. T. Baker (Korngröße 40-60  $\mu\text{m}$ ), teilweise mithilfe eines CombiFlash Sq 16x Geräts der Firma Isco.

Präparativ-HPL-chromatographische Trennungen erfolgten an einer Nucleodur C18 Gravity 5 $\mu$  Säule der Firma Macherey-Nagel auf einem Agilent (1100 Series) Chromatographen. Die Detektion erfolgte bei 215 nm. Folgende Laufmittelsysteme wurden verwendet:

**P1:** Flussrate: 25 mL/min; Gradient: 0 min, (ACN + 0.1 % TFA/H<sub>2</sub>O + 0.1 % TFA): 50:50; 3 min: 50:50; 22 min: 100:0.

**P2:** Flussrate: 27 mL/min; Gradient: (ACN + 0.1 % Ameisensäure/H<sub>2</sub>O + 0.1 % Ameisensäure): 0 min: 5:95; 1 min: 5:95; 8 min: 20:80; 8.1 min: 100:0.

**P3:** Flussrate: 15 mL/min; Gradient: (ACN + 0.1 % TFA/H<sub>2</sub>O + 0.1 % TFA): 0 min: 30:70; 3 min: 30:70; 22 min: 100:0.

**P4:** Flussrate: 27 mL/min; Gradient: (ACN + 0.1 % TFA/H<sub>2</sub>O + 0.1 % TFA): 0 min: 80:20; 3 min: 80:20; 22 min: 100:0.

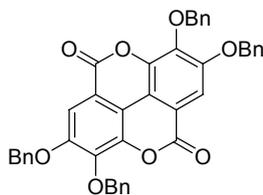
Das Lösungsmittel der Produktfraktionen wurde anschließend durch Lyophilisieren entfernt.

Für die gaschromatographische Analyse wurde ein Gaschromatograph HP6890 gekoppelt mit einem Massendetektor HP5973 der Firma Hewlett-Packard ausgestattet mit einer Kapillarsäule HP-5TA (0.33  $\mu$ m, 25 m  $\times$  0.2 mm ID) der Firma Macherey-Nagel und Helium als Trägergas verwendet (Splitverhältnis 1:50, Flußrate 99.7 mL/min). Temperaturprogramm: 0–1 min: 100 °C, 1–6 min: Aufheizgeschwindigkeit 40 °C/min bis 300 °C, 6–11 min: 300 °C.

Alle Chemikalien wurden bei den Firmen Acros, Aldrich, Avocado, Biosolve, Fluka und Lancaster erworben. Mit Ausnahme von DMF wurden alle verwendeten Lösungsmittel vor Gebrauch destilliert und bei Bedarf absolutiert.<sup>[267]</sup> Absolutes DMF wurde von der Firma Fluka erworben.

### 6.3.2 Versuche zur Synthese von Ulocladolderivaten

#### *Tetrabenzylellagsäure (70)*



Zu einer Suspension von 4.0 g (13.2 mmol) bei 100 °C im Hochvakuum getrockneter Ellagsäure, 15.2 g (110 mmol) K<sub>2</sub>CO<sub>3</sub> und 1.2 g (7.23 mmol) KI in 60 mL Acetophenon wurden 16 mL (139 mmol) Benzylchlorid getropft. Unter dauerndem, kräftigem Rühren

wurde das Reaktionsgemisch auf 140 °C gebracht und bis zum Nachlassen der CO<sub>2</sub>-Entwicklung (20 h) auf dieser Temperatur gehalten. Am Ende wurde kurz auf 170 °C erhitzt und rasch abgesaugt. Der Rückstand bestehend aus KCl, K<sub>2</sub>CO<sub>3</sub> und nicht umgesetzter Ellagsäure wurde noch einmal mit heißem Acetophenon ausgegossen. Im sich abkühlenden Filtrat kristallisierte Tetrabenzylellagsäure aus und wurde durch Filtration abgetrennt. Der Filterkuchen wurde mit kaltem Diethylether, Methanol und Wasser gewaschen. Anschließend wurde aus Acetophenon umkristallisiert.

**Ausbeute:** 2.4 g (3.6 mmol, 55%, elfenbeinfarbener Feststoff)

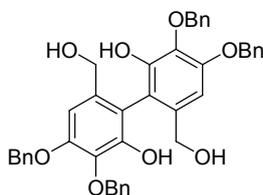
**Smp.:** 263 °C (267 °C<sup>[268]</sup>)

**R<sub>f</sub>** = 0.56 (CH<sub>2</sub>Cl<sub>2</sub>/Cyclohexan 5:2 (v/v))

**MS** (MALDI-TOF, DHB): ber. [M+Na]<sup>+</sup> 685.2, gef. 685.6

**HR-MS** (FAB, 3-NBA) für C<sub>42</sub>H<sub>31</sub>O<sub>8</sub> [M+H]<sup>+</sup>: ber.: 663.2019; gef.: 663.2042

**2,2'-Dihydroxy-3,3',4,4'-tetrakis(benzyloxy)-1,1'-biphenyl-6,6'-dimethanol (71)**



Eine Suspension von 3.0 g (4.5 mmol) Tetrabenzylellagsäure in 75 mL THF wurde zusammen mit 0.85 g (21 mmol) Lithiumaluminiumhydrid unter einer Argonatmosphäre bei Raumtemperatur über Nacht gerührt. Nach wäßriger Hydrolyse des überschüssigen Lithiumaluminiumhydrids wurde THF unter vermindertem Druck abdestilliert. Die wäßrige Phase wurde in 100 mL Ethylacetat aufgenommen und mit gesättigter Kochsalzlösung gewaschen. Die organische Phase wurde über Na<sub>2</sub>SO<sub>4</sub> getrocknet und das Lösungsmittel unter vermindertem Druck entfernt. Es wurde in Ethanol umkristallisiert.

**Ausbeute:** 2.22 g (3.31 mmol, 74%, weißer Feststoff)

**Smp.:** 141 °C

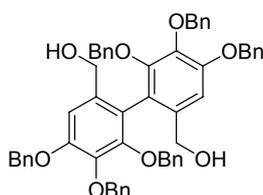
**R<sub>f</sub>** = 0.36 (Ethylacetat/Cyclohexan 3:2 (v/v))

**$^1\text{H-NMR}$**  (400 MHz,  $\text{CDCl}_3$ ):  $\delta = 7.34\text{--}7.20$  (m, 20 H, arom.), 6.74 (s, 2H, arom.), 5.67 (s(br), 2H, OH), 5.14 (d,  $J=11.1$  Hz, 2H,  $\text{CH}_2$ ), 5.12 (s, 4H,  $\text{CH}_2$ ), 5.02 (d,  $J=11.1$  Hz, 2H,  $\text{CH}_2$ ), 4.08 (d, 2H,  $J=11.9$  Hz,  $\text{CH}_2$ ), 4.04 (d, 2H,  $J=11.7$  Hz,  $\text{CH}_2$ )

**$^{13}\text{C-NMR}$**  (100.6 MHz,  $\text{CDCl}_3$ ):  $\delta = 151.5, 147.5, 137.2, 136.9, 136.4, 134.1, 128.9, 128.8, 127.8, 114.1, 106.5, 75.5, 71.0$

**HR-MS** (FAB, 3-NBA) für  $\text{C}_{42}\text{H}_{38}\text{O}_8$   $[\text{M}]^+$ : ber.: 670.2567; gef.: 670.2562

**2,2',3,3',4,4'-Hexakis(benzyloxy)-1,1'-biphenyl-6,6'-dimethanol (72)**



Eine Suspension von 3.98 g (5.93 mmol) 2,2'-Dihydroxy-3,3',4,4'-tetrakis(benzyloxy)-1,1'-biphenyl-6,6'-dimethanol, 20.0 g (145 mmol) wasserfreiem  $\text{K}_2\text{CO}_3$  und 15.0 mL (126.0 mmol) Benzylbromid in 200 mL trockenem Aceton wurde für 3 h unter Rückfluß unter einer Argonatmosphäre erhitzt. Das Lösungsmittel wurde unter vermindertem Druck abdestilliert und der Rückstand durch Säulenchromatographie mit zunächst reinem Cyclohexan zur Entfernung überschüssigen Benzylbromids und anschließend mit Aceton/Cyclohexan 2:5 (v/v) als Eluens aufgereinigt.

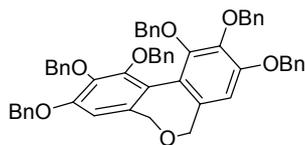
**Ausbeute:** 4.29 g (5.04 mmol, 85%, weißer Feststoff)

$R_f = 0.48$  (Aceton/Cyclohexan 2:3 (v/v))

**$^1\text{H-NMR}$**  (400 MHz,  $\text{CDCl}_3$ ):  $\delta = 7.52\text{--}6.81$  (m, 30 H, arom.), 7.04 (s, 2H, arom.), 5.22 (d,  $J=11.7$  Hz, 2H,  $\text{CH}_2$ ), 5.18 (d,  $J=11.5$  Hz, 2H,  $\text{CH}_2$ ), 5.05 (d,  $J=11.7$  Hz, 2H,  $\text{CH}_2$ ), 5.00 (d, 2H,  $J=11.6$  Hz,  $\text{CH}_2$ ), 4.95 (d, 2H,  $J=11.0$  Hz,  $\text{CH}_2$ ), 4.62 (d, 2H,  $J=11.1$  Hz,  $\text{CH}_2$ ), 4.18 (d, 2H,  $J=12.3$  Hz,  $\text{CH}_2$ ), 4.15 (d, 2H,  $J=12.1$  Hz,  $\text{CH}_2$ ), 2.71 (s(br), 2H, OH)

**$^{13}\text{C-NMR}$**  (100.6 MHz,  $\text{CDCl}_3$ ):  $\delta = 153.1, 150.8, 141.5, 137.4, 137.0, 136.6, 129.0, 128.8, 128.5, 128.4, 127.8, 127.7, 122.6, 110.9, 75.7, 71.3$

**MS** (MALDI-TOF, DHB): ber.  $[\text{M}+\text{Na}]^+$  873.3, gef. 873.6, ber.  $[\text{M}+\text{K}]^+$  889.3, gef. 889.6

**5,7-Dihydro-1,2,3,9,10,11-hexakis(benzyloxy)dibenz[*c,e*]oxepin (73)**

Eine Lösung von 3.07 g (3.61 mmol) 2,2',3,3',4,4'-Hexakis(benzyloxy)-1,1'-biphenyl-6,6'-dimethanol und 200 mg (1.04 mmol) *p*-Toluolsulfonäure-Monohydrat in 175 mL trockenem Benzol wurde 1 h unter Rückfluß unter einer Argonatmosphäre erhitzt. Das Lösungsmittel wurde unter vermindertem Druck entfernt, und der Rückstand durch Säulenchromatographie mit Aceton/Cyclohexan 1:9 (v/v) als Eluens aufgereinigt. Für die biologische Evaluierung wurde zusätzlich mittels präparativer HPLC (Methode **P1**) aufgereinigt.

**Ausbeute:** 2.20 g (2.64 mmol, 73%, beiger Feststoff)

**R<sub>f</sub>** = 0.28 (Aceton/Cyclohexan 2:7 (v/v))

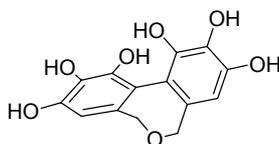
**Smp.:** 122 °C

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>): δ = 7.49-6.96 (m, 30H, arom.), 6.69 (s, 2H, arom.), 5.20 (d, *J*=11.7 Hz, 2H, CH<sub>2</sub>), 5.14 (d, *J*=10.7 Hz, 2H, CH<sub>2</sub>), 5.12 (d, *J*=11.5 Hz, 2H, CH<sub>2</sub>), 5.03 (d, *J*=9.96 Hz, 4H, CH<sub>2</sub>), 4.92 (d, *J*=11.3 Hz, 2H, CH<sub>2</sub>), 4.08 (d, *J*=11.3 Hz, 2H, CH<sub>2</sub>), 3.55 (d, *J*=11.1 Hz, 2H, CH<sub>2</sub>)

**<sup>13</sup>C-NMR** (100.6 MHz, CDCl<sub>3</sub>): δ = 152.6, 150.1, 142.2, 137.8, 137.5, 137.0, 128.7, 128.6, 128.5, 128.4, 128.0, 127.9, 127.7, 127.6, 124.9, 110.1, 75.6, 71.3

**MS** (MALDI-TOF, DHB): ber. [M+Na]<sup>+</sup> 855.3, gef. 855.5, ber. [M+K]<sup>+</sup> 871.3, gef. 871.5

**HR-MS** (FAB, 3-NBA) für C<sub>56</sub>H<sub>48</sub>O<sub>7</sub> [M]<sup>+</sup>: ber. 832.3400, gef. 832.3405; C<sub>56</sub>H<sub>49</sub>O<sub>7</sub> [M+H]<sup>+</sup>: ber. 833,3478, gef. 833.3464; C<sub>56</sub>H<sub>48</sub>O<sub>7</sub>Na [M+Na]<sup>+</sup> ber. 855.3298, gef. 855.3279

**5,7-Dihydro-1,2,3,9,10,11-hexahydroxydibenz[*c,e*]oxepin (74)**

Eine Suspension von 600 mg (0.72 mmol) 5,7-Dihydro-1,2,3,9,10,11-hexakis(benzyloxy)-dibenz[*c,e*]oxepin und 120 mg Pd-C in 45 mL Methanol/Ethylacetat (2:1) wurde dreimal evakuiert und mit Wasserstoff belüftet und anschließend 2 h bei Raumtemperatur unter einer Wasserstoffatmosphäre gerührt. Die Reaktionsmischung wurde anschließend durch Celite filtriert und das Lösungsmittel unter vermindertem Druck abdestilliert. Die Aufreinigung erfolgt durch Säulenchromatographie mithilfe des Isco CombiFlash Sq 16x Geräts (Detektion: 254 nm; Flussrate: 30 mL/min; Gradient (Essigester + 5 % (v/v) Methanol/Cyclohexan): 0 min: 0:100, 10 min: 3:97, 35 min: 20:80, 45 min: 50:50, 49 min: 100:0). Für die biologische Evaluierung wurde zusätzlich mittels präparativer HPLC (Methode **P2**) aufgereinigt.

**Ausbeute:** 120 mg (0.41 mmol, 57%, weißer Feststoff)

**R<sub>f</sub>** = 0.48 (Methanol/Ethylacetat 1:9 (v/v))

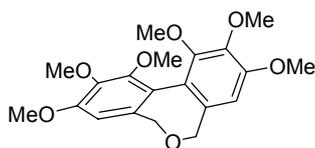
**Smp.:** 240 °C (Zersetzung)

**<sup>1</sup>H-NMR** (400 MHz, CD<sub>3</sub>OD): δ = 6.56 (s, 2H, arom.), 4.90 (s(br)), 4.25 (d, *J*=11.0 Hz, 2H, CH<sub>2</sub>), 3.98 (d, *J*=11.0 Hz, 2H, CH<sub>2</sub>)

**<sup>13</sup>C-NMR** (100.6 MHz, CD<sub>3</sub>OD): δ = 146.6, 136.2, 128.3, 119.5, 111.2, 69.0

**MS** (MALDI-TOF, DHB): ber. [M-H]<sup>-</sup> 291.1, gef. 291.6

#### **5,7-Dihydro-1,2,3,9,10,11-hexamethoxydibenz[*c,e*]oxepin (75)**



Eine Suspension von 120 mg (0.41 mmol) 5,7-Dihydro-1,2,3,9,10,11-hexahydroxydibenz[*c,e*]oxepin, 580 mg (4.20 mmol) wasserfreiem K<sub>2</sub>CO<sub>3</sub> und 1.2 mL (12.7 mmol) Dimethylsulfat in 50 mL trockenem Aceton wurde 2 h unter Rückfluß erhitzt. Das Lösungsmittel wurde unter vermindertem Druck entfernt, und der Rückstand durch Säulenchromatographie mit Ethylacetat/Cyclohexan 2:3 (v/v) als Eluens aufgereinigt. Zusätzlich wurde für die biologische Evaluierung eine Aufreinigung über präparative HPLC vorgenommen (Methode **P3**).

**Ausbeute:** 89 mg (0.24 mmol, 59%, gelber Feststoff)

$R_f = 0.42$  (Ethylacetat/Cyclohexan 2:3 (v/v))

**Smp.:** 151 °C

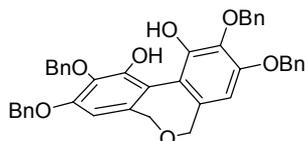
$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta = 6.72$  (s, 2H, arom.), 4.39 (d,  $J=11.1$  Hz, 2H,  $\text{CH}_2$ ), 4.06 (d,  $J=11.4$  Hz, 2H,  $\text{CH}_2$ ), 3.93 (s, 6H,  $\text{CH}_3$ ), 3.92 (s, 6H,  $\text{CH}_3$ ), 3.73 (s, 6H,  $\text{CH}_3$ )

$^{13}\text{C-NMR}$  (100.6 MHz,  $\text{CDCl}_3$ ):  $\delta = 153.5, 151.5, 142.5, 131.0, 123.3, 108.1, 67.7, 61.2, 60.9, 56.2$

**GC-MS:**  $t_R = 6.43$  min,  $[\text{M}]^+$  376

**HR-MS** (FAB, 3-NBA) für  $\text{C}_{20}\text{H}_{24}\text{O}_7$   $[\text{M}]^+$ : ber.: 376.4004, gef.: 376.1506;  $\text{C}_{20}\text{H}_{25}\text{O}_7$   $[\text{M}+\text{H}]^+$ : ber.: 377.1600, gef.: 377.1588;  $\text{C}_{20}\text{H}_{24}\text{O}_7\text{Na}$   $[\text{M}+\text{Na}]^+$ : ber.: 399.1420, gef.: 399.1454

**5,7-Dihydro-1,11-dihydroxy-2,3,9,10-tetrakis(benzyloxy)dibenz[*c,e*]oxepin (76)**



Eine Lösung von 0.50 g (0.75 mmol) 2,2'-Dihydroxy-3,3',4,4'-tetrakis(benzyloxy)-1,1'-biphenyl-6,6'-dimethanol und 60 mg (0.31 mmol) *p*-Toluolsulfonäure-Monohydrat in 25 mL trockenem Benzol wurde 1.5 h unter Rückfluß unter einer Argonatmosphäre erhitzt. Das Lösungsmittel wurde unter vermindertem Druck entfernt, und der Rückstand durch Umkristallisation aus Ethanol gereinigt. Für die biologische Evaluierung wurde zusätzlich mittels präparativer HPLC (Methode **P3**) aufgereinigt.

