

3D-QSAR-Untersuchungen
an antimalaria-aktiven
Naphthylisochinolin-Alkaloiden

Dissertation zur Erlangung des
naturwissenschaftlichen Doktorgrades der
Bayerischen Julius-Maximilians-Universität Würzburg

vorgelegt von

Christian Rummey

aus

Würzburg

Würzburg 2002

Eingereicht am:

bei der Fakultät für Chemie und Pharmazie

1. Gutachter:

2. Gutachter:

der Dissertation

1. Prüfer:

2. Prüfer:

der mündlichen Prüfung

Tag der mündlichen Prüfung:

Doktorurkunde ausgehändigt am:

Die vorliegende Arbeit wurde in der Zeit von
Mai 1998 bis Oktober 2002
am Institut für Organische Chemie
der Bayerischen Julius-Maximilians-Universität
Würzburg angefertigt.

Herrn Prof. Dr. G. Bringmann danke ich sehr herzlich für die Unterstützung bei der
Durchführung dieser Arbeit, sein großes Engagement beim Ausbau der
theoretischen Abteilung unserer Gruppe und vor allem für die mir überlassenen
wissenschaftlichen Freiräume.

Inhaltsverzeichnis

1	Einleitung	1
2	Theoretische Grundlagen	7
2.1	Der Datensatz	9
2.2	Aufstellen von 3D-Strukturen	9
2.3	Das Alignment	9
2.3.1	FLEXS	10
2.3.2	GASP	13
2.4	Die Berechnung der CoMSIA-Felder	15
2.5	Die PLS-Analyse	19
2.5.1	Vorgehensweise	19
2.5.2	Die Kreuz- oder Gegenvalidierung	21
2.5.3	Berechnung des abschließenden Modells	22
2.5.4	Algorithmen zur Einsparung von Rechenzeit	23
3	Kenntnisstand	24
3.1	Malaria: Erreger und Bekämpfung	24
3.1.1	Plasmodien-Arten und deren Lebenszyklus	24
3.1.2	Der Wirkmechanismus von Chloroquin	26
3.1.3	Die Chloroquin-Problematik	27
3.1.4	Neue Entwicklungen	28
3.2	QSAR-Untersuchungen an Naphthylisochinolin-Alkaloiden	29
3.2.1	Biologische Tests auf Antimalaria-Aktivität	29
3.2.2	Alignmentstrategien	30
3.2.3	CoMFA-Modelle	30

3.2.4	Resultate der Pharmacophoranalysen und neue Strukturvor- schläge	31
3.3	Der Datensatz	32
4	Ergebnisse	43
4.1	Allgemeine Vorbemerkungen	43
4.1.1	Flexible Ringsysteme	43
4.1.2	GASP-Alignments als Grundlage für CoMSIA-Analysen	44
4.1.3	Behandlung unterschiedlicher Kupplungspositionen	45
4.2	Eingabe der 3D-Strukturen	46
4.3	Das Alignment	46
4.3.1	FLEXS	46
4.3.2	GASP	57
4.3.3	Manuelle Modifikationen der Alignments	61
4.4	Statistische Auswertung	63
4.4.1	Outlier innerhalb der Gruppen I bis IX	63
4.4.2	Monomere Verbindungen (Gruppe X)	68
4.4.3	Resultate der PLS-Analysen	69
4.4.4	Schlußfolgerungen	71
4.4.5	Optimierung des CoMSIA-Modells	77
4.4.6	Eigenschaften des endgültigen CoMSIA-Modells	82
4.4.7	Evaluierung des Modells anhand eines Testdatensatzes	86
4.5	Interpretation der CoMSIA-Kontur-Maps	89
4.5.1	Die Wasserstoffbrücken-Donor- und -Akzeptor-Felder	89
4.5.2	Das Hydrophobie-Feld	93
4.5.3	Elektrostatisches und sterisches Feld	94
4.5.4	Zusammenfassung der Feldeigenschaften: Ein Pharmacophor- Modell	95
4.6	Die Aktivität neuer Verbindungen	98
4.6.1	Mögliche strukturelle Vereinfachungen	98
4.6.2	Weitere im Isochinolinteil vereinfachte Strukturen	102
4.6.3	Neu isolierte Verbindungen	104
4.7	Vergleich mit Chloroquin und ähnlichen Wirkstoffen	106
5	Zusammenfassung	110

6 Summary	117
Literaturverzeichnis	124
A Verwendete Abkürzungen	134
B Technische Details	136
B.1 Soft- und Hardware	136
B.2 FLEXS-Script	137

Kapitel 1

Einleitung

Die Malaria ist heute die gefährlichste Infektionskrankheit. Diese Tatsache wird in vielen Industrienationen kaum wahrgenommen – häufig erst, wenn prophylaktische Maßnahmen aufgrund einer Reise in ein potentiell gefährliches Gebiet notwendig werden. Daß dies in vielen Fällen zu spät oder nicht ausreichend geschieht, wird belegt durch die Tatsache, daß sogar in Deutschland jährlich etwa 10-20 Menschen an einer Malariainfektion sterben.^[1] Insgesamt lebt heute annähernd 40% der Weltbevölkerung in Gebieten mit dem Risiko einer Malaria-Infektion (Abb. 1).

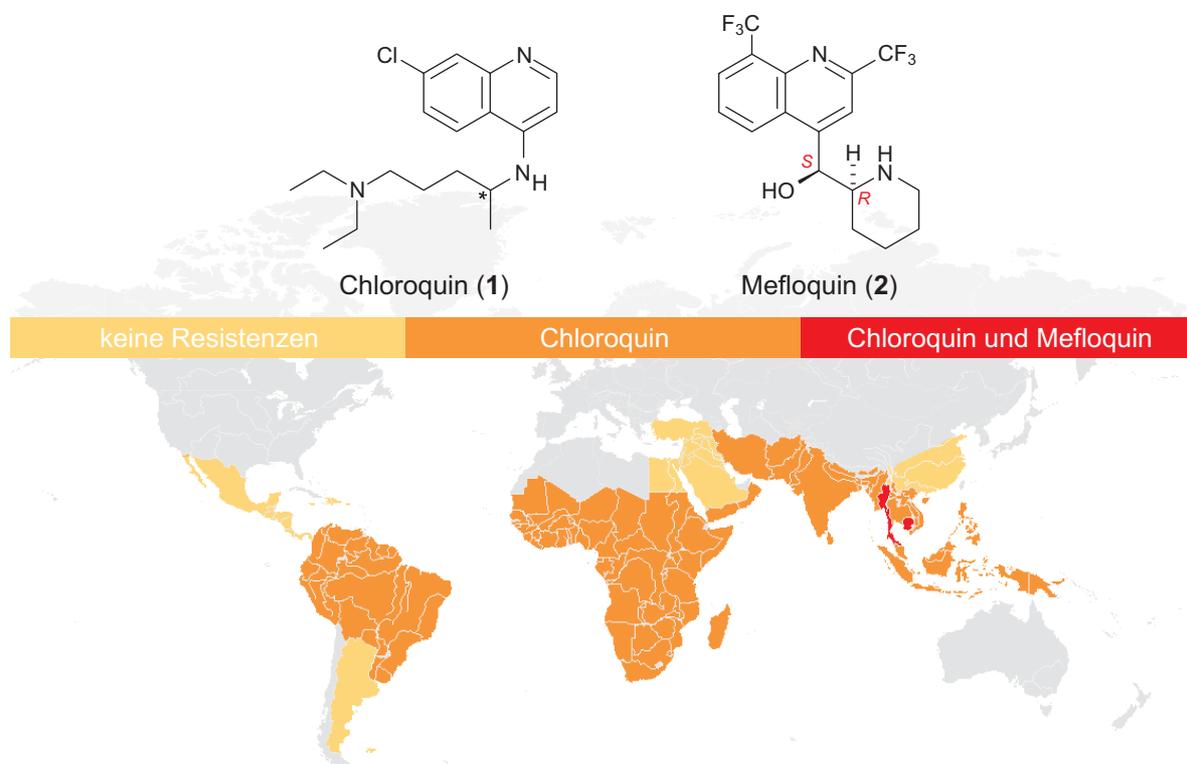


Abb. 1. Weltweite Malaria-Verbreitungsgebiete, differenziert nach Resistenzen gegen die gängigen Wirkstoffe Chloroquin (1) und Mefloquin (2)^[2]

Bei weltweit etwa 500 Millionen Neuinfektionen sterben pro Jahr durchschnittlich zwei Millionen Menschen an dieser Tropenkrankheit.^[1] Die meisten Todesfälle treten in afrikanischen Entwicklungsländern auf, Hauptbetroffene sind dabei vor allem Kinder unter 5 Jahren; neben der allgemeinen Verbreitung von Malaria veranschaulicht Abb. 1 auf der vorherigen Seite auch die inzwischen häufig auftretenden Resistenzen gegen zwei der wichtigsten derzeit verwendeten Medikamente. Vor allem der massenhafte Einsatz von Chloroquin (1) zu prophylaktischen Zwecken in Afrika Ende der 50er Jahre des vergangenen Jahrhunderts führte inzwischen zur Verbreitung Chloroquin-resistenter Malaria-Stämme in fast ganz Afrika.

Aus diesem Grund besteht nach wie vor die dringende Notwendigkeit, alternative Leitstrukturen gegen die Malaria zu suchen. Eine vielversprechende Substanzklasse sind die Naphthylisochinolin-Alkaloide, deren Antimalaria-Wirkeigenschaften seit Mitte der 90er Jahre bekannt sind.^[3,4] Zwei sehr aktive Vertreter, Dioncopeltin A (3) und Dioncophyllin C (4), wurden aus der tropischen Liane *Triphyophyllum peltatum* isoliert^[5,6] und zeigen stellvertretend die strukturell ungewöhnlichen Eigenschaften dieser Substanzklasse (siehe Abb. 2):

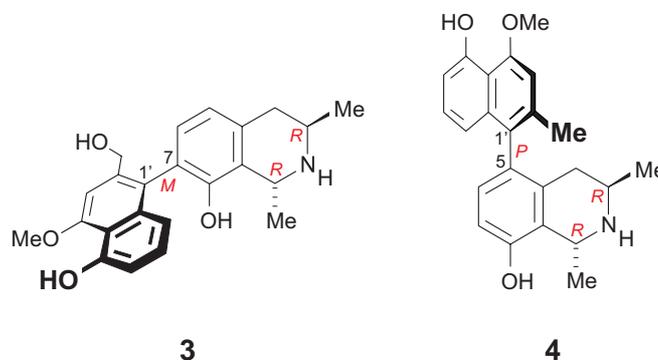


Abb. 2. Zwei der aktivsten Vertreter der Naphthylisochinolin-Alkaloide: Dioncopeltin A (3) und Dioncophyllin C (4); das für die dreidimensionale Form bestimmende Strukturelement ist eine bei den meisten Vertretern stereogene Biarylachse

Naphthylisochinolin-Alkaloide zeigen ein interessantes Spektrum an Bioaktivitäten, vor allem gegen die Erreger gefährlicher Tropenkrankheiten wie beispielsweise der Schlafkrankheit^[3] (*Trypanosoma brucei rhodesiense*), der visceralen Leishmaniose^[7] (*Leishmania donovani*) oder der Chagas-Krankheit^[7] (*T. cruzi*). Am differenziertesten untersucht werden konnte bisher die Antimalaria-Aktivität, die zunächst

anhand von *In-vitro*-Tests an verschiedenen Parasiten, wie *P. falciparum*, *P. chabaudi chabaudi* und auch an dem für den Menschen ungefährlichen *P. berghei* nachgewiesen wurde.

Die beiden in Abb. 2 gezeigten Verbindungen gehören mit IC₅₀-Werten^a von 5 ng/ml (3) und 6 ng/ml (4) zu den *in vitro* bisher vielversprechendsten Substanzen gegen *P. falciparum*. Inzwischen konnten auch *in vivo* hervorragende Resultate erzielt werden: Mit *P. berghei* infizierte OF-1-Mäuse zeigten nach der Behandlung mit Di-oncophyllin C (4) schon innerhalb weniger Tage keine Krankheitssymptome mehr, die Parasitämie sank unter die Nachweisgrenze.^[8]

Informationen über mögliche biologische Targets liegen bereits vor: Naphthylisochinolin-Alkaloide zeigen sowohl gegen erythrocytische^[9] wie auch gegen exoerythrocytische^[10] Formen gute Aktivitäten. Neben Untersuchungen zur Phasensensitivität^[11,12] gibt es Hinweise auf eine Hemmung des Häm-Abbaus,^[13] ähnlich wie er auch für Chloroquin (1) diskutiert wird.^[14]

Die Isolierung, Strukturaufklärung und Synthese der Vertreter dieser Naturstoffklasse steht im Mittelpunkt des Interesses unserer Arbeitsgruppe. Zentraler Aspekt ist dabei die Erschließung neuer Methoden auf unterschiedlichsten Teilgebieten: Moderne analytische Verfahren wie die Kopplung von HPLC mit MS/MS, NMR und der CD-Spektroskopie (sog. 'hyphenated methods'^[15,16,17]) sowie computergestützte Verfahren zur Strukturaufklärung^[18,19] gehen dabei einher mit der Entwicklung und Anwendung von neuen Synthesestrategien wie dem 'Lacton-Konzept'^[20,21] oder der oxidativen Biarylkupplung.^[22]

Diese enge Zusammenarbeit verschiedener Disziplinen 'unter einem Dach' hat einen großen Beitrag dazu geleistet, daß heute etwa 100 Vertreter der Naphthylisochinolin-Alkaloide bekannt und zum Teil totalsynthetisch erschlossen sind.^[22,23] Der Einfluß bestimmter funktioneller Gruppen auf die biologische Aktivität wurde durch zahlreiche partialsynthetische Modifikationen genauer untersucht.^[10] Auf diese Weise konnte ein einzigartiger Datensatz erschlossen werden, der einen großen Teil der strukturellen Diversität der Naphthylisochinolin-Alkaloide abdeckt.

^a IC₅₀: Inhibitory Concentration 50; die Konzentration einer Substanz, die das Wachstum der Parasiten um 50% senkt.

Aufbauend auf einen derartigen Datensatz könnten nun computergestützte QSAR^a-Methoden eine Korrelation zwischen der dreidimensionalen Struktur und der biologischen Aktivität herstellen und auf diese Weise zu einer systematischen Optimierung beitragen. Vitt^[24] und Schmitt^[25] in unserem Arbeitskreis versuchten daher bereits, mit Hilfe der CoMFA^b-Methode^[26] derartige QSAR-Modelle aufzustellen. Während der praktischen Umsetzung machten sich allerdings einige methodische Unzulänglichkeiten bemerkbar: So betrachtet CoMFA lediglich sterische und elektrostatische Wechselwirkungen zwischen Ligand und Rezeptor, was vermutlich für eine hinreichende Beschreibung des Bindungsmodus von Naphthylisochinolin-Alkaloiden nicht ausreicht. Zudem müssen sog. *cut-off*-Energien definiert werden, was möglicherweise der Grund dafür war, daß entscheidende Bereiche im Raum nicht berücksichtigt werden konnten.^c

Ferner zeigten die zwar für den Chemiker reizvollen, weil einzigartigen strukturellen Eigenschaften schnell die Grenzen der computerchemischen Modelle auf: Die durch unterschiedliche Kupplungspositionen bedingten dreidimensionalen Formen sind nicht trivial aneinander auszurichten (zu 'alignen', von engl. 'alignment'), ein Schritt, der für die Resultate einer CoMFA-Analyse entscheidend ist. Die Verwendung von GASP,^d das in der Lage ist, automatische Alignment-Vorschläge zu berechnen,^[27,28] war nur sehr eingeschränkt möglich, da die damalige Version es nicht erlaubte, axialchirale Verbindungen korrekt zu berücksichtigen. Die resultierenden CoMFA-Modelle wiesen zwar beachtliche statistische Eigenschaften auf,^[3,29] basierten aber jeweils nur auf sehr kleinen Datensätzen mit demzufolge eingeschränkter struktureller Diversität.

Inzwischen gibt es gerade auf dem Gebiet des molekularen Alignments erhebliche methodische Weiterentwicklungen. Das Programm GASP erlaubt es in der aktualisierten Version nun, die Beweglichkeit einzelner Torsionswinkel so einzuschränken, daß axialchirale Verbindungen korrekt behandelt werden können. Sehr vielversprechend erschien des weiteren das Tool FLEXS,^e das von Lemmen^[30,31] entwickelt wurde. Es basiert auf Algorithmen, die sich bereits innerhalb des sehr populären

^a Quantitative Structure-Activity Relationships

^b Comparative Molecular Field Analysis

^c Entsprechende Einzelheiten werden in Abschnitt 2.4 genauer erläutert.

^d Genetic Algorithm Similarity Program

^e Von Flexible Superposition. Zu Beginn dieser Arbeit befand sich FLEXS noch in der Beta-Testphase, inzwischen ist es fertiggestellt und kommerziell erhältlich.

Docking-Programms FLEXX^[32,33] bewährt hatten.

Der entscheidende Vorteil dieser Anwendungen gegenüber dem bisher verfolgten manuellen Alignment ist die direkte Verknüpfung von Konformationsenergien und der Qualität des Alignments in einer einzigen Bewertungsfunktion. Auf diese Weise kann die Flexibilität der Strukturen sehr viel umfassender berücksichtigt werden, als es aufbauend auf einer klassischen Konformationsanalyse, per manuellem Alignment, möglich ist.

Die genannten Schwierigkeiten, die bei der Anwendung der CoMFA-Methode auftraten, können inzwischen ebenfalls umgangen werden. Das CoMSIA^a-Verfahren betrachtet zusätzlich zu sterischen und elektrostatischen Wechselwirkungen auch Wasserstoffbrückenbindungen und die Hydrophobie der Liganden. Diese zusätzliche Differenzierung erlaubt eine wesentlich flexiblere Anpassung an unterschiedliche Fragestellungen.^[34] Darüber hinaus wurde die mathematische Implementierung der Energiefunktionen in einer Weise verändert, die die Verwendung von *cut-off*-Werten überflüssig macht.

Diese methodischen Fortschritte sollten es ermöglichen, ein verbessertes QSAR-Modell zu erstellen, das die Zusammenhänge zwischen den guten Antimalaria-Aktivitäten und den diversen Strukturelementen der Naphthylisochinolin-Alkaloide sehr viel genauer als bisher beschreiben kann.

Darüber hinaus erschien es äußerst reizvoll zu prüfen, inwieweit axialchirale Strukturen mit Hilfe von FLEXS und GASP automatisch aneinander ausgerichtet werden können, bisher sind derartige Untersuchungen in der Literatur nicht beschrieben.

Vor diesem Hintergrund ergaben sich für die vorliegende Arbeit folgende Aufgabenstellungen:

1. Erstellung eines QSAR-Modells zur Vorhersage der Antimalaria-Aktivität von Naphthylisochinolin-Alkaloiden mit Hilfe der CoMSIA-Methode;
2. Anwendung neuer, flexibler Ansätze zur Aufstellung des Alignments von axialchiralen Verbindungen;

^a Comparative Molecular Similarity Indices Analysis

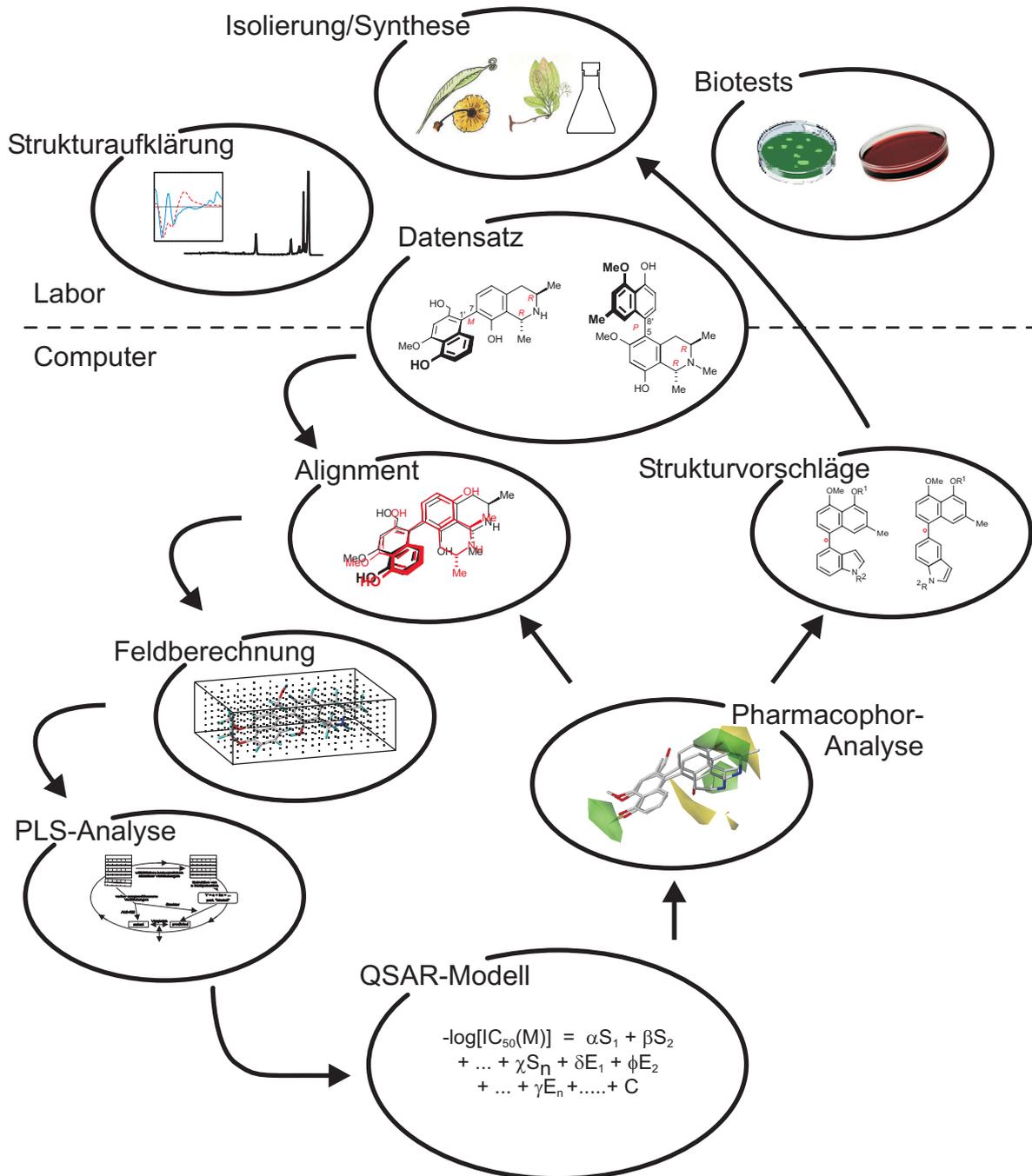
3. Eingrenzung der für die antiplasmodiale Aktivität (in positiver oder negativer Weise) entscheidenden funktionellen Gruppen und Aufstellung eines Pharmacophor-Modells;
4. Daraus resultierend Erarbeitung von Vorschlägen zu möglichen Vereinfachungen des komplizierten Grundgerüsts von Naphthylisochinolin-Alkaloiden (beispielsweise Reduktion der stereogenen Elemente), mit der Absicht, neue Impulse für Synthesen und eine zielgerichtete Optimierung der biologischen Aktivität zu geben;
5. Schließlich Prüfung der Vergleichbarkeit der Aktivitäten anderer Substanzklassen innerhalb eines QSAR-Modells mit Naphthylisochinolin-Alkaloiden, mit dem Ziel, Hinweise auf ähnliche Wirkmechanismen zu finden.

Kapitel 2

Theoretische Grundlagen

QSAR-Methoden ermöglichen es, eine direkte Beziehung zwischen der Struktur einer Verbindung und ihrer biologischen Wirksamkeit herzustellen. Für die Formulierung eines solchen Zusammenhangs müssen zunächst die Strukturen in einer mathematisch erfaßbaren Form dargestellt werden. Dazu werden sog. 'Deskriptoren' verwendet, die es erlauben, eine chemische Struktur mathematisch zu beschreiben. Anschließend werden innerhalb dieses numerischen Ausdrucks Korrelationen mit der biologischen Aktivität gesucht.