**Ausbeute:** 0.38 g (0.58 mmol, 77%, weißer Feststoff)

$R_f = 0.52$  (Ethylacetat/Cyclohexan 2:3 (v/v))

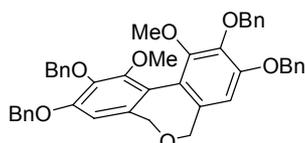
**Smp.:** 130 °C

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta = 7.48-7.32$  (m, 20H, arom.), 6.72 (s, 2H, arom.), 6.42 (s(br), 2H, OH), 5.22-5.13 (m, 8H,  $\text{CH}_2$ ), 4.34 (d,  $J=11.1$  Hz, 2H,  $\text{CH}_2$ ), 4.04 (d,  $J=11.1$  Hz, 2H,  $\text{CH}_2$ )

$^{13}\text{C-NMR}$  (100.6 MHz,  $\text{CDCl}_3$ ):  $\delta = 151.9, 146.3, 137.3, 136.8, 136.1, 131.7, 128.7, 128.6, 128.5, 128.2, 127.6, 117.5, 108.0, 75.7, 71.3$

**HR-MS** (FAB, 3-NBA) für  $\text{C}_{42}\text{H}_{37}\text{O}_7$   $[\text{M}+\text{H}]^+$ : ber.: 653.2539, gef.: 653.2585

**5,7-Dihydro-1,11-dimethoxy-2,3,9,10-tetrakis(benzyloxy)dibenz[*c,e*]oxepin (77)**



Eine Suspension von 100 mg (0.15 mmol) 5,7-Dihydro-1,11-dihydroxy-2,3,9,10-tetrakis-(benzyloxy)dibenz[*c,e*]oxepin, 85 mg (0.61 mmol) wasserfreiem  $\text{K}_2\text{CO}_3$  und  $88 \mu\text{L}$  (0.92 mmol) Dimethylsulfat in 5 mL trockenem THF wurde 2 h unter Rückfluß unter einer Argonatmosphäre erhitzt. Das Lösungsmittel wurde unter vermindertem Druck entfernt, und der Rückstand durch Säulenchromatographie mithilfe des Isco CombiFlash Sq 16x Geräts (Detektion: 254 nm; Flussrate: 25 mL/min; Gradient (Essigester/Cyclohexan): 0 min: 0:100, 6 min: 0:100, 26 min: 7:93, 36 min: 15:85, 46 min: 30:70, 47 min: 100:0) aufgereinigt. Zusätzlich wurde für die biologische Evaluierung eine Aufreinigung über präparative HPLC vorgenommen (Methode **P4**).

**Ausbeute:** 70 mg (0.10 mmol, 67%, weißer Feststoff)

$R_f = 0.75$  (Ethylacetat/Cyclohexan 2:3 (v/v))

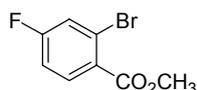
**Smp.:** 134 °C

$^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ ):  $\delta = 7.44\text{-}7.30$  (m, 20H, arom.), 6.79 (s, 2H, arom.), 5.18-5.07 (m, 8H,  $\text{CH}_2$ ), 4.35 (d,  $J=11.3$  Hz, 2H,  $\text{CH}_2$ ), 4.05 (d,  $J=10.5$  Hz, 2H,  $\text{CH}_2$ ), 3.73 (s, 6H,  $\text{CH}_3$ )

$^{13}\text{C-NMR}$  (100.6 MHz,  $\text{CDCl}_3$ ):  $\delta = 152.8, 151.8, 141.6, 137.7, 136.9, 131.0, 128.6, 128.3, 128.1, 128.0, 127.6, 109.8, 75.4, 71.2, 67.6, 61.2$

**MS** (MALDI-TOF, DHB): ber.  $[\text{M}+\text{Na}]^+$  703.3, gef. 703.8, ber.  $[\text{M}+\text{K}]^+$  719.2, gef. 719.8

**HR-MS** (FAB, 3-NBA) für  $\text{C}_{44}\text{H}_{40}\text{O}_7$   $[\text{M}]^+$ : ber. 680.2774, gef. 680.2798 ;  $\text{C}_{44}\text{H}_{41}\text{O}_7$   $[\text{M}+\text{H}]^+$ : ber. 681.2852; gef. 681.2829

**Methyl-2-brom-4-fluorbenzoat (79)**

Eine Lösung von 1.0 g (4.57 mmol) 2-Brom-4-fluorbenzoesäure und 0.05 mL konzentrierter Schwefelsäure in 30 mL wasserfreiem Methanol wurde über Nacht zum Rückfluß erhitzt. Die Reaktionslösung wurde mit 2N Natriumhydroxidlösung neutralisiert und das Lösungsmittel unter vermindertem Druck abdestilliert. Das Rohprodukt wurde in Ethylacetat gelöst und mit Wasser und gesättigter NaCl-Lösung gewaschen. Die organische Phase wurde über Na<sub>2</sub>SO<sub>4</sub> getrocknet und das Lösungsmittel unter vermindertem Druck entfernt. Das Rohprodukt wurde säulenchromatographisch mithilfe des Isco CombiFlash Sq 16x Geräts (Detektion: 254 nm; Flussrate: 30 mL/min; Gradient (Essigester/Cyclohexan): 0 min: 0:100, 10 min: 3:97, 35 min: 20:80, 45 min: 50:50, 49 min: 100:0) aufgereinigt.

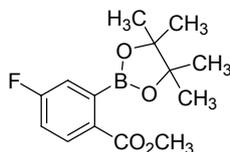
**Ausbeute:** 0.97 g (4.16 mmol, 91%, hellbeiges Öl)

**R<sub>f</sub>** = 0.44 (Ethylacetat/Cyclohexan 1:2 (v/v))

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>): δ = 7.89-7.86 (m, 1H, arom.), 7.43-7.40 (m, 1H, arom.), 7.09-7.05 (m, 1H, arom.), 3.93 (s, 3H, CH<sub>3</sub>)

**<sup>13</sup>C-NMR** (100.6 MHz, CDCl<sub>3</sub>): δ = 165.8, 165.3, 162.8, 133.6, 133.5, 128.3, 128.2, 123.4, 123.3, 122.2, 122.0, 114.8, 114.6, 52.7

**GC-MS:** t<sub>R</sub> = 2.91 min, [M]<sup>+</sup> 232

**Methyl-4-fluor-2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzoat (80)**

Eine Suspension von 350 mg (1.5 mmol) Methyl-2-brom-4-fluorbenzoat, 419 mg (1.65 mmol) Bispinacolatodiboron, 11 mg (0.05 mmol) Pd(OAc)<sub>2</sub> und 442 mg (4.5 mmol) KOAc in 6 mL DMF wurde zunächst 30 min lang entgast. Anschließend wurde unter einer Argonatmosphäre für 3 h auf 85 °C erhitzt. Nach dem Abkühlen wurden zur

Reaktionsmischung 20 mL Wasser und 20 mL Ethylacetat hinzugefügt und über Celite filtriert. Es wurde mit Ethylacetat extrahiert, die organische Phase über  $\text{MgSO}_4$  getrocknet und das Lösungsmittel unter vermindertem Druck abdestilliert. Das Rohprodukt wurde säulenchromatographisch mithilfe des Isco CombiFlash Sq 16x Geräts (Detektion: 254 nm; Flussrate: 35 mL/min; Gradient (Essigester/Cyclohexan): 0 min: 0:100, 10 min: 0:100, 40 min: 10:90, 60 min: 30:70, 70 min: 50:50, 71 min: 100:0) aufgereinigt.

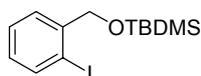
**Ausbeute:** 123 mg (0.44 mmol, 29%, weißer Feststoff)

**R<sub>f</sub>** = 0.35 (Ethylacetat/Cyclohexan 1 :2 (v/v))

**<sup>1</sup>H-NMR** (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 7.97-7.94 (m, 1H, arom.), 7.16-7.13 (m, 1H, arom.), 7.09-7.04 (m, 1H, arom.), 3.90 (s, 3H,  $\text{CH}_3$ ), 1.42 (s, 12H,  $\text{CH}_3$ )

**GC-MS:**  $t_R$  = 4.17 min,  $[\text{M}]^+$  281

### *2-Iodbenzyloxy-tert-butylidimethylsilan (82)*



Zu einer Lösung von 1.0 g (4.3 mmol) 2-Iodbenzylalkohol und 0.73 g Imidazol (10.8 mmol) in 3 mL DMF wurden bei 0 °C unter einer Argonatmosphäre 0.98 g (6.5 mmol) TBDMS-Chlorid gegeben. Die Lösung wurde über Nacht bei Raumtemperatur gerührt. Zur Hydrolyse überschüssigen TBDMS-Chlorids wurden 30 mL gesättigte  $\text{NH}_4\text{Cl}$ -Lösung zu der Reaktionsmischung gegeben. Es wurde dreimal mit je 30 mL Diethylether extrahiert. Die vereinigten organischen Phasen wurden mit gesättigter  $\text{NaHCO}_3$ - und  $\text{NaCl}$ -Lösung gewaschen und anschließend über  $\text{MgSO}_4$  getrocknet. Das Lösungsmittel wurde unter vermindertem Druck abdestilliert und das Rohprodukt säulenchromatographisch mithilfe des Isco CombiFlash Sq 16x Geräts (Detektion: 254 nm; Flussrate: 30 mL/min; Gradient (Essigester/Cyclohexan): 0 min: 0:100, 15 min: 0:100, 45 min: 10:90, 55 min: 50:50, 65 min: 100:0) aufgereinigt.

**Ausbeute:** 1.48 g (4.25 mmol, 99%, farbloses Öl)

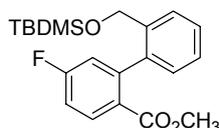
**R<sub>f</sub>** = 0.51 (Ethylacetat/Cyclohexan 2:3 (v/v))

**$^1\text{H-NMR}$**  (400 MHz,  $\text{CDCl}_3$ ):  $\delta = 7.79\text{-}7.76$  (m, 1H, arom.),  $7.53\text{-}7.50$  (m, 1H, arom.),  $7.39\text{-}7.34$  (m, 1H, arom.),  $6.98\text{-}6.96$  (m, 1H, arom.),  $4.63$  (s, 2H,  $\text{CH}_2\text{-O}$ ),  $0.98$  (s, 9H,  $\text{CH}_3$ ),  $0.14$  (s, 6H,  $\text{CH}_3$ )

**$^{13}\text{C-NMR}$**  (100.6 MHz,  $\text{CDCl}_3$ ):  $\delta = 143.1, 138.9, 128.7, 128.4, 127.6, 96.0, 69.6, 26.2, -5.1$

**GC-MS:**  $t_{\text{R}} = 4.29$  min,  $[\text{M}-\text{CH}_3]^+$  333,  $[\text{M}-t\text{Bu}]^+$  291

***Methyl-4-fluor-2-(2-tert-butyldimethylsilyloxymethyl)phenylbenzoat (83)***



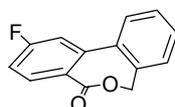
Eine Suspension von 100 mg (0.35 mmol) Methyl-4-fluor-2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzoat, 125 mg (0.35 mmol) 2-Iodbenzyloxy-*tert*-butyldimethylsilan und 173 mg (0.53 mmol)  $\text{Cs}_2\text{CO}_3$  in 3 mL DMF wurde zunächst 30 min lang entgast. Sodann wurden 15 mg (0.011 mmol)  $\text{Pd}(\text{PPh}_3)_4$  zugegeben und die Reaktionsmischung unter einer Argonatmosphäre für 1 h auf  $85\text{ }^\circ\text{C}$  erhitzt. Nach dem Abkühlen der Reaktionsmischung wurden 10 mL Wasser und 10 mL Ethylacetat hinzugefügt und über Celite filtriert. Das Filtrat wurde mit Ethylacetat extrahiert, die vereinigten organischen Phasen über  $\text{MgSO}_4$  getrocknet und das Lösungsmittel unter vermindertem Druck abdestilliert. Das Rohprodukt wurde säulenchromatographisch mithilfe des Isco CombiFlash Sq 16x Geräts (Detektion: 254 nm; Flussrate: 30 mL/min; Gradient (Essigester/Cyclohexan): 0 min: 0:100, 15 min: 0:100, 45 min: 10:90, 55 min: 50:50, 65 min: 100:0) aufgereinigt.

**Ausbeute:** 78 mg (0.21 mmol, 60%, weißer Feststoff)

**$R_f$**  = 0.72 (Ethylacetat/Cyclohexan 1:2 (v/v))

**GC-MS:**  $t_{\text{R}} = 5.28$  min,  $[\text{M}-t\text{-Bu}]^+$  317

***2-Fluordibenz[*c,e*]oxepin-5(7H)-on (69)***



Eine Lösung von 78 mg (0.21 mmol) Methyl-4-fluor-2-(2-*tert*-butyldimethylsilyloxymethyl)-phenylbenzoat und 0.42 mmol Tetrabutylammoniumfluorid (420  $\mu$ L einer 1 M Lösung in THF) in 2.6 mL THF wurde über Nacht bei Raumtemperatur gerührt. Dann wurden 10 mL Wasser hinzugefügt, das THF bei vermindertem Druck entfernt, und anschließend mit Ethylacetat extrahiert. Die vereinigten organischen Phasen wurden über MgSO<sub>4</sub> getrocknet und das Lösungsmittel unter vermindertem Druck abdestilliert. Das Rohprodukt wurde säulenchromatographisch mithilfe des Isco CombiFlash Sq 16x Geräts (Detektion: 254 nm; Flußrate: 25 mL/min; Gradient (Essigester/Cyclohexan): 0 min: 0:100, 10 min: 0:100, 30 min: 10:90, 40 min: 20:80, 60 min: 50:50, 61 min: 100:0) aufgereinigt. Zusätzlich wurde für die biologische Evaluierung eine Aufreinigung über präparative HPLC vorgenommen (Methode **P3**).

**Ausbeute:** 33 mg (0.14 mmol, 67%, weißer Feststoff)

**R<sub>f</sub>** = 0.42 (Ethylacetat/Cyclohexan 1:2 (v/v))

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.81 (dd,  $J=8.8$  Hz,  $J=5.8$  Hz, 1H, arom.), 7.42-6.97 (m<sub>verd</sub>, 6H), 4.81 (d,  $J=18.3$  Hz, 2H, CH<sub>2</sub>)

**<sup>13</sup>C-NMR** (100.6 MHz, CDCl<sub>3</sub>):  $\delta$  = 169.2, 166.1, 163.6, 140.1, 138.0, 134.9, 134.8, 133.8, 133.7, 130.4, 129.4, 128.8, 127.0, 116.0, 115.8, 115.5, 115.3, 69.3

**GC-MS:**  $t_R$  = 5.19 min, [M]<sup>+</sup> 228

**HR-MS** (FAB, 3-NBA) für C<sub>14</sub>H<sub>10</sub>N<sub>2</sub>F [M+H]<sup>+</sup>: ber.: 229.0659; gef.: 229.0649

## 6.4 Chemoinformatik

Dieser Teil der Arbeit wurde zusammen mit Dr. Peter Ertl und Dr. Ansgar Schuffenhauer, beide Novartis Institutes for Biomedical Research, Basel, realisiert.

### 6.4.1 Primäre molekulare Prozessierung des Naturstoff-Datensatzes

Das CRC Dictionary of Natural Products (Chapman & Hall/CRC, Taylor and Francis Books), v12.2.2004 (DNP), das 161 278 Naturstoff-Datensätze enthält, war die Grundlage für die Analyse der Naturstoffe.

Für alle Molekülmanipulationen wurden die molekularen Prozessierwerkzeuge der Firma Molinspiration (Molinspiration Cheminformatics, Bratislava, Slowakische Republik, <http://www.molinspiration.com>) verwendet.

Zunächst wurden die Daten vom MDL SDF-Format in Daylight SMILES-Strings (Simplified Molecular Input Line Entry Specification)<sup>[236]</sup> konvertiert. Bei der Konversion wurden auch fehlerhafte Einträge aus dem Datensatz entfernt. Die nachfolgenden Befehle wurden von der MS-DOS-Kommandozeile aus ausgeführt.

```
java -jar mitools.jar -f <filename>.sdf -out smi > a
```

Eine weitere Standardisierung erfolgte durch die Normalisierung von Ladungen und durch das Entfernen von Gegenionen und kleineren Molekülen (z.B. Wasser und mit den Naturstoffen assoziierte Salze).

```
java -jar mitools.jar -f a -standardize > b
```

### 6.4.2 Extraktion der Grundgerüste

Zur Identifizierung der Grundgerüste wurden acyclische Substituenten aus den ringenthaltenden Naturstoffen entfernt. Cyclische Substituenten wurden als Teil des Grundgerüsts beibehalten.

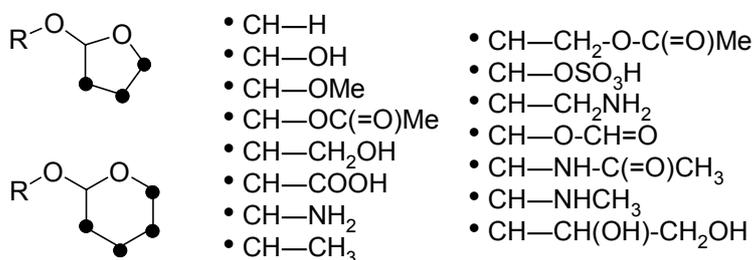
```
java -jar shredder.jar -f b -scaffolds > c
```

Grundgerüste sind als Einheiten von Ringsystemen<sup>[241]</sup> einschließlich exocyclischer Doppelbindungen und etwaiger Verbindungsketten zwischen den Ringen zu verstehen.

### 6.4.3 *In Silico*-Deglykosylierung der Naturstoffe

Aufgrund von Redundanzen durch unterschiedliche Glykosylierungsmuster wurden die Naturstoffe des Datensatzes vor de einer *In silico*-Deglykosylierung unterzogen. Die Diversität der Glykosylierungsmuster, die häufig aus mehreren Zuckermomeren bestanden, ließ eine Beschreibung durch vordefinierte Substrukturen nicht zu. Dies wurde zusätzlich dadurch erschwert, daß die Zuckereinheiten oft mehrere Verknüpfungspunkte aufwiesen und so komplizierte baumartige Substrukturen bildeten.

Zur Lösung dieser Aufgabe mußte ein spezieller rechnergestützter Algorithmus in Java entwickelt werden. Dieser wurde in Verbindung mit einem kommerziell erhältlichen Programm (Molinspiration Cheminformatics, Bratislava, Slowakische Republik, <http://www.molinspiration.com>), das auf rekursiven SMARTs basierende Substruktursuchen erlaubt, eingesetzt. Es wurde zunächst eine Substruktursuche nach substituierten Furanose- und Pyranose-Zuckern formuliert. Jede Kombination der in der untenstehenden Abbildung formulierten Zuckersubstituenten wurde als rekursiver SMART ausgedrückt.<sup>[269]</sup> SMART-Strings stellen eine Weiterentwicklung der SMILES-Strings für die Substruktur-Suche dar. Die SMART-Darstellungweise erlaubt es, komplexe Substrukturabgleiche zu formulieren.



Verschieden substituierte Pyranose- und Furanose-Einheiten, die durch den *In-silico*-Deglykosylierungsprozeß identifiziert wurden. Die Punkte in den Strukturformeln können einen der neben den Strukturen aufgeführten Substituenten beschreiben.

```
*****
[O] 1- [CH] (O [&]) -
[CR2; $( [CH2] ), $( * [OH] ), $( * [O] [CH3] ), $( * [O] C (=O) [CH3] ), $( * [CH2] [OH] ), $( * [C] (
=O) [OH] ), $( * [NH2] ), $( * [CH3] ), $( * [CH2] OC (=O) [CH3] ), $( * OS (=O) (=O) [OH] ), $( * [CH
2] [NH2] ), $( * [CH2] O [CH] (=O) ), $( * [NH] C (=O) [CH3] ), $( * [NH] [CH3] ), $( * [CH] ( [OH] ) [
CH2] ( [OH] ) ) ] -
[CR2; $( [CH2] ), $( * [OH] ), $( * [O] [CH3] ), $( * [O] C (=O) [CH3] ), $( * [CH2] [OH] ), $( * [C] (
=O) [OH] ), $( * [NH2] ), $( * [CH3] ), $( * [CH2] OC (=O) [CH3] ), $( * OS (=O) (=O) [OH] ), $( * [CH
2] [NH2] ), $( * [CH2] O [CH] (=O) ), $( * [NH] C (=O) [CH3] ), $( * [NH] [CH3] ), $( * [CH] ( [OH] ) [
CH2] ( [OH] ) ) ] -
[CR2; $( [CH2] ), $( * [OH] ), $( * [O] [CH3] ), $( * [O] C (=O) [CH3] ), $( * [CH2] [OH] ), $( * [C] (
=O) [OH] ), $( * [NH2] ), $( * [CH3] ), $( * [CH2] OC (=O) [CH3] ), $( * OS (=O) (=O) [OH] ), $( * [CH
2] [NH2] ), $( * [CH2] O [CH] (=O) ), $( * [NH] C (=O) [CH3] ), $( * [NH] [CH3] ), $( * [CH] ( [OH] ) [
CH2] ( [OH] ) ) ] -
[CR2; $( [CH2] ), $( * [OH] ), $( * [O] [CH3] ), $( * [O] C (=O) [CH3] ), $( * [CH2] [OH] ), $( * [C] (
```

```

=O) [OH] ), $ (* [NH2] ), $ (* [CH3] ), $ (* [CH2] OC (=O) [CH3] ), $ (* OS (=O) (=O) [OH] ), $ (* [CH
2] [NH2] ), $ (* [CH2] O [CH] (=O) ), $ (* [NH] C (=O) [CH3] ), $ (* [NH] [CH3] ), $ (* [CH] ( [OH] ) [
CH2] ( [OH] ) ) ] 1
[O] 1 - [CH] ( O [&] ) -
[CR2; $ ( [CH2] ), $ (* [OH] ), $ (* [O] [CH3] ), $ (* [O] C (=O) [CH3] ), $ (* [CH2] [OH] ), $ (* [C] (
=O) [OH] ), $ (* [NH2] ), $ (* [CH3] ), $ (* [CH2] OC (=O) [CH3] ), $ (* OS (=O) (=O) [OH] ), $ (* [CH
2] [NH2] ), $ (* [CH2] O [CH] (=O) ), $ (* [NH] C (=O) [CH3] ), $ (* [NH] [CH3] ), $ (* [CH] ( [OH] ) [
CH2] ( [OH] ) ) ] -
[CR2; $ ( [CH2] ), $ (* [OH] ), $ (* [O] [CH3] ), $ (* [O] C (=O) [CH3] ), $ (* [CH2] [OH] ), $ (* [C] (
=O) [OH] ), $ (* [NH2] ), $ (* [CH3] ), $ (* [CH2] OC (=O) [CH3] ), $ (* OS (=O) (=O) [OH] ), $ (* [CH
2] [NH2] ), $ (* [CH2] O [CH] (=O) ), $ (* [NH] C (=O) [CH3] ), $ (* [NH] [CH3] ), $ (* [CH] ( [OH] ) [
CH2] ( [OH] ) ) ] -
[CR2; $ ( [CH2] ), $ (* [OH] ), $ (* [O] [CH3] ), $ (* [O] C (=O) [CH3] ), $ (* [CH2] [OH] ), $ (* [C] (
=O) [OH] ), $ (* [NH2] ), $ (* [CH3] ), $ (* [CH2] OC (=O) [CH3] ), $ (* OS (=O) (=O) [OH] ), $ (* [CH
2] [NH2] ), $ (* [CH2] O [CH] (=O) ), $ (* [NH] C (=O) [CH3] ), $ (* [NH] [CH3] ), $ (* [CH] ( [OH] ) [
CH2] ( [OH] ) ) ] 1
*****

```

Als rekursive SMARTs formulierte unterschiedlich substituierte Pyranose- und Furanose-Einheiten

Beim Deglykosylierungsprozeß wurden die Naturstoffmoleküle zunächst dahingehend geprüft, ob sie Furanose- oder Pyranose-Ringe enthalten, um lediglich die Glykoside den SMART-basierten Substruktur-Suchen zu unterziehen und so den Prozeß zu beschleunigen. Jede identifizierte Zuckereinheit in den Glykosiden wurde entfernt, indem die *O*-Zucker-Bindung durch eine Hydroxylgruppe ersetzt wurde. Das resultierende Molekül wurde erneut dem Suchalgorithmus solange unterworfen, bis keine Zuckereinheiten mehr identifiziert werden konnten. Es wurden folglich die (unterschiedlich substituierten) Zuckerreste von den Naturstoffglykosiden sukzessive *in silico* „abhydrolysiert“ und so das Aglykon herausgeschält.

Der für die *In silico*-Deglykosylierung formulierte Algorithmus ruft im wesentlichen den ‚substructure matcher‘ der Firma Molinspiration auf und ist im folgenden als Java-basierter Pseudocode wiedergegeben:

```

*****
for each (molecule) {
while(molecule_contains_sugar_unit) {
    replace_sugar_unit_by_OH;
}
}
*****

```

#### 6.4.4 Baumdiagramm der Naturstoffgrundgerüste

Zur hierarchischen Gliederung der Naturstoffe wurden die Naturstoffgrundgerüste einer Baumanalyse unterzogen. Hierzu wurde ein spezieller rechnergestützter Algorithmus entwickelt, der das Elter-Ringgerüst eines jeden individuellen Ringgerüsts identifiziert. Das

Elter-Ringgerüst sollte eine Substruktur des Ringgerüsts des jeweiligen Abkömmlings darstellen. Die Anordnung erfolgte auf verschiedenen hierarchischen Ebenen, wobei die Einzelringe als einfachste Ringgerüste gleichsam die Wurzeln des Naturstoff-Ringgerüstbaums bildeten. Die Einzelringe wurden wiederum nach Heteroatom und Ringgröße in drei Hauptklassen chemisch ähnlicher Cluster zusammengefaßt: Carbocyclen, *N*- und *O*-Heterocyclen. Für die Identifizierung des Elter-Ringgerüsts wurden einige Priorisierungsregeln bestimmt:

1. Das Elter-Ringgerüst sollte eine Substruktur des Ringgerüsts des Abkömmlings darstellen.
2. Ringbindungen durften nicht aufgebrochen werden.
3. Wenn mehrere mögliche Elter-Kandidaten vorlagen, dann wurde das Elter-Ringgerüst so gewählt, daß es die maximale Anzahl von Heteroatomen beinhaltete.
4. Wenn Regel 3 nicht griff, wurde der Elter mit dem größeren Ringgerüst gewählt.
5. Wenn nach Regel 3 und 4 keine Zuordnung herbeigeführt werden konnte, so wurde das in Naturstoffen häufiger vorkommende Ringgerüst als Elter bestimmt.

Diese Regeln werden durch den folgenden Algorithmus, dargestellt als Java-basierter Pseudocode, zum Ausdruck gebracht:

```
*****
parent = null;
// loop through all available candidates to select the parent
for each (candidate) {
    // do not allow ring opening in the children scaffold
    if (candidate.max_ring_size > child.max_ring_size) continue;
    // parent should be the largest candidate (if other criteria fit)
    if (candidate.number_of_atoms > parent.number_of_atoms + 2) {
        parent = candidate;
        continue;
    }
    if (parent == null) parent = candidate;
    // select parent with maximal number of ring bonds
    if (candidate.nonring_bond_count > parent.nonring_bond_count)
        continue;
    // take parent with more heteroatoms
    if (candidate.number_of_heteroatoms > parent.number_of_heteroatoms) {
        parent = candidate;
        continue;
    }
    // if still more possibilities, take the most common scaffold
```

```
    if (candidate.frequency > parent.frequency) {  
        parent = candidate;  
        continue;  
    }  
}
```

\*\*\*\*\*

#### 6.4.5 Statistische Auswertung des Strukturbaums

Die Verteilung der Naturstoffe nach ihrer Anzahl wurde gegen die hierarchischen Ringebenen des Strukturbaums (gegen die Anzahl der Ringe) aufgetragen.

Darüber hinaus wurde das van der Waals-Volumen der deglykosylierten und nach der Anzahl der Ringe im Ringgerüst sortierten Naturstoffe berechnet. Die Anzahl der Naturstoffe, gruppiert nach der Anzahl der Ringe im Grundgerüst, wurde gegen das van der Waals-Volumen aufgetragen. Die Berechnung des van der Waals-Volumens erfolgte mittels eines Java-basierten Programms der Firma Novartis. Grundlage der Berechnung waren 3D-Strukturen der Naturstoffe, die unter Anwendung des Programms CORINA 3.1 (Molecular Networks, Erlangen, Germany, <http://www.mol-net.de/>) aus den SMILES-Strings der Moleküle generiert wurden. Eine analoge Analyse der Volumenverteilung wurde an ca. 30 000 Arzneistoffen des World Drug Index (Thomson Derwent, <http://thomsonderwent.com>) vorgenommen.

#### 6.4.6 Annotation des Strukturbaums

Aus dem BSRC-Feld des DNP wurde der Genus des Quellorganismus extrahiert und taxonomisch nach der ITIS-Datenbank (Integrated Taxonomic Information System, <http://www.itis.usda.gov/>) klassifiziert. Für in ITIS nicht gelistete Spezies wurde die taxonomische Zuordnung nach dem NCBI Taxonomy Browser (<http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/>) vorgenommen. Das so ermittelte Taxon wurde in den taxonomischen Baum nach ITIS auf der niedrigst möglichen taxonomischen Stufe integriert.

Die kanonischen SMILES-Strings<sup>[236]</sup> ohne Stereochemie der Strukturen im DNP und im MDL Drug Data Report (MDDR) (<http://www.mdl.com>) wurden miteinander verglichen. Aus diesem Vergleich wurde die Schnittmenge zwischen beiden Datenbanken bestimmt und diese mit der Information über biologische Aktivität aus dem MDDR annotiert.

Alle Annotationen und die statistische Auswertung wurden unter Verwendung der PIPELINEPILOT-Software (<http://www.scitegic.com>) realisiert.

## 7 Literaturverzeichnis

- [1] S. Class, *Chem. Eng. News* **2002**, *80*, 39-49.
- [2] S. L. Schreiber, *Science* **2000**, *287*, 1964-1969.
- [3] A. Golebiowski, S. R. Klopfenstein, D. E. Portlock, *Curr. Opin. Chem. Biol.* **2001**, *5*, 273-284.
- [4] J. S. Mason, N. A. Hermsmeier, *Curr. Opin. Chem. Biol.* **1999**, *3*, 342-349.
- [5] R. W. Spencer, *Biotechnol. Bioeng.* **1998**, *61*, 61-67.
- [6] W. P. Walters, A. Murcko, M. A. Murcko, *Curr. Opin. Chem. Biol.* **1999**, *3*, 384-387.
- [7] Ajay, W. P. Walters, M. A. Murcko, *J. Med. Chem.* **1998**, *41*, 3314-3324.
- [8] J. Sadowski, H. Kubinyi, *J. Med. Chem.* **1998**, *41*, 3325-3329.
- [9] A. K. Ghose, V. N. Viswanadhan, J. J. Wendoloski, *J. Comb. Chem.* **1999**, *1*, 55-68.
- [10] M. L. Lee, G. Schneider, *J. Comb. Chem.* **2001**, *3*, 284-289.
- [11] R. Breinbauer, I. R. Vetter, H. Waldmann, *Angew. Chem. Int. Ed.* **2002**, *41*, 2879-2890.
- [12] M. A. Koch, R. Breinbauer, H. Waldmann, *Biol. Chem.* **2003**, *384*, 1265-1272.
- [13] M. A. Koch, H. Waldmann, in *Chemogenomics in Drug Discovery: A Medicinal Chemistry Perspective* (Eds.: H. Kubinyi, G. Müller), Wiley-VCH, Weinheim, **2004**, pp. 377-403.
- [14] B. E. Evans, K. E. Rittle, M. G. Bock, R. M. Dipardo, R. M. Freidinger, W. L. Whitter, G. F. Lundell, D. F. Veber, P. S. Anderson, R. S. L. Chang, V. J. Lotti, D. J. Cerino, T. B. Chen, P. J. Kling, K. A. Kunkel, J. P. Springer, J. Hirshfield, *J. Med. Chem.* **1988**, *31*, 2235-2246.
- [15] D. Brohm, S. Metzger, A. Bhargava, O. Müller, F. Lieb, H. Waldmann, *Angew. Chem. Int. Ed.* **2002**, *41*, 307-311.
- [16] D. Brohm, N. Philippe, S. Metzger, A. Bhargava, O. Müller, F. Lieb, H. Waldmann, *J. Am. Chem. Soc.* **2002**, *124*, 13171-13178.

- [17] N. V. Grishin, *J. Struct. Biol.* **2001**, *134*, 167-185.
- [18] K. H. Bleicher, H. J. Bohm, K. Muller, A. I. Alanine, *Nat. Rev. Drug Discov.* **2003**, *2*, 369-378.
- [19] L. J. Gershell, J. H. Atkins, *Nat. Rev. Drug Discov.* **2003**, *2*, 321-327.
- [20] G. Wess, *Drug Discov. Today* **2002**, *7*, 533-535.
- [21] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Deliv. Rev.* **1997**, *23*, 3-25.
- [22] J. Zuegge, U. Fechner, O. Roche, N. J. Parrott, O. Engkvist, G. Schneider, *QSAR Comb. Sci.* **2002**, *21*, 249-256.
- [23] O. Roche, G. Trube, J. Zuegge, P. Pflimlin, A. Alanine, G. Schneider, *ChemBioChem* **2002**, *3*, 455-459.
- [24] J. S. Mason, A. C. Good, E. J. Martin, *Curr. Pharm. Des.* **2001**, *7*, 567-597.
- [25] C. Bissantz, G. Folkers, D. Rognan, *J. Med. Chem.* **2000**, *43*, 4759-4767.
- [26] A. N. Lupas, C. P. Ponting, R. B. Russell, *J. Struct. Biol.* **2001**, *134*, 191-203.
- [27] S. Govindarajan, R. Recabarren, R. K. Goldstein, *Proteins* **1999**, *35*, 408-414.
- [28] C. P. Ponting, J. Schultz, R. R. Copley, M. A. Andrade, P. Bork, in *Advances in Protein Chemistry, Vol. 54*, **2000**, pp. 185-244.
- [29] A. G. Murzin, S. E. Brenner, T. Hubbard, C. Chothia, *J. Mol. Biol.* **1995**, *247*, 536-540.
- [30] A. Andreeva, D. Howorth, S. E. Brenner, T. J. Hubbard, C. Chothia, A. G. Murzin, *Nucleic Acids Res.* **2004**, *32 Database issue*, D226-229.
- [31] C. A. Orengo, D. T. Jones, J. M. Thornton, *Nature* **1994**, *372*, 631-634.
- [32] C. Chothia, *Nature* **1992**, *357*, 543-544.
- [33] C. T. Zhang, *Protein Eng.* **1997**, *10*, 757-761.
- [34] Z. X. Wang, *Protein Eng.* **1998**, *11*, 621-626.
- [35] C. O. Zhang, C. DeLisi, *J. Mol. Biol.* **1998**, *284*, 1301-1305.

- [36] Y. I. Wolf, N. V. Grishin, E. V. Koonin, *J. Mol. Biol.* **2000**, *299*, 897-905.
- [37] A. F. W. Coulson, J. Moulton, *Proteins* **2002**, *46*, 61-71.
- [38] E. V. Koonin, Y. I. Wolf, G. P. Karev, *Nature* **2002**, *420*, 218-223.
- [39] C. Chothia, A. M. Lesk, *EMBO J.* **1986**, *5*, 823-826.
- [40] D. J. Newman, G. M. Cragg, K. M. Snader, *J. Nat. Prod.* **2003**, *66*, 1022-1037.
- [41] M. A. Koch, L.-O. Wittenberg, S. Basu, D. A. Jeyaraj, E. Gourzoulidou, K. Reinecke, A. Odermatt, H. Waldmann, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16721-16726.
- [42] T. Henkel, R. M. Brunne, H. Müller, F. Reichel, *Angew. Chem. Int. Ed.* **1999**, *38*, 643-647.
- [43] M. Feher, J. M. Schmidt, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 218-227.
- [44] K. H. Bleicher, *Curr. Med. Chem.* **2002**, *9*, 2077-2084.
- [45] T. J. Mitchison, *Chem. Biol.* **1994**, *1*, 3-6.
- [46] S. L. Schreiber, *Chem. Eng. News* **2003**, *81*, 51-+.
- [47] D. Falb, S. Jindal, *Curr. Opin. Drug Discov. Devel.* **2002**, *5*, 532-539.
- [48] A. M. Piggott, P. Karuso, *Comb. Chem. High Throughput Screen.* **2004**, *7*, 607-630.
- [49] E. Jacoby, A. Schuffenhauer, P. Floersheim, *Drug News Perspect.* **2003**, *16*, 93-102.
- [50] S. V. Frye, *Chem. Biol.* **1999**, *6*, R3-7.
- [51] M. M. G. M. Thunnissen, P. Nordlund, J. Z. Haeggstrom, *Nat. Struct. Biol.* **2001**, *8*, 131-135.
- [52] P. C. Rudberg, F. Tholander, M. Andberg, M. M. Thunnissen, J. Z. Haeggstrom, *J. Biol. Chem.* **2004**, *279*, 27376-27382.
- [53] L. Orning, G. Krivi, F. A. Fitzpatrick, *J. Biol. Chem.* **1991**, *266*, 1375-1378.
- [54] W. Yuan, B. Munoz, C. H. Wong, J. Z. Haeggstrom, A. Wetterholm, B. Samuelsson, *J. Med. Chem.* **1993**, *36*, 211-220.
- [55] I. R. Ollmann, J. H. Hogg, B. Munoz, J. Z. Haeggstrom, B. Samuelsson, C. H. Wong, *Bioorg. Med. Chem.* **1995**, *3*, 969-995.

- [56] J. H. Hogg, I. R. Ollmann, J. Z. Haeggstrom, A. Wetterholm, B. Samuelsson, C. H. Wong, *Bioorg. Med. Chem.* **1995**, *3*, 1405-1415.
- [57] M. Q. Zhang, *Curr. Med. Chem.* **1997**, *4*, 67-78.
- [58] M. M. G. M. Thunnissen, B. Andersson, B. Samuelsson, C. H. Wong, J. Z. Haeggstrom, *FASEB J.* **2002**, *16*, 1648-1650.
- [59] R. Natesh, S. L. Schwager, E. D. Sturrock, K. R. Acharya, *Nature* **2003**, *421*, 551-554.
- [60] H. M. Holden, D. E. Tronrud, A. F. Monzingo, L. H. Weaver, B. W. Matthews, *Biochemistry* **1987**, *26*, 8542-8553.
- [61] Y. Kakuta, L. G. Pedersen, C. W. Carter, M. Negishi, L. C. Pedersen, *Nat. Struct. Biol.* **1997**, *4*, 904-908.
- [62] Y. Kakuta, E. V. Petrotchenko, L. C. Pedersen, M. Negishi, *J. Biol. Chem.* **1998**, *273*, 27325-27330.
- [63] T. C. Norman, N. S. Gray, J. T. Koh, P. G. Schultz, *J. Am. Chem. Soc.* **1996**, *118*, 7430-7431.
- [64] N. S. Gray, L. Wodicka, A. Thunnissen, T. C. Norman, S. J. Kwon, F. H. Espinoza, D. O. Morgan, G. Barnes, S. LeClerc, L. Meijer, S. H. Kim, D. J. Lockhart, P. G. Schultz, *Science* **1998**, *281*, 533-538.
- [65] H. J. Müller-Dieckmann, G. E. Schulz, *J. Mol. Biol.* **1994**, *236*, 361-367.
- [66] H. J. Müller-Dieckmann, G. E. Schulz, *J. Mol. Biol.* **1995**, *246*, 522-530.
- [67] J. I. Armstrong, A. R. Portley, Y. T. Chang, D. M. Nierengarten, B. N. Cook, K. G. Bowman, A. Bishop, N. S. Gray, K. M. Shokat, P. G. Schultz, C. R. Bertozzi, *Angew. Chem. Int. Ed.* **2000**, *39*, 1303-1306.
- [68] D. E. Verdugo, M. T. Cancilla, X. Ge, N. S. Gray, Y. T. Chang, P. G. Schultz, M. Negishi, J. A. Leary, C. R. Bertozzi, *J. Med. Chem.* **2001**, *44*, 2683-2686.
- [69] E. Chapman, S. Ding, P. G. Schultz, C. H. Wong, *J. Am. Chem. Soc.* **2002**, *124*, 14524-14525.
- [70] M. Robinson-Rechavi, H. Escrivá Garcia, V. Laudet, *J. Cell. Sci.* **2003**, *116*, 585-586.
- [71] M. Schapira, *Curr. Cancer Drug Targets* **2002**, *2*, 243-256.