Die in dieser Arbeit verwendete CoMSIA-Analyse erlaubt es, ohne jegliche Kenntnis des Wirkorts Struktur-Wirkungs-Beziehungen aufzustellen. Eine Übersicht des prinzipiellen Ablaufes dieses Verfahrens zeigt Schema 1 auf der nächsten Seite. Die Methode geht aus von einem experimentell erarbeitetem Datensatz, bestehend aus dreidimensionalen Strukturen und den biologischen Aktivitäten der Verbindungen. Nachdem die Strukturen aneinander ausgerichtet (überlagert) wurden (siehe Schema 1 auf der nächsten Seite: 'Alignment'), wird eine Art Negativ-Abdruck jeder Substanz generiert ('Feldberechnung'); anschließend wird versucht, aus dem Vergleich dieser Negative auf die für die Aktivität ausschlaggebenden Merkmale zu schließen (mit Hilfe der 'PLS-Analyse'). Entscheidend ist dabei der *Vergleich*, denn ein einzelner Abdruck alleine erlaubt zunächst keinerlei Rückschlüsse. Erst die bei der Gegenüberstellung auftretenden Unterschiede erlauben eine Korrelation. Die Suche nach einem sinnvollen Alignment ist dabei meist der schwierigste Teil, insbesondere wenn der zugrunde liegende Datensatz sehr unterschiedliche Strukturprinzipien enthält.



Schema 1: Schematischer Ablauf einer QSAR-Untersuchung: Die Strukturen in Verbindung mit den biologischen Aktivitäten bilden den Trainingsdatensatz; nach der Erstellung des QSAR-Modells (Alignment, Feldberechnung, PLS-Analyse), kann mit dessen Hilfe das Alignment nochmals überprüft und gegebenenfalls verbessert werden; mit dem endgültigen Modell können neue Verbindungen oder Modifikationen vorgeschlagen werden

2.1 Der Datensatz

Der einer QSAR-Untersuchung zugrunde liegende Datensatz muß mehrere Bedingungen erfüllen:

1. Alle Verbindungen müssen dem gleichen Wirkmechanismus folgen, d. h. gleiche biologische Targets besitzen.
2. Die Testwerte sollten eine möglichst hohe Varianz aufweisen, d. h. der Aktivitätsbereich aller getesteten Verbindungen sollte mindestens drei Größenordnungen^[26] umspannen.^a Eine möglichst gleichmäßige Verteilung über den gesamten Wertebereich ist dabei von Vorteil.
3. Die strukturellen Unterschiede innerhalb des Datensatzes sollten nicht zu stark variieren, da sonst kein sinnvolles Alignment mehr möglich ist.

2.2 Aufstellen von 3D-Strukturen

Liegen keinerlei Informationen über einen Protein-Ligand-Komplex vor, müssen alle 3D-Strukturen computerchemisch berechnet werden, wobei eine große Anzahl von Möglichkeiten^b zur Verfügung steht. Viel wichtiger als eine sehr exakte Berechnung der Struktur (beispielsweise der Bindungslängen) ist für eine QSAR-Untersuchung allerdings die Konformation der Liganden, die der aktiven Konformation möglichst ähnlich sein sollte. Je besser die realen Verhältnisse (Protein-Ligand-Wechselwirkungen, Konformationsenergien) dargestellt werden, desto exakter sollte auch das erstellte QSAR-Modell diese wiedergeben können.

2.3 Das Alignment

Die Möglichkeit eines automatisierten Alignments, das auch die konformative Flexibilität der Liganden berücksichtigt, wurde in letzter Zeit eingehend diskutiert.^[35,36]

^a Die aktivste Verbindung sollte mindestens 1000mal aktiver sein als die am wenigsten aktive Substanz.

^b Möglich sind Kraftfeld-, semiempirische bis hin zu aufwendigen *ab-initio*-Methoden.

Zahlreiche Verfahren wurden entwickelt für die Berechnung von Konformeren kleiner, sogenannter 'drug-size'-Moleküle, wie das Programm MIMUMBA^[37] oder das aktuellere DISTILL-Verfahren^[38]. Eine abschließende manuelle Prüfung der Resultate ist momentan aber noch bei allen Verfahren angeraten.

Für diese Arbeit wurden zwei sehr unterschiedliche Methoden ausgesucht, die in der Lage sein sollten, die Flexibilität kleiner Moleküle während des Alignments direkt zu berücksichtigen: Zum einen war dies FLEXS, programmiert von Lemmen *et al.*,^[30,31] zum anderen kam das Programm GASP von Jones *et al.*^[27] zum Einsatz, das einen genetischen Algorithmus zur Berechnung von Konformeren verwendet, die gleichzeitig auf molekulare Ähnlichkeiten untersucht werden. Beide Programme sind kommerziell erhältlich, als Modul der Modelling-Software SYBYL,^[38] wobei FLEXS erst kürzlich voll integriert wurde.^a

2.3.1 FLEXS^[30]

Das Programm FLEXS basiert auf Algorithmen, die bereits in dem Docking-Tool FLEXX^[32,33] verwendet wurden. Prinzipiell versuchen beide, Ligand-Protein-Wechselwirkungen zu simulieren. Während FLEXX Wechselwirkungen von Liganden mit bekannten Proteinen modelliert, verwendet FLEXS einen Referenz-Liganden als eine Art 'inverse' Bindungstasche und ist damit prädestiniert für das nötige Alignment als Grundlage einer CoMSIA-Analyse. Bei der Entwicklung stand der geringe Ressourcenverbrauch im Vordergrund, damit in kurzer Zeit eine große Zahl an Verbindungen berechnet werden kann.

Der FLEXS-Prozess behandelt stets nur zwei Strukturen, d. h. er richtet einen Test-Liganden ('test ligand') an einem sogenannten Referenz-Liganden ('reference ligand') aus. Dabei werden folgende Voraussetzungen zugrunde gelegt:^[30]

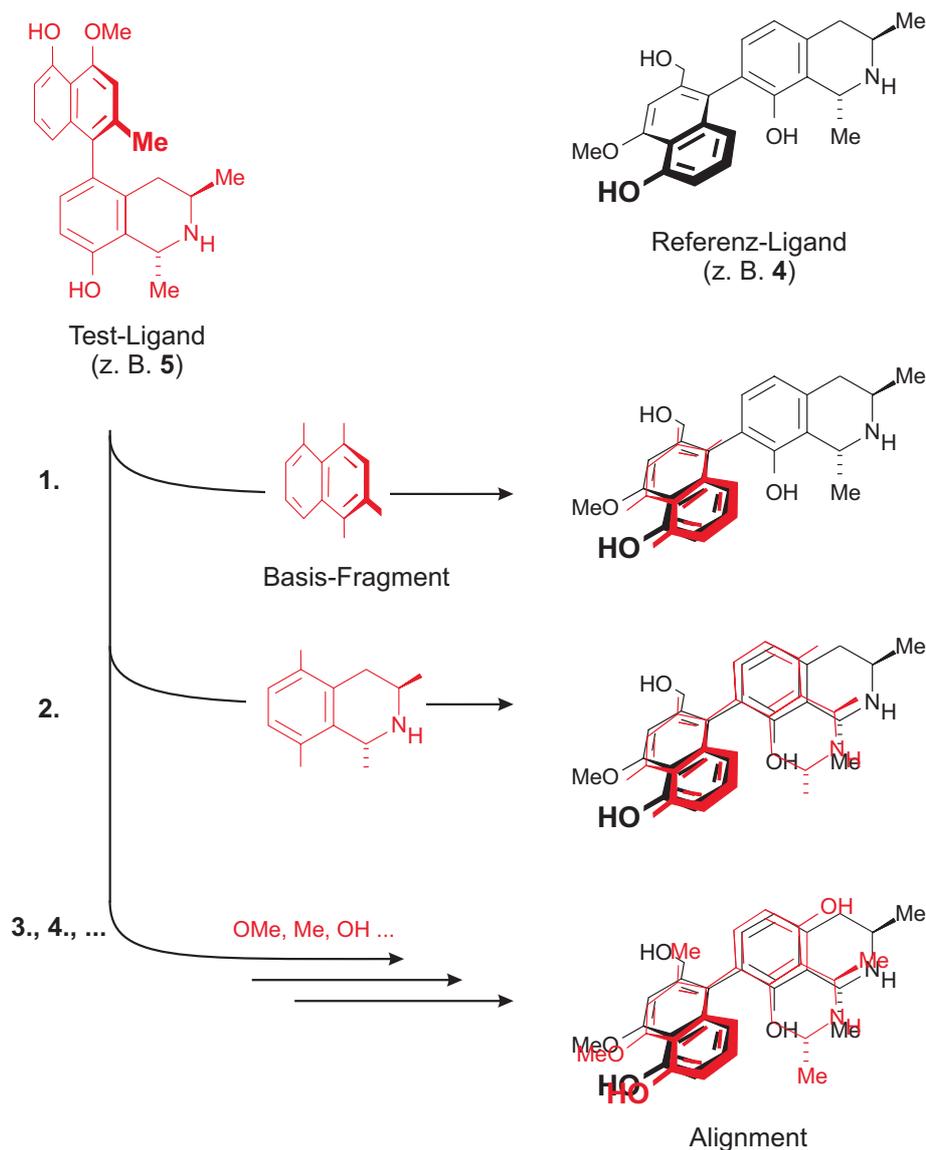
1. Referenz- und Test-Ligand nehmen im Raum die größtmögliche Überlappung ein, außerdem wechselwirken beide mit den gleichen funktionellen Gruppen in der Rezeptortasche.

^a Zu Beginn dieser Arbeit befand sich FLEXS noch in der Beta-Testphase, die hier beschriebenen Ergebnisse wurden alle mit der aktuellen Kommandozeilen-Version erarbeitet; diese unterscheidet sich lediglich durch die Bedienung von der in SYBYL implementierten Version.

2. Es werden generell nur zwei Liganden gleichzeitig untersucht.
3. Die Anzahl der Freiheitsgrade wird auf die Torsionswinkel der Bindungen sowie die Rotation und Translation der Strukturen beschränkt.
4. Alle Atome des Referenz-Liganden bleiben während des gesamten Alignment-Prozesses fixiert.

Bei der Auswahl des Referenz-Liganden müssen einige Überlegungen berücksichtigt werden: die gewählte Konformation sollte größtmögliche Ähnlichkeit mit der Rezeptortasche aufweisen (siehe oben). Sehr vorteilhaft sind hier insbesondere relativ starre Liganden mit guter Aktivität, da hier bereits das Vorliegen einer aktiven Konformation gewährleistet ist. Anschließend werden nacheinander alle Test-Liganden am Referenz-Liganden ausgerichtet. Dabei wird der entsprechende Test-Ligand zunächst in einzelne Fragmente zerlegt, wobei jede drehbare Bindung^a als Fragmentierungsgrenze dient (Schema 2 auf der nächsten Seite).

^a Betroffen sind alle Bindungen, an denen keine H-Atome beteiligt sind (alle nicht rotationsinvarianten Bindungen).



Schema 2: Der FLEXS-Algorithmus am Beispiel von zwei Naphthylisochinolin-Alkaloiden (3 und 4)

Jetzt wird ein Basis-Fragment gewählt und an einer geeigneten Stelle im Referenz-Liganden plziert. Anschließend werden die einzelnen Fragmente sukzessive wieder an die betreffende Stelle angehängt. Dabei rotiert man die geknüpfte Bindung so lange, bis eine größtmögliche Überlappung mit dem Referenz-Liganden gegeben ist. Auf diese Weise werden bis zu mehreren hundert Alignments pro Test-Ligand erzeugt, die dann mit Hilfe einer Wertungs-Funktion ('scoring function') verglichen und gewichtet werden.

Diese Funktion setzt sich im Wesentlichen aus zwei Bestandteilen zusammen: Einer-

seits aus intermolekularen Wechselwirkungen, die von beiden Liganden ausgeübt werden können. Dabei wird die Ähnlichkeit funktioneller Gruppen und die Möglichkeit zur Wechselwirkung mit einer potentiellen Rezeptor-Gruppe abgeschätzt. Andererseits werden Überlappungsvolumina verschiedener Dichtefunktionen berechnet, wie z. B. das elektrostatische Potential oder das van-der-Waals-Volumen.^[30]

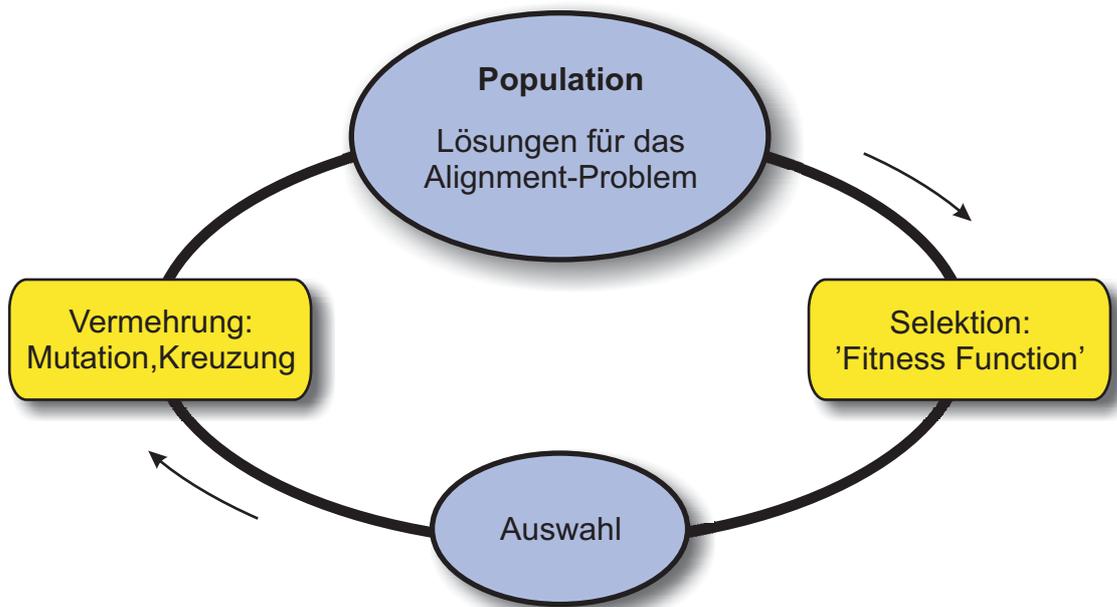
Ein Problem ergibt sich bei der Betrachtung nicht-aromatischer Ringsysteme, eine Fragmentierung im beschriebenen Sinne ist hier nicht möglich; für diesen Fall nutzt FLEXS eines der Programme CORINA^[39,40] oder SCA;^[41] beide sind gängige Tools für die effiziente Berechnung von Konformeren kleiner Ringsysteme.

2.3.2 GASP^[27,28]

GASP ist ein genetischer Algorithmus, der speziell für die Überlagerung von kleinen Molekülen entwickelt wurde.

2.3.2.1 Genetische Algorithmen

Genetische Algorithmen versuchen, die Vorteile natürlicher Selektion zu nutzen. Bezogen auf ein Alignment bedeutet dies, Orientierung und Konformation von Molekülen optimal an bestimmte Ähnlichkeitsbedingungen zwischen den Strukturen anzupassen. Eine genetische Operation beginnt immer mit einer zufällig erzeugten Ansammlung von Problemlösungen. Eine solche Lösung wird als Chromosom bezeichnet, einen ganzen Satz an Lösungen nennt man eine Population (siehe Schema 3 auf der nächsten Seite).



Schema 3: Schematische Darstellung eines genetischen Algorithmus. Ein Umlauf entspricht einer Generation

Während der sog. Selektion wird allen Lösungen der Population eine gewisse Überlebenswahrscheinlichkeit zugewiesen, die sich nach deren Qualität richtet. Entsprechend dieser Wahrscheinlichkeit scheiden einige Lösungen aus und auf die so erhaltene Auswahl werden anschließend genetische Operationen (Kreuzung und Mutation) angewendet und wiederum eine neue Population vervollständigt. Einen solchen Umlauf bezeichnet man als eine Generation.

Sobald die Lösungen innerhalb der Population ein bestimmtes Kriterium erfüllen (eine bestimmte 'Fitness' aufweisen), wird der Lauf abgebrochen.

2.3.2.2 Alignment per genetischem Algorithmus

Bezogen auf ein molekulares Alignment stellen die Elemente der Population Lösungen für das Alignment-Problem dar, also ausgerichtete Konformere der behandelten Strukturen.

Die während der Selektion berechnete Fitness-Funktion wird aus mehreren Bestandteilen abgeleitet. Dies sind im Wesentlichen die Anzahl und Ähnlichkeit überlagerter pharmacophorer Elemente, die Überlappung des Molekülvolumens der ausge-

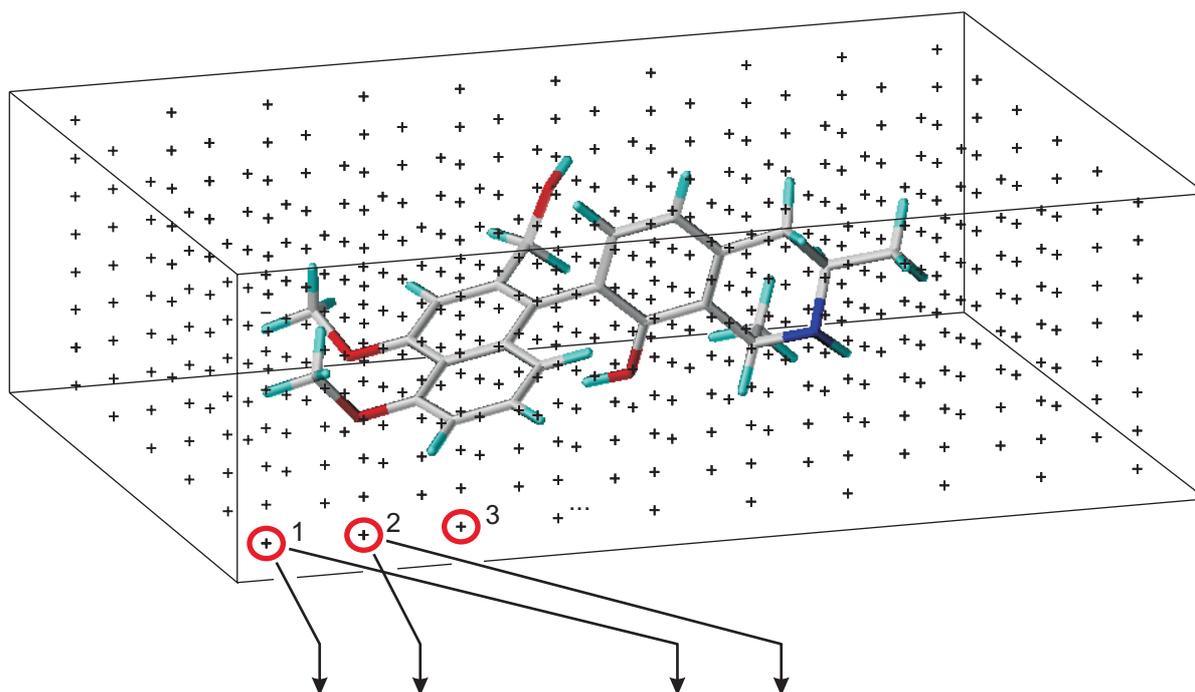
richteten Konformere und deren interne van-der-Waals-Energie.^[27,28]

Bei der Verwendung von GASP (wie auch bei FLEXS) ist es wichtig, den einzelnen Atomen korrekte SYBYL-Atom-Typen zuzuordnen. GASP fügt außerdem an bestimmten strukturellen Elementen Dummy-Atome hinzu,^a die nach dem Alignment-Prozeß wieder entfernt werden müssen, ansonsten könnten sie während der CoMSIA-Feldberechnung zu Störungen führen.

2.4 Die Berechnung der CoMSIA-Felder

Nachdem die Strukturen aneinander ausgerichtet wurden, folgt die Berechnung der Ähnlichkeits-Indizes. Die Strukturen werden dazu an definierten Gitterpunkten mit einer Sonde 'abgetastet', d. h. man berechnet an jedem Gitterpunkt die Summe der Wechselwirkungsenergien aller Atome mit der Sonde. Die Gitterpunkte werden innerhalb einer Box vorher festgelegt (siehe Abb. 3). Damit alle relevanten Einflüsse erfaßt werden, wird diese Box üblicherweise so gewählt, daß der Abstand der Wand zu den Molekülen jeweils 4 Å nicht unterschreitet.

^a Beispielsweise wird im Schwerpunkt von aromatischen Ringen ein Dummy-Atom plziert.



$$-\log(IC_{50}) = s_1S_1 + s_2S_2 + \dots + s_nS_n + e_1E_1 + e_2E_2 + \dots + e_nE_m + \dots + C$$

Abb. 3. Berechnung von Wechselwirkungsenergien zwischen Sonde und Molekül ($S_1, E_1, S_2, E_2, \dots$) an definierten Gitterpunkten

Die ursprüngliche Idee, die 3D-Struktur von kleinen Molekülen auf diese Weise mathematisch zu kodieren, geht auf Cramer^[26,42] zurück, wobei elektrostatische und sterische Wechselwirkungen berechnet wurden. Als Funktion verwendet man dabei ein 12-6-Lennard-Jones-Potential für die sterischen Wechselwirkungen (S)

$$S = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right], \quad (1)$$

und analog ein Coulomb-Potential (E) für die Elektrostatik eingesetzt:

$$E = \frac{q_1 \cdot q_2}{4\pi\epsilon_0} \cdot \frac{1}{r}. \quad (2)$$

Weil diese Funktionen bei kleinen Radien sehr steil ansteigen (Abb. 4), müssen sog. *cut-off*-Werte eingeführt werden; Energiewerte, die eine bestimmte Barriere überschreiten, setzt man auf einen konstanten Wert. So wurden zu hohe Einzelenergien verhindert, was während der PLS-Analyse zu Problemen geführt hätte. Allerdings konnten so Effekte in bestimmten Bereichen (sehr nahe an den Atomen) nicht mehr berücksichtigt werden.

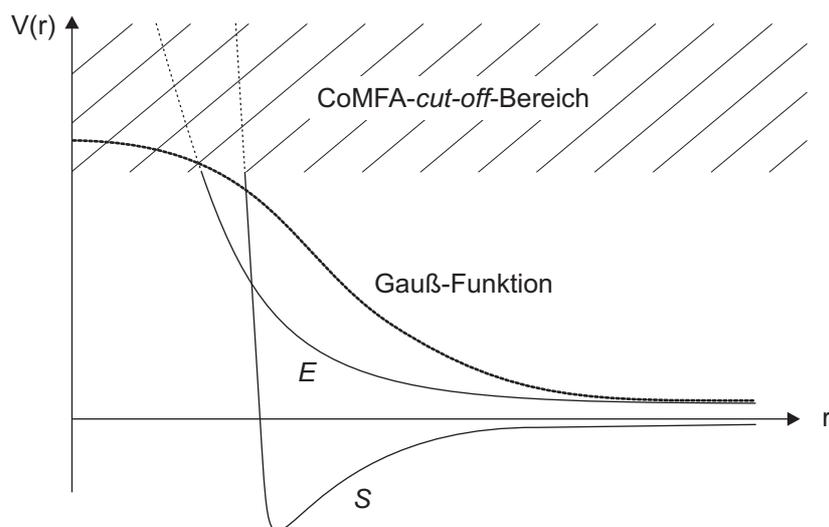


Abb. 4. Zu große Energiewerte aufgrund von zu steil verlaufenden Energiepotentialen (S und E) werden bei der CoMFA-Methode auf einen konstanten Wert gesetzt (gestrichelte Linienabschnitte werden nicht berücksichtigt); CoMSIA umgeht dies durch Verwendung von Gauß-Funktionen

Eine weiteres Problem bei der Verwendung von Lennard-Jones- und Coulomb-Potentialen war die Tendenz der sterischen Energien, die elektrostatischen zu dominieren. Elektrostatische Effekte konnten so nicht mehr erkannt werden. Aus diesem Grund wurden Skalierungsfaktoren eingeführt, die die Energiewerte aneinander anpassen sollten.

Sowohl dieses als auch das vorher beschriebene *cut-off*-Problem wird in CoMSIA durch die Verwendung von Gauß-Funktionen für die Abstandsabhängigkeit der Energien umgangen (Abb. 4 und Gleichung 3). Die erhaltenen Energiewerte haben allerdings nicht mehr den physikalischen Bezug wie noch die CoMFA-Felder und werden demzufolge als sog. 'Ähnlichkeits-Indizes' ('Similarity Indices') bezeichnet.

Die CoMSIA-Ähnlichkeits-Indizes werden nach folgender Gleichung berechnet:

$$A_{f,g}^q(j) = - \sum_i \omega_{probe,g} \omega_{ig} e^{\alpha r_{iq}^2} \quad (3)$$

Dabei gelten folgende Definitionen:

- A ist der Ähnlichkeits-Index am Gitterpunkt q , summiert über alle Atome i des Moleküls j .