- [72] M. Schapira, B. M. Raaka, H. H. Samuels, R. Abagyan, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1008-1013.
- [73] T. Claudel, E. Sturm, F. Kuipers, B. Staels, *Expert Opin. Investig. Drugs* **2004**, *13*, 1135-1148.
- [74] A. C. W. Pike, A. M. Brzozowski, R. E. Hubbard, T. Bonn, A. G. Thorsell, O. Engstrom, J. Ljunggren, J. K. Gustafsson, M. Carlquist, *EMBO J.* **1999**, *18*, 4608-4618.
- [75] Z. C. Dang, V. Audinot, S. E. Papapoulos, J. A. Boutin, C. Lowik, *J. Biol. Chem.* **2003**, *278*, 962-967.
- [76] L. Van Gaal, A. J. Scheen, *Diabetes Metab. Res. Rev.* **2002**, *18 Suppl 2*, S1-4.
- [77] K. C. Nicolaou, J. A. Pfefferkorn, A. J. Roecker, G. Q. Cao, S. Barluenga, H. J. Mitchell, *J. Am. Chem. Soc.* **2000**, *122*, 9939-9953.
- [78] M. Downes, M. A. Verdecia, A. J. Roecker, R. Hughes, J. B. Hogenesch, H. R. Kast-Woelbern, M. E. Bowman, J. L. Ferrer, A. M. Anisfeld, P. A. Edwards, J. M. Rosenfeld, J. G. A. Alvarez, J. P. Noel, K. C. Nicolaou, R. M. Evans, *Mol. Cell* **2003**, *11*, 1079-1092.
- [79] R. T. Nolte, G. B. Wisely, S. Westin, J. E. Cobb, M. H. Lambert, R. Kurokawa, M. G. Rosenfeld, T. M. Willson, C. K. Glass, M. V. Milburn, *Nature* **1998**, *395*, 137-143.
- [80] K. C. Nicolaou, J. A. Pfefferkorn, H. J. Mitchell, A. J. Roecker, S. Barluenga, G. Q. Cao, R. L. Affleck, J. E. Lillig, *J. Am. Chem. Soc.* **2000**, *122*, 9954-9967.
- [81] K. C. Nicolaou, R. M. Evans, A. J. Roecker, R. Hughes, M. Downes, J. A. Pfefferkorn, *Org. Biomol. Chem.* **2003**, *1*, 908-920.
- [82] H. Gronemeyer, J. A. Gustafsson, V. Laudet, *Nat. Rev. Drug Discov.* **2004**, *3*, 950-964.
- [83] I. M. Billas, T. Iwema, J. M. Garnier, A. Mitschler, N. Rochel, D. Moras, *Nature* **2003**, *426*, 91-96.
- [84] M. Farnegardh, T. Bonn, S. Sun, J. Ljunggren, H. Ahola, A. Wilhelmsson, J. A. Gustafsson, M. Carlquist, *J. Biol. Chem.* **2003**, *278*, 38821-38828.

- [85] N. L. Urizar, A. B. Liverman, D. T. Dodds, F. V. Silva, P. Ordentlich, Y. Z. Yan, F. J. Gonzalez, R. A. Heyman, D. J. Mangelsdorf, D. D. Moore, *Science* **2002**, 296, 1703-1706.
- [86] A. A. Bogan, F. E. Cohen, T. S. Scanlan, *Nat. Struct. Biol.* **1998**, 5, 679-681.
- [87] L. Lo Conte, B. Ailey, T. J. Hubbard, S. E. Brenner, A. G. Murzin, C. Chothia, *Nucleic Acids Res.* **2000**, 28, 257-259.
- [88] C. A. Orengo, A. D. Michie, S. Jones, D. T. Jones, M. B. Swindells, J. M. Thornton, *Structure* **1997**, 5, 1093-1108.
- [89] L. Holm, C. Sander, *Science* **1996**, 273, 595-602.
- [90] L. Holm, C. Ouzounis, C. Sander, G. Tuparev, G. Vriend, *Protein Sci.* **1992**, 1, 1691-1698.
- [91] L. Holm, C. Sander, *Nucleic Acids Res.* **1994**, 22, 3600-3609.
- [92] L. Holm, C. Sander, *Nucleic Acids Res.* **1996**, 24, 206-209.
- [93] L. Holm, C. Sander, *Nucleic Acids Res.* **1997**, 25, 231-234.
- [94] L. Holm, C. Sander, *Nucleic Acids Res.* **1998**, 26, 316-319.
- [95] T. J. Hubbard, B. Ailey, S. E. Brenner, A. G. Murzin, C. Chothia, *Nucleic Acids Res.* **1999**, 27, 254-256.
- [96] L. Holm, C. Sander, *J. Mol. Biol.* **1993**, 233, 123-138.
- [97] I. N. Shindyalov, P. E. Bourne, *Protein Eng.* **1998**, 11, 739-747.
- [98] C. Guda, S. Lu, E. D. Scheeff, P. E. Bourne, I. N. Shindyalov, *Nucleic Acids Res.* **2004**, 32, W100-103.
- [99] M. Shatsky, O. Dror, D. Schneidman-Duhovny, R. Nussinov, H. J. Wolfson, *Nucleic Acids Res.* **2004**, 32, W503-507.
- [100] J. Shapiro, D. Brutlag, *Protein Sci.* **2004**, 13, 278-294.
- [101] J. Shapiro, D. Brutlag, *Nucleic Acids Res.* **2004**, 32, W536-W541.
- [102] W. Wriggers, K. Schulten, *Proteins* **1997**, 29, 1-14.
- [103] W. R. Taylor, C. A. Orengo, *J. Mol. Biol.* **1989**, 208, 1-22.

- [104] S. Subbiah, D. V. Laurents, M. Levitt, *Curr. Biol.* **1993**, *3*, 141-148.
- [105] G. J. Kleywegt, *Acta Crystallogr. D Biol. Crystallogr.* **1996**, *52*, 842-857.
- [106] E. Krissinel, K. Henrick, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2256-2268.
- [107] R. Day, D. A. Beck, R. S. Armen, V. Daggett, *Protein Sci.* **2003**, *12*, 2150-2160.
- [108] X. Yuan, C. Bystroff, *Bioinformatics* **2005**, *21*, 1010-1019.
- [109] A. S. Yang, B. Honig, *J. Mol. Biol.* **2000**, *301*, 665-678.
- [110] A. Zemla, *Nucleic Acids Res.* **2003**, *31*, 3370-3374.
- [111] A. S. Yang, *Bioinformatics* **2002**, *18*, 1658-1665.
- [112] E. B. Fauman, J. P. Cogswell, B. Lovejoy, W. J. Rocque, W. Holmes, V. G. Montana, H. Piwnica-Worms, M. J. Rink, M. A. Saper, *Cell* **1998**, *93*, 617-625.
- [113] I. Blomberg, I. Hoffmann, *Mol. Cell. Biol.* **1999**, *19*, 6183-6194.
- [114] J. Bartek, J. Lukas, *FEBS Lett.* **2001**, *490*, 117-122.
- [115] D. F. McCain, I. E. Catrina, A. C. Hengge, Z. Y. Zhang, *J. Biol. Chem.* **2002**, *277*, 11190-11200.
- [116] Z. Y. Zhang, *Curr. Opin. Chem. Biol.* **2001**, *5*, 416-423.
- [117] J. L. Sussman, M. Harel, F. Frolow, C. Oefner, A. Goldman, L. Toker, I. Silman, *Science* **1991**, *253*, 872-879.
- [118] B. Ibach, E. Haen, *Curr. Pharm. Des.* **2004**, *10*, 231-251.
- [119] B. R. Walker, J. R. Seckl, *Expert Opin. Ther. Targets* **2003**, *7*, 771-783.
- [120] H. Jornvall, B. Persson, M. Krook, S. Atrian, R. Gonzalez-Duarte, J. Jeffery, D. Ghosh, *Biochemistry* **1995**, *34*, 6003-6013.
- [121] R. A. Schweizer, A. G. Atanasov, B. M. Frey, A. Odermatt, *Mol. Cell Endocrinol.* **2003**, *212*, 41-49.
- [122] H. Masuzaki, J. S. Flier, *Curr. Drug Targets Immune Endocr. Metabol. Disord.* **2003**, *3*, 255-262.

- [123] P. Alberts, L. Engblom, N. Edling, M. Forsgren, G. Klingstrom, C. Larsson, Y. Ronquist-Nii, B. Ohman, L. Abrahmsen, *Diabetologia* **2002**, *45*, 1528-1532.
- [124] S. A. Ross, E. A. Gulve, M. Wang, *Chem. Rev.* **2004**, *104*, 1255-1282.
- [125] T. C. Sandeep, J. L. Yau, A. M. MacLulich, J. Noble, I. J. Deary, B. R. Walker, J. R. Seckl, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6734-6739.
- [126] M. I. New, R. C. Wilson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12790-12797.
- [127] B. Boeckmann, A. Bairoch, R. Apweiler, M. C. Blatter, A. Estreicher, E. Gasteiger, M. J. Martin, K. Michoud, C. O'Donovan, I. Phan, S. Pilbout, M. Schneider, *Nucleic Acids Res.* **2003**, *31*, 365-370.
- [128] K. Ginalski, A. Elofsson, D. Fischer, L. Rychlewski, *Bioinformatics* **2003**, *19*, 1015-1018.
- [129] A. Tramontano, V. Morea, *Proteins* **2003**, *53 Suppl. 6*, 352-368.
- [130] L. Rychlewski, L. Jaroszewski, W. Li, A. Godzik, *Protein Sci.* **2000**, *9*, 232-241.
- [131] L. J. McGuffin, D. T. Jones, *Bioinformatics* **2003**, *19*, 874-881.
- [132] D. Fischer, J. Pas, L. Rychlewski, *Bioinformatics* **2004**, *20*, 2482-2484.
- [133] A. Sali, T. L. Blundell, *J. Mol. Biol.* **1993**, *234*, 779-815.
- [134] B. Wallner, A. Elofsson, *Protein Sci.* **2003**, *12*, 1073-1086.
- [135] R. W. Hooft, G. Vriend, C. Sander, E. E. Abola, *Nature* **1996**, *381*, 272.
- [136] R. A. Laskowski, M. W. Macarthur, D. S. Moss, J. M. Thornton, *J. Appl. Cryst.* **1993**, *26*, 283-291.
- [137] D. J. Hosfield, Y. Wu, R. J. Skene, M. Hilgers, A. Jennings, G. P. Snell, K. Aertgeerts, *J. Biol. Chem.* **2005**, *280*, 4639-4648.
- [138] M. A. Lyon, A. P. Ducruet, P. Wipf, J. S. Lazo, *Nat. Rev. Drug Discov.* **2002**, *1*, 961-976.
- [139] L.-O. Wittenberg, Dissertation, Universität Dortmund, **2004**.
- [140] D. Brohm, S. Metzger, A. Bhargava, O. Müller, F. Lieb, H. Waldmann, *Angew. Chem. Int. Ed.* **2001**, *41*, 307-311.

- [141] M. Ihara, S. Suzuki, N. Taniguchi, K. Fukumoto, *J. Chem. Soc. Perkin Trans. I* **1993**, 2251-2258.
- [142] M. Frigerio, M. Santagostino, S. Sputore, G. Palmisano, *J. Org. Chem.* **1995**, *60*, 7272-7276.
- [143] M. R. Kernan, D. J. Faulkner, *J. Org. Chem.* **1988**, *53*, 2773-2776.
- [144] E. J. Corey, B. E. Roberts, *J. Am. Chem. Soc.* **1997**, *119*, 12425-12431.
- [145] S. P. Brown, N. C. Goodwin, D. W. C. MacMillan, *J. Am. Chem. Soc.* **2003**, *125*, 1192-1194.
- [146] B. Baratte, L. Meijer, K. Galaktionov, D. Beach, *Anticancer Res.* **1992**, *12*, 873-880.
- [147] G. L. Ellman, K. D. Courtney, V. Andres, R. M. Featherstone, *Biochem. Pharmacol.* **1961**, *7*, 88-&.
- [148] J. M. Contreras, I. Parrot, W. Sippl, Y. M. Rival, C. G. Wermuth, *J. Med. Chem.* **2001**, *44*, 2707-2718.
- [149] P. Arnold, S. Tam, L. Yan, M. E. Baker, F. J. Frey, A. Odermatt, *Mol. Cell Endocrinol.* **2003**, *201*, 177-187.
- [150] S. L. McGovern, B. T. Helfand, B. Feng, B. K. Shoichet, *J. Med. Chem.* **2003**, *46*, 4265-4272.
- [151] S. P. Gunasekera, P. J. McCarthy, M. KellyBorges, E. Lobkovsky, J. Clardy, *J. Am. Chem. Soc.* **1996**, *118*, 8759-8760.
- [152] A. Maelicke, E. X. Albuquerque, *Drug Discov. Today* **1996**, *1*, 53-59.
- [153] J. Kaur, M. Q. Zhang, *Curr. Med. Chem.* **2000**, *7*, 273-294.
- [154] R. B. Russell, P. D. Sasieni, M. J. E. Sternberg, *J. Mol. Biol.* **1998**, *282*, 903-918.
- [155] A. Achari, D. O. Somers, J. N. Champness, P. K. Bryant, J. Rosemond, D. K. Stammers, *Nat. Struct. Biol.* **1997**, *4*, 490-497.
- [156] D. P. Richey, G. M. Brown, *J. Biol. Chem.* **1969**, *244*, 1582-1592.
- [157] T. Shiota, C. M. Baugh, R. Jackson, R. Dillard, *Biochemistry* **1969**, *8*, 5022-5028.

- [158] D. W. Banner, A. C. Bloomer, G. A. Petsko, D. C. Phillips, C. I. Pogson, I. A. Wilson, P. H. Corran, A. J. Furth, J. D. Milman, R. E. Offord, J. D. Priddle, S. G. Waley, *Nature* **1975**, 255, 609-614.
- [159] M. Henn-Sax, B. Hocker, M. Wilmanns, R. Sterner, *Biol. Chem.* **2001**, 382, 1315-1320.
- [160] J. J. Almrud, M. A. Oliveira, A. D. Kern, N. V. Grishin, M. A. Phillips, M. L. Hackert, *J. Mol. Biol.* **2000**, 295, 7-16.
- [161] L. K. Jackson, H. B. Brooks, A. L. Osterman, E. J. Goldsmith, M. A. Phillips, *Biochemistry* **2000**, 39, 11247-11257.
- [162] A. M. Baca, R. Sirawaraporn, S. Turley, W. Sirawaraporn, W. G. Hol, *J. Mol. Biol.* **2000**, 302, 1193-1212.
- [163] N. V. Grishin, A. L. Osterman, H. B. Brooks, M. A. Phillips, E. J. Goldsmith, *Biochemistry* **1999**, 38, 15174-15184.
- [164] G. Domagk, *Dtsch. Med. Wochenschr.* **1935**, LXI, 250.
- [165] J. Trefouel, J. Trefouel, F. Nitti, D. Bovet, *C. R. Seances Soc. Biol.* **1935**, 120, 756.
- [166] C. J. Bacchi, H. C. Nathan, S. H. Hutner, P. P. McCann, A. Sjoerdsma, *Science* **1980**, 210, 332-334.
- [167] F. A. Kuzoe, *Acta Trop.* **1993**, 54, 153-162.
- [168] E. W. Gerner, F. L. Meyskens, Jr., *Nat. Rev. Cancer* **2004**, 4, 781-792.
- [169] A. E. Pegg, L. M. Shantz, C. S. Coleman, *J. Cell Biochem. Suppl.* **1995**, 22, 132-138.
- [170] L. K. Jackson, E. J. Goldsmith, M. A. Phillips, *J. Biol. Chem.* **2003**, 278, 22037-22043.
- [171] L. K. Jackson, M. A. Phillips, *Curr. Top. Med. Chem.* **2002**, 2, 425-438.
- [172] P. P. McCann, A. E. Pegg, *Pharmacol. Ther.* **1992**, 54, 195-215.
- [173] B. W. Metcalf, P. Bey, C. Danzin, M. J. Jung, P. Casara, J. P. Vevvert, *J. Am. Chem. Soc.* **1978**, 100, 2551-2553.

- [174] P. Bey, C. Danzin, V. Vandorsselaer, P. Mamont, M. Jung, C. Tardif, *J. Med. Chem.* **1978**, *21*, 50-55.
- [175] C. Danzin, P. Casara, N. Claverie, B. W. Metcalf, *J. Med. Chem.* **1981**, *24*, 16-20.
- [176] J. Stanek, J. Frei, H. Mett, P. Schneider, U. Regenass, *J. Med. Chem.* **1992**, *35*, 1339-1344.
- [177] G. Aizencang, R. B. Frydman, S. Giorgieri, L. Sambrotta, L. Guerra, B. Frydman, *J. Med. Chem.* **1995**, *38*, 4337-4341.
- [178] A. S. Bachmann, P. Matile, A. J. Slusarenko, *Physiol. Mol. Plant Pathol.* **1998**, *53*, 287-299.
- [179] D. B. Langley, M. D. Templeton, B. A. Fields, R. E. Mitchell, C. A. Collyer, *J. Biol. Chem.* **2000**, *275*, 20012-20019.
- [180] L. Badolo, J. Dubois, M. Helson-Cambier, M. Hanocq, *Eur. J. Pharmacol.* **1998**, *342*, R1-R2.
- [181] H. Zollner, *Handbook of Enzyme Inhibitors*, VCH Verlagsgesellschaft, Weinheim, **1993**.
- [182] I. Schomburg, A. Chang, C. Ebeling, M. Gremse, C. Heldt, G. Huhn, D. Schomburg, *Nucleic Acids Res.* **2004**, *32*, D431-433.
- [183] U. Meyer, A. Schuhmann, C. Friedrich, *Pharm. Unserer Zeit* **2002**, *31*, 242-250.
- [184] M. F. Chan, I. Okun, F. L. Stavros, E. Hwang, M. E. Wolff, V. N. Balaji, *Biochem. Biophys. Res. Commun.* **1994**, *201*, 228-234.
- [185] P. D. Stein, J. T. Hunt, D. M. Floyd, S. Moreland, K. E. Dickinson, C. Mitchell, E. C. Liu, M. L. Webb, N. Murugesan, J. Dickey, et al., *J. Med. Chem.* **1994**, *37*, 329-331.
- [186] C. P. Rains, S. Noble, D. Faulds, *Drugs* **1995**, *50*, 137-156.
- [187] A. J. Levi, A. M. Fisher, L. Hughes, W. F. Hendry, *Lancet* **1979**, *2*, 276-278.
- [188] C. Pholpramool, S. Ruchirawat, V. Verawatnapakul, C. Paovalo, L. M. Lewin, *J. Reprod. Fertil.* **1991**, *92*, 169-178.
- [189] P. Coffino, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 4421-4423.

- [190] S. E. Jeong, Y. Lee, J. H. Hwang, D. C. Knipple, *J. Exp. Biol.* **2001**, *204*, 3935-3942.
- [191] A. Osterman, N. V. Grishin, L. N. Kinch, M. A. Phillips, *Biochemistry* **1994**, *33*, 13662-13667.
- [192] R. L. Strausberg, S. L. Schreiber, *Science* **2003**, *300*, 294-295.
- [193] P. Burkhard, P. Dominici, C. Borri-Voltattorni, J. N. Jansonius, V. N. Malashkevich, *Nat. Struct. Biol.* **2001**, *8*, 963-967.
- [194] N. V. Grishin, *J. Mol. Biol.* **1999**, *291*, 239-247.
- [195] C. Fan, P. C. Moews, C. T. Walsh, J. R. Knox, *Science* **1994**, *266*, 439-443.
- [196] T. K. Ritter, C. H. Wong, *Angew. Chem. Int. Ed.* **2001**, *40*, 3508-3533.
- [197] X. Zhu, J. L. Kim, J. R. Newcomb, P. E. Rose, D. R. Stover, L. M. Toledo, H. Zhao, K. A. Morgenstern, *Structure Fold. Des.* **1999**, *7*, 651-661.
- [198] G. Chiosis, M. Vilenchik, J. Kim, D. Solit, *Drug Discov. Today* **2004**, *9*, 881-888.
- [199] M. Jahnz, M. A. Medina, P. Schwille, *ChemBioChem* **2005**, *6*, 1-7.
- [200] A. Tanitame, Y. Oyamada, K. Ofuji, M. Fujimoto, K. Suzuki, T. Ueda, H. Terauchi, M. Kawasaki, K. Nagai, M. Wachi, J. Yamagishi, *Bioorg. Med. Chem.* **2004**, *12*, 5515-5524.
- [201] M. Matsushita, K. D. Janda, *Bioorg. Med. Chem.* **2002**, *10*, 855-867.
- [202] L. Brino, A. Urzhumtsev, M. Mousli, C. Bronner, A. Mitschler, P. Oudet, D. Moras, *J. Biol. Chem.* **2000**, *275*, 9468-9475.
- [203] W. M. Obermann, H. Sondermann, A. A. Russo, N. P. Pavletich, F. U. Hartl, *J. Cell Biol.* **1998**, *143*, 901-910.
- [204] K. D. Corbett, J. M. Berger, *EMBO J.* **2003**, *22*, 151-163.
- [205] T. Tanaka, S. K. Saha, C. Tomomori, R. Ishima, D. Liu, K. I. Tong, H. Park, R. Dutta, L. Qin, M. B. Swindells, T. Yamazaki, A. M. Ono, M. Kainosho, M. Inouye, M. Ikura, *Nature* **1998**, *396*, 88-92.
- [206] J. Hermoso, D. Pignol, B. Kerfelec, I. Crenon, C. Chapus, J. C. Fontecilla-Camps, *J. Biol. Chem.* **1996**, *271*, 18007-18016.

- [207] Y. Devedjiev, Z. Dauter, S. R. Kuznetsov, T. L. Jones, Z. S. Derewenda, *Structure Fold. Des.* **2000**, *8*, 1137-1146.
- [208] S. J. Kridel, F. Axelrod, N. Rozenkrantz, J. W. Smith, *Cancer Res.* **2004**, *64*, 2070-2075.
- [209] B. Chakravarty, Z. Gu, S. S. Chirala, S. J. Wakil, F. A. Quioco, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15567-15572.
- [210] F. E. Koehn, G. T. Carter, *Nat. Rev. Drug Discov.* **2005**, *4*, 206-220.
- [211] M. Vieth, R. E. Higgs, D. H. Robertson, M. Shapiro, E. A. Gragg, H. Hemmerle, *Biochim. Biophys. Acta* **2004**, *1697*, 243-257.
- [212] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, *Science* **2002**, *298*, 1912-1934.
- [213] M. E. Noble, J. A. Endicott, L. N. Johnson, *Science* **2004**, *303*, 1800-1805.
- [214] R. L. Dow, T. T. Chou, B. M. Bechle, C. Goddard, E. R. Larson, *J. Med. Chem.* **1994**, *37*, 2224-2231.
- [215] Y. Kashiwada, L. Huang, L. M. Ballas, J. B. Jiang, W. P. Janzen, K. H. Lee, *J. Med. Chem.* **1994**, *37*, 195-200.
- [216] Z. Zhao, W. H. Leister, R. G. Robinson, S. F. Barnett, D. Defeo-Jones, R. E. Jones, G. D. Hartman, J. R. Huff, H. E. Huber, M. E. Duggan, C. W. Lindsley, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 905-909.
- [217] C. Sun, Y. Newbatt, L. Douglas, P. Workman, W. Aherne, S. Linardopoulos, *J. Biomol. Screen.* **2004**, *9*, 391-397.
- [218] Y. Suzuki, K. Nakano, M. Sugiyama, J. Imagawa, *J. Cardiovasc. Pharmacol.* **2004**, *44*, 329-334.
- [219] L. Alberghina, F. Chiaradonna, M. Vanoni, *ChemBioChem* **2004**, *5*, 1322-1333.
- [220] E. K. Han, T. McGonigal, J. Wang, V. L. Giranda, Y. Luo, *Anticancer Res.* **2004**, *24*, 3899-3905.
- [221] B. B. Dokken, J. A. Sloniger, E. J. Henriksen, *Am. J. Physiol. Endocrinol. Metab.* **2005**.

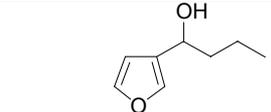
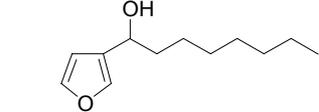
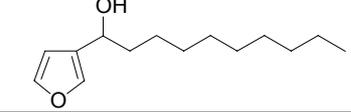
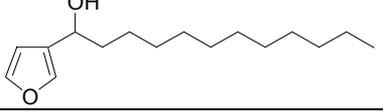
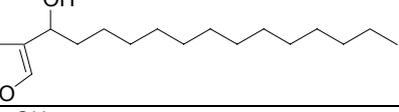
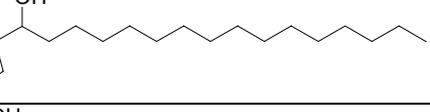
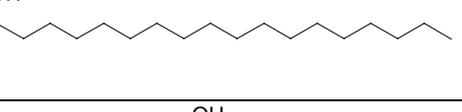
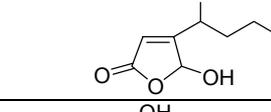
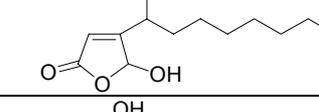
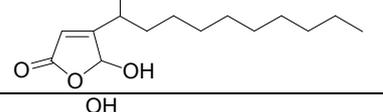
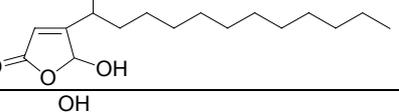
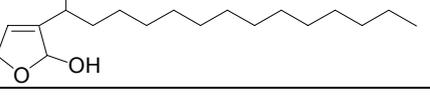
- [222] V. M. Macaulay, *Novartis Found. Symp.* **2004**, 262, 235-243; discussion 243-236, 265-238.
- [223] L. Resnick, M. Fennell, *Drug Discov. Today* **2004**, 9, 932-939.
- [224] S. Kumar, J. Boehm, J. C. Lee, *Nat. Rev. Drug Discov.* **2003**, 2, 717-726.
- [225] P. F. Cirillo, C. Pargellis, J. Regan, *Curr. Top. Med. Chem.* **2002**, 2, 1021-1035.
- [226] R. K. Vadlamudi, R. Kumar, *Cancer Metastasis Rev.* **2003**, 22, 385-393.
- [227] A. E. Gururaj, S. K. Rayala, R. Kumar, *Breast Cancer Res.* **2005**, 7, 5-12.
- [228] N. Takai, R. Hamanaka, J. Yoshimatsu, I. Miyakawa, F. Eckerdt, J. Yuan, K. Strebhardt, *Oncogene* **2005**, 24, 287-291.
- [229] F. Eckerdt, J. Yuan, K. Strebhardt, *Oncogene* **2005**, 24, 267-276.
- [230] M. Warmuth, R. Damoiseaux, Y. Liu, D. Fabbro, N. Gray, *Curr. Pharm. Des.* **2003**, 9, 2043-2059.
- [231] G. S. Martin, *Nat. Rev. Mol. Cell Biol.* **2001**, 2, 467-475.
- [232] M. Popkov, N. Jendreyko, D. B. McGavern, C. Rader, C. F. Barbas, 3rd, *Cancer Res.* **2005**, 65, 972-981.
- [233] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, 267, 727-748.
- [234] C. M. Dobson, *Nature* **2004**, 432, 824-828.
- [235] J. Clardy, C. Walsh, *Nature* **2004**, 432, 829-837.
- [236] D. Weininger, *J. Chem. Inf. Comput. Sci.* **1988**, 28, 31-36.
- [237] R. D. Brown, Y. C. Martin, *J. Chem. Inf. Comput. Sci.* **1996**, 36, 572-584.
- [238] H. Matter, T. Potter, *J. Chem. Inf. Comput. Sci.* **1999**, 39, 1211-1225.
- [239] R. P. Sheridan, S. K. Kearsley, *Drug Discov. Today* **2002**, 7, 903-911.
- [240] J. Fejzo, C. A. Lepre, J. W. Peng, G. W. Bemis, Ajay, M. A. Murcko, J. M. Moore, *Chem. Biol.* **1999**, 6, 755-769.
- [241] G. W. Bemis, M. A. Murcko, *J. Med. Chem.* **1996**, 39, 2887-2893.

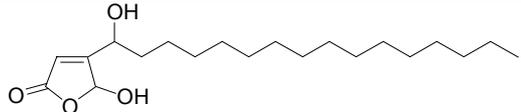
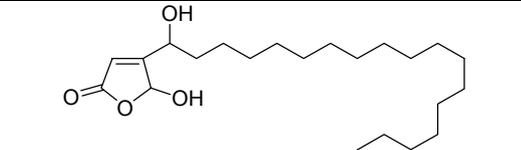
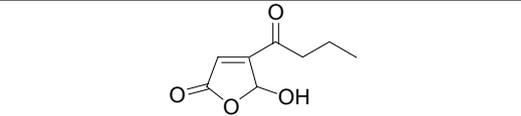
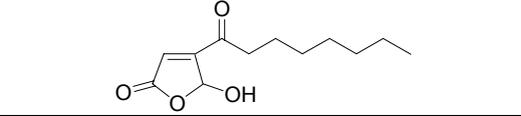
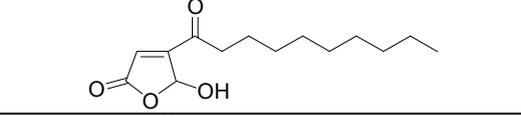
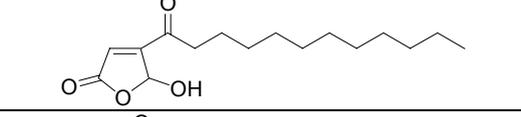
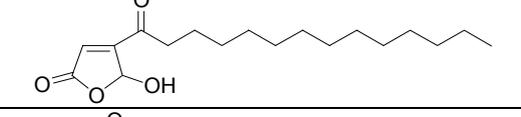
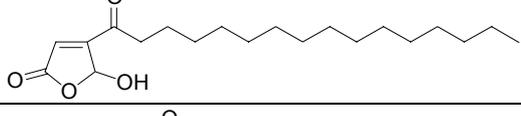
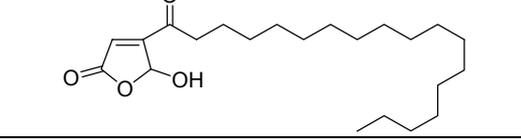
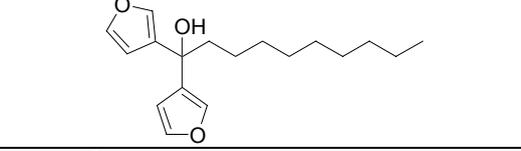
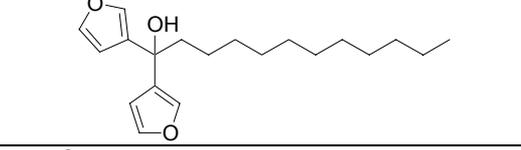
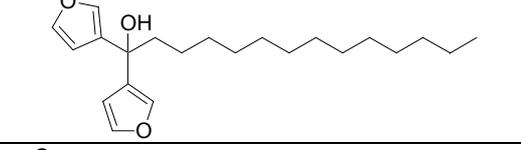
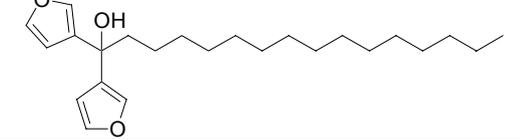
- [242] V. Kren, L. Martinková, *Curr. Med. Chem.* **2001**, *8*, 1303-1328.
- [243] R. M. Gleadow, I. E. Woodrow, *J. Chem. Ecol.* **2002**, *28*, 1301-1313.
- [244] U. Wittstock, N. Agerbirk, E. J. Stauber, C. E. Olsen, M. Hippler, T. Mitchell-Olds, J. Gershenzon, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4859-4864.
- [245] D. Butina, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 747-750.
- [246] R. A. Jarvis, E. A. Patrick, *IEEE Trans. Comput.* **1973**, *C-22*, 1025-1034.
- [247] S. Schmitt, D. Kuhn, G. Klebe, *J. Mol. Biol.* **2002**, *323*, 387-406.
- [248] E. J. Friderichs, I. Graudums, W. Winter, E. Frankus, W. W. A. Straßburger, European Patent EP0780369, **1997**.
- [249] D. Lednicer, P. F. VonVoigtlander, D. E. Emmert, *J. Med. Chem.* **1981**, *24*, 404-408.
- [250] M. Scheck, Dissertation, Universität Dortmund, **2005**.
- [251] U. Eder, G. Sauer, R. Weichert, *Angew. Chem. Int. Ed.* **1971**, *10*, 496.
- [252] Z. G. Hajos, D. R. Parrish, *J. Org. Chem.* **1974**, *39*, 1615-1621.
- [253] L. A. Thompson, J. A. Ellman, *Tetrahedron Lett.* **1994**, *35*, 9333-9336.
- [254] A. Odermatt, P. Arnold, A. Stauffer, B. M. Frey, F. J. Frey, *J. Biol. Chem.* **1999**, *274*, 28762-28770.
- [255] K. L. Carey, S. A. Richards, K. M. Lounsbury, I. G. Macara, *J. Cell Biol.* **1996**, *133*, 985-996.
- [256] A. Odermatt, P. Arnold, F. J. Frey, *J. Biol. Chem.* **2001**, *276*, 28484-28492.
- [257] A. G. Rebuffat, S. Tam, A. R. Nawrocki, M. E. Baker, B. M. Frey, F. J. Frey, A. Odermatt, *Mol. Cell Endocrinol.* **2004**, *214*, 27-37.
- [258] R. A. Sayle, E. J. Milner-White, *Trends Biochem. Sci.* **1995**, *20*, 374.
- [259] M. Hendlich, F. Rippmann, G. Barnickel, *J. Mol. Graph. Model.* **1997**, *15*, 359-363, 389.
- [260] A. C. Wallace, R. A. Laskowski, J. M. Thornton, *Protein Eng.* **1995**, *8*, 127-134.

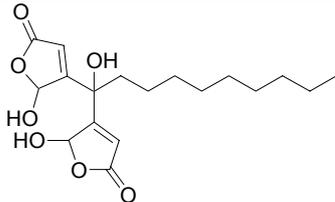
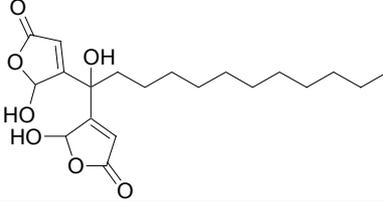
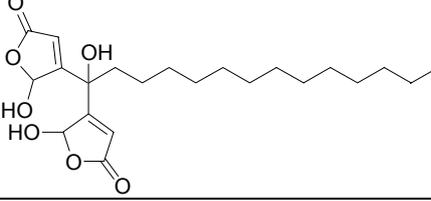
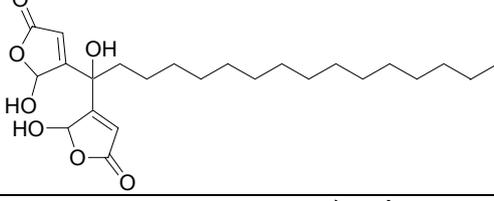
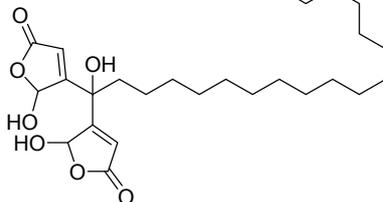
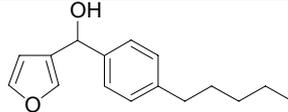
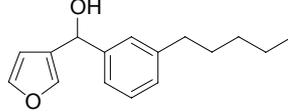
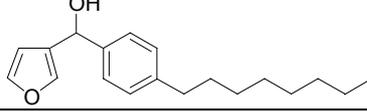
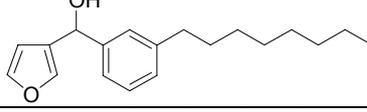
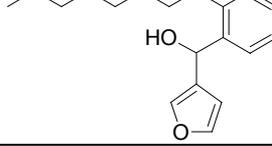
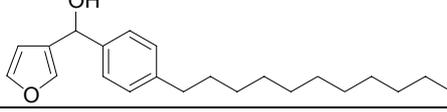
- [261] I. Nobeli, J. B. O. Mitchell, A. Alex, J. M. Thornton, *J. Comput. Chem.* **2001**, *22*, 673-688.
- [262] H. Gohlke, M. Hendlich, G. Klebe, *J. Mol. Biol.* **2000**, *295*, 337-356.
- [263] G. Escher, I. Galli, B. S. Vishwanath, B. M. Frey, F. J. Frey, *J. Exp. Med.* **1997**, *186*, 189-198.
- [264] J. M. Plummer, M. J. Greenberg, H. K. Lehman, J. A. Watts, *Biochem. Pharmacol.* **1983**, *32*, 151-158.
- [265] T. Carbonell, I. Masip, F. Sanchez-Baeza, M. Delgado, E. Araya, O. Llorens, F. Corcho, J. J. Perez, E. Perez-Paya, A. Messeguer, *Mol. Divers.* **2000**, *5*, 131-143.
- [266] H. Dvir, H. L. Jiang, D. M. Wong, M. Harel, M. Chetrit, X. C. He, G. Y. Jin, G. L. Yu, X. C. Tang, I. Silman, D. L. Bai, J. L. Sussman, *Biochemistry* **2002**, *41*, 10810-10818.
- [267] D. D. Perrin, W. L. F. Armstrong, *Purification of Laboratory Chemicals*, Pergamon Press, Oxford, **1988**.
- [268] O. T. Schmidt, H. Voigt, W. Puff, R. Koster, *Justus Liebigs Ann. Chem.* **1954**, *586*, 165-178.
- [269] C. A. James, D. Weininger, J. Delany, Daylight Chemical Information Systems, Inc., **2004**.