- $\omega_{probe,g}$ ist die entsprechende Eigenschaft des Sonden-Atoms; diese sind Radius 1 Å, Ladung +1, Hydrophobie +1, Wasserstoffbrücken bindend +1 und Wasserstoffbrücken akzeptierend +1.
- ω_{ig} ist der tatsächliche Wert der Funktion g für Atom i (siehe unten).
- r_{iq} ist der Abstand der Sonde am Gitterpunkt q zu Atom i des Test-Moleküls j .
- α ist der sog. Abschwächungs-('attenuation')-Faktor, der standardmäßig zwischen 0.2 und 0.4 gewählt wird.^[43] Größere Werte resultieren in steileren Gauß-Kurven und damit einer starken Reduzierung der abstandsabhängigen Effekte der molekularen Ähnlichkeit. Globale Eigenschaften treten in den Hintergrund, und lokale Merkmale werden genauer dargestellt (weniger 'verschmiert'). Die physikalisch-chemische Bedeutung dieses Wertes liegt in der Veränderung des Verhältnisses von lokalisierten (z. B. einzelne funktionelle Gruppen) und globalen Eigenschaften der betrachteten Substanzen (Form des ganzen Moleküls). Der optimale Wert hängt von der aktuellen Problemstellung ab, Standard ist 0.3.^[43]

Es stehen fünf Funktionen (g) zu Verfügung: Sterische und elektrostatische Wechselwirkungen, Wasserstoffbrücken Donor- und Akzeptor-Funktionen sowie die Hydrophobie. Bei der Auswahl der Funktionen wurde vor allem darauf geachtet, alle für die Ligand-Protein-Wechselwirkung wichtigen Beiträge zu berücksichtigen.^[34]

Die tatsächlichen Funktionswerte für jedes Atom stammen aus folgenden Quellen:^[38]

- Die Werte für das **sterische Feld** entsprechen der dritten Potenz des Atomradius.
- Das **elektrostatische Feld** basiert auf entsprechend vorher berechneten Atomladungen. Üblicherweise werden Gasteiger-Hückel-^[44] oder semiempirische AM1-Ladungen verwendet.
- Für die Simulation von **Wasserstoffbrücken-Donor- und -Akzeptor-Wechselwirkungen** werden Pseudoatome an möglichen Wasserstoffbrücken-Donor-

bzw. -Akzeptor-Punkten erstellt,^a die entsprechende Andockstellen in einem potentiellen Rezeptor-Protein simulieren.

- Die Werte für das **Hydrophobiefeld** basieren auf Arbeiten von Viswanadhan.^[47] Es handelt sich dabei um Lipophilie-Partialwerte für Molekülfragmente oder einzelne Atome, die durch statistische Regressionsverfahren aus experimentellen Daten abgeleitet wurden.

2.5 Die PLS^b-Analyse

2.5.1 Vorgehensweise

Wie in Abb. 3 auf Seite 16 angedeutet, versucht man, die an den Gitterpunkten (1 bis m) berechneten Ähnlichkeits-Indizes (S , E)^c mit der biologischen Aktivität (den IC_{50} -Werten) der entsprechenden Substanz zu korrelieren:

$$-\log(IC_{50}) = s_1 S_1 + s_2 S_2 + \dots + s_m S_m + e_1 E_1 + e_2 E_2 + \dots + e_m E_m + \dots + C$$

oder zusammengefaßt

$$-\log(IC_{50}) = \sum_{i=1}^m (s_i S_i + e_i E_i) + C. \quad (4)$$

Dabei sind s_i und e_2 die zu bestimmenden Regressionskoeffizienten und C eine Fehlerkonstante. Die biologische Aktivität wird dabei aus statistischen und energetischen Gründen^[26] in einer logarithmischen Skala ausgedrückt.

Wird für jede Verbindung (1, 2, ..., N) im Trainingsdatensatz eine Gleichung 4 aufgestellt, ergibt sich mit $-\log(IC_{50}) = pIC_{50} = Y$ ein unterbestimmtes^d Gleichungssy-

^a Ähnlich den in dem Programm DISCO verwendeten 'Extension-Points'.^[45,46]

^b Partial Least Squares

^c Der Übersichtlichkeit halber werden hier nur zwei (S , E) der fünf CoMSIA-Indizes verwendet.

^d Es gibt wesentlich mehr Variablen (Koeffizienten an den einzelnen Gitterpunkten) als Gleichungen (jeweils eine pro Verbindung).

stem:

$$\begin{aligned}
 Y(1) &= \sum_{i=1}^m (s_i S_{1i} + e_i E_{1i}) + C \\
 Y(2) &= \sum_{i=1}^m (s_i S_{2i} + e_i E_{2i}) + C \\
 &\quad \vdots \\
 Y(N) &= \sum_{i=1}^m (s_i S_{Ni} + e_i E_{Ni}) + C
 \end{aligned}$$

Dieses kann man als Matrixgleichung formulieren:

$$\begin{pmatrix} Y(1) \\ Y(2) \\ \vdots \\ Y(N) \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \begin{pmatrix} S_{11} S_{12} \dots S_{1m} \\ S_{21} S_{22} \dots S_{2m} \\ \vdots \\ S_{n1} S_{n2} \dots S_{nm} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \begin{pmatrix} E_{11} E_{12} \dots E_{1m} \\ E_{21} E_{22} \dots E_{2m} \\ \vdots \\ E_{n1} E_{n2} \dots E_{nm} \end{pmatrix} + \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{pmatrix}, \quad (5)$$

oder, wiederum zusammengefaßt, als

$$\mathbf{Y} = \mathbf{bX} + \mathbf{C}. \quad (6)$$

Dabei nennt man \mathbf{Y} auch *Aktivitäts-Matrix*, \mathbf{X} die *Struktur-Matrix* (oder *Descriptor-Matrix*), \mathbf{b} enthält die Regressionskoeffizienten. Das von Cramer^[26] verwendete Verfahren, dieses Gleichungssystem näherungsweise zu lösen, die PLS-Analyse, geht auf den dem Mathematiker H. Wold^[48,49] zurück und wurde später von S. Wold^[50] auf chemometrische Probleme angewendet.

Die PLS-Analyse ist ein iteratives Verfahren, bei dem aus der Struktur-Matrix \mathbf{X} sukzessive Untermatrizen \mathbf{K} ('Komponenten') extrahiert werden, die mit der Aktivitäts-Matrix \mathbf{Y} korrelieren. Aus diesen Matrizen wird ein neues Gleichungssystem aufgestellt (analog zu Gleichung 6):

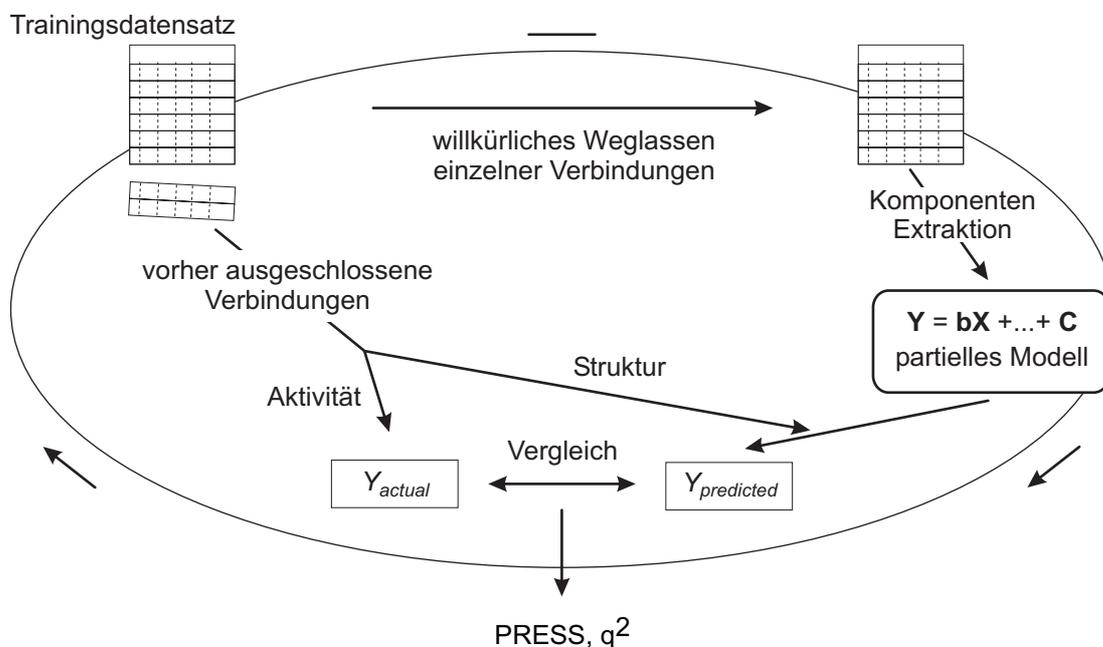
$$\mathbf{Y}_{\text{pred.}} = \mathbf{b}_k \mathbf{X}_{\text{pred.}} + \mathbf{C}_{\text{pred.}}, \quad (7)$$

wobei \mathbf{X} zur *Vorhersage-Matrix* $\mathbf{X}_{\text{pred.}}$ und \mathbf{Y} zur *Antwort-Matrix* $\mathbf{Y}_{\text{pred.}}$ wird. Je mehr Komponenten verwendet werden, um so besser gibt $\mathbf{Y}_{\text{pred.}}$ die ursprüngliche Matrix \mathbf{Y} aus Gleichung 6 wieder, gleichbedeutend mit steigender Vorhersagekraft des QSAR-Modells. Extrahiert man zu viele Komponenten \mathbf{K} , so paßt man das Modell zu sehr an die Trainings-Daten an und läuft Gefahr, nur noch Rauschen anstelle von brauchbarer Information zu reproduzieren; man spricht von sog. 'Overfitting', die

Vorhersagekraft sinkt wieder. Um die optimale Anzahl an Komponenten zu finden, bedient man sich des Prinzips der 'Crossvalidation'.

2.5.2 Die Kreuz- oder Gegenvalidierung

Während einer Kreuzvalidierung (von engl. crossvalidation) extrahiert man nacheinander Komponenten aus der Struktur-Matrix und prüft für jede weitere die Vorhersagekraft des Modells. Die optimale Zahl k an Komponenten ist gefunden, wenn zusätzliche Komponenten nur noch zu schlechteren Vorhersagen führen. Schema 4 zeigt den prinzipiellen Ablauf.



Schema 4: Ablauf einer Kreuzvalidierung

Bei der Crossvalidation wird zunächst willkürlich ein Teil des Trainingsdatensatzes ausgeklammert und basierend auf den restlichen Verbindungen ein Modell erstellt. Mit Hilfe dieses partiellen Modells werden anschließend die zunächst weggelassenen Verbindungen vorhergesagt; ist dies für jede Verbindungen genau ein Mal geschehen, werden die Kenngrößen q^2 und PRESS (siehe unten) berechnet und anschließend der komplette Prozeß für weitere Komponenten wiederholt, bis der q^2 -Wert nicht mehr ansteigt

Für jede Komponentenzahl k wird die Summe der Abweichungen der Vorhersagen

im Quadrat, engl. PRESS ('Predictive Error Sum of Squares') berechnet:

$$\text{PRESS} = \sum_{i=1}^n (Y_{i,\text{pred.}} - Y_{i,\text{exp.}})^2, \quad (8)$$

Wobei $Y_{i,\text{exp.}}$ die tatsächlichen und $Y_{i,\text{pred.}}$ die jeweils berechneten Aktivitäten sind.

Die statistische Signifikanz des abschließenden Modells bewertet man anhand des q^2 -Wertes:

$$q^2 = 1 - \frac{\text{PRESS}}{\sum_{i=1}^n (Y_{i,\text{exp.}} - \bar{Y})^2} \quad (9)$$

Dieser Wert variiert von $-\infty$ bis $+1$, wobei $+1$ einem perfekten Modell entspräche. Negative q^2 -Werte bedeuten, daß die Vorhersage des Modells schlechter ist als der einfache Mittelwert über alle Testwerte (letzteres entspräche $q^2 = 0$). Generell sollten brauchbare Vorhersagen ab einem q^2 -Wert von etwa 0.5–0.6 möglich sein.

Ist eine Korrelation innerhalb der Daten vorhanden, so steigt normalerweise der q^2 -Wert für die ersten Komponenten an (PRESS fällt); wird die optimale Komponentenzahl überschritten, sinkt er wieder (PRESS steigt).

Es gibt für CoMFA-/CoMSIA-Analysen die Faustregel,^[51,52] daß die Anzahl der Hauptkomponenten k etwa im Bereich von 10% der Anzahl der betrachteten Verbindungen sein sollte.

2.5.3 Berechnung des abschließenden Modells

Ist die optimale Komponentenzahl k (maximaler q^2 , minimaler PRESS) gefunden, wird das endgültige QSAR-Modell auf Basis aller vorhandenen Verbindungen berechnet (ohne Crossvalidation). Für solches Modell berechnet man den r^2 -Wert (anstelle von q^2 für kreuzvalidierte Vorhersagen). Dabei ist r^2 stets höher als q^2 ; Verbindungen, die zur Erstellung eines Modells verwendet wurden, können von diesem natürlich besser vorhergesagt werden, als 'externe' Datenpunkte.

Die tatsächliche Vorhersagekraft eines Modells kann man an anhand von Vorhersagen eines Testdatensatzes erkennen; man berechnet dazu den sogenannten 'predictive r^2 ', in dieser Arbeit abgekürzt als $r_{\text{pred.}}^2$; dieser Wert ist analog zu q^2 definiert, nur nicht bezogen auf Testwerte im Trainingsdatensatz, sondern auf einen externen

Testdatensatz.

2.5.4 Algorithmen zur Einsparung von Rechenzeit

Während der in dieser Arbeit durchgeführten PLS-Analysen wurden zwei gängige Verfahren verwendet, die die Rechenzeit enorm verkürzen können. Nötig wird dies vor allem während einer Kreuzvalidierung, da hier sehr viele einzelne PLS-Analysen durchgeführt werden (so viele wie Substanzen im Trainingsdatensatz).

- 'Column Filtering': Bestimmte Spalten der Struktur-Matrix (Gleichung 6 auf Seite 20), innerhalb derer die Varianz der Energie unter einem bestimmten Wert ('minimum Sigma') liegt, werden von der PLS-Analyse ausgeschlossen. Der Einfluß derartiger Spalten ist normalerweise ohnehin sehr gering.
- Der SAMPLS-Algorithmus^[53] ist eine Modifikation der gewöhnlichen PLS-Analyse, gegenüber der er im Ergebnis keinerlei Vorteile erbringt. Lediglich die Rechenzeit ist um einige Größenordnungen geringer, und die erhaltenen q^2 -Werte sind nahezu gleich, so daß man den q^2_{SAMPLS} -Wert als eine Art 'schnelle Vorschau' auf das Ergebnis einer normalen PLS-Analyse sehen kann.^[54] Eine genaue Analyse der Abweichungen einzelner Verbindungen ist mit dem Programm SYBYL momentan allerdings nicht möglich.

Kapitel 3

Kenntnisstand

3.1 Malaria: Erreger und Bekämpfung

3.1.1 Plasmodien-Arten und deren Lebenszyklus

Es gibt vier Arten von Malaria-Parasiten, die den Menschen befallen. Die möglichen Auswirkungen sind dabei sehr unterschiedlich: Infektionen mit *Plasmodium vivax*, *P. ovale* und *P. malariae* enden nur selten tödlich, die bei weitem höchste Mortalität tritt bei *P. falciparum* auf, dem Erreger der Malaria tropica.^[55] Weitere Arten befallen beispielsweise ausschließlich Nagetiere (*P. berghei*) oder Primaten (*P. knowlesi*); die Analogien in den Entwicklungszyklen der unterschiedlichen Arten sind groß, so daß den Menschen nicht befallende Stämme oft für Modellstudien verwendet werden.^[56]

Plasmodien sind äußerst 'evasive' Parasiten, d. h. sie weichen den Abwehrmechanismen des Wirts aus. Die meiste Zeit verbringen sie innerhalb der Erythrocyten und der Leber, unerreichbar für das Immunsystem.^[57]

Der Lebenszyklus von Plasmodien ist in Abb. 5 auf der nächsten Seite schematisch dargestellt.

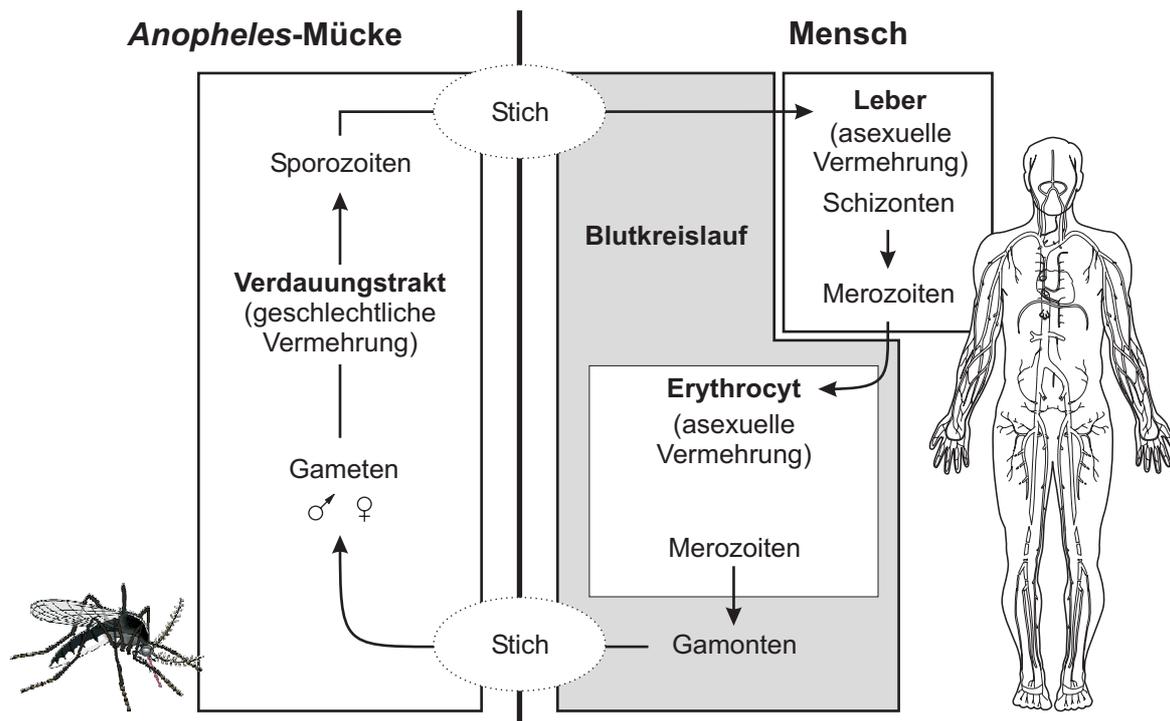


Abb. 5. Schematische Darstellung des Lebenszyklus von Plasmodien^[57]

Durch den Stich einer infizierten *Anopheles*-Mücke wird die Krankheit in Form von Sporozoiten auf den Menschen übertragen. Die Erreger gelangen aus der Speicheldrüse des Überträgers in den Blutkreislauf und dringen innerhalb weniger Minuten in die Leber des Wirts ein. Dort vollziehen sie bereits nach einigen Tagen (*P. falciparum*) zahlreiche Teilungen (asexuelle Vermehrung), bei anderen Parasitenarten (*P. vivax*, *ovale* und *malariae*) kann dem eine Ruhephase von bis zu mehreren Jahren vorangehen. Die gereiften Schizonten, sogenannte Merozoiten, können abschließend innerhalb der Blutbahn Erythrocyten befallen. Innerhalb der roten Blutkörperchen findet wiederum eine asexuelle Vermehrungsphase statt, wobei in Intervallen von 48-72 Stunden neue Merozoiten freigesetzt werden, die weitere Erythrocyten befallen. Dieser Vorgang ist verantwortlich für die periodisch wiederkehrenden Fieberschübe bei Malaria-Patienten. Einige der Merozoiten entwickeln sich in der Folge zu Gamonten, die wiederum von einer Mücke aufgenommen werden können und dort zu männlichen oder weiblichen Gameten reifen. Im Verdauungstrakt der Mücke findet die geschlechtliche Fortpflanzungsphase statt. Das Ergebnis ist die Bildung von Sporozoiten, die wiederum auf neue Wirte übertragen werden können. So schließt sich der Zyklus.^[57]

Dieser komplexe Lebenszyklus bietet prinzipiell verschiedene Angriffspunkte für die Bekämpfung von Malaria. Neben der medikamentösen Behandlung ist dies vor allem die Einschränkung der Ausbreitung der *Anopheles*-Mücke durch den Einsatz von Pestiziden. So war beispielsweise der großflächigen Einsatz von DDT^a entscheidend dafür verantwortlich, daß Malaria in Europa und Russland praktisch nicht mehr vorkommt.^[58]

Mittlerweile ist DDT allerdings aufgrund von toxischen Effekten auf Warmblütern weltweit geächtet,^[58] in Deutschland sogar verboten. Eine Maßnahme die, wegen der enormen Vorteile die die Anwendung von DDT erbracht hat, allerdings nicht unumstritten ist.^[59,60]

3.1.2 Der Wirkmechanismus von Chloroquin (1)

Die Wirkungsweise von Chloroquin (1) (vgl. Abb. 1 auf Seite 1) ist im Detail noch immer Gegenstand von Diskussionen.^[14] Prinzipiell sind die wichtigsten Vorgänge aber bekannt: Chloroquin (1) wirkt lediglich auf das erythrocytische Stadium der Plasmodien. Diese ernähren sich innerhalb der roten Blutkörperchen durch die Aufnahme von Hämoglobin aus dem Cytoplasma der Wirtszelle (siehe Abb. 6 links).

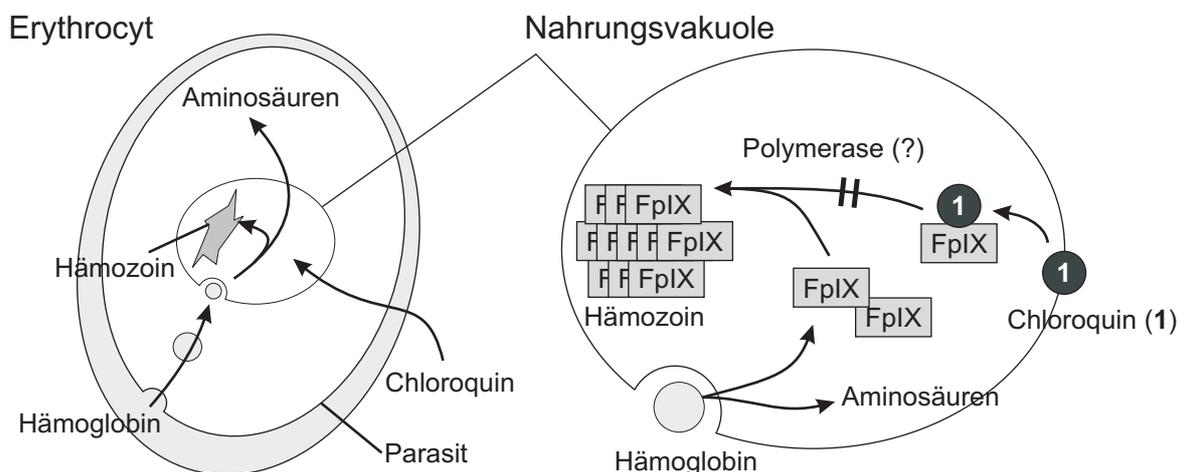


Abb. 6. Der Wirkmechanismus von Chloroquin (1) (siehe Text)

Das Hämoglobin wird in der Nahrungsvakuole abgebaut, wobei als 'Abfallprodukt' freies Häm (Ferriprotoporphyrin IX, FpIX) entsteht (Abb. 6 rechts). Dieses ist für den

^a 1,1,1-Trichlor-2,2-bis(*p*-chlorphenyl)ethan (DDT, Gerasol[®])

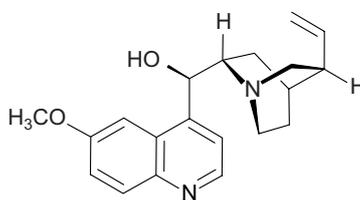
Parasiten toxisch, beispielsweise können entstehende aktivierte Sauerstoffspezies zelluläre Membranen zerstören. Aus diesem Grund polymerisiert der Parasit Häm zu unlöslichen Kristallen, dem Hämozoin (Malaria-Pigment), vermutlich mit Hilfe einer Polymerase (Abb. 6 auf der vorherigen Seite). Chloroquin (1) reichert sich innerhalb der Nahrungsvakuole an, und es wird angenommen, daß es durch Komplexbildung mit Häm die Polymerisation zu Hämozoin verhindert und der Parasit so an seinem eigenen Abfallprodukt zugrunde geht.^[14] Ein Teil des entstandenen freien Häms verlässt die Nahrungsvakuole und wird in der Folge durch Glutathion abgebaut. Auch dieser Prozess wird durch Chloroquin (1) inhibiert.^[61]

Chloroquin-resistente Stämme verhindern die Anreicherung von Chloroquin (1) in der Nahrungsvakuole, und gehen so der Wirkung aus dem Weg.^[14,62]

Die Entwicklung von Chloroquin (1), die Entstehung der weitverbreiteten Resistenzen und alternative Ansätze werden im folgenden Abschnitt kurz zusammengefaßt.