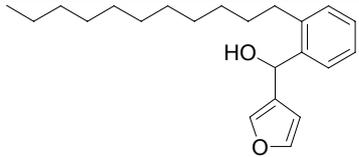
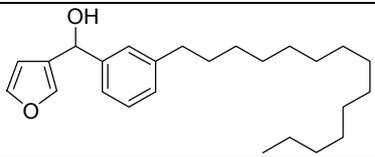
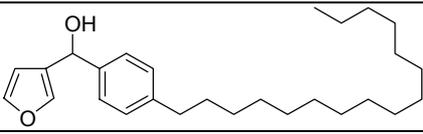
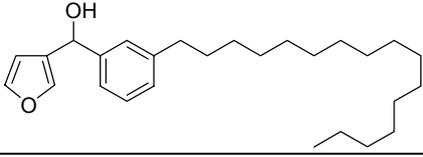
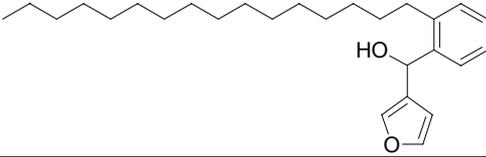
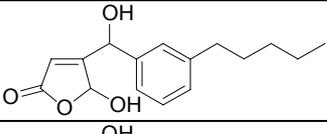
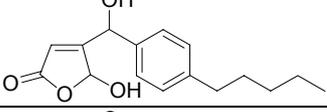
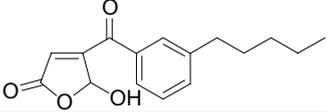
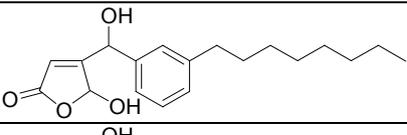
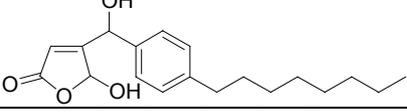
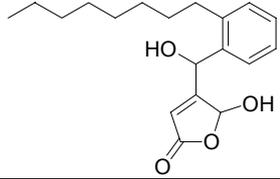
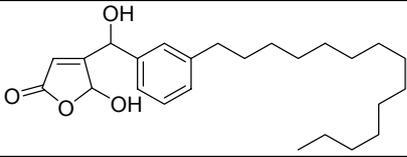
## 8 Anhang

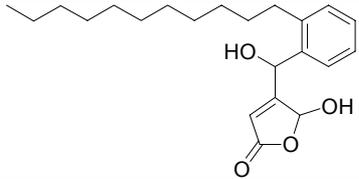
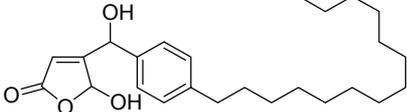
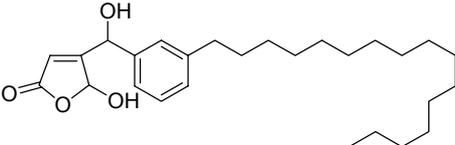
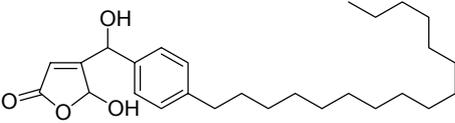
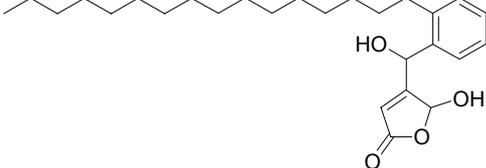
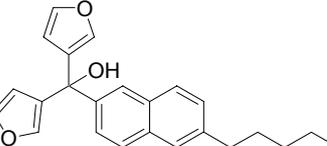
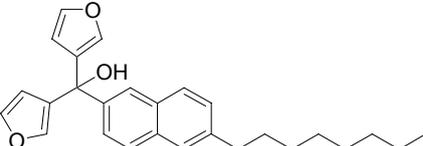
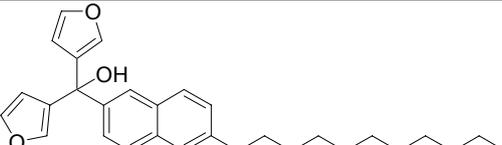
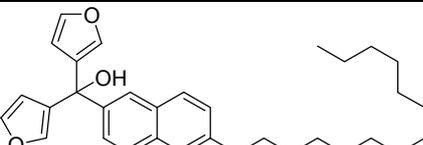
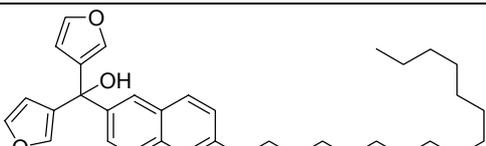
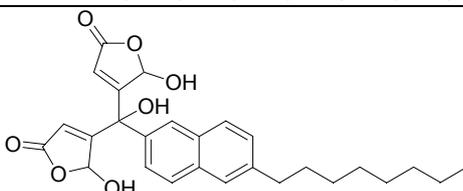
### 8.1 Anhang I: Von Dysidiolid abgeleitete Bibliothek

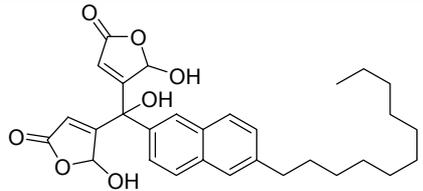
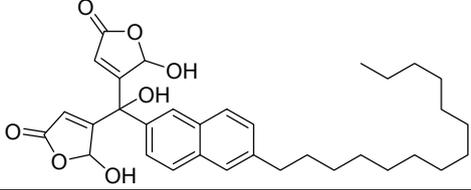
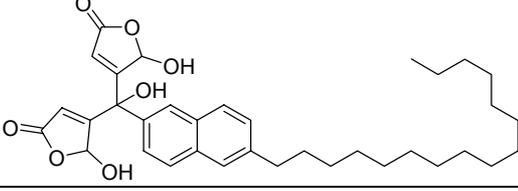
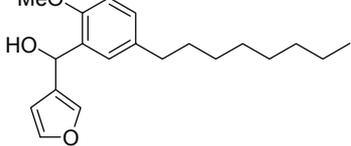
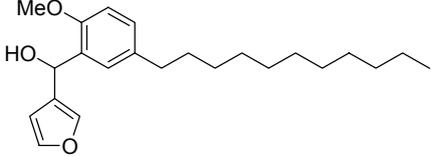
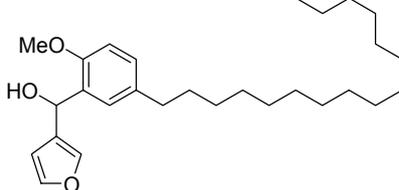
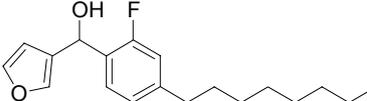
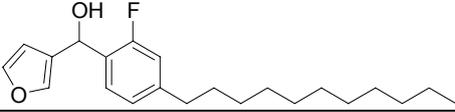
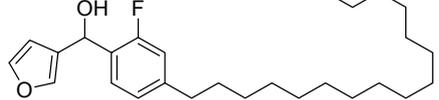
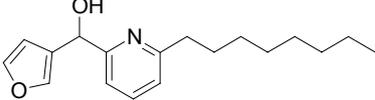
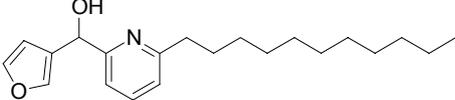
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
1		> 100	> 20	> 20	> 20
2		> 100	> 20	> 20	> 20
3		> 100	> 20	> 20	> 20
4		> 100	> 20	> 20	> 20
5		> 100	> 20	> 20	> 20
6		> 100	> 20	> 20	> 20
7		> 100	> 20	> 20	> 20
8		> 100	> 20	> 20	> 20
9		26	> 20	> 20	> 20
10		8.4	> 20	> 20	> 20
11		2.3	> 20	20.4±1.5	16.7±0.4
12		6.4±2.4	> 20	45±15	10.2±0.2

LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
13		7	> 20	> 20	> 20
14		2	> 20	> 20	> 20
15		> 100	> 20	> 20	> 20
16		> 100	> 20	> 20	> 20
17		18	> 20	> 20	> 20
18		5.9±1.7	> 20	84±12	12.3±1.0
19		4.5	> 20	> 20	> 20
20		> 100	> 20	> 20	> 20
21		> 100	> 20	> 20	> 20
22		> 100	> 20	> 20	> 20
23		> 100	> 20	> 20	> 20
24		> 100	> 20	> 20	> 20
25		> 100	> 20	> 20	> 20

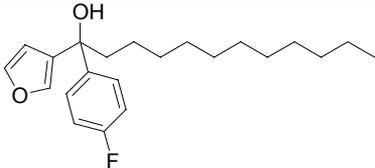
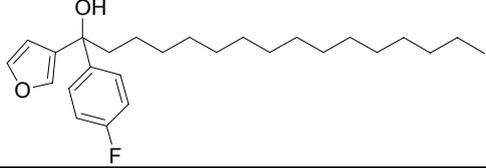
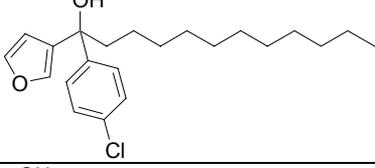
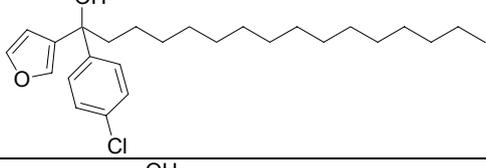
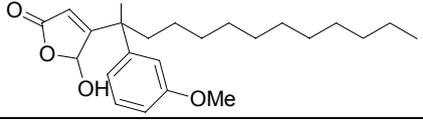
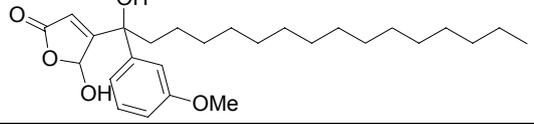
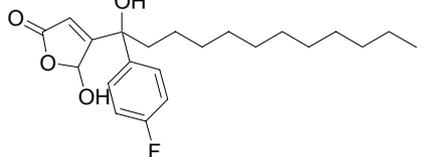
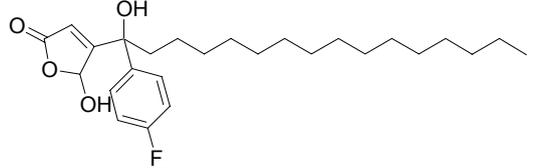
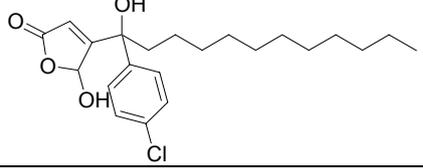
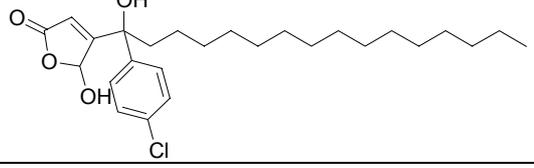
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
26		> 100	> 20	> 20	> 20
27		5.3	> 20	> 20	> 20
28		3	> 20	> 20	> 20
29		6	> 20	> 20	> 20
30		1.4±0.5	> 20	> 20	> 20
31		2	> 20	> 20	> 20
32		> 100	> 20	> 20	> 20
33		> 100	> 20	> 20	> 20
34		> 100	> 20	> 20	> 20
35		> 100	> 20	> 20	> 20
36		> 100	> 20	> 20	> 20

LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
37		> 100	> 20	> 20	> 20
38		> 100	> 20	> 20	> 20
39		> 100	> 20	> 20	> 20
40		> 100	> 20	> 20	> 20
41		> 100	> 20	> 20	> 20
42		> 100	> 20	> 20	> 20
43		8	> 20	> 20	> 20
44		13	> 20	> 20	> 20
45		23	> 20	> 20	> 20
46		8.5	> 20	> 20	> 20
47		6.4	> 20	> 20	> 20
48		5.4	> 20	> 20	> 20

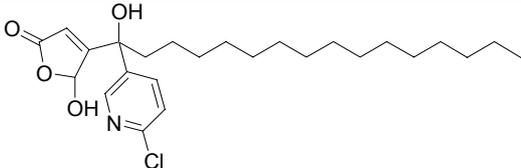
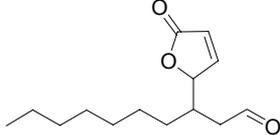
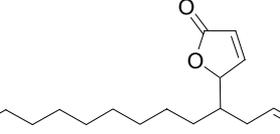
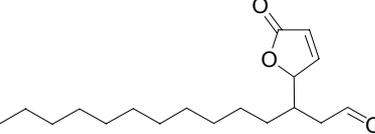
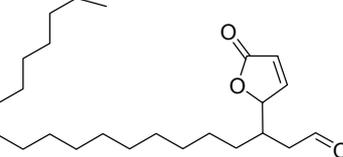
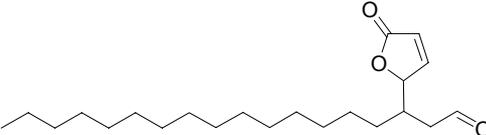
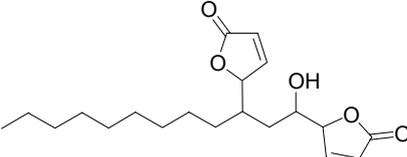
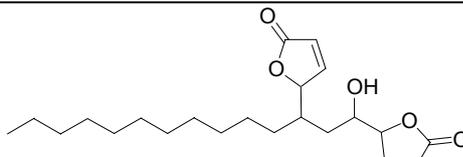
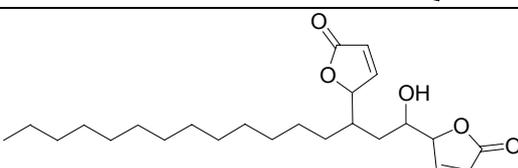
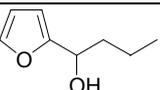
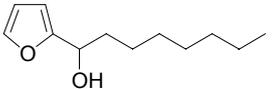
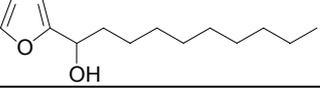
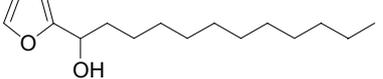
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
49		3.5	> 20	> 20	> 20
50		5.7	> 20	> 20	> 20
51		4.8	> 20	> 20	> 20
52		4	> 20	> 20	> 20
53		1.7	> 20	> 20	> 20
54		> 100	> 20	> 20	> 20
55		> 100	> 20	> 20	> 20
56		> 100	> 20	> 20	> 20
57		> 100	> 20	> 20	> 20
58		> 100	> 20	> 20	> 20
59		> 100	> 20	> 20	> 20

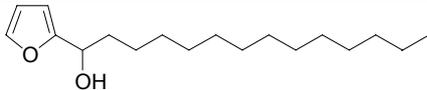
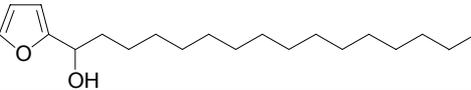
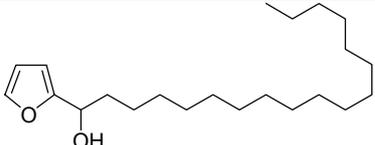
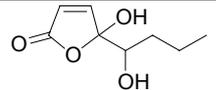
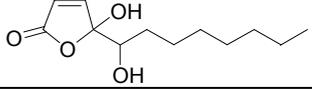
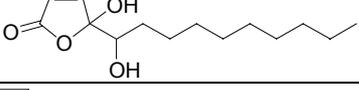
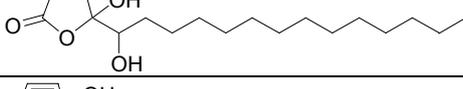
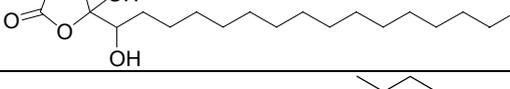
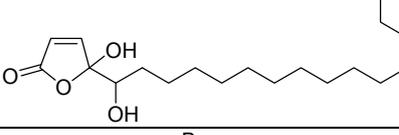
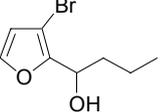
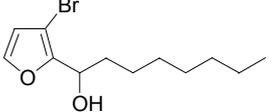
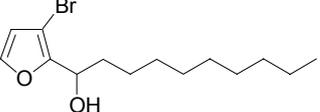
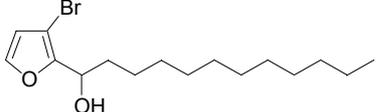
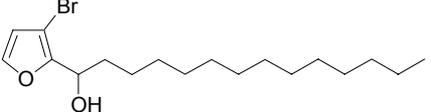
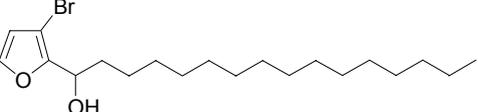
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
60		4.8	> 20	> 20	> 20
61		0.35±0.18	> 20	14±3	2.4±0.3
62		10	> 20	> 20	> 20
63		3.9	> 20	> 20	> 20
64		> 100	> 20	> 20	> 20
65		> 100	> 20	> 20	> 20
66		1.8±0.7	> 20	19±3	6.7±0.9
67		> 100	> 20	> 20	> 20
68		> 100	> 20	> 20	> 20
69		> 100	> 20	> 20	> 20
70		> 100	> 20	> 20	> 20

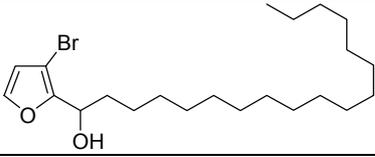
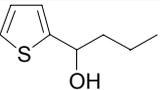
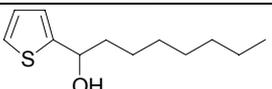
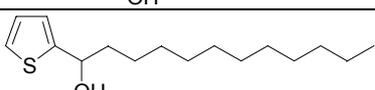
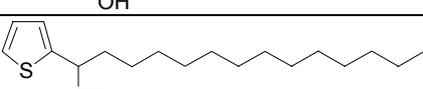
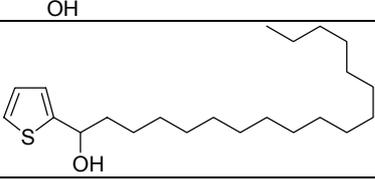
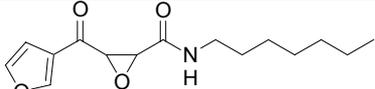
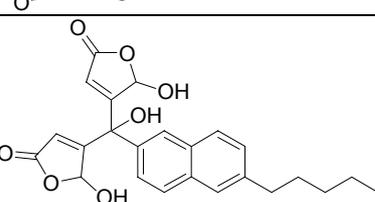
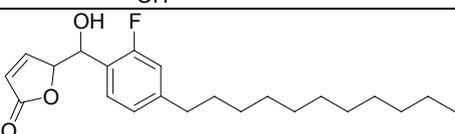
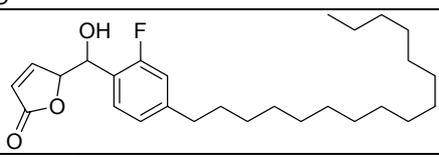
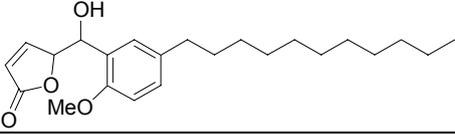
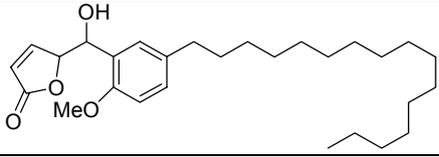
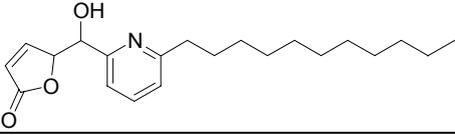
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
71		> 100	> 20	> 20	> 20
72		> 100	> 20	> 20	> 20
73		5.4	> 20	> 20	> 20
74		2.0±1.0	> 20	13.2±0.5	13.9±0.8
75		2.2	> 20	> 20	> 20
76		1.9±0.7	> 20	11.2±0.3	13.5±1.0
77		4	> 20	> 20	> 20
78		> 100	> 20	19±3	5.3±1.1
79		14	> 20	> 20	> 20
80		12	> 20	> 20	> 20
81		11	> 20	> 20	> 20
82		> 100	> 20	> 20	> 20

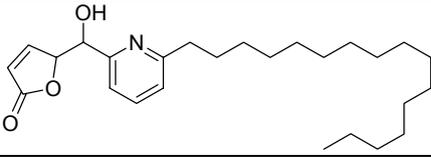
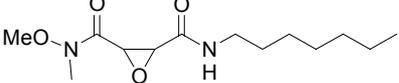
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
83		> 100	> 20	> 20	> 20
84		> 100	> 20	> 20	> 20
85		> 100	> 20	> 20	> 20
86		> 100	> 20	> 20	> 20
87		> 100	> 20	> 20	> 20
88		1.6±0.6	4.5±0.3	10±1	14±2
89		4.5±2.1	21±3	> 20	> 20
90		1.5±0.2	2.1±0.5	13±3	34±4
91		2.9±1.1	24±6	> 20	> 20
92		1.2±0.2	16±2	93±10	23±8

LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
93		2.3±0.5	1.3±0.1	7.8±1.8	2.8±0.4
94		> 100	> 20	> 20	> 20
95		> 100	> 20	> 20	> 20
96		> 100	> 20	> 20	> 20
97		> 100	> 20	> 20	> 20
98		> 100	> 20	> 20	> 20
99		> 100	> 20	> 20	> 20
100		30	> 20	> 20	> 20
101		19	> 20	> 20	> 20
102		35	> 20	> 20	> 20
103		12	> 20	> 20	> 20

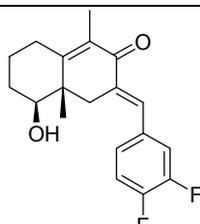
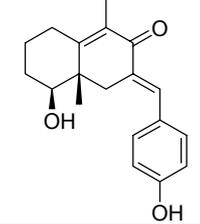
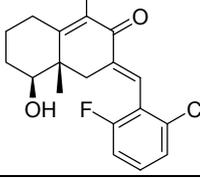
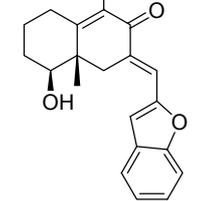
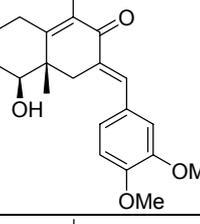
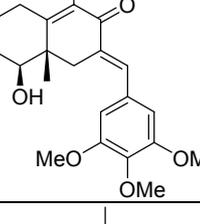
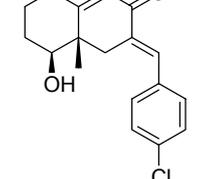
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
104		2.8±0.3	15±1	> 20	> 20
105		5.1±2.1	23±5	> 20	> 20
106		7.9±0.8	> 20	5.8±0.7	52±9
107		3.1±0.3	> 20	3.8±0.4	39±6
108		10.7	> 20	> 20	> 20
109		> 100	> 20	90	117
110		> 100	> 20	88	85
111		43	> 20	38	83
112		45	> 20	10±2	95±4
113		30	> 20	> 20	> 20
114		> 100	> 20	> 20	> 20
115		> 100	> 20	> 20	> 20
116		> 100	> 20	> 20	> 20

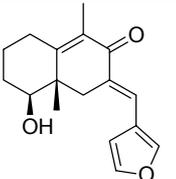
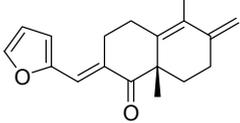
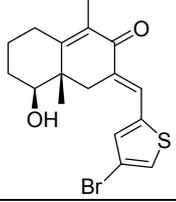
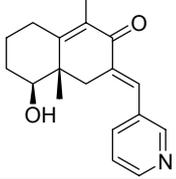
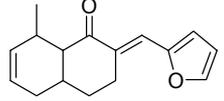
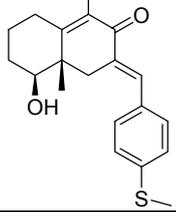
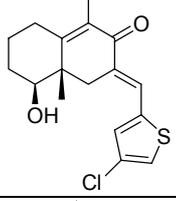
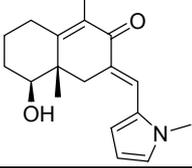
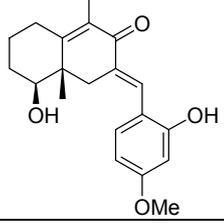
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
117		> 100	> 20	> 20	> 20
118		> 100	> 20	> 20	> 20
119		> 100	> 20	> 20	> 20
120		> 100	> 20	> 20	> 20
121		> 100	> 20	> 20	> 20
122		20	> 20	> 20	> 20
123		12	> 20	> 20	> 20
124		13	> 20	> 20	> 20
125		14	> 20	> 20	> 20
126		3.7	> 20	> 20	> 20
127		> 100	> 20	> 20	> 20
128		> 100	> 20	> 20	> 20
129		> 100	> 20	> 20	> 20
130		> 100	> 20	> 20	> 20
131		> 100	> 20	> 20	> 20

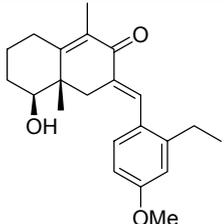
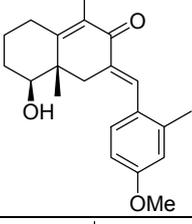
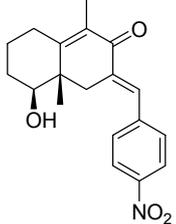
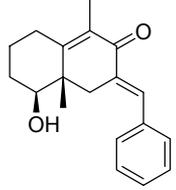
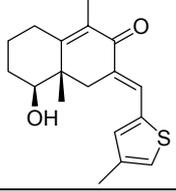
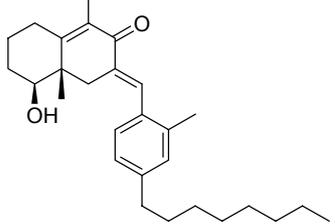
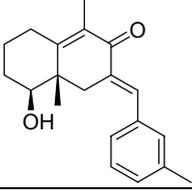
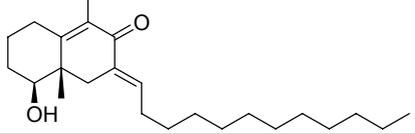
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
132		> 100	> 20	> 20	> 20
133		> 100	> 20	> 20	> 20
134		> 100	> 20	> 20	> 20
135		> 100	> 20	> 20	> 20
136		> 100	> 20	> 20	> 20
137		> 100	> 20	> 20	> 20
138		> 100	> 20	> 20	> 20
139		90	> 20	> 20	> 20
140		> 100	> 20	> 20	> 20
141		> 20	> 20	> 20	> 20
142		40	> 20	> 20	> 20
143		100	> 20	> 20	> 20
144		> 100	> 20	> 20	> 20

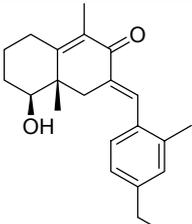
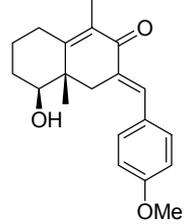
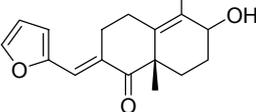
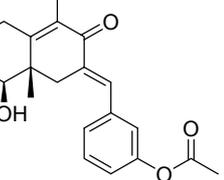
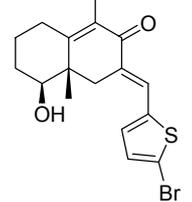
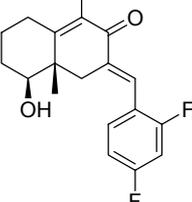
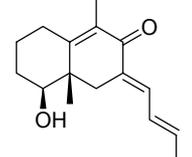
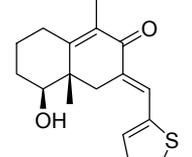
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
145		> 100	> 20	> 20	> 20
146		50	> 20	> 20	> 20

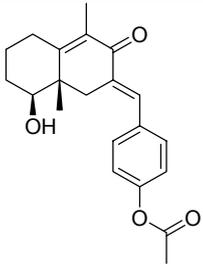
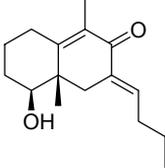
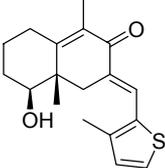
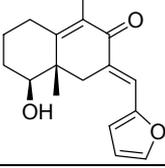
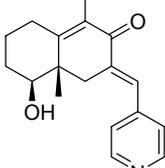
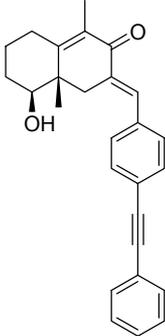
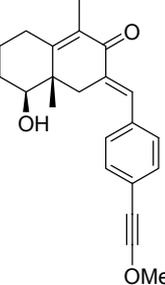
## 8.2 Anhang II: Von Dysidiolid und Glycyrrhetinsäure abgeleitete Bibliothek

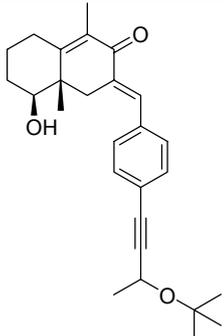
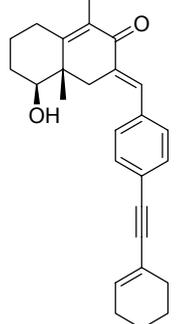
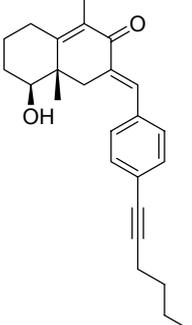
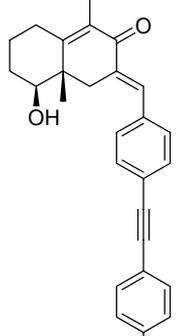
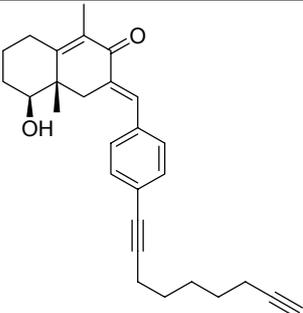
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
1		> 100	> 20	1.5±0.5	> 30
2		> 100	> 20	> 10	> 30
3		> 100	> 20	> 10	> 30
4		17±12	> 20	3.1±0.5	> 30
5		> 100	> 20	> 10	> 30
6		> 100	> 20	> 10	> 30
7		> 100	> 20	> 10	> 30

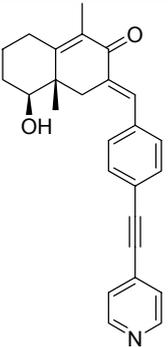
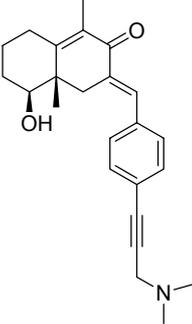
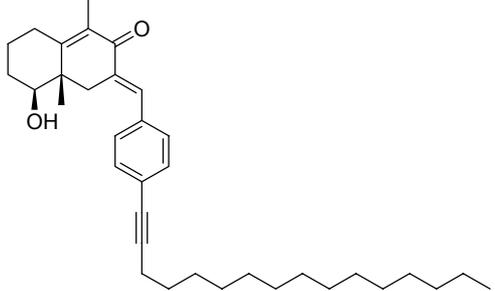
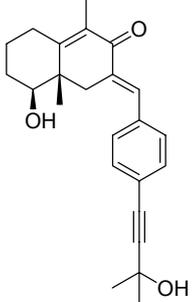
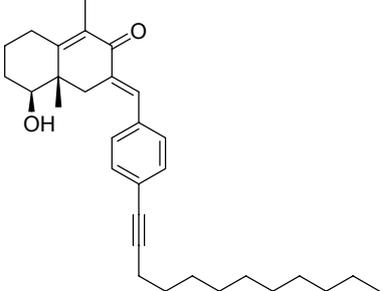
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
8		44±34	> 20	> 10	> 30
9		18±11	> 20	> 10	> 30
10		28±15	> 20	> 10	> 30
11		> 100	> 20	2.0±0.3	> 30
12		21±18	> 20	> 10	> 30
13		> 100	> 20	2.1±0.1	> 30
14		31±23	> 20	> 10	> 30
15		> 100	> 20	> 10	> 30
16		> 100	> 20	3.6±0.3	> 30

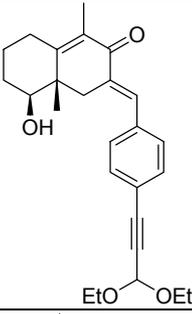
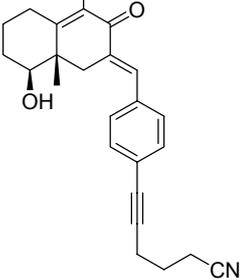
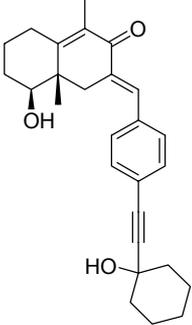
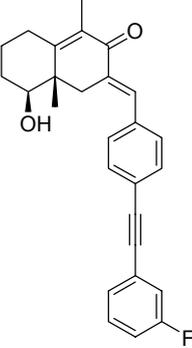
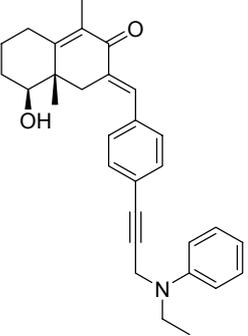
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
17		> 100	> 20	6.5±0.7	> 30
18		> 100	> 20	> 10	> 30
19		> 100	> 20	> 10	> 30
20		> 100	> 20	6.1±0.4	> 30
21		> 100	> 20	> 10	> 30
22		> 100	> 20	> 10	> 30
23		> 100	> 20	> 10	> 30
24		97±32	> 20	6.8±0.7	> 30

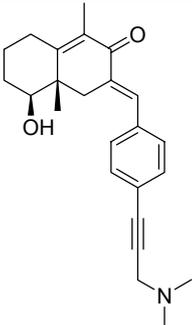
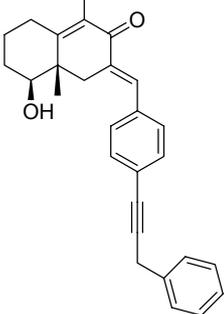
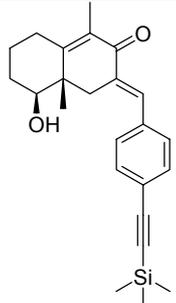
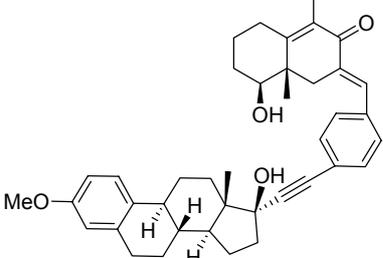
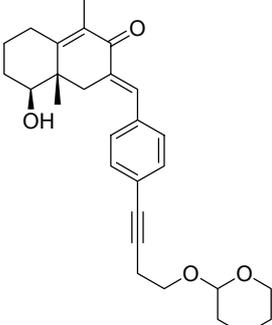
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
25		> 100	> 20	4.3±0.2	> 30
26		19±7	> 20	4.7±0.7	> 30
27		> 100	> 20	> 10	> 30
28		> 100	> 20	3.1±0.1	29 ± 5
29		> 100	> 20	> 10	> 30
30		> 100	> 20	> 10	> 30
31		25±17	> 20	2.8±0.1	5.2±1.0
32		> 100	> 20	> 10	> 30

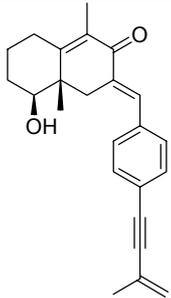
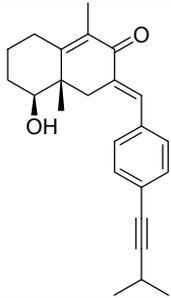
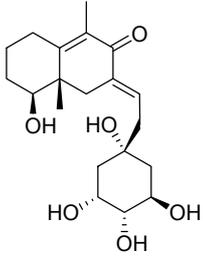
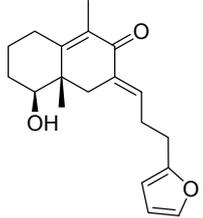
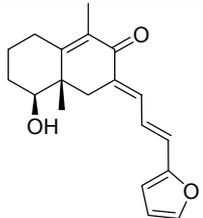
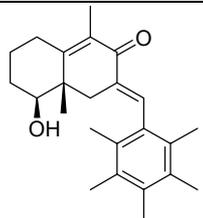
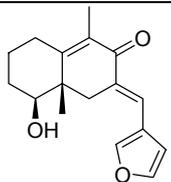
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
33		> 100	> 20	2.9 $\pm$ 0.1	29 $\pm$ 7
34		15 $\pm$ 6	5.3 $\pm$ 1.1	0.31 $\pm$ 0.03	6.6 $\pm$ 0.9
35		> 100	> 20	> 10	> 30
36		> 100	> 20	> 10	> 30
37		> 100	> 20	> 10	> 30
38		> 100	> 20	> 10	> 30
39		> 100	> 20	> 10	> 30

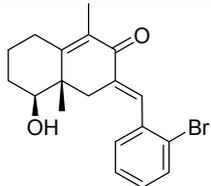
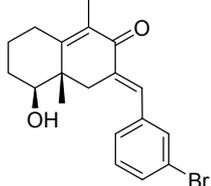
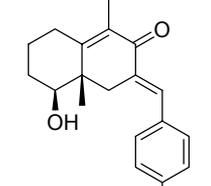
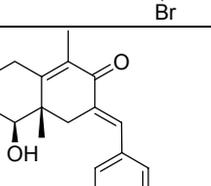
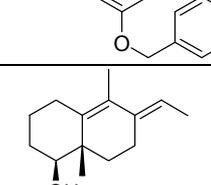
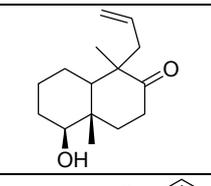
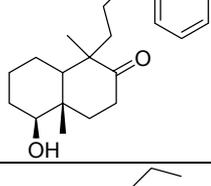
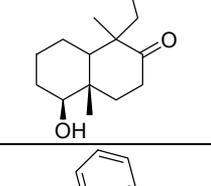
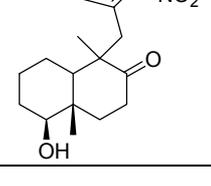
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
40		5.9±2.4	> 20	> 10	> 30
41		> 100	> 20	> 10	> 30
42		> 100	> 20	> 10	> 30
43		> 100	> 20	> 10	> 30
44		2.6±0.8	> 20	> 10	> 30

LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
45		> 100	> 20	> 10	> 30
46		> 100	> 20	> 10	> 30
47		> 100	> 20	> 10	> 30
48		7.0 $\pm$ 2.2	> 20	> 10	> 30
49		17 $\pm$ 7	> 20	> 10	> 30

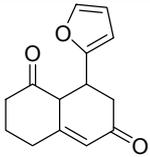
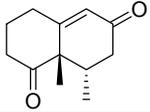
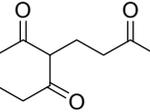
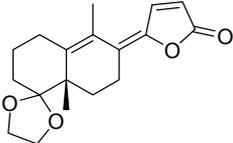
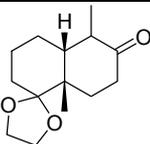
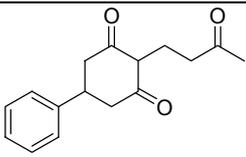
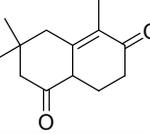
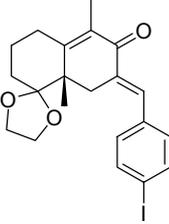
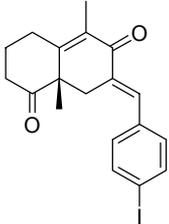
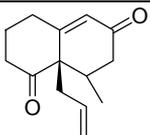
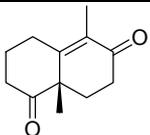
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
50		5.6 $\pm$ 5.1	> 20	> 10	> 30
51		15 $\pm$ 8	> 20	> 10	> 30
52		13 $\pm$ 6	> 20	> 10	> 30
53		1.2 $\pm$ 0.5	15.9 $\pm$ 6.7	> 10	> 30
54		7.1 $\pm$ 3.1	> 20	> 10	> 30