3.1.3 Die Chloroquin-Problematik

Bis zur Herstellung synthetischer Arzneimittel war das einzig bekannte Malaria-Medikament Chinin (5) (siehe Abb. 7),^[63] ein wirkungsvolles Antipyretikum,^a das inzwischen für die Malariaphylaxe als zu toxisch eingeschätzt wird und deshalb heute nur bei schweren Malariafällen eingesetzt wird.^[14] Bedingt durch die eingeschränkte Anwendung, oft in Kombination mit anderen Wirkstoffen, sind allerdings kaum Resistenzen bekannt.



Chinin (5)

Abb. 7. Chinin (5) aus den Chinarindenbäumen *Cinchona officinalis* und *C. succirubra*

Ausgehend von Chinin (5) wurde versucht, unter Erhalt der Aktivität weniger toxische Analoga zu entwickeln.^[63] Das erfolgreichste so entstandene Derivat ist Chlo-

^a Fiebersenkendes Mittel

roquin (**1**) (siehe Abb. 8), das bereits 1934 als 'Resochin[®]' von der Bayer AG synthetisiert wurde^a und lange weltweit das Hauptmedikament gegen akute Malaria war. Chloroquin (**1**) kann oral appliziert werden, ist relativ gut verträglich, und die vergleichsweise einfache chemische Struktur ermöglicht eine Produktion im Tonnenmaßstab bei geringen Herstellungskosten. Aufgrund des langjährigen großflächigen Einsatzes sind in Afrika Chloroquin-Resistenzen inzwischen weit verbreitet (vgl. dazu Abb. 1 auf Seite 1). In Kenia wurde 1999 der Einsatz von **1** aufgrund dieser Problematik sogar verboten.^[65]

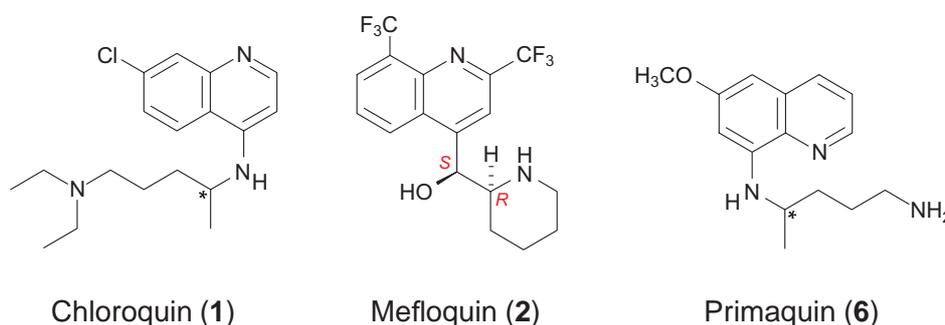


Abb. 8. Die Wirkstoffe einiger gängiger Malaria-Medikamente: Chloroquin (**1**) (Resochin[®]/Nivaquine[®]), Mefloquin (**2**) (Lariam[®], Mephaquin[®]) und Primaquin[®] (**6**); die Verbindungen **1** und **6** werden therapeutisch als Racemat eingesetzt

Mefloquin (**2**) und Primaquin (**6**) sind weitere Beispiele für Chinin-ähnliche Wirkstoffe. Für beide werden annähernd gleiche Wirkmechanismen diskutiert, auch wenn es im Detail Abweichungen gibt.^[14] Allerdings nimmt auch hier die Verbreitung von Resistenzen rasch zu (vgl. dazu Abb. 1 auf Seite 1).

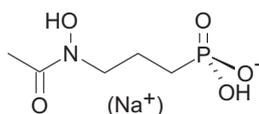
3.1.4 Neue Entwicklungen

Gerade in der jüngsten Vergangenheit gibt es wieder neue vielversprechende Ansätze für alternative Therapiemöglichkeiten.

Ein aktuelles Beispiel ist der strukturell sehr einfache Wirkstoff **7** (s. Abb. 9), der bereits 1999 von Jomaa^[66] beschrieben wurde. Seine Aktivität basiert auf der Blockie-

^a Chloroquin wurde zunächst als zu giftig für den Menschen befunden und deshalb lange zurückgehalten. Erst als das Patent nach dem zweiten Weltkrieg an amerikanische Firmen überging, fand unter Beteiligung der Weltgesundheitsorganisation (WHO) eine breite Vermarktung statt.^[64]

zung eines vom Menschen nicht genutzten, alternativen^a Isoprenoid-Stoffwechsels. Dieser Ansatz bietet allgemein immense Möglichkeiten bei der Entwicklung neuer Antiinfektiva.^[67] Es ist geplant,^[68] diese Strategie auf andere Krankheiten wie beispielsweise Tuberkulose, bestimmte Allergien und sogar Multiple Sklerose auszuweiten.



7

Abb. 9. Natrium-3-(N-acetyl-N-hydroxyamino)-propylphosphonat (7)

Eine andere Tatsache bietet ebenfalls große Hoffnung auf entscheidenden Fortschritt im Kampf gegen die Malaria: Parallel zum Abschluß dieser Arbeit erschien eine Reihe von Veröffentlichungen^[69,70,71,72] anlässlich der Sequenzierung der Genome von *P. falciparum*^[73] und des wichtigsten Überträgers, *Anopheles gambiae*.^[74] Dieser Schritt wird in Zukunft einen verbesserten Einblick in die involvierten physiologischen Prozesse erlauben und die Identifizierung neuer Angriffspunkte für Medikamente ermöglichen.^[75]

3.2 QSAR-Untersuchungen an Naphthylisochinolin-Alkaloiden

3.2.1 Biologische Tests auf Antimalaria-Aktivität

In Zusammenarbeit mit der Arbeitsgruppe François (Instituut voor Tropische Geneeskunde, Antwerpen), konnte ein Datensatz von etwa 50 antimalaria-aktiven Naphthylisochinolin-Alkaloiden aufgebaut werden.^[76] Getestet wurde die Wirkung der Substanzen auf infizierte Blutkulturen, wobei die IC₅₀-Werte über die Einbaurate einer radioaktiv markierten Purinbase in die DNA der Erreger bestimmt wurde.^b

^a Meint hier: Alternativ zum vom Menschen genutzten und sehr gut bekannten Mevalonat-Weg.

^b Inzwischen können in Antwerpen keine Tests mehr vorgenommen werden. Die vorliegende Arbeit basiert auf Testwerten einer Kooperation mit Dr. R. Kaminsky bzw. Dr. R. Brun am Schweizerischen Tropeninstitut in Basel; siehe dazu Abschnitt 3.3)

Aufbauend auf diesen Datensatz wurden von Vitt^[24] und Schmitt^[25] in unserem Arbeitskreis erste 3D-QSAR-Untersuchungen durchgeführt.

3.2.2 Alignmentstrategien

Das Alignment basierte auf vorher durch Konformationsanalysen bestimmten Minimumstrukturen. Anhand dieser Konformere wurde versucht, die potentiell pharmacophoren Elemente (z. B. freie Hydroxyfunktionen) der Liganden per RMS-fit^a zu überlagern. Bereits hier zeigte sich, daß räumlich sehr ähnliche Strukturtypen gut überlagert werden können, Verbindungen mit unterschiedlicher Achsenkonfiguration hingegen bereiteten Probleme. Die räumlich etwas anders aufgebauten 7,6'-gekuppelten Verbindungen führten zu Schwierigkeiten, da zunächst kein schlüssiges Alignment gefunden werden konnte.^[24]

Von Schmitt^[25] wurde erstmals das Programm GASP^[27,28] zum automatischen Erstellen eines Alignments eingesetzt; weil dabei allerdings die Axialchiralität der Naphthylisochinolin-Alkaloide nicht berücksichtigt werden konnte, wurden lediglich einzelne Vorschläge daraus manuell umgesetzt, führten aber dennoch zu erkennbaren Verbesserungen der Resultate.

3.2.3 CoMFA-Modelle

Die Berechnung der CoMFA-Felder erfolgte jeweils unter Verwendung von Standardparametern, lediglich die *cut-off*-Werte (vgl. Abb. 4 auf Seite 17) wurden von üblicherweise 20 auf 5 kcal/mol herabgesetzt, da so ein erheblich besserer q^2 -Wert erzielt werden konnte. In Tabelle 2 auf der nächsten Seite sind die Ergebnisse der jeweils besten erhaltenen CoMFA-Modelle zusammengefaßt.

^a RMS: Root Mean Square; die Überlagerung von zwei Strukturen erfolgt hier durch definierte Atompaaire zwischen den Strukturen, wobei die Wurzel des mittleren Abstandsquadrates minimiert wird.

Tab. 2. Ergebnisse der von Vitt und Schmitt aufgestellten CoMFA-Modelle.

	Vitt ^[24]	Schmitt ^[25]
Verbindungen im Datensatz	32	44
davon im Modell berücksichtigt	16	34
q^2	0.610	0.760
k	5	6

Das von Vitt^[24] erstellte Modell wurde auf der Basis eines Trainingsdatensatzes von 32 Verbindungen erstellt, letztlich wurden für das Modell aber nur 16 Strukturen berücksichtigt, eine höhere Zahl führte schnell zu geringeren q^2 -Werten. Der Ausschluß von mehr als 50% der Verbindungen des Basisdatensatzes deutet darauf hin, daß mit diesem Modell – wenn überhaupt – nur in einem sehr begrenzten Rahmen Vorhersagen möglich sind.

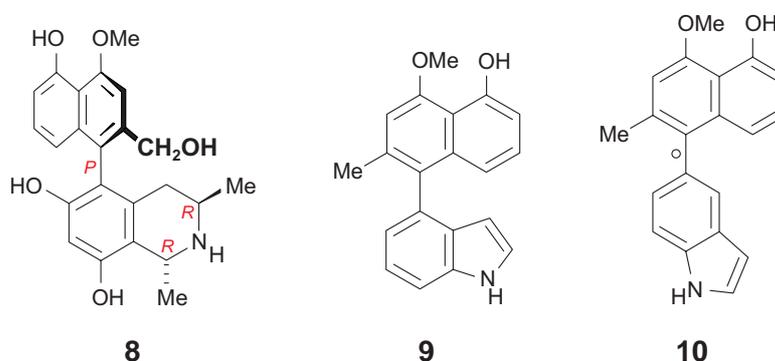
Den Arbeiten von Schmitt^[25] lag ein um einige Verbindungen erweiterter Datensatz zugrunde, und durch das verbesserte Alignment konnte ein optimiertes Modell erhalten werden. Aber auch hier mußten 10 Verbindungen (23%) ausgeschlossen werden, wofür im Einzelnen keine Begründung gegeben wurde.

3.2.4 Resultate der Pharmacophoranalysen und neue Strukturvorschläge

Auf der Basis der erstellten QSAR-Modelle wurden Rückschlüsse auf möglicherweise entscheidende pharmacophore Gruppen gezogen und auch neue Strukturen vorgeschlagen:^[25]

- Das Stickstoff-Atom in der einen Hälfte des Biaryls und eine freie OH-Gruppe (gegenüber der Achse) in der anderen sind für eine gute biologische Aktivität essentiell.
- Ein signifikanter Einfluß der Methylgruppen an den Positionen 1 und 3 im Isochinolinteil konnte nicht gefunden werden, somit zeigte auch die Konfiguration der Stereozentren keinen Einfluß.

- Ein modifiziertes Dioncophyllin C mit zwei zusätzlichen Hydroxyfunktionen zeigte im Vergleich zur Stammverbindung gesteigerte Aktivität (siehe Abb. 10).



IC ₅₀ [ng/ml]	8	9	10
berechnet	12	73	47
experimentell	-	415	348

Abb. 10. Auf Basis des CoMFA-Modells getroffene Strukturvorschläge (**8**, **9** und **10**)

- Die Substitution der Isochinolinhälfte von Dioncophyllin C durch Indol, Indolin oder Iso-Indolin führt nur zu geringer Abnahme der Aktivität.

Die Verbindungen **9** und **10** wurden inzwischen von Saeb^[77] synthetisiert, wobei die vorher berechneten sehr guten Aktivitäten experimentell jedoch nicht bestätigt werden konnten. Der Grund für die hohen Abweichungen der Vorhersagen liegt möglicherweise in der zu vereinfachten Darstellung der elektronischen Verhältnisse des Stickstoff-Atoms. Die in einer CoMFA-Analyse ausschließlich betrachteten sterischen und elektrostatischen Wechselwirkungen reichen hier offenbar nicht aus.

3.3 Der Datensatz

Wie bereits erwähnt entstanden die in dieser Arbeit verwendeten Testwerte in Kooperation mit Dr. Ronald Kaminsky und Dr. Reto Brun am Schweizerischen Tropeninstitut in Basel, da in der Arbeitsgruppe François (Instituut voor Tropische Geneeskunde, Antwerpen) keine Tests mehr durchgeführt werden konnten.

Die Antimalaria-Aktivitäten der untersuchten Strukturen wurden an zwei unterschiedlichen Zellstämmen untersucht. Einerseits war dies der 'NF54'-Stamm,^[78] ge-

gen den alle derzeit bekannten Medikamente Wirkung zeigen, und andererseits der Chloroquin-resistente 'K1'-Stamm.^[79]

In dieser Arbeit wurden die Testwerte gegen den K1-Stamm verwendet, da hier vollständigere Zahlen vorlagen. Zudem ist natürlich die Wirkung gegen Chloroquin-resistente Erreger viel bedeutender einzuschätzen.

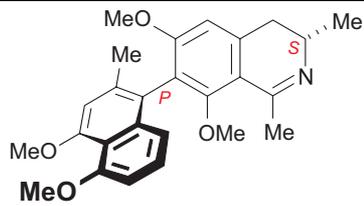
In diesem Abschnitt werden alle Verbindungen und ihre gemessenen Aktivitäten aufgelistet, wie sie für den Trainingsdatensatz verwendet wurden. Eine detaillierte Beschreibung der strukturellen Eigenschaften folgt im anschließenden Ergebnisteil (Kapitel 4.3 auf Seite 46). Für eine differenziertere Beschreibung der Alignments und der bei der Durchführung aufgetretenen Schwierigkeiten wurden die Strukturen in mehrere Gruppen aufgeteilt. Diese Vorgehensweise war schon vorher^[24,25] verfolgt worden.

Die IC_{50} -Werte sind in logarithmischen Einheiten angegeben (pIC_{50}), da diese auch bei der Berechnung herangezogen wurden. Sie variieren für Naphthylisochinolin-Alkaloide von 1.90 für Habropetalin A (**11**) (siehe Seite 35) und -1.05 für Ancistrobertsonin A (**12**) (Seite 36); positive Werte zeigen also gute, negative Werte schlechte Aktivitäten an. Da IC_{50} -Werte in ng/ml oft geläufiger sind, werden diese in Klammer zusätzlich angeführt. Jeweils angegebene Literatur bezieht sich auf die Isolierung oder die Synthese der Substanzen und nicht auf die Antimalaria-Aktivitäten.

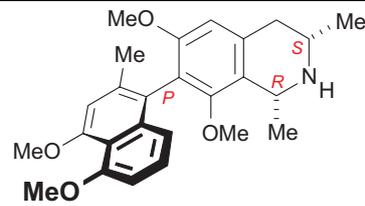
Die Tabellen enthalten auch die später ausgeschlossenen Outlier, die mit einem entsprechenden Hinweis versehen wurde. Es wurden folgende Abkürzungen verwendet: Tos = p-Toluolsulfonyl; Bn = Benzyl; ein \circ an der Biarylachse zeigt eine bei Raumtemperatur frei rotierbare Biarylachse.

Anmerkung des Verfassers: Die folgende Seiten (S. 34 – 42) können aus dem Dokument herausgeklappt werden, damit die Strukturen während der Diskussion der Outlier und vor allem der Interpretation der CoMSIA-Felder leichter zu finden sind.

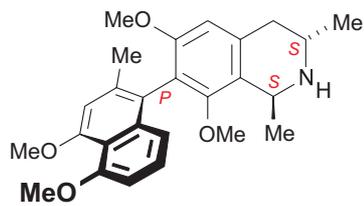
Tab. 3. Gruppen I und II : 7,1'- und 7,8' gekuppelte Biaryle.



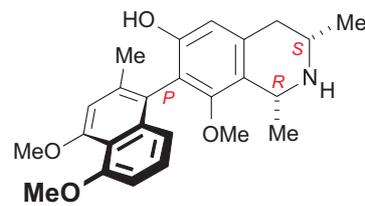
Ancistrocladisin^[80,81]
(13), -0.236 (723 ng/ml)



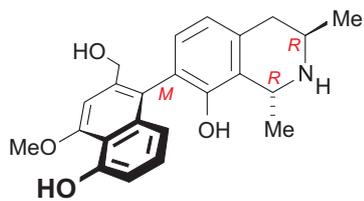
cis-1,2-Dihydroancistrocladisin^[82]
(14), -0.327 (894 ng/ml)



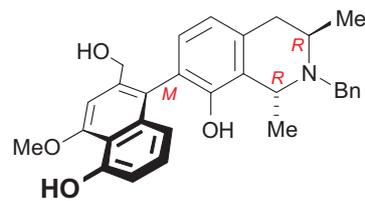
trans-1,2-Dihydroancistrocladisin^[82]
(15), -0.265 (776 ng/ml)



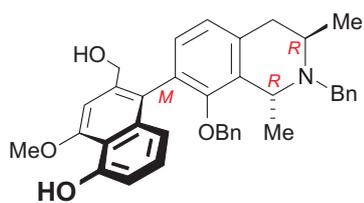
Ancistrobertsonin D^[82]
(16), -0.639 (1773 ng/ml)



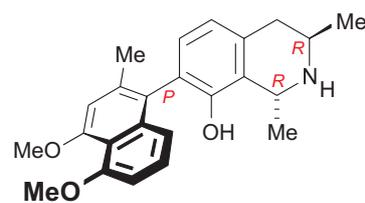
Dioncopeltin A^[6,83]
(3), 1.880 (5 ng/ml)



N-Benzyl Dioncopeltin A^[83]
(17), 0.299 (236 ng/ml)



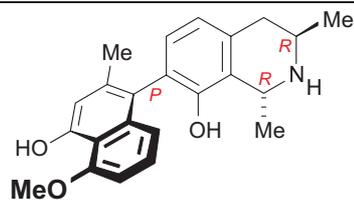
N-8-*O*-Dibenzyl Dioncopeltin A^[83]
(18), -0.001 (561 ng/ml)



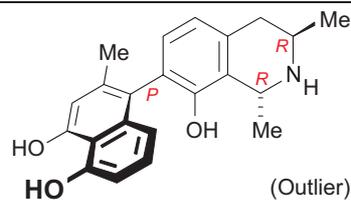
Dioncophyllin A^[84]
(19), 0.419 (144 ng/ml)

(Fortsetzung auf der nächsten Seite)

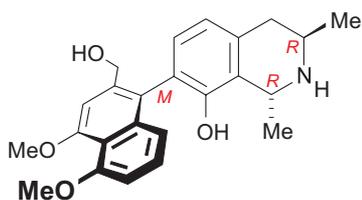
Tab. 3. (Fortsetzung)



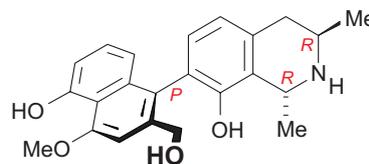
5'-*O*-Demethyldioncophyllin A^[85,86]
(20), 0.381 (151 ng/ml)



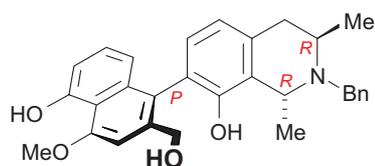
4'-*O*-5'-*O*-Demethyldioncophyllin A^[77]
(21), -0.343 (770 ng/ml)



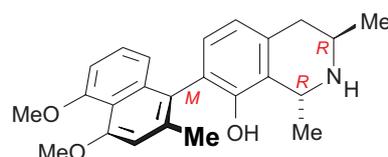
Habropetalin A^[15]
(11), 1.896 (5 ng/ml)



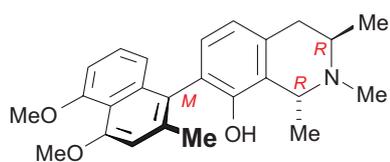
7-*epi*-Dioncopeltin A^[83]
(22), 0.538 (110 ng/ml)



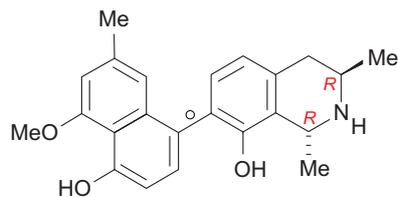
N-Benzyl-7-*epi*-dioncopeltin A^[83]
(23), 0.257 (260 ng/ml)



7-*epi*-Dioncophyllin A^[87]
(24), 0.516 (115 ng/ml)



N-Methyl-7-*epi*-dioncophyllin A^[88]
(25), -0.172 (582 ng/ml)



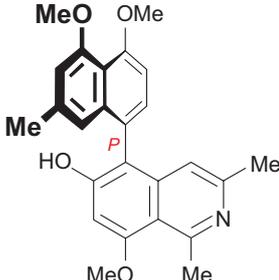
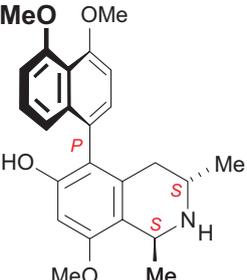
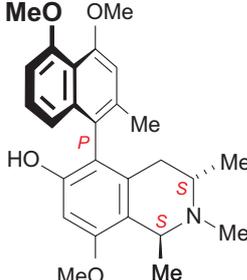
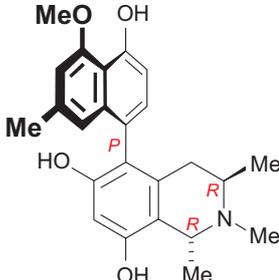
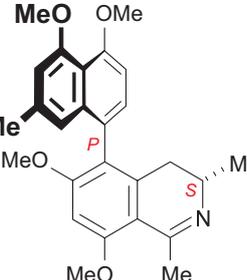
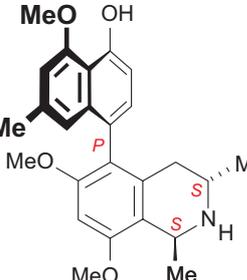
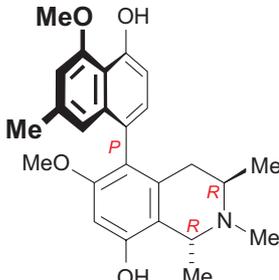
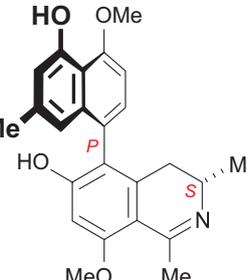
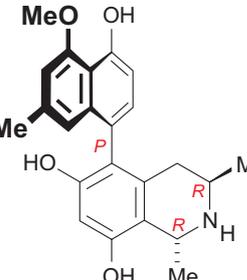
Dioncophyllin D^[89]
(26), 0.797 (58 ng/ml)

Tab. 4. Gruppen III und IV : 5,1'- und 5,8' gekuppelte Biaryle.

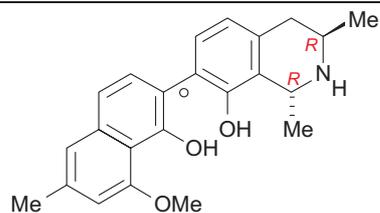
<p>Ancistrobertsonin B^[82] (27), -0.868 (3215 ng/ml)</p>	<p>Ancistrolidikokin B^[90] (28), 0.226 (234 ng/ml)</p>	<p>Ancistrolidikokin C^[90] (29), -0.356 (924 ng/ml)</p>
<p>Dioncophyllin C^[5,91] (4), 1.782 (6 ng/ml)</p>	<p>N-Methyldioncophyllin C^[76] (30), 0.453 (133 ng/ml)</p>	<p>Hamatein^[77] (31), -0.261 (736 ng/ml)</p>
<p>Korupensamin B^[4,92] (32), 0.936 (44 ng/ml)</p>	<p>Ancistrobertsonin A^[93] (12), -1.051 (4743 ng/ml)</p>	<p>Ancistrobertsonin C^[82] (33), -0.347 (969 ng/ml)</p>

(Fortsetzung auf der nächsten Seite)

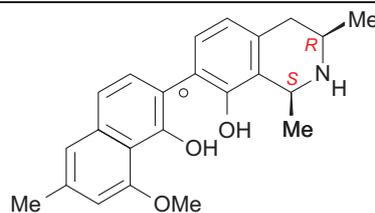
Tab. 4. (Fortsetzung)

 <p>Ancistrocladein^[94] (34), -0.277 (764 ng/ml)</p>	 <p>Ancistrocladin^[95] (35), -0.542 (1421 ng/ml)</p>	 <p>N-Methylancistrocladin (36), -0.329 (899 ng/ml)</p>
 <p>Ancistrocongolol A^[96] (37), 0.300 (197 ng/ml)</p>	 <p>Ancistroealain A^[7] (38), -0.401 (1056 ng/ml)</p>	 <p>Ancistroealain B^[7] (39), -0.102 (515 ng/ml)</p>
 <p>Ancistrolikokin A^[90] (40), 0.464 (140 ng/ml)</p>	 <p>Ancistrolikokin D (41), -0.305 (790 ng/ml)</p>	 <p>Korupensamin A^[4] (42), 0.738 (72 ng/ml)</p>

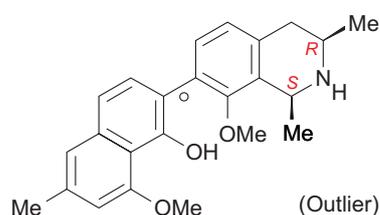
Tab. 5. Gruppe V: 7,6'-gekoppelte Biaryle.



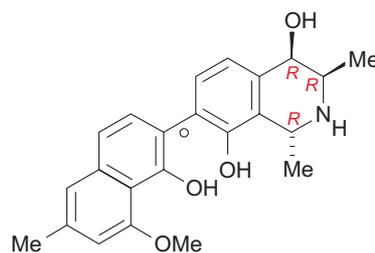
Dioncophyllin B^[97,89]
(43), 0.636 (84 ng/ml)



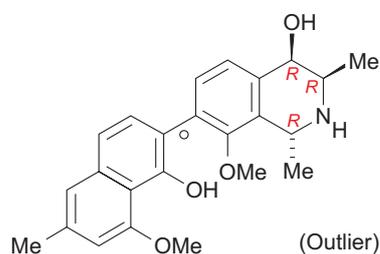
1-*epi*-Dioncophyllin B^[89]
(44), 0.370 (155 ng/ml)



8-*O*-Methyl-1-*epi*-dioncophyllin B^[89]
(45), 1.431 (14 ng/ml)

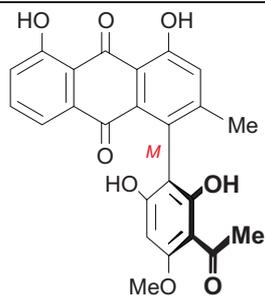


Dioncophyllinol B^[89]
(46), 1.074 (32 ng/ml)

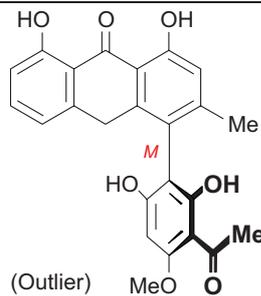


8-*O*-Methyldioncophyllinol B^[89]
(47), -0.221 (654 ng/ml)

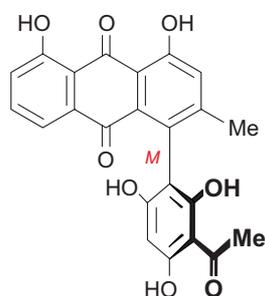
Tab. 6. Gruppe VI: Knipholon und Derivate.



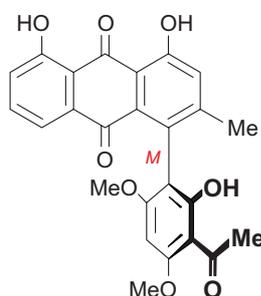
Knipholon^[98,99,100]
(48), -0.189 (672 ng/ml)



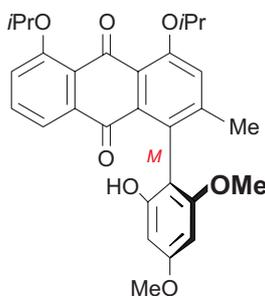
Knipholon-Anthron^[98,101]
(49), 0.450 (149 ng/ml)



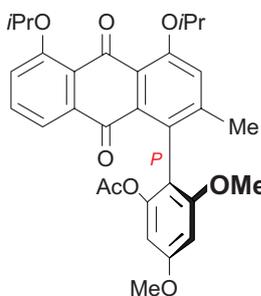
4-O-Demethylknipholon^[98,102]
(50), -0.071 (512 ng/ml)



Knipholon-6-methylether^[98,99,103]
(51), -0.374 (1060 ng/ml)

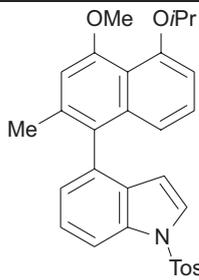
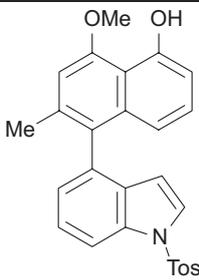
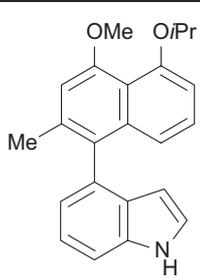
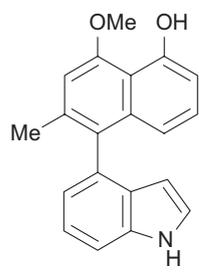
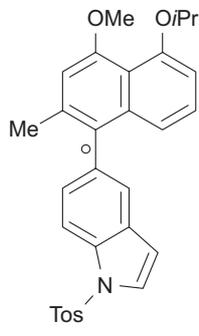
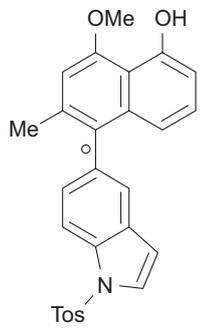
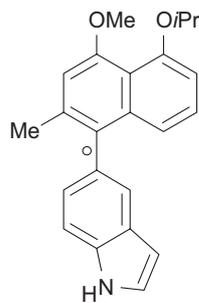
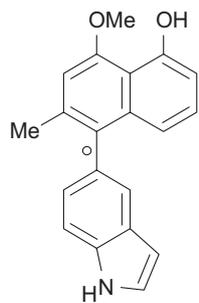


52^[104]
 -0.178 (739 ng/ml)

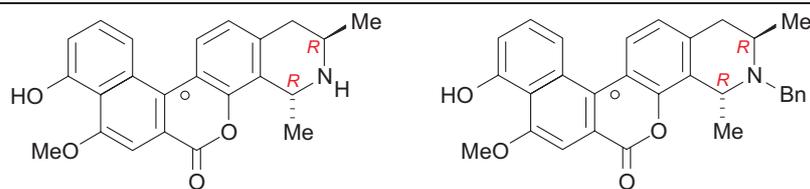


53^[104]
 -0.121 (704 ng/ml)

Tab. 7. Gruppe VI: 4,1'-gekoppelte Naphthylindole und Gruppe VII (5,1'-gekoppelte Naphthylindole).

 <p>54^[77]</p> <p>0.935 (58 ng/ml)</p>	 <p>55^[77]</p> <p>0.375 (193 ng/ml)</p>	 <p>56^[77]</p> <p>-0.080 (415 ng/ml)</p>
 <p>9^[77]</p> <p>-0.093 (376 ng/ml)</p>	 <p>57^[77]</p> <p>0.473 (168 ng/ml)</p>	 <p>58^[77]</p> <p>0.299 (230 ng/ml)</p>
 <p>59^[77]</p> <p>0.036 (318 ng/ml)</p>	 <p>10^[77]</p> <p>-0.060 (348 ng/ml)</p>	

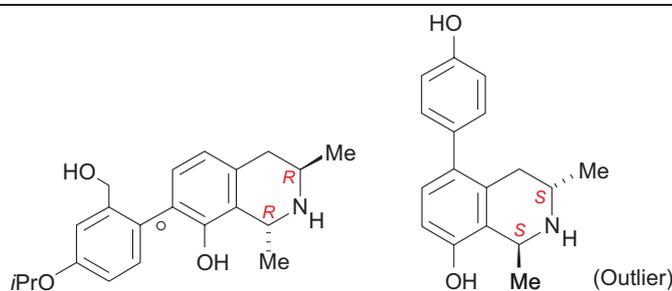
Tab. 8. Gruppe IX: Biaryl-Lactone.



Dioncolacton A^[6,83]
(60), -0.085 (457 ng/ml)

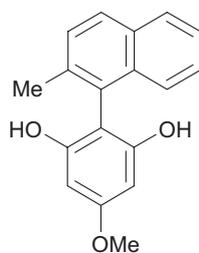
N-Benzyl Dioncolacton A^[83]
(61), -0.392 (1147 ng/ml)

Tab. 9. Einzelne Verbindungen.



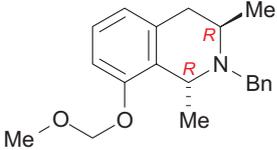
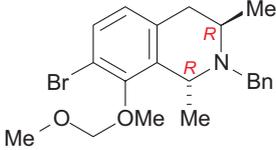
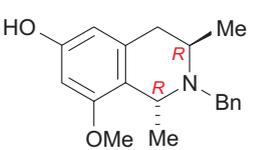
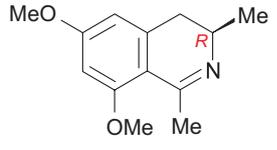
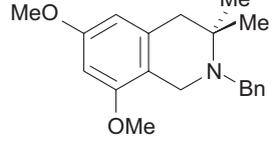
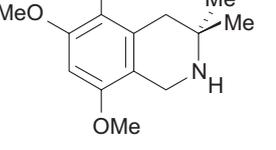
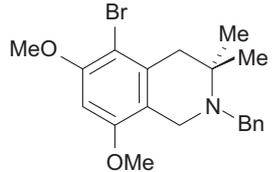
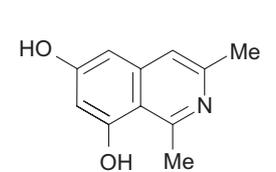
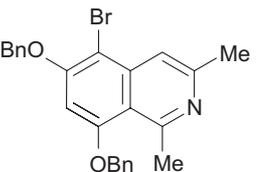
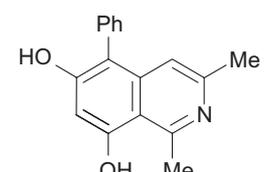
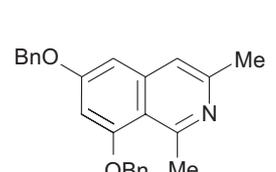
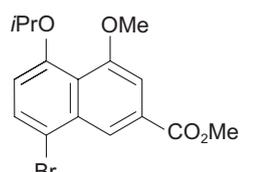
62^[105]
 1.454 (12 ng/ml)

63^[106]
 -0.973 (2531 ng/ml)
 (Outlier)



64^[104]
 -1.070 (3294 ng/ml)

Tab. 10. Gruppe X.

 <p>65^[107] 0.363 (135 ng/ml)</p>	 <p>66^[107] 0.047 (350 ng/ml)</p>	 <p>67^[108] 1.011 (29 ng/ml)</p>
 <p>68^[109] -1.240 (4059 ng/ml)</p>	 <p>69^[109] -0.652 (1396 ng/ml)</p>	 <p>70^[109] -0.321 (628 ng/ml)</p>
 <p>71^[109] -0.036 (424 ng/ml)</p>	 <p>72^[110] -0.101 (239 ng/ml)</p>	 <p>73^[111] -0.647 (1991 ng/ml)</p>
 <p>74^[110] -0.857 (1910 ng/ml)</p>	 <p>75^[111] -0.617 (1528 ng/ml)</p>	 <p>76^[112] -0.103 (448 ng/ml)</p>

Kapitel 4

Ergebnisse

4.1 Allgemeine Vorbemerkungen

Ein Schwerpunkt dieser Arbeit lag auf der Etablierung eines Alignment-Konzeptes, das in der Lage ist, die Flexibilität der eingesetzten Strukturen möglichst umfassend und direkt während des Ausrichtungsprozesses zu berücksichtigen. Die gewählten Programme FLEXS und GASP versuchen (auf sehr unterschiedliche Weise), den relevanten Konformationsraum der Strukturen durch Modulation der einzelnen Torsionswinkel zu erfassen, wobei die jeweiligen Konformationsenergien in die Bewertungsfunktionen für die erhaltenen Lösungen mit einfließen. Eine vorangestellte vollständige Konformationsanalyse der untersuchten Strukturen war somit praktisch überflüssig.

4.1.1 Flexible Ringsysteme

Strukturelemente, die eine unabhängige Variation der Torsionswinkel nicht erlauben, stellen für FLEXS und GASP ein Problem dar. Aliphatische oder teilweise aliphatische Ringsysteme können mit den entsprechenden Algorithmen nicht behandelt werden.^a

FLEXS ruft in derartigen Fällen externe Programme wie SCA^[41] oder das in dieser Arbeit verwendete CORINA^[39,40] auf. Während des FLEXS-Prozesses wird dabei die Konnektivität der entsprechenden Atome an CORINA übergeben, das aus dieser Angabe adäquate Konformere der Substruktur generiert und diese an FLEXS

^a Aromatische Ringe werden als völlig starr angesehen.

zurückgibt. Dort werden wiederum mehrere Lösungen gemäß ihrer energetischen Lage verwertet.

Das Programm GASP sieht eine solche Lösung nicht vor, sondern besitzt einen eigenen Algorithmus, der sich aber für die hier behandelte Problemstellung als nicht brauchbar erwiesen hat^a und deshalb nicht verwendet wurde.

4.1.2 GASP-Alignments als Grundlage für CoMSIA-Analysen

Die Verwendung von GASP zur Erzeugung des Alignments für eine CoMSIA-Analyse ist nicht unproblematisch. Der eigentliche Zweck von GASP ist es, eine Pharmakophor-Hypothese für wirksame Liganden mit unbekanntem biologischem Target zu bestätigen oder zu widerlegen. Der Algorithmus versucht, pharmacophore Elemente verschiedener Liganden räumlich in Einklang zu bringen und auf diesem Weg Rückschlüsse auf einen möglichen (gemeinsamen) Rezeptor und dessen Konformation zu schließen.

Dies bedeutet nicht notwendigerweise, daß GASP-Alignments für CoMSIA-Untersuchungen geeignet sind.^b Die CoMSIA-Methode korreliert strukturelle Unterschiede der Liganden mit einer biologischen Eigenschaft. Deshalb ist es notwendig, daß gemeinsame Strukturelemente sehr genau ausgerichtet werden, damit sie sich gewissermaßen gegenseitig aufheben. GASP hingegen versucht pharmacophore Gruppen in Einklang zu bringen, ohne dabei auf exakte strukturelle Gemeinsamkeiten Rücksicht zu nehmen.

Das Problem kann am Beispiel von Benzol demonstriert werden: GASP richtet aromatische Ringe lediglich anhand ihrer Schwerpunkte aus: Zwei um 30° gegeneinander verdrehte Benzolringe sind für GASP perfekt ausgerichtet, denn die Schwerpunkte liegen direkt übereinander (Abb. 11). Für CoMSIA hingegen 'sehen' beide Ringe keinesfalls gleich aus.

^a Damit die Torsionswinkel aliphatischer Ringe unabhängig voneinander betrachtet werden können, bietet GASP die Möglichkeit, eine Bindung des entsprechenden Rings zu spalten und die resultierenden offenen Ketten zu alignen. Für eine CoMSIA-Analyse ist dieses Verfahren aber unbrauchbar.

^b Dennoch bewirbt die Firma Tripos Inc. das Programm GASP durchaus als Grundlage für derartige Untersuchungen.^[113]

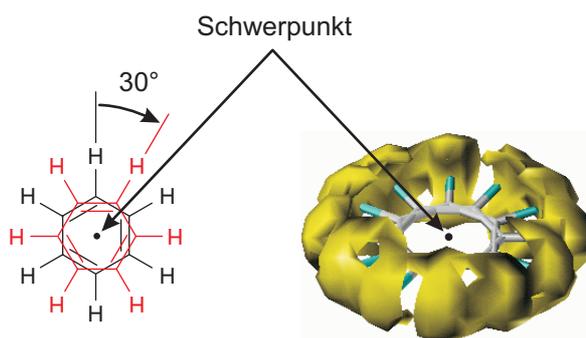


Abb. 11. Zwei überlagerte, um 30° gegeneinander verdrehte Benzolringe; gelbe Konturen kennzeichnen Bereiche, die jeweils von nur einem der beiden Ringe belegt werden, entsprechend dem Unterschied in den sterischen Feldern beider Ringe, den die CoMSIA-Methode 'sieht'

Für Naphthylisochinolin-Alkaloide könnte dies bedeuten, daß möglicherweise die aromatischen Substrukturen, die eine Art 'Spacer' zwischen den verschiedenen Sauerstoffatomen und dem Stickstoff bilden, möglicherweise nur schlecht ausgerichtet werden. Dies ist aber, aufgrund der unterschiedlichen Kupplungstypen ohnehin nur bedingt möglich, so daß GASP durchaus geeignet sein sollte, ein brauchbares Alignment zu erstellen.

4.1.3 Behandlung unterschiedlicher Kupplungspositionen

Die Schwierigkeiten beim Alignment liegen vor allem begründet in den unterschiedlichen räumlichen Anordnungen der beiden Biarylhälften zueinander. Man könnte dieses Problem leicht umgehen, indem man die verschiedenen Kupplungstypen nebeneinander behandelt und für jeweils kleinere Datensätze eigene Modelle erstellt. Auf diese Weise würden allerdings teilweise sehr kleine Gruppen entstehen, die nur eine sehr geringe Varianz ihrer Aktivitäten aufwiesen.^a Zudem wären derartig eingeschränkte Modelle kaum in der Lage, Aufschlüsse über Form und Eigenschaften eines möglichen Rezeptors zu geben, da sie jeweils nur einen sehr begrenzten Strukturraum betrachten würden.

^a Bezogen auf den kompletten Datensatz beträgt die Aktivitätsspanne zwischen aktivster und inaktivster Verbindung 2.95 logarithmische Einheiten, was gerade ausreicht,^[26] brauchbare QSAR-Modelle zu erstellen.

4.2 Eingabe der 3D-Strukturen

Alle Strukturen wurden zunächst mit Hilfe des SKETCH-Modules in SYBYL eingegeben und anschließend unter Verwendung des Tripos-Kraftfeldes minimiert. Die für die spätere Berechnung der CoMSIA-Felder nötigen Ladungen wurden mit Hilfe des Gasteiger-Hückel-Algorithmus^[44] berechnet.

4.3 Das Alignment

Die Programme FLEXS und GASP ermöglichen eine weitgehende Automatisierung des Alignment-Prozesses. Der folgende Teil beschreibt die Bereitstellung und die visuellen Eigenschaften der erzeugten Alignments, zunächst ohne eine vergleichende statistische Bewertung. Diese folgt nach der Diskussion der Outlier in Abschnitt 4.4.4.1 (Seite 71).

4.3.1 FLEXS

4.3.1.1 Vorbemerkungen

Grundlage des FLEXS-Alignments ist die Wahl eines Referenz-Liganden. Da dieser während des gesamten Alignments starr bleibt und alle anderen Substanzen im Datensatz an ihm ausgerichtet werden, fungiert er als eine Art Grundgerüst für das komplette Alignment.

Das für die räumliche Form entscheidende Strukturelement von Naphthylisochinolin-Alkaloiden ist die trotz (meist) konfigurativer Stabilität sehr flexible Biarylachse. In einem solchen Fall können in ihrer Beweglichkeit eingeschränkte Verbindungen Hinweise auf die optimale Konformation geben ('active-analogue approach'^[114]). Solche Beispiele liegen momentan aber nicht vor.

Da es derzeit keine strukturellen Informationen über den Wirkort gibt, wurde die zu Beginn dieser Arbeit aktivste Verbindung Dioncopeltin A (**3**) (Seite 41) in ihrer energetisch günstigsten Konformation als Referenz-Ligand verwendet.

4.3.1.2 Durchführung

FLEXS (in der verwendeten Version 1.8^[115]) ist ein Kommandozeilen-Programm, das mit Hilfe von Batch-Skripten automatisiert werden kann. Auf diese Weise konnte das eigentliche Alignment fast vollständig automatisiert werden, das verwendete Script ist im Anhang (Seite 137) abgedruckt. Der Ablauf entspricht im Wesentlichen dem von den Autoren des Programms empfohlenen Verfahren. Hinzugefügt wurde lediglich der Befehl 'OPTIMIZE', wodurch die Überlappungsvolumina der Gauß-Funktionen zusätzlich optimiert wurden.^[30] Danach wurde ein erneutes Sortieren der erhaltenen Alignments ('SORTE E_TOTAL') nötig. Anschließend wurde der FLEXS-interne 'CLUSTER'-Algorithmus verwendet, der die Anzahl der Alignments pro Struktur reduziert, indem er sehr ähnliche Konformere gruppiert und jeweils nur das günstigste einer solchen Serie ausgibt. Diese Reduktion erleichterte den Überblick über die erhaltenen Resultate.

Methodisch bedingt ist der FLEXS-Algorithmus, der die Strukturen zunächst fragmentiert und anschließend wieder zusammensetzt (siehe Schema 2 auf Seite 12), nicht in der Lage, axialchirale Verbindungen korrekt zu behandeln. Daß dies aber unbedingt notwendig war, zeigten beispielsweise die Aktivitäten von Dioncopeltin A (**3**) (Seite 34) und 7-*epi*-Dioncopeltin A (**22**) (Seite 35), die um über eine Zehnerpotenz voneinander abweichen, obwohl die Verbindungen sich lediglich in der Achsenkonfiguration unterscheiden.

Da die Berechnung von axialchiralen Verbindungen mit FLEXS also zur Ausgabe von zunächst beiden Atropdiastereomeren führte, wurden die Konformere mit falscher Achsenkonfiguration von Hand aussortiert und verworfen.

Auf diese Weise wurden etwa 10-50 (korrekt konfigurierte) Alignment-Möglichkeiten^a pro jeweiligem Test-Ligand erhalten.

4.3.1.3 Diskussion der Alignments

Gruppen I und II. Da der Referenz-Ligand aus Gruppe I stammt, sind die Alignments dieser Gruppe trivial; die Grundgerüste der Verbindungen sind identisch und

^a Im Einzelfall waren es bis zu mehreren Hundert, wobei durch zusätzliches Clustern weitere Einschränkungen möglich waren.

werden entsprechend übereinander gelegt. Liegt eine andere Achsenkonfiguration vor (Gruppe II), ist diese Bedingung zwar nicht mehr gegeben, trotzdem können auch diese annähernd gut aneinander ausgerichtet werden, wie Abb. 12 zeigt.

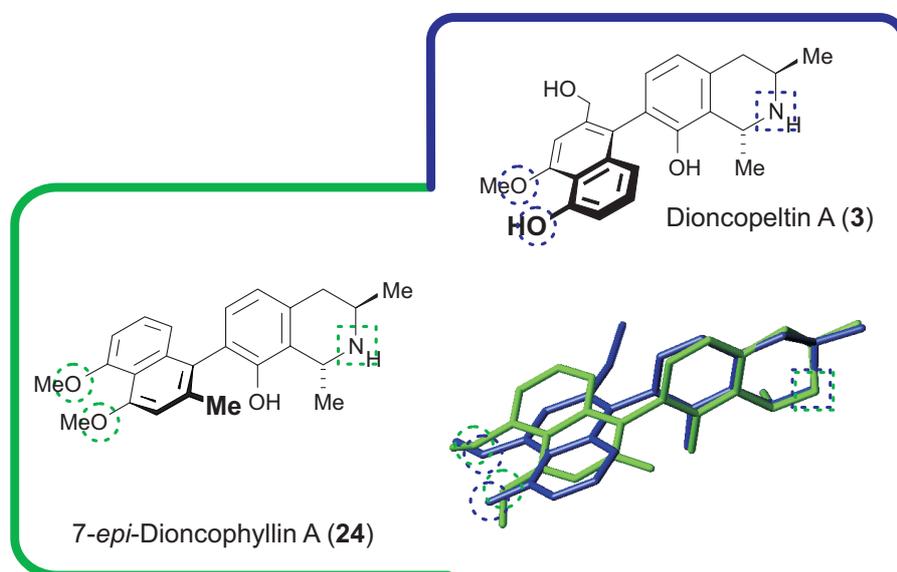


Abb. 12. Das Alignment der Verbindungen aus Gruppe II, am Beispiel von 7-*epi*-Dioncophyllin A (24); die gepunkteten Kästchen markieren die Stickstoffatome, die Kreise zeigen die Sauerstoffe im Naphthalinteil (in der Farbe der jeweiligen Struktur; Wasserstoff-Atome werden aus Gründen der Übersichtlichkeit in den 3D-Strukturen nicht gezeigt)

Gruppen III und IV. Man findet für diese 5,1'- bzw. 5,8'-gekuppelten Verbindungen generell drei Möglichkeiten für ein Alignment, trotz der zunächst sehr unterschiedlich erscheinenden dreidimensionalen Formen. Abb. 13 auf der nächsten Seite zeigt diese drei Fälle zunächst für Dioncophyllin C (4), repräsentativ für Gruppe III. Die übrigen FLEXS-Vorschläge unterscheiden sich von diesen nur noch durch geringfügige Verschiebungen der Strukturen gegeneinander und durch etwas andere Torsionswinkel an den Hydroxy- und Methoxyfunktionen.

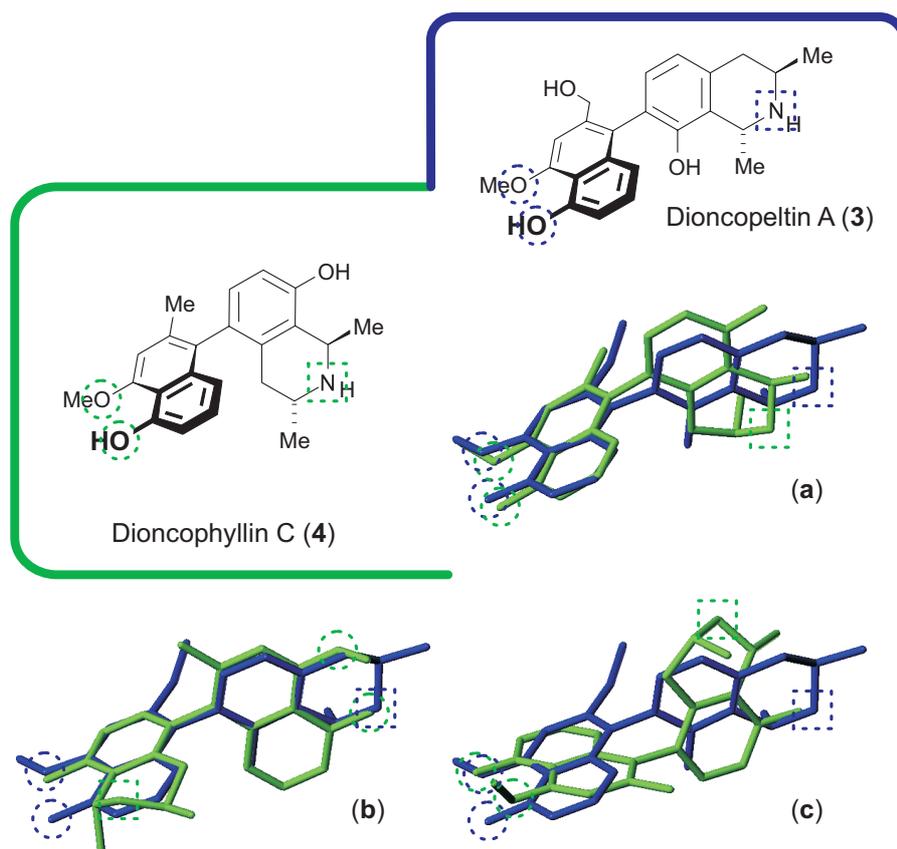


Abb. 13. Die drei möglichen Alignments für Gruppe III (a bis c) von Dioncophyllin C (4) (grün) mit Dioncopeltin A (3) (blau)

In Abb. 13 ist dargestellt, daß trotz der unterschiedlichen Kupplungstypen ein fast optimales Alignment der *N*-Funktion im Isochinolinteil und gleichzeitig der beiden Sauerstoff-Funktionen im Naphthylteil möglich ist (Fall a). Die Möglichkeiten b und c erlauben nur eine schlechtere Überlagerung der gekennzeichneten funktionellen Gruppen.

Dies gilt ebenso für Gruppe IV, wobei hier die Achsenkonfigurationen von Test- und Referenz-Ligand nicht übereinstimmen.^a Abb. 14 auf der nächsten Seite zeigt dies anhand des Alignments von Ancistrocongolin A (37) mit Dioncopeltin A, wobei die Lösungen zunächst visuell nicht so günstig wirken wie noch in Abb. 13.

^a 'Gleiche' Achsenkonfiguration bezieht sich hier auf die Lage der beiden Biarylhälften im Raum; unterschiedliche Substitutionsmuster bedingen, daß in diesem Sinne 'gleich' nicht immer auch den gleichen Achsendeskriptor (*M/P*) impliziert.

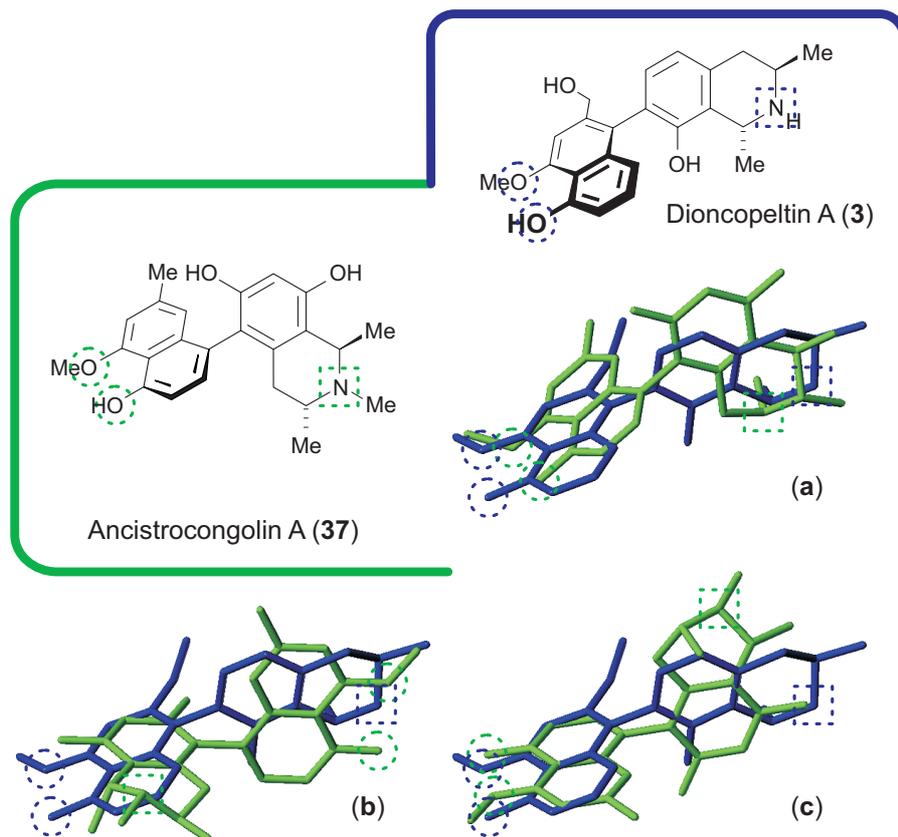


Abb. 14. Die drei FLEXS-Alignments (a bis c) von Ancistrocogolin A (37) (grün) mit Dioncopeltin A (3) (blau)

Allgemein kann man die gezeigten drei verschiedenen Möglichkeiten wie folgt beschreiben:

- **a:** Die Stickstoffatome der Isochinolinteile sowie die Sauerstoff-Funktionen in den Naphthylteilen sind optimal ausgerichtet.
- **b:** Die Naphthylteile der Test-Liganden werden auf den Isochinolinteil des Referenz-Liganden gelegt (und umgekehrt).
- **c:** Die Test-Liganden werden (verglichen mit **a**) um 180 Grad gedreht. Während beide Sauerstoff-Funktionen in den Naphthylteilen ausgerichtet bleiben, ist dies für die Stickstoff-Atome nicht mehr möglich.

Die von FLEXS berechnete Prioritätsreihenfolge ergab jedoch für die Strukturen der Gruppen III und IV keine generell bevorzugte Möglichkeit. Für das FLEXS-

Alignment wurden zunächst die jeweils vom Programm am besten eingeschätzten Konformere verwendet.

Gruppe V. 7,6'-gekuppelte Naphthylisochinolin-Alkaloide^a können aufgrund der räumlichen Anordnung der beiden Hälften zueinander nicht optimal an **3** ausgerichtet werden (siehe Abb. 15). FLEXS legte in diesem Fall die Isochinolinhälften fast exakt übereinander und vernachlässigte den Naphthylteil.

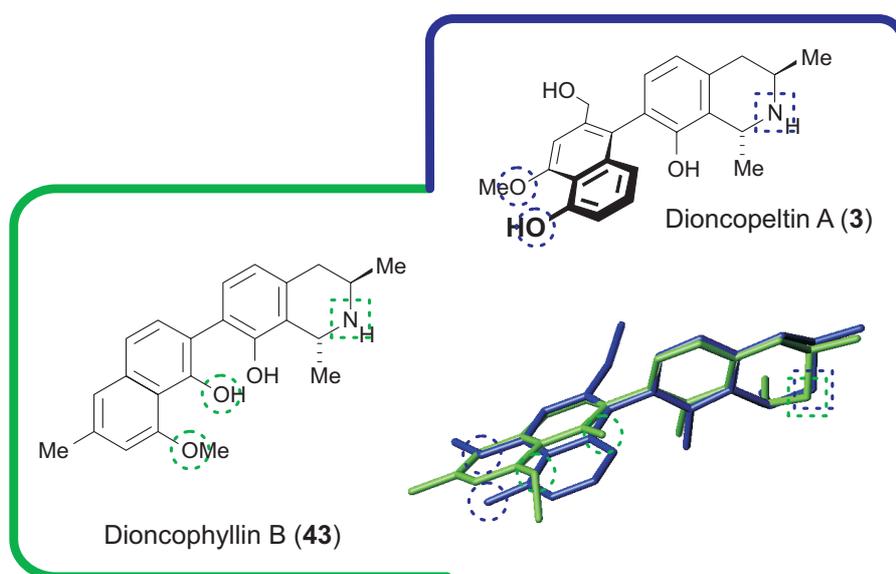


Abb. 15. Das Alignment der 7,6'-gekuppelten Naphthylisochinolin-Alkaloide (Gruppe V) am Beispiel von Dioncophyllin B (**43**)

Gruppe VI. Die Phenylanthrachinone nehmen im hier behandelten Datensatz strukturell offensichtlich eine besondere Rolle ein. Die Stammverbindung Knipholon (**48**, siehe Seite 39) wurde bereits 1984 aus den Wurzeln der Fackellilie *Kniphofia foliosa* (Asphodelaceae) isoliert,^[100] die Aufklärung der absoluten Stereostruktur gelang allerdings erst später, durch quantenchemische Berechnung der CD-Spektren in unserem Arbeitskreis.^[98] Im Rahmen dieser Untersuchungen wurde ebenfalls die antiplasmodiale Aktivität verschiedener Vertreter dieses Strukturtyps gefunden. Das Knipholon-Anthron (**49**, siehe Seite 39) fiel hier besonders auf, zeigte aber gleichzeitig eine nicht unerhebliche Cytotoxizität.^[117]

^a Verbindungen des ebenfalls bekannten 7,3'-Kupplungstypes^[116] sind diesen sehr ähnlich, allerdings im Trainingsdatensatz bisher nicht enthalten.

Durch die strukturellen Unterschiede erschien es zunächst fraglich, ob diese Substanzen in das Modell aufgenommen werden konnten. Dennoch erschien der Versuch lohnend, da man auf diese Weise, auch für den Fall, daß die Aktivitäten nicht korrelieren, zumindest Indizien für die Ähnlichkeit des Wirkmechanismus erhalten könnte.

Das Alignment (siehe Abb. 16) stellte sich als relativ unkompliziert heraus, denn FLEXS lieferte für alle Derivate eindeutige Plazierungen: Die Anthrachinonhälfte von **48** wurde auf den Naphthylteil von **3** plaziert, wodurch die Phenylhälfte auf den aromatischen Teil der Isochinolinhälfte von **3** gesetzt wurde. In einem solchen Alignment konnten neben zwei aromatischen Ringen auch drei Sauerstoff-Funktionalitäten gut überlagert werden.

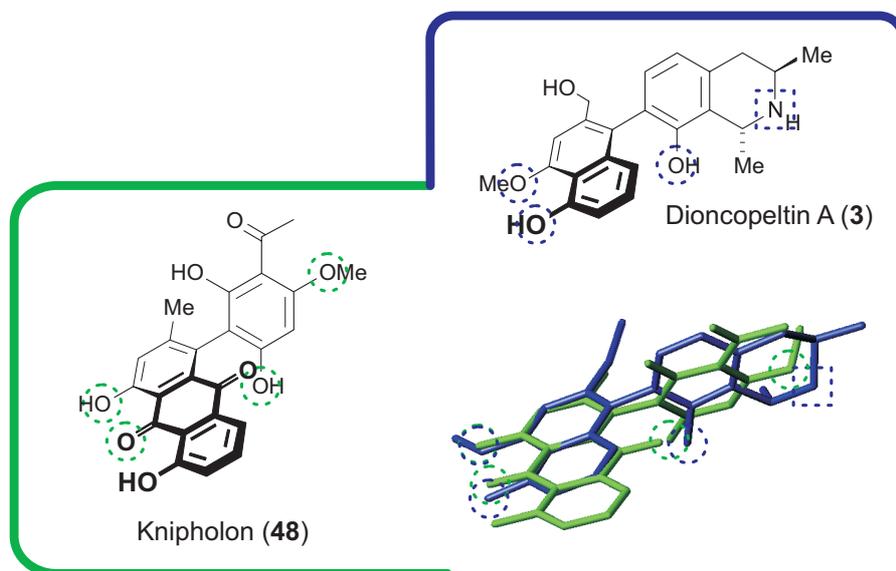


Abb. 16. Das Alignment der Phenylanthrachinone (Gruppe VI): Knipholon (**48**) (grün) ausgerichtet an Dioncopeltin A (**3**) (blau)

Gruppen VII und VIII. Diese Naphthylindole (siehe S. 40) wurden von Saeb^[77] synthetisiert, der Strukturvorschlag geht zurück auf Ergebnisse früherer CoMFA-Modelle^[24,25] (vgl. Abschnitt 3.2.4). Gerade weil die dort vorhergesagten Aktivitäten experimentell nicht bestätigt werden konnten, stellte diese Substanzklasse eine interessante Bereicherung des Datensatzes dar. Für die in 4-Position gekuppelten Verbindungen dieses Typs wurde konfigurative Labilität angenommen, obwohl dies bisher experimentell nicht nachgewiesen wurde.

Das Alignment erfolgte wiederum eindeutig; die potentiellen pharmacophoren Elemente konnten für beide Kupplungstypen sehr gut an **3** ausgerichtet werden (siehe Abb. 17).

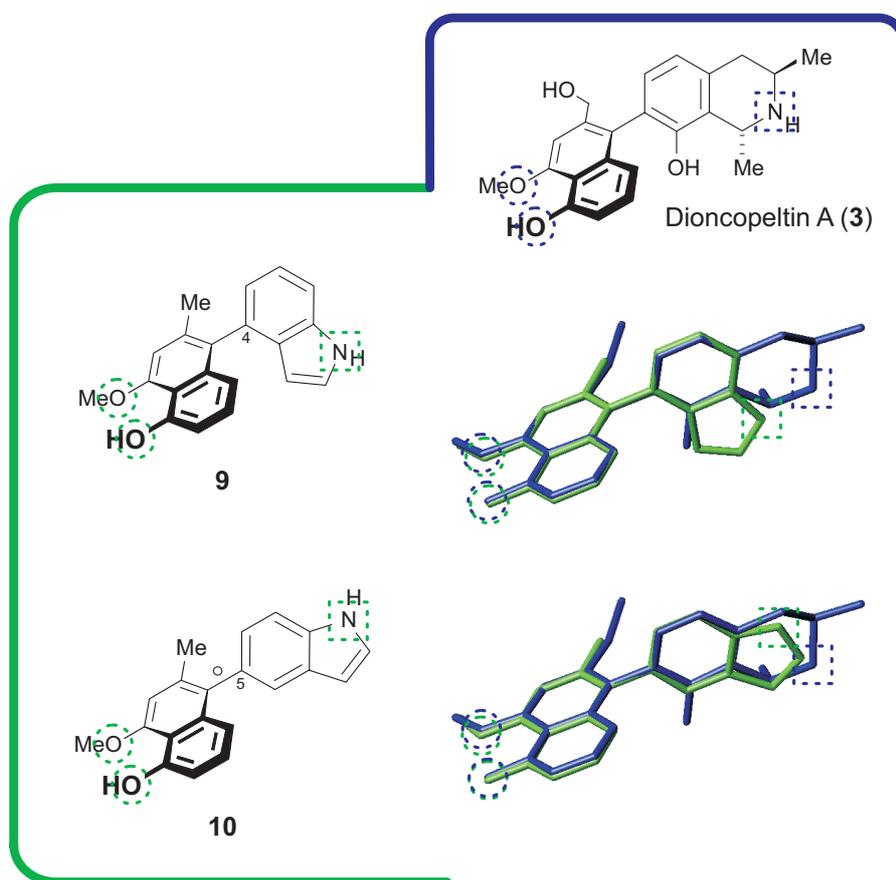


Abb. 17. Die Alignments der 4,1' und 5,1' gekuppelten Naphthylindole (Gruppen VII und VIII) anhand von **9** und **10** (jeweils grün), ausgerichtet an **3** (blau)

Gruppe IX. Durch die Lactonbrücke ist die Beweglichkeit dieser Strukturen deutlich anders geartet als bei **3**; Dioncolacton A (**60**) und *N*-Benzyldioncolacton A (**61**) (siehe Seite 41) sind nicht axialchiral.^a Eine optimale Überlappung der beiden Biarylhälften ist aufgrund der gewählten Konformation von **3** nicht möglich, FLEXS bevorzugt eine optimale Überlagerung der Naphthylteile auf Kosten der Isochinolinhälften (Abb. 18).

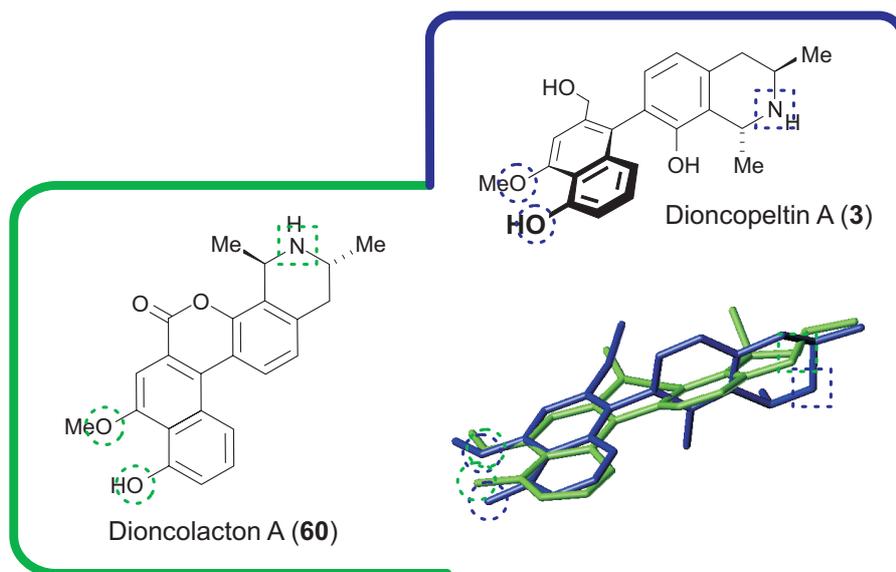


Abb. 18. Die Ausrichtung der Biaryllactone (Gruppe IX) am Beispiel von Dioncolacton A (**60**)

^a Die konfigurative Labilität derartiger Lactone bei Raumtemperatur ist bekannt und konnte sowohl experimentell als auch durch quantenchemischen Untersuchungen^[118,24] nachgewiesen werden.

Einzelne Verbindungen. Der Datensatz enthielt drei weitere Verbindungen mit Biarylachse, die keiner der gezeigten Gruppen zugeordnet werden konnten: Bedingt durch die Kupplungspositionen ähnelt aber **62** sehr stark den Gruppen I und II (siehe Abb. 19), **63** entsprechend den Gruppen III und IV. Die Plazierungen erinnern demzufolge sehr an die dort gefundenen.

Verbindung **64** wurde, entsprechend den aromatischen Ringen, ebenfalls plausibel platziert.

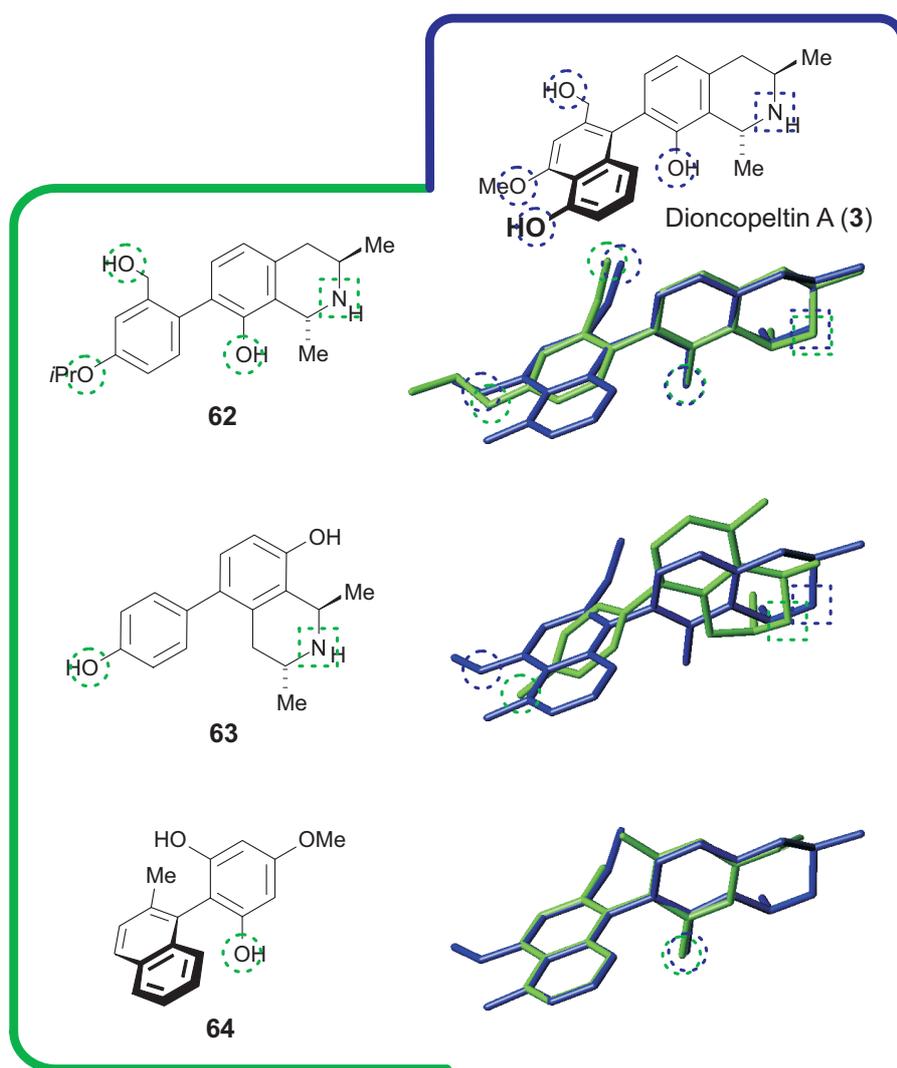


Abb. 19. Das Alignment von **62**, **63** und **64**, jeweils mit Dioncopeltin A (**3**) als Referenz

Gruppe X. Auch kleinere Strukturen (bezogen auf den Referenz-Liganden) konnten mit FLEXS ausgerichtet werden (siehe Abb. 20 auf der nächsten Seite). Dabei

ergaben sich jedoch, bedingt durch die Größenunterschiede, viele Platzierungsmöglichkeiten, die zudem zwischen unterschiedlichen Liganden stark voneinander abwichen. Da die Bewertungsfunktion von FLEXS hier also keine einheitlichen Resultate lieferte, war nur sehr schwer zu entscheiden, welche der Lösungen gewählt werden sollte.

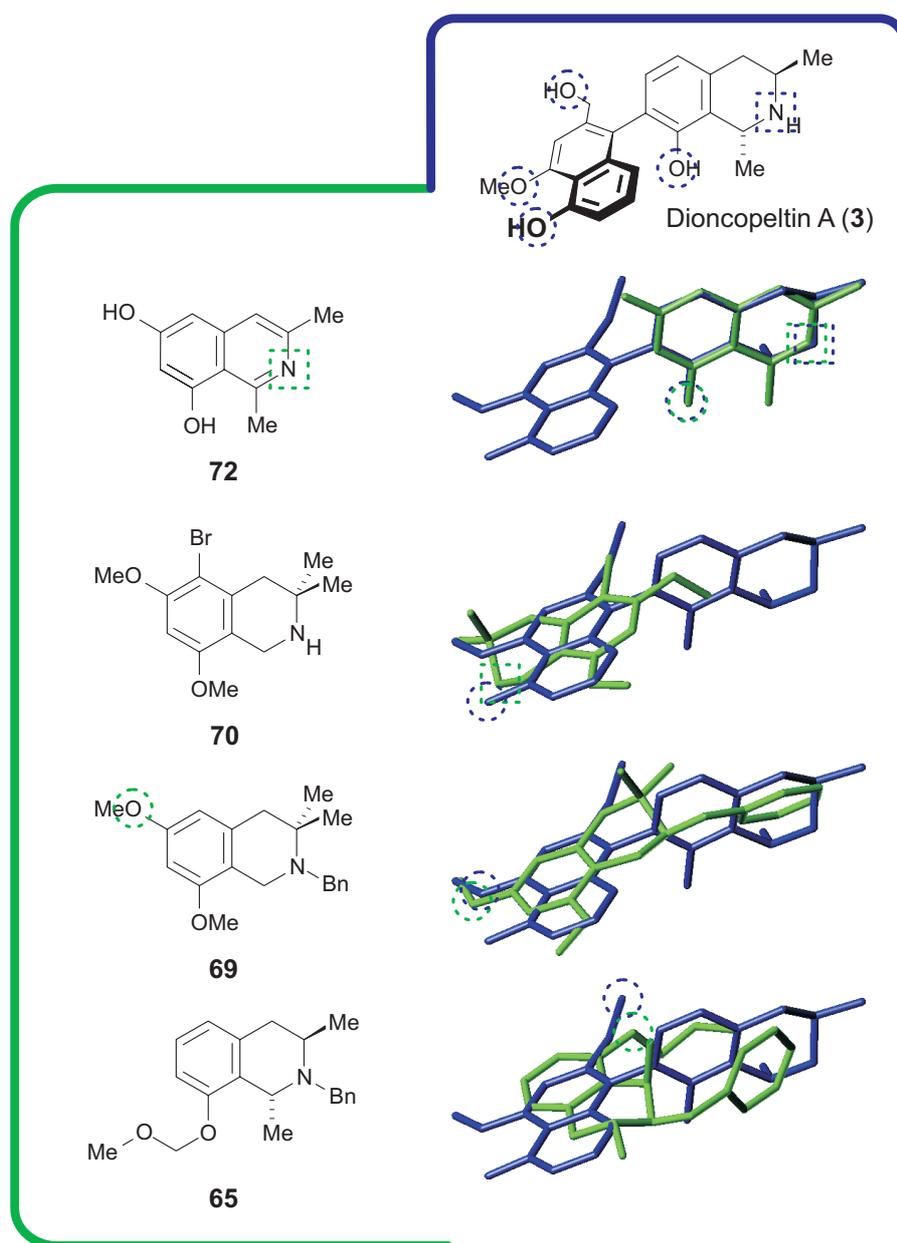


Abb. 20. Beispielhafte Alignments von Verbindungen aus Gruppe X

Die 'Monomeren Molekülhälften' wurden teilweise auf entsprechende Partialstrukturen des Referenz-Liganden gesetzt (z. B. 72 in Abb. 20), teilweise aber auch in der

anderen Biarylhälfte (**70**). Die Alignments von benzylsubstituierten Verbindungen wirken zum Teil sehr gut, wie z. B. für **69**, in anderen Fällen aber führt dies auch zu offensichtlich sehr schlechten Lösungen, wie bei **65**.

Es erschien zwar schon an dieser Stelle zweifelhaft, ob derartig ausgerichtete Strukturen einen positiven Beitrag zu einem QSAR-Modell (mit dem Ziel, Naphthylisochinolin-Alkaloide zu beschreiben) liefern können, dennoch wurde zunächst versucht (Abschnitt 4.4.2), diese Verbindungen mit zu integrieren.

4.3.2 GASP

4.3.2.1 Vorbemerkungen

Ein prinzipieller Vorteil des GASP-Ansatzes gegenüber FLEXS ist die Möglichkeit, mehr als zwei Verbindungen gleichzeitig zu alignen. Mit der Anzahl der rotierbaren Bindungen steigen allerdings die Hardwareanforderungen exponentiell an, so daß es nicht möglich ist, den gesamten Datensatz von insgesamt über 60 Verbindungen in einem einzigen GASP-Lauf zu berechnen.^a

Im vorhergehenden Abschnitt konnte gezeigt werden, daß Verbindungen gleichen Kupplungstyps, unabhängig von der Achsenkonfiguration, problemlos aneinander ausgerichtet werden können. Der Schlüssel zum Erfolg liegt offensichtlich an der Art der Ausrichtung von Verbindungen mit unterschiedlichen Kupplungspositionen, also hauptsächlich^b zwischen Verbindungen mit den Kupplungsposition 5 oder 7 im Isochinolinteil. Die Art der Kupplung im Naphthylteil und damit die Position der Methylgruppe ist für die Form des Moleküls nur von untergeordneter Bedeutung.

Deshalb wurde zunächst folgende Vorgehensweise gewählt: Die Konformationen zweier Vertreter des jeweiligen Kupplungstyps sollten iterativ mit GASP aneinander angepaßt werden. Die Wahrscheinlichkeit, daß ein solches Alignment der beiden aktivsten Vertreter, Dioncopeltin A (**3**) und Dioncophyllin C (**4**), die Verhältnisse

^a Momentan (GASP 2.0) sind laut Manual^[119] etwa 4–5 Verbindungen mit mittlerer Anzahl an Freiheitsgraden möglich. Einige Tests zeigten, daß schon bei 4 Naphthylisochinolin-Alkaloiden die Grenze erreicht ist.

^b Von den vergleichsweise wenigen 7,6' gekuppelten Verbindungen im Datensatz abgesehen.

am Wirkort sehr gut wiedergibt, sollte relativ hoch sein und damit eine sehr gute Basis für ein Alignment der übrigen Verbindungen darstellen.

4.3.2.2 Die Berücksichtigung der Axialchiralität mit GASP

GASP erlaubt (seit Version 2.0) die Definition von sogenannten Constraints, wodurch die freie Beweglichkeit an einzelnen Torsionswinkeln begrenzt und somit Axialchiralität berücksichtigt werden kann, ohne mühsames Aussortieren falsch konfigurierter Alignments, wie es bei FLEXS nötig war.^a Leider traten jedoch schon bei ersten Versuchen erhebliche Probleme mit dieser Funktion auf. Es stellte sich heraus, daß die erhaltenen Torsionswinkel von der Definition der Constraints abhängt und die erhaltenen Geometrien energetisch äußerst ungünstig waren.

Für eine genauere Untersuchung dieses Ergebnisses wurden für **3** und **4** die Potentiale, die sich bei der Rotation um die Biarylachse ergeben, semiempirisch^b berechnet. Dabei wurde nur der Bereich der korrekten Achsenkonfigurationen herangezogen, die erhaltenen Kurven sind in Abb. 21 auf der nächsten Seite dargestellt. Die Positionen der aus unterschiedlichen GASP-Läufen erhaltenen Konformere wurden gekennzeichnet, einerseits ohne Constraints berechnet, andererseits mit einer eingeschränkten Beweglichkeit, die den beiden tatsächlichen Achsenkonfigurationen entsprach.

^a Die Möglichkeit, Constraints zu definieren, war früher^[24,25] als entscheidendes Kriterium für den Einsatz von GASP definiert worden.

^b Die Berechnung erfolgte unter Verwendung von VAMP^[120] und der AM1-Methode. Die Winkel wurden in 2°-Schritten berechnet.

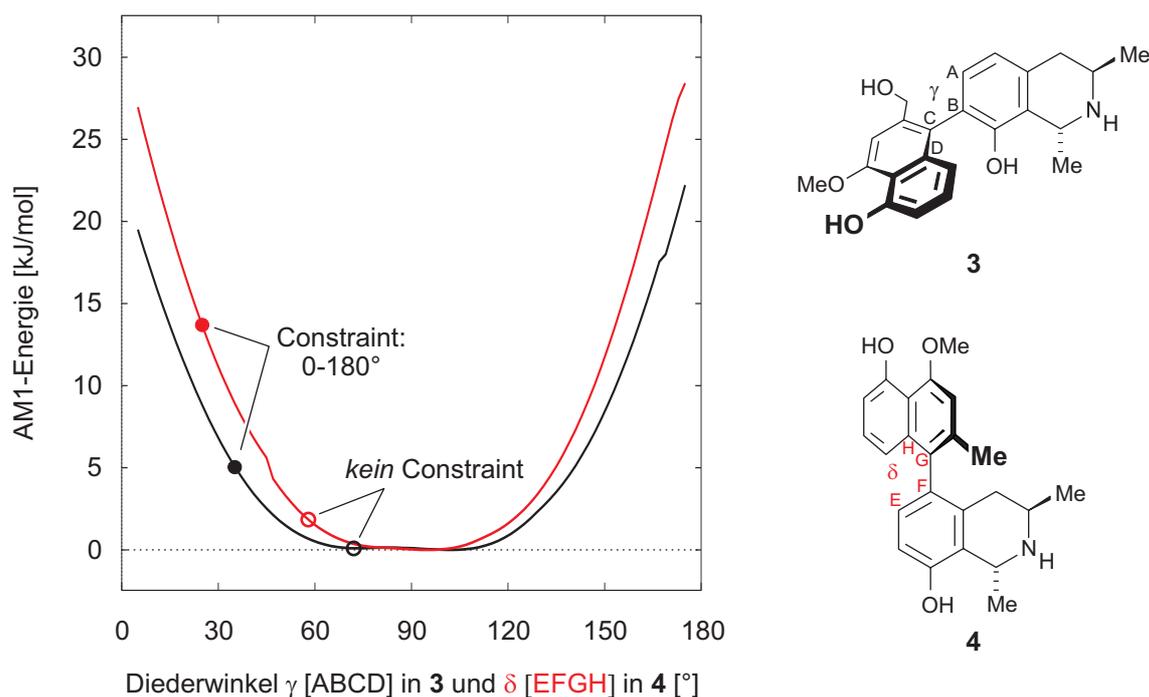


Abb. 21. Änderung der potentiellen Energie (relativ zum Minimum) während der Rotation um die Biarylachsen von Dioncopeltin A (**3**) (Winkel γ [ABCD]) und Dioncophyllin C (**4**) (Winkel δ [EFGH]);

Obwohl das energetisch sehr viel günstigere Alignment auch den geforderten Bedingungen entsprach, lieferte der GASP-Lauf mit Constraint ein energetisch unbrauchbares Resultat: Die zusätzliche Energie, die durch den ungünstigen Diederwinkel an der Biarylachse aufgebracht werden müßte, beträgt (laut AM1-Rechnung) für **3** etwa 5 kcal/mol und für **4** sogar fast 12 kcal/mol. Das Problem wurde noch deutlicher durch weitere GASP-Läufe mit anderen Constraints, deren Resultate in Tabelle 11 zusammengefaßt sind.

Tab. 11. Torsionswinkel (Definitionen s. Abb. 21) an der Biarylachse in verschiedenen GASP-Alignments, in Abhängigkeit von den definierten Constraints (die beiden in Abb. 21 markierten Fälle sind fett gedruckt).

Constraint	60–120°	50–150°	40–140°	30–150°	20–160°	0–180°	-
Verbindung							
3 (γ)	60.8	51.0	41.2	31.5	35.5	35.5	72.0
4 (δ)	60.8	51.0	41.1	35.5	25.7	25.7	58.0

Die gefundenen Werte (Tabelle 11) zeigten, daß die Torsionswinkel stets sehr nahe an der definierten Grenze der Constraints lagen.^a

Energetisch derartig ungünstige Konformere sind für eine CoMSIA-Analyse nicht verwertbar. Dies gilt aber auch für ein GASP-Alignment, das die Axialchiralität nicht berücksichtigt. Auch ein Aussortieren der falsch konfigurierten Strukturen, wie es noch bei FLEXS durchgeführt wurde, war hier nicht möglich, denn GASP benötigt im Vergleich zu FLEXS wesentlich mehr Ressourcen und lieferte in vertretbarer Zeit nur wenige Alignment-Vorschläge, die sich zudem im Regelfall nur marginal unterschieden. Sie enthielten fast immer nur eines der jeweils zwei möglichen Atropdiastereomere.

Da die Achsenkonfiguration aber entscheidend für diese QSAR-Untersuchung ist, mußte eine relativ rigide Einschränkung vorgenommen werden: Die einzige Möglichkeit, mit GASP dennoch ein brauchbares Alignment zu erzeugen, war eine Fixierung der Torsionswinkel an der Achse (bei axialchiralen Verbindungen) auf 90°. Der Konformationsraum der Strukturen wurde so allerdings erheblich eingeschränkt.

4.3.2.3 Durchführung

Die auf Seite 57 beschriebene Idee eines grundlegenden Alignments von **3** und **4** mußte aufgrund dieser Ergebnisse fallengelassen werden. Wie schon bei FLEXS (Abschnitt 4.3.1) wurde Dioncopeltin A als Referenz-Ligand gewählt. Es wurden drei Alignment-Vorschläge für jede Verbindung berechnet und der jeweils höchste Fitness-Score für die folgenden PLS-Analysen verwendet.

GASP verändert vor der Identifizierung der pharmacophoren Elemente die SYBYL-Atomtypen in bestimmten Substrukturen. Beispielsweise werden in Amininen entsprechende N.3-Atomtypen in N.4-Atome (quartäre Stickstoffe) umgewandelt und protoniert. Diese und einige ähnliche Modifikationen, beispielsweise die Platzierung von Dummy-Atomen zur Kennzeichnung der Schwerpunkte in aromatischen Ringen, wurden vor der CoMSIA-Feldberechnung manuell rückgängig gemacht. Eventuelle Störungen, beispielsweise durch unterschiedlich geladene Strukturen, wäh-

^a Dieses Problem wurde mit den Entwicklern von GASP und der Firma Tripos Inc. ausführlich diskutiert; eine Verbesserung des Programms konnte aber in akzeptabler Zeit nicht zur Verfügung oder auch nur in Aussicht gestellt werden.

rend der Berechnung der CoMSIA-Energien sollten so vermieden werden.

Die erhaltenen GASP-Alignments zeigten grundsätzlich keine neuen Alignments, sondern entsprachen den schon durch FLEXS berechneten Möglichkeiten: Von den Naphthylisochinolin-Alkaloiden wurden die meisten entsprechend Möglichkeit (a) (vgl. Abb. 13 auf Seite 49 und Abb. 14 auf Seite 50) ausgerichtet. Die Plazierungen von Ancistrocladein (34) und Hamatein (31) entsprachen Möglichkeit (b), und schließlich wurden Korupensamin B (32), Ancistrocladin (35), Korupensamin A (42), N-Methylancistrocladin (36), Ancistrolikokin D (41), Ancistroealain A (38), Ancistrobertsonin C (33) und Ancistrobertsonin A (12) nach Möglichkeit (c) ausgerichtet.

Aufgrund der Analogien zu den FLEXS-Plazierungen soll hier auf eine genauere Diskussion verzichtet werden.

4.3.3 Manuelle Modifikationen der Alignments

Da FLEXS sehr viele und auch unterschiedliche Alignment-Vorschläge für jede Verbindung lieferte, erschien es aussichtsreich, auch einige der alternativen Möglichkeiten in Betracht zu ziehen.

4.3.3.1 Berücksichtigung des Isochinolin-Stickstoffs (FLEXS^{II})

Ein wichtiger Erfahrungswert im Zusammenhang mit antiplasmodialen Aktivitäten in unserem Arbeitskreis ist die Tatsache, daß eine Abschirmung der Stickstoff-Funktion durch aliphatische Gruppen (Methyl, Benzyl^a) oder eine Veränderung der elektronischen Eigenschaften beispielsweise durch Acylierung^b im Isochinolinteil normalerweise die *In-vitro*-Aktivität vermindert, in den meisten Fällen sogar dramatisch.

In dem automatisch erzeugten FLEXS-Alignment waren diese funktionellen Gruppen in einigen Fällen nicht korrekt ausgerichtet. Deshalb erschien es aussichtsreich, alle Fälle (b) und (c) (siehe Abb. 13 auf Seite 49 und Abb. 14 auf Seite 50) zu verwer-

^a Siehe beispielsweise 3 und 17 (Seite 34) oder 22 und 23 (Seite 35).

^b Dieser Befund basiert noch auf einem älteren Biotests;^[76] in dem hier verwendeten Testsystem wurden bisher keine acylierten Verbindungen untersucht.

fen und nur noch Alignments des Typs (a) zuzulassen.

Diese manuell veränderte Version des ursprünglichen FLEXS-Alignments wird im Folgenden mit FLEXS^{II} bezeichnet, wobei darin die Ausrichtungen von insgesamt neun Strukturen, nämlich Ancistrolikokin B (28), Ancistrolikokin C (29), Ancistrobertsonin B (27), Korupensamin B (32), Ancistrocladin (35), N-Methylancistrocladin (36), Ancistroealain A (38), Ancistrobertsonin A (12) und Ancistrobertsonin C (33) verändert wurden. Die Plazierungen der übrigen Verbindungen wurden für FLEXS^{II} nicht modifiziert.

4.3.3.2 Manuelle Anpassung einiger Torsionswinkel (FLEXS^{III})

Die Substitutionsmuster der Sauerstoff-Funktionen von Naphthylisochinolin-Alkaloiden sind äußerst vielfältig, und demzufolge gibt es praktisch keine Verbindung, die alle möglichen funktionellen Gruppen besitzt und für jede eine Referenz-Stellung bietet.

Das Problem soll an einem Beispiel verdeutlicht werden: Der Referenz-Ligand trägt in 5'-Position eine Hydroxyfunktion (siehe Abb. 22), die eine Wasserstoffbrücke zum Sauerstoffatom der Methoxygruppe in 4'-Position ausbildet. Trägt ein Test-Ligand in 5'-Position etwa eine OMe-Funktion (Abb. 22), so kann dieser Rest nicht korrekt ausgerichtet werden, da im Referenz-Liganden das entsprechende Gegenstück fehlt.^a

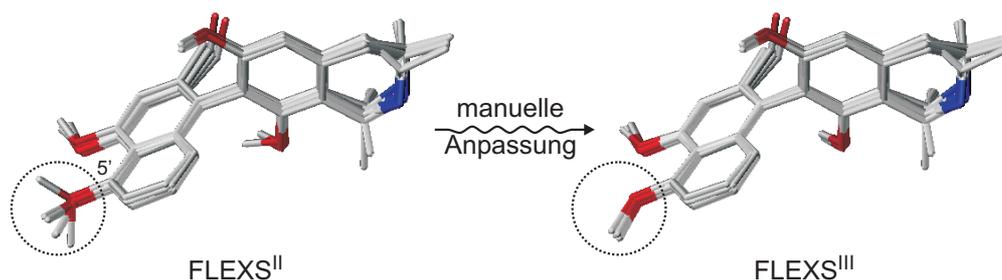


Abb. 22. Manuelle Korrektur falscher Alignments von Gruppen, die im Referenz-Liganden nicht vorliegen, wie z. B. die 5'-Methoxyfunktion (gepunkteter Kreis);

^a Die Hydroxyfunktion kann wegen der Wasserstoffbrücke und der dadurch bedingten Lage nicht als passende Referenz dienen.

Derartige Ungenauigkeiten führen zu erhöhtem Rauschen innerhalb der Daten und vermindern den q^2 -Wert und die Vorhersagekraft des QSAR-Modells.^[43] Ist allerdings der Datensatz an einer solchen Position nicht divergent genug, so hätte das unzulängliche Alignment dieser Gruppe auch auf die PLS-Analyse keinen negativen Einfluß.

Zur Vermeidung derartiger negativer Effekte auf die spätere PLS-Analyse wurden solche Torsionswinkel per Hand modifiziert und so ein drittes Alignment (FLEXS^{III}) erstellt.

4.4 Statistische Auswertung

In Abschnitt 4.3.1 wurde gezeigt, daß das Alignment der kleineren Verbindungen (Gruppe **X**) nicht zu eindeutigen Ergebnissen führte. Deshalb wurden diese Strukturen zunächst nicht berücksichtigt, und erst später versucht, sie in das Modell aufzunehmen.

Die Identifizierung von Outliern wird hier zunächst anhand des FLEXS^{II}-Alignments beschrieben, da sich dieser Ansatz als überlegen herausgestellt hat. Der statistische Vergleich mit GASP und den anderen Alignments auf FLEXS-Basis ist erst nach der Bereinigung des Datensatzes von Outliern sinnvoll, weshalb diese hier vorangestellt wurde. Der Vergleich der Alignments folgt in Kapitel 4.4.3 auf Seite 69.

4.4.1 Outlier innerhalb der Gruppen I bis IX

Bei allen kreuzvalidierten PLS-Analysen in dieser Arbeit wurde 'Column Filtering' auf 2.0 kcal/mol gesetzt.^a Die erste PLS-Analyse basierte auf 58 Verbindungen, 55 aus den Gruppen **I–IX**, sowie den 'gruppenlosen' Strukturen **62**, **63** und **64**. Das Ergebnis war ein q^2 -Wert von 0.239, bei 17 Hauptkomponenten.

Einige Verbindungen zeigten dabei enorm hohe Abweichungen zwischen vorhergesagten und experimentellen Werten. Die Entscheidung, ob eine Verbindung als Outlier aus dem Trainingsdatensatz genommen werden muß, darf aber nicht al-

^a Erläuterung auf Seite 23

lein aufgrund dieser Residuen getroffen werden. Vielmehr müssen strukturelle Besonderheiten oder andere außergewöhnliche Eigenschaften gesucht werden, die zu einer alternativen Wirkungsweise führen könnten. Auch Verbindungen, die zu Zersetzung neigen, sind potentielle Fehlerquellen. Werden keine Ursachen für die Abweichung gefunden, liegt möglicherweise ein falsches Alignment zugrunde.

Innerhalb der Gruppen I–IX wurden fünf Outlier identifiziert und aus dem Trainingsdatensatz ausgeschlossen. Tabelle 12 zeigt die jeweils berechneten Abweichungen und die Resultate der sukzessive durchgeführten PLS-Analysen; mögliche Gründe für das 'nicht Passen' dieser Verbindungen werden im Anschluß diskutiert.

Tab. 12. Abweichungen der als Outlier klassifizierten Strukturen, basierend auf dem FLEXS^{II}-Alignment; eine positive Abweichung bedeutet, daß die experimentelle Aktivität besser ist als die vom Modell vorhergesagte.

		Eigenschaften der Modelle					
Verbindungen im Modell		58	57	56	55	54	53
	q^2	0.239	0.539	0.608	0.652	0.720	0.787
	k	17	20	15	18	16	9
Outlier	pIC ₅₀ (exp.)	Abweichungen der Vorhersage					
45	1.431	2.117	-	-	-	-	-
63	-0.973	-1.735	-1.553	-	-	-	-
47	-0.221	-1.146	0.774	-0.858	-	-	-
49	0.450	0.931	0.739	0.628	1.124	-	-
21	-0.343	-0.114	0.011	-0.541	-0.841	-0.701	-

4.4.1.1 8-O-Methyl-1-*epi*-dioncophyllin B (45) und 8-O-Methyldioncophyllin B (47)

Sowohl **45** als auch **47** (siehe Tabelle 5 auf Seite 38), also zwei der fünf Verbindungen aus Gruppe V, zeigen deutliche Abweichungen, was auf ein falsches Alignment der kompletten Gruppe hinweisen könnte. Dagegen spricht allerdings, daß sowohl FLEXS als auch GASP sehr ähnliche und qualitativ für alle fünf Verbindungen gleiche Plazierungen lieferten.

In früheren Arbeiten^[25] konnte gezeigt werden, daß die Sauerstoff-Funktion an Position 5' von Dioncophyllin B (**43**) (stellvertretend für Gruppe V) mit der Hydroxy-

methylgruppe in Position 2' von Dioncopeltin A (**3**) überlagert werden kann. Diese Wechselwirkung wird aber weder von FLEXS noch von GASP vorgeschlagen. Beide plazierten die Hydroxyfunktion auf dem aromatischen Naphthalinrest von **3**. Angesichts der beiden Outlier wurde die früher gefundene Möglichkeit hier genauer untersucht, anhand von manuell erstellten, alternativen Alignmentmöglichkeiten für Gruppe V (siehe Abb. 23).

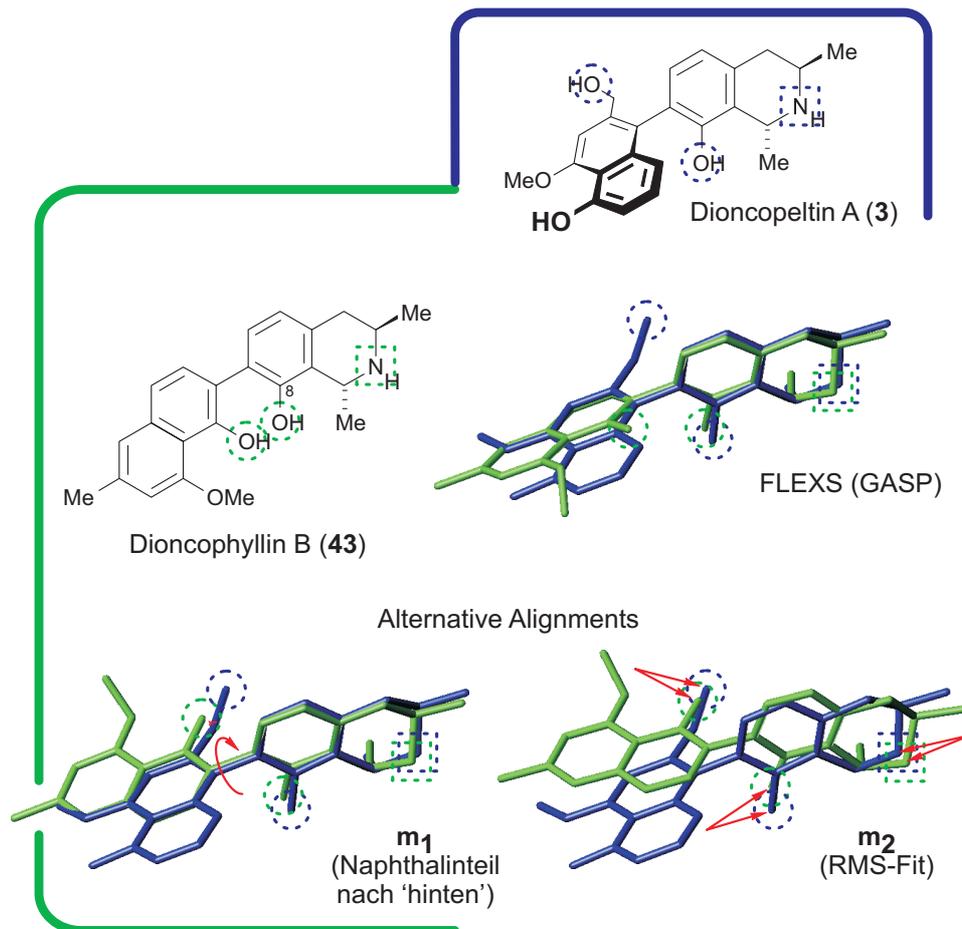


Abb. 23. Manuelle Modifikationen des Alignments der Gruppe V; die durchgeführte Rotation sowie die für den RMS-Fit herangezogenen Atome sind durch rote Pfeile markiert

Weil 7,6'-gekuppelte Naphthylisochinolin-Alkaloide bei Raumtemperatur nicht atropstabil sind, kann die Achsenkonfiguration ohne Einschränkungen modifiziert werden. Zunächst wurde der Diederwinkel an der Biarylachse so verändert, daß die Hydroxyfunktion auf der 'richtigen' Seite lag (siehe 'm(anuell)₁', Abb. 23 unten).

Weil dieses Alignment noch nicht optimal erschien, wurde zusätzlich ein RMS-Fit

durchgeführt. Dazu herangezogen wurden jeweils die Stickstoff-Atome sowie die Sauerstoffe in Position 8 des Isochinolinteils und die entsprechenden Sauerstoffatome im Naphthalinteil (siehe \mathbf{m}_2 in Abb. 23).

Die erhaltenen Resultate der anschließenden PLS-Analyse sind in Tabelle 13 aufgeführt.

Tab. 13. Kreuzvalidierte q^2 -Werte für die manuell modifizierten Alignments der Gruppe V (vgl. Abb. 23); die übrigen drei Outlier (s. unten) wurden hier bereits nicht mehr berücksichtigt.

	FLEXS ^{II}	\mathbf{m}_1	\mathbf{m}_2
q^2	0.569	0.558	0.524
k	8	8	6

Da beide Modifikationen keinen positiven Einfluß auf den q^2 -Wert zeigten, wurde für die weiteren Untersuchungen das ursprüngliche FLEXS^{II}-Alignment beibehalten, und die beiden Verbindungen **45** und **47** mußten als Outlier aus dem Trainingsdatensatz ausgeschlossen werden.

Für diese Vorgehensweise sprach ebenfalls, daß **45** und **47** sich lediglich durch eine Hydroxyfunktion an C-4 und die Konfiguration an C-1 des Isochinolinteils unterscheiden (siehe Seite 38), dennoch aber sehr verschiedene Aktivitäten aufweisen. Sollte dies tatsächlich auf diese geringen strukturellen Unterschiede zurückzuführen sein, war das Modell offensichtlich nicht in der Lage, dies zu differenzieren.

4.4.1.2 Weitere Outlier

- Das Phenyltetrahydroisochinolin **63** (siehe Seite 41) ist vollständig *O*-demethyliert und beinhaltet ein ungeschütztes, phenyloges Hydrochinonsystem, zumal ohne sterische Hinderung der Achse. Die Verbindung kann deshalb wahrscheinlich sehr leicht, z. B. (unter Einebnung) zum Diphenochinon oxidiert werden;^[106] laut Tabelle 12 auf Seite 64 sollte **63** um 1.74 (bzw. 1.55) Größenordnungen aktiver sein, als die Substanz tatsächlich getestet wurde. Dies sprach dafür, daß ein großer Teil der Probe sich vor dem Test bereits zersetzt hatte.

- Knipholon-Anthron (**49**) ist die einzige Verbindung aus Gruppe VI, die ein reduziertes (hydrodeoxygeniertes) Anthrachinonsystem (im Tautomerengleichgewicht mit einem extrem elektronenreichen Trihydroxyanthracen) aufweist (Seite 39). Dies läßt, wie bei Verbindung **63**, seine Empfindlichkeit gegen Oxidation^[121] verstehen, könnte aber auch auf einen andersgearteten Wirkmechanismus hindeuten. Für das Herausnehmen aus dem Trainingsdatensatz sprach außerdem die hohe Cytotoxizität,^[117] die zu einer Verfälschung des Testwertes geführt haben könnte.
- 4'-O-5'-O-Demethyldioncophyllin A (**21**) zeigte bereits während der Aufreinigung^[77] enorme Empfindlichkeit gegenüber Luftsauerstoff, möglicherweise aufgrund der beiden freien Hydroxyfunktionen im Naphthylteil. Die Abweichung von 0.70 (Tabelle 12 auf Seite 64) wäre zwar noch tolerierbar, da es aber zudem fraglich war, ob die getestete Probe ausschließlich das korrekte Atropdiastereomer enthielt,^[77] wurde die Verbindung ausgeschlossen.

4.4.1.3 Verifizierung der Outlier mit Hilfe des GASP-Alignments

Die analogen Berechnungen wurden schließlich unter Verwendung des GASP-Alignments verifiziert. Da die q^2 -Werte hier insgesamt etwas niedriger waren, wichen auch die Vorhersagen nicht so sehr vom Experiment ab wie noch bei den FLEXS-Modellen. Dennoch bestätigten die Ergebnisse die mit FLEXS gefundenen Outlier (Tabelle 14).

Tab. 14. Abweichungen der Vorhersagen für die als Outlier klassifizierte Strukturen, basierend auf dem GASP-Alignment.

		Eigenschaften der Modelle					
Verbindungen im Modell		58	57	56	55	54	53
q^2		0.263	0.472	0.545	0.598	0.613	0.686
k		5	6	6	6	6	5
Outlier	pIC ₅₀ (exp.)	Abweichungen der Vorhersage					
63	-0.973	-2.439	-	-	-	-	-
45	1.431	1.231	1.079	-	-	-	-
47	-0.221	-0.736	-0.786	-1.050	-	-	-
21	-0.343	-0.727	-0.975	-0.929	-0.956	-	-
49	0.450	0.580	0.824	0.698	0.688	0.663	-

4.4.2 Monomere Verbindungen (Gruppe X)

Nachdem die Outlier aus den Gruppen I-IX ausgeschlossen worden waren, wurde jetzt versucht, auch die Strukturen aus Gruppe X, die weder von FLEXS noch von GASP einheitlich ausgerichtet wurden (siehe Abschnitt 4.3.1 auf Seite 46), in das Modell aufzunehmen.

Tabelle 15 zeigt die Eigenschaften der erstellten Modelle und die gefundenen Abweichungen für alle hinzugefügten Strukturen, jeweils für beide Alignments.

Tab. 15. Abweichungen ('Residuals') aller Verbindungen aus Gruppe X.

		Eigenschaften der Modelle					
Alignment		FLEXS ^{II}			GASP		
Verb. im Modell		65	64	63	65	64	63
q^2		0.601	0.627	0.677	0.486	0.560	0.600
k		7	8	7	6	5	
Outlier	pIC ₅₀ (exp.)	Abweichungen der Vorhersage					
67	1.011	1.410	-	-	1.617	-	-
68	-1.240	-0.803	-1.268	-	-1.106	-1.073	-
Übrige Verbindungen aus Gruppe X							
65	0.363	0.443	0.544	0.331	0.825	0.794	0.870
66	0.047	0.315	0.986	0.993	-0.135	0.172	-0.189
69	-0.652	-0.485	-0.363	-0.449	-0.209	-0.170	-0.284
70	-0.321	-0.018	-0.081	-0.164	0.755	0.717	-0.063
71	-0.036	0.413	0.477	0.511	0.307	0.641	0.464
72	-0.101	-0.193	-0.228	-0.137	0.784	0.734	0.542
73	-0.647	0.330	0.327	0.350	0.427	0.149	-0.198
74	-0.857	-1.095	-0.877	-0.871	-1.145	-1.092	-0.967
75	-0.617	-0.752	-0.623	-0.620	-0.058	0.307	0.087
76	-0.103	0.093	0.072	0.052	0.375	0.752	0.488

Für einen q^2 -Wert auf einem akzeptablen Niveau mußten zwei Verbindungen (**67** und **68**, siehe Seite 42) ausgeschlossen werden. Offensichtliche strukturelle Gründe, die dieses Vorgehen unterstützen konnten, wurden hier nicht gefunden. Deshalb und weil die Abweichungen für alle 63 Verbindungen im Datensatz keine extremen Werte mehr zeigten, wurden keine weiteren Verbindungen ausgeschlossen.

4.4.3 Resultate der PLS-Analysen

Nach der Identifizierung der Outlier wurden die Resultate von GASP und FLEXS sowie die Ergebnisse der manuell veränderten FLEXS-Alignments genauer verglichen. Zu diesem Zweck wurden neben der 'leave-one-out'-Methode auch Gruppen-Kreuzvalidierungen ('n-fold-crossvalidations') durchgeführt.

Dabei wird bei jedem Durchlauf der Datensatz in n Gruppen aufgeteilt; anschlie-

ßend versucht man, die Aktivitäten jeweils einer dieser Gruppe vorherzusagen. So bedeutet $n = 5$, daß 20% der Verbindungen während eines Vorhersagezyklus herausgenommen werden. Besonders bei größeren Datensätzen ist es wichtig, diese Vorgehensweise ebenfalls zu verfolgen, da hier die Leave-one-out- q^2 -Werte (im Folgenden q_{LOO}^2 genannt) oft zu optimistisch ausfallen und zu einer Überschätzung der Vorhersagekraft des Modells führen.^a In dieser Arbeit wurden Gruppen-Kreuzvalidierungen für $n = 3, 5$, und 10 durchgeführt (q^2 -Werte für die Gruppen-Kreuzvalidierungen werden als Folgenden q_n^2 benannt). Weil der q_n^2 -Wert natürlich sehr stark von der Gruppeneinteilung abhängt, wurden die Berechnung mehrmals durchgeführt und dabei die Gruppen jeweils zufällig gewählt. In dieser Arbeit wurden die Kreuzvalidierungen in Gruppen jeweils 100mal wiederholt.

In den Tabellen 16 und 17 auf der nächsten Seite sind die Ergebnisse vergleichend zusammengefaßt. Dabei wurde sowohl der Datensatz mit ausschließlich Biarylen (53 Verbindungen aus den Gruppen I–IX) als auch der Gesamtdatensatz (63 Verbindungen, zusätzlich Gruppe X) basierend auf jeweils allen 4 Alignments berechnet.

Tab. 16. Resultate der 'Leave-one-out'-Crossvalidation ($n = 1$).

Alignment		Gruppen I–IX	Gruppen I–X
FLEXS	q_{LOO}^2	0.623	0.555
	k^α	8	6
FLEXS ^{II}	q_{LOO}^2	0.787	0.704
	k	9	8
FLEXS ^{III}	q_{LOO}^2	0.783	-
	k	8	
GASP	q_{LOO}^2	0.686	0.600
	k	5	6

^a Optimale Anzahl an Hauptkomponenten.

^a Bei größeren Datensätzen kommt es unweigerlich zu Redundanzen, d. h. zu einigen Substanzen gibt es Analoga im Datensatz, mit gleichen oder sehr ähnlichen Eigenschaften. Diese können während der Kreuzvalidierung wesentlich genauer vorhergesagt werden, als es ohne ihren 'Zwillingsbruder' im Datensatz möglich gewesen wäre. Demzufolge erhält man durch diese Art der Kreuzvalidierung oft zu optimistische Ergebnisse.^[122] Das Problem der Redundanz ist dennoch zweiseitig: Wäre wirklich jedes entscheidende Merkmal nur ein einziges Mal im Datensatz vorhanden, könnte ein Gesamtmodell sehr schnell durch nur einen unpräzisen experimentellen Wert gestört werden.

Tab. 17. Resultate der Gruppen-Kreuzvalidierungen.

Alignment	n	Gruppen I-IX			Gruppen I-X		
		10	5	3	10	5	3
FLEXS	mittlerer $q_n^2{}^\alpha$	0.606	0.572	0.518	0.564	0.537	0.498
	höchster q_n^2	0.685	0.685	0.673	0.646	0.647	0.637
	niedrigster q_n^2	0.481	0.410	0.253	0.463	0.369	0.154
	SA ^{β}	0.039	0.058	0.089	0.036	0.055	0.088
FLEXS ^{II}	mittlerer q_n^2	0.757	0.714	0.651	0.675	0.640	0.579
	höchster q_n^2	0.812	0.804	0.789	0.732	0.731	0.733
	niedrigster q_n^2	0.631	0.517	0.353	0.590	0.138	0.231
	SA	0.035	0.055	0.080	0.031	0.066	0.090
FLEXS ^{III}	mittlerer q_n^2	0.749	0.720	0.651			
	höchster q_n^2	0.799	0.806	0.794			
	niedrigster q_n^2	0.523	0.507	0.382			
	SA	0.037	0.051	0.083			
GASP	mittlerer q_n^2	0.668	0.630	0.577	0.600	0.554	0.496
	höchster q_n^2	0.735	0.757	0.742	0.663	0.672	0.635
	niedrigster q_n^2	0.538	0.373	0.211	0.404	0.298	0.277
	SA	0.040	0.069	0.086	0.040	0.063	0.086

^{α} Mittelwert aus 100 Läufen.

^{β} Standardabweichung.

4.4.4 Schlußfolgerungen

4.4.4.1 Optimaler Alignmentansatz

Die Kreuzvalidierungen (Tabellen 16 und 17) zeigten, daß optimale Resultate mit dem FLEXS^{II}-Alignment erhalten wurden. Der q_{LOO}^2 -Wert für diesen Ansatz lag bei 0.787, der mittlere q_n^2 (für 10 Gruppen) bei beachtlichen 0.757. Die Schwankung dieses Wertes während der 100 Wiederholungen zwischen 0.631 und 0.812 wies auf gewisse strukturelle Redundanzen im Datensatz hin. Besonders der mittlere q_n^2 -Wert von 0.651 für $n = 3$ zeigte die Überlegenheit gegenüber dem GASP- und dem ursprünglichen FLEXS-Alignment, die für $n = 3$ allesamt mittlere q_n^2 -Werte deutlich unter 0.6 lieferten.

Die Resultate des FLEXS^{III}-Ansatzes waren geringfügig schlechter ($q_{LOO}^2 = 0.783$, mittlerer $q_n^2 = 0.749$, $n = 10$), danach folgte das GASP-Alignment mit schon deutlich abfallenden Werten ($q_{LOO}^2 = 0.686$, mittlerer $q_n^2 = 0.668$, $n = 10$), überraschenderweise aber noch besser als das unmodifizierte FLEXS-Alignment.

Die manuelle Veränderung des automatisch erhaltenen FLEXS-Alignments unter Berücksichtigung des Isochinolin-Stickstoff-Atoms (vgl. Abschnitt 4.3.3.1 auf Seite 61) führte also zu einer deutlichen Verbesserung des Modells, wodurch die Wichtigkeit des Stickstoff-Atoms in diesem System deutlich wird.

Die Anpassung der Torsionswinkel (FLEXS^{II} nach FLEXS^{III}) brachte keine Verbesserungen. Dies kann zwei Gründe haben:

- Die entsprechenden Gruppen haben keinen gravierenden Einfluß auf die biologische Aktivität.
- Die Effekte in diesen Regionen können vom Modell nicht wiedergegeben werden, da der gesamte Datensatz dort keine ausreichende Diversität aufweist.^a

Ein direkter Vergleich von FLEXS und GASP ist aufgrund der sehr unterschiedlichen Techniken, die die beiden Programme verwenden, nur bedingt möglich. Die in Abschnitt 4.1.2 (Seite 44) dargestellte Problematik bei der Verwendung von GASP in Verbindung mit CoMSIA stellt dabei nur einen Teilaspekt dar; letztlich entscheidend sind die Besonderheiten der betrachteten Verbindungen und die daraus resultierenden Alignmentmöglichkeiten. Das folgende Fazit bezieht sich deshalb in erster Linie auf die anhand von Naphthylisochinolin-Alkaloiden gemachten Erfahrungen und ist nur bedingt auf andere Substanzklassen übertragbar.

1. Betrachtet man die reine CPU-Zeit, so ist FLEXS substantiell schneller als GASP. Abhängig von den verwendeten Liganden wurden mit FLEXS zwischen 50 und einigen hundert Alignments erhalten und nach entsprechendem Clustering konnten anschließend sehr schnell mehrere verschiedene Lösungsvorschläge erhalten werden, während von GASP in mehr als der doppelten Zeit nur wenige Alignments berechnet werden konnten.

^a Da die CoMSIA-Methode strukturelle Unterschiede korreliert, können in Regionen, die keine Diversität aufweisen, auch keine Rückschlüsse gezogen werden. Siehe dazu auch die Fußnote auf Seite 97.

2. Nicht nur die höhere Anzahl, sondern auch die größere Diversität der Lösungen ermöglichte bei FLEXS mehr Flexibilität bei der manuellen Selektion der vielversprechendsten Alignments. Die wenigen GASP-Lösungen waren in fast allen Fällen untereinander sehr ähnlich (sie entsprachen meist nur einer der drei von FLEXS gefundenen Möglichkeiten, vgl. Abb. 13 auf Seite 49).
3. Auf der anderen Seite kann GASP mehrere Verbindungen gleichzeitig behandeln, während FLEXS stets nur zwei Verbindungen vergleicht; leider konnte von diesem Vorteil wegen der beschriebenen Schwierigkeiten bisher noch kein Gebrauch gemacht werden.

Die notwendige rigide Einschränkung der Beweglichkeit, die während des GASP-Alignments vorgenommen werden mußte (Abschnitt 4.3.2), machte letztlich die Unterlegenheit dieses Verfahrens aus. Der Einsatz von zukünftigen, fehlerbereinigten Versionen von GASP könnte möglicherweise zu verbesserten Ergebnissen führen.

4.4.4.2 Einfluß anderer Substanzklassen auf das Modell

Die Phenylanthrachinone aus Gruppe VI (siehe Seite 39), aber auch die Naphthylindole aus den Gruppen VII und VIII konnten in das Modell mit aufgenommen werden. Wertvoll für das Modell sind derartige Strukturen aber nur, wenn es anschließend in der Lage ist, neue Verbindungen verlässlich vorherzusagen. Man kann dies untersuchen, indem man Modelle erstellt, die zunächst bestimmte Strukturtypen nicht 'kennen' und die entsprechende Vorhersagekraft prüfen. Die Vorhersagen für eine neue Substanzklasse sollten sich verbessern, wenn für die Erstellung des Modells auch derartige Verbindungen verwendet werden.

Knipholon (48) und seine Derivate (Gruppe VI)

Innerhalb dieser Strukturklasse mußte bereits Knipholon-Anthron (49) als Outlier ausgeschlossen werden, wofür möglicherweise die Cytotoxizität der Verbindung verantwortlich sein könnte.

Für die Erstellung der Modelle wurde wiederum das FLEXS^{II}-Alignment verwendet. Für das K1-Modell (siehe Tabelle 18 auf der nächsten Seite) wurden alle fünf Substanzen aus Gruppe VI entfernt. Für K2 hingegen wurden dem Trainingsdaten-

satz wieder drei zufällig ausgewählte Phenylanthrachinone (**48**, **49**, **51**, **52**) hinzugefügt.

Tab. 18. Die Eigenschaften der Modelle mit modifiziertem Trainingsdatensatz, ohne Phenylanthrachinone (K1) und mit drei der fünf Strukturen aus Gruppe VI (K2).

Modell	K1	K2
Verbindungen im Modell	48	51
q^2	0.775	0.780
k	8	7

Tabelle 19 zeigt die mit diesen Modellen berechneten Vorhersagen für die übrigen beiden Strukturen aus Gruppe VI

Tab. 19. Anweichungen der mit dem K1- und dem K2-Modell vorhergesagten pIC_{50} -Werte für Verbindungen aus Gruppe VI.

	$pIC_{50}(\text{exp.})$	Abweichungen	
		K1	K2
4-O-Demethylknipholon (50)	-0.071	0.428	0.396
53	-0.121	0.237	0.330

Die Vorhersage der beiden Modelle zeigten etwa gleiche Qualität. Dies bedeutet entweder, daß die im zweiten Modell berücksichtigten Strukturen keinen Einfluß auf das Modell nahmen, oder daß es bereits vorher in der Lage war, die Aktivitäten dieser Substanzklasse innerhalb gewisser Fehlergrenzen zu berechnen.

Da die Berücksichtigung dieser Strukturen offensichtlich keinen negativen Einfluß auf die Vorhersagekraft des Modells haben, wurden sie für den abschließenden Trainingsdatensatz herangezogen.

Die Naphthylindole (Gruppen VII und VIII)

Analog zum vorhergehenden Abschnitt wurden auch hier zwei neue Modelle (I1, I2) erstellt und ihre Vorhersagekraft verglichen. Zufällig ausgewählt für den Trainingsdatensatz von I2 wurden **54**, **55**, **9**, **57**, **59** und **10**, die Ergebnisse zeigen die Tabellen 20 und 21.

Tab. 20. Die Eigenschaften von Modellen, basierend auf Trainingsdatensätzen ohne (I1) und mit sechs der acht Strukturen aus den Gruppen VII und VIII (I2).

Modell	I1	I2
Verbindungen im Modell	45	51
q^2	0.732	0.759
k	8	7

Tab. 21. Vorhersagen des Modells für Verbindungen aus den Gruppen VII und VIII, in Abhängigkeit vom verwendeten Trainingsdatensatz.

		Abweichungen	
	pIC ₅₀ (exp.)	I1	I2
56	-0.080	-0.383	-0.217
58	0.299	0.859	0.307

Die zunächst großen Abweichungen für das I1-Modell erklären die Tatsache, daß diese Verbindungen zunächst mit einem CoMFA-Modell^[25] als sehr aktiv vorhergesagt wurden, dies dann im experiment aber nicht nachvollzogen werden konnte. Wie in Ref.^[25] ist auch hier das nicht 'trainierte' Modell nicht in der Lage, die Aktivitäten richtig vorherzusagen. Anders als im vorherigen Abschnitt bei