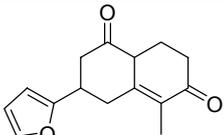
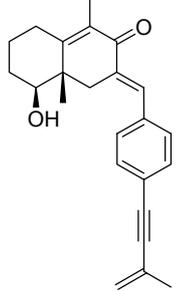
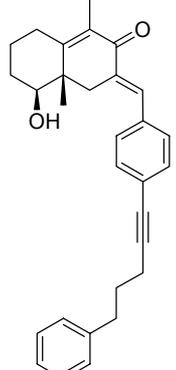
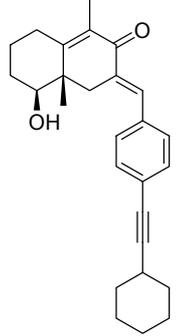
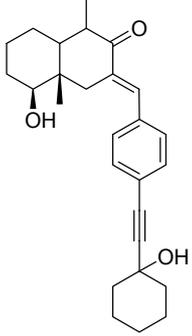
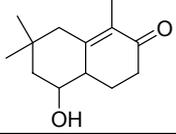
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
55		8.3±2.4	> 20	> 10	> 30
56		> 100	> 20	> 10	> 30
57		20±11	> 20	> 10	> 30
58		9.2±4.8	> 20	> 10	> 30
59		8.6±6.2	> 20	> 10	> 30

LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
60		> 100	> 20	> 10	> 30
61		> 100	> 20	> 10	> 30
62		> 100	> 20	> 10	> 30
63		64±32	> 20	> 10	> 30
64		43±22	> 20	> 10	> 30
65		> 100	> 20	> 10	> 30
66		70±25	> 20	> 10	> 30

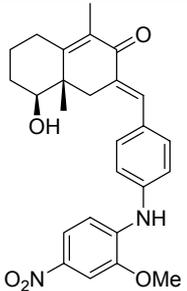
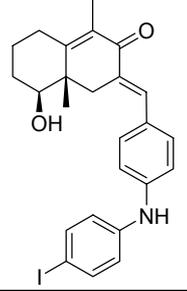
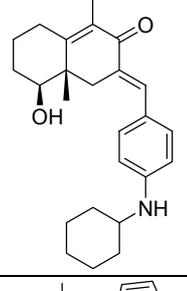
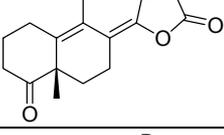
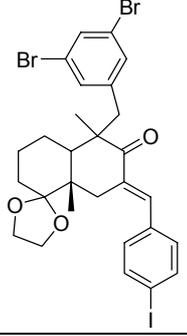
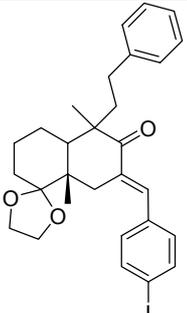
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
67		> 100	> 20	> 10	> 30
68		43 $\pm$ 14	> 20	> 10	> 30
69		60 $\pm$ 20	> 20	13.3 $\pm$ 0.1	> 30
70		> 100	> 20	> 10	> 30
71		71 $\pm$ 24	> 20	> 10	> 30
72		64 $\pm$ 17	> 20	> 10	> 30
73		28 $\pm$ 20	> 20	> 10	> 30
74		31 $\pm$ 13	> 20	> 10	> 30
75		> 100	> 20	4.3 $\pm$ 0.5	> 30

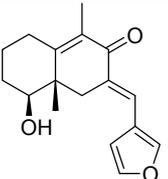
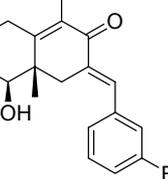
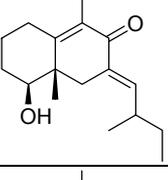
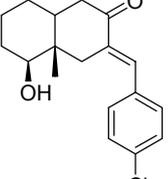
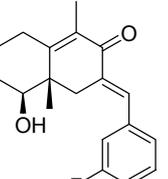
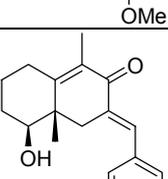
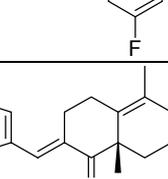
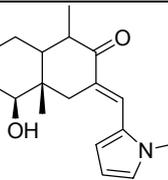
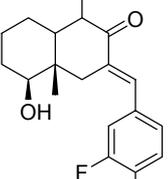
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
76		> 100	> 20	> 10	> 30
77		42±8	> 20	> 10	> 30
78		22±7	> 20	> 10	> 30
79		> 100	> 20	> 10	> 30
80		> 100	> 20	> 10	> 30
81		> 100	> 20	> 10	> 30
82		> 100	> 20	> 10	> 30
83		> 100	> 20	> 10	> 30

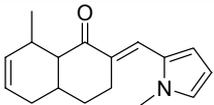
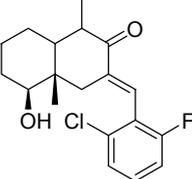
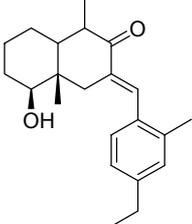
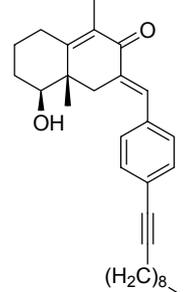
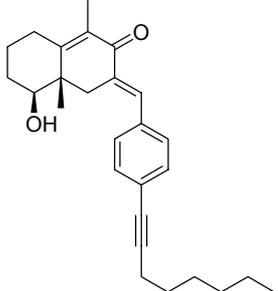
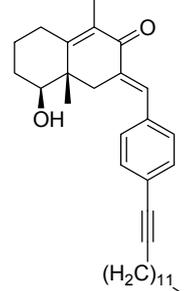
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
84		> 100	> 20	> 10	> 30
85		14.7±0.3	> 20	> 10	> 30
86		16±2	> 20	> 10	> 30
87		22±3	> 20	> 10	> 30
88		> 100	> 20	> 10	> 30
89		> 100	> 20	> 10	> 30
90		> 100	> 20	> 10	> 30
91		> 100	> 20	> 10	> 30
92		> 100	> 20	> 10	> 30
93		> 100	> 20	> 10	> 30
94		> 100	> 20	> 10	> 30

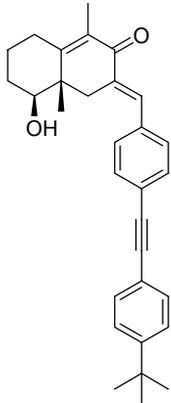
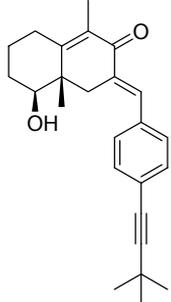
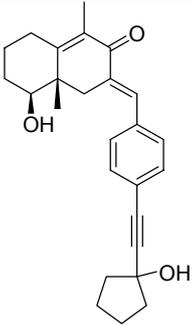
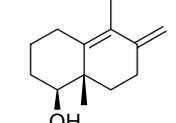
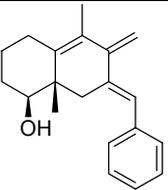
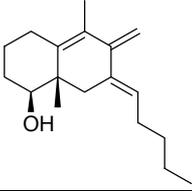
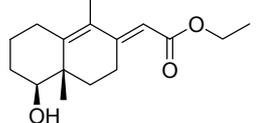
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
95		> 100	> 20	2.8±0.3	> 30
96		18±3	> 20	> 10	> 30
97		15±1	> 20	> 10	> 30
98		> 100	> 20	2.4±0.2	> 30
99		> 100	> 20	> 10	> 30
100		> 100	> 20	> 10	> 30

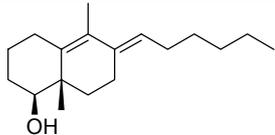
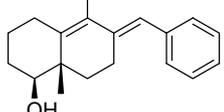
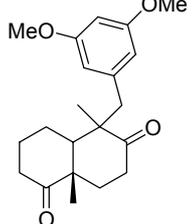
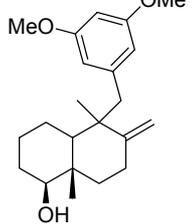
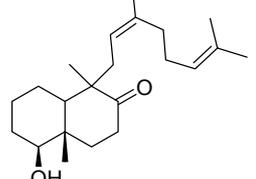
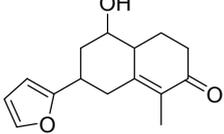
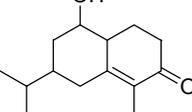
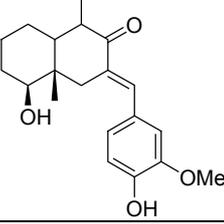
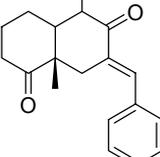
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
101		> 100	> 20	> 10	> 30
102		> 100	> 20	> 10	> 30
103		> 100	> 20	> 10	> 30
104		> 100	> 20	1.6±0.3	> 30
105		> 100	> 20	6.6±0.7	> 30
106		> 100	> 20	> 10	> 30
107		> 100	> 20	> 10	> 30
108		> 100	> 20	> 10	> 30

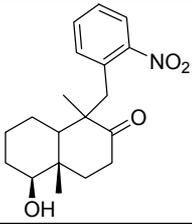
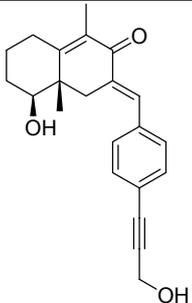
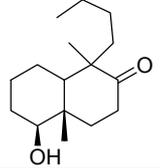
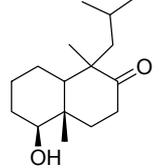
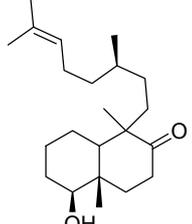
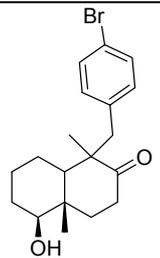
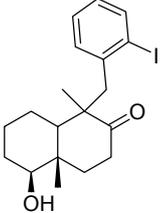
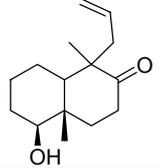
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [μM]	AChE IC <sub>50</sub> [μM]	11βHSD1 IC <sub>50</sub> [μM]	11βHSD2 IC <sub>50</sub> [μM]
109		> 100	> 20	> 10	> 30
110		> 100	> 20	> 10	> 30
111		> 100	> 20	1.6±0.2	> 30
112		18±3	> 20	> 10	> 30
113		> 100	> 20	2.2±0.3	> 30
114		> 100	> 20	> 10	> 30

LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
115		> 100	> 20	> 10	> 30
116		> 100	> 20	> 10	> 30
117		> 100	> 20	> 10	> 30
118		39	> 20	2.5 $\pm$ 0.4	> 30
119		> 100	> 20	5.8 $\pm$ 0.4	> 30
120		> 100	> 20	0.74 $\pm$ 0.11	> 30
121		13 $\pm$ 2	> 20	> 10	> 30
122		12 $\pm$ 3	> 20	> 10	> 30
123		> 100	> 20	> 10	> 30

LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
124		> 100	> 20	> 10	> 30
125		> 100	> 20	> 10	> 30
126		> 100	> 20	> 10	> 30
127		> 100	> 20	> 10	> 30
128		6.9 $\pm$ 1.1	> 20	> 10	> 30
129		> 100	> 20	> 10	> 30

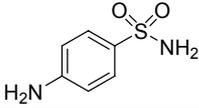
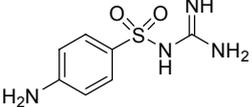
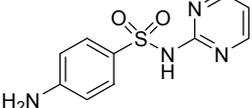
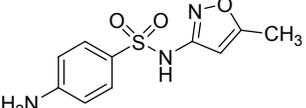
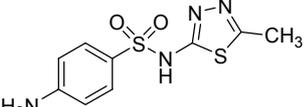
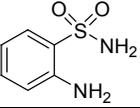
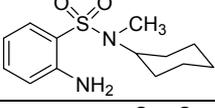
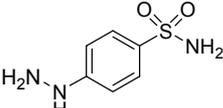
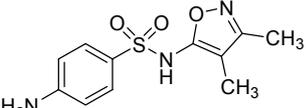
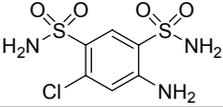
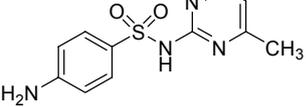
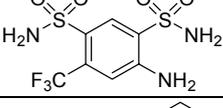
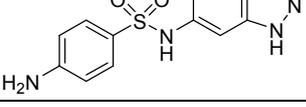
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
130		> 100	> 20	> 10	> 30
131		> 100	> 20	> 10	> 30
132		> 100	> 20	> 10	> 30
133		> 100	> 20	> 10	> 30
134		> 100	> 20	0.35±0.04	> 30
135		> 100	> 20	2.0±0.3	> 30
136		> 100	> 20	> 10	> 30

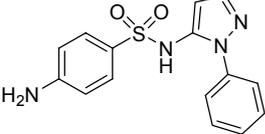
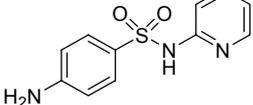
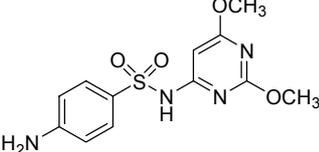
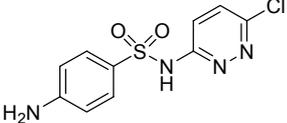
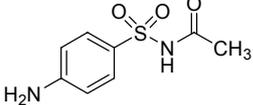
LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
137		> 100	> 20	> 10	> 30
138		4.9±0.9	> 20	> 10	> 30
139		5.7±1.2	> 20	> 10	> 30
140		> 100	> 20	> 10	> 30
141		> 100	> 20	0.63±0.12	> 30
142		> 100	> 20	> 10	> 30
143		> 100	> 20	> 10	> 30
144		> 100	> 20	> 10	> 30
145		> 100	> 20	> 10	> 30

LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
146		> 100	> 20	> 10	> 30
147		> 100	> 20	> 10	> 30
148		> 100	> 20	> 10	> 30
149		> 100	> 20	> 10	> 30
150		> 100	> 20	> 10	> 30
151		> 100	> 20	> 10	> 30
152		> 100	> 20	9.1±0.9	> 30
153		> 100	> 20	> 10	> 30

LfdNr	Struktur	Cdc25A IC <sub>50</sub> [ $\mu$ M]	AChE IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD1 IC <sub>50</sub> [ $\mu$ M]	11 $\beta$ HSD2 IC <sub>50</sub> [ $\mu$ M]
154		> 100	> 20	4.5±0.2	> 30
155		> 100	> 20	> 10	> 30
156		> 100	> 20	> 10	> 30
157		> 100	> 20	1.0±0.1	2.0±0.4
158		> 100	> 20	> 10	> 30
159		> 100	> 20	> 10	> 30
160		> 100	3.7±0.4	> 10	> 30
161		> 100	> 20	> 10	> 30
162		> 100	> 20	> 10	> 30

## 8.3 Anhang III: Kollektion von Benzolsulfonamid-Derivaten

LfdNr	Struktur	<i>t</i> ODC IC <sub>50</sub> [ $\mu$ M]	<i>h</i> ODC IC <sub>50</sub> [ $\mu$ M]
1		> 1000	> 1000
2		> 1000	> 1000
3		> 1000	> 1000
4		> 1000	> 1000
5		> 1000	> 1000
6		> 1000	> 1000
7		> 1000	> 1000
8		81±5	66±6
9		> 1000	> 1000
10		> 1000	> 1000
11		> 1000	> 1000
12		> 1000	> 1000
13		> 1000	> 1000

LfdNr	Struktur	tODC IC <sub>50</sub> [ $\mu$ M]	hODC IC <sub>50</sub> [ $\mu$ M]
14		> 1000	> 1000
15		> 1000	> 1000
16		> 1000	> 1000
17		> 1000	> 1000
18		> 1000	> 1000



## Dank

Besonders herzlich bedanke ich mich bei Herrn Professor Herbert Waldmann, der mich bei all meinen Vorhaben nicht nur stets unterstützte, sondern mir auch die Freiräume für eine kreative Umsetzung einräumte.

Allen Mitstreitern der Arbeitsgruppe danke ich für das angenehme Arbeitsklima und die stete Hilfsbereitschaft, allen voran meinem Labornachbarn, Herrn Dipl.-Chem. Lars Arve.

Zu großem Dank bin ich meinen Kooperationspartnern, Herrn Dr. Peter Ertl, Herrn PD Dr. Alex Odermatt und Herrn Dr. Ansgar Schuffenhauer verpflichtet, die durch ihr Engagement mustergültige Kooperationen ermöglichten.

Frau Heike Rimpel, Frau Samra Kavazovic und insbesondere Frau Walburga Hecker, die mit großem Engagement und kritischer Sorgfalt bei den biochemischen Assays mitwirkte, danke ich für die technische Unterstützung. Frau Dr. Petra Janning sei für ihre kompetenten Ratschläge in Fragen der Analytik und Herrn PD Dr. Heino Prinz für die hilfreichen und unterhaltsamen Diskussionen zu bioanalytischen Fragestellungen gedankt. Frau Dr. Ingrid Vetter bin ich für ihre Diskussionsbereitschaft bezüglich (bio-)informatischer Fragestellungen, Herrn Dr. Christoph Schwittek für Tips rund um Unix und Frau Gesine Schulte und ihrem Team für die Unterstützung bei der Erstellung von Bildmaterial dankbar.

Meiner Mutter und meinem „Nachfolger im Amt“, Herrn Dipl.-Chem. Stefan Wetzels, danke ich für die zügige und kritische Durchsicht des Manuskripts.

Der Studienstiftung des deutschen Volkes sei nicht nur für die großzügige finanzielle Unterstützung, sondern vor allem für die ideelle Förderung gedankt. Diese Förderung ermöglichte mir wertvolle Einblicke auch jenseits von Wissenschaft und Forschung. Meinem Vertrauensdozenten, Herrn Professor Dietrich Wegener, danke ich für die exzellente Betreuung.

Schließlich gilt mein Dank meiner Familie und besonders meiner Freundin Claudia, die mich bei dieser Arbeit tatkräftig unterstützte und mir immer liebevoll zur Seite stand.

# Curriculum Vitae

## Marcus Alexander Koch

Geburtsdatum, -ort: 17. August 1975, Albstadt-Ebingen

Staatsangehörigkeiten: deutsch, französisch

Familienstand: ledig

## Schulische Ausbildung

09/1990–07/1991 Europäische Schule, Brüssel, Belgien

09/1986–06/1995 Hohenzollern-Gymnasium, Sigmaringen, Abitur

## Wehrdienst

07/1995–09/1996 Sanitätsdienst der Bundeswehr, Kempten, München

## Universitäre Ausbildung

10/1996–03/2001 Studium der Pharmazie an der Albert-Ludwigs-Universität, Freiburg

09/1998 Erster Abschnitt der Pharmazeutischen Prüfung

03/2001 Zweiter Abschnitt der Pharmazeutischen Prüfung

06/2002 Dritter Abschnitt der Pharmazeutischen Prüfung

06/2002 Approbation als Apotheker

12/2001–04/2005 Anfertigung der Doktorarbeit unter der Betreuung von Prof. Dr. Herbert Waldmann an der Universität Dortmund und am Max-Planck-Institut für Molekulare Physiologie, Dortmund, zum Thema: *Protein- und Naturstoffstruktur als Leitprinzipien für die Entwicklung von Verbindungsbibliotheken*

Dortmund, im Mai 2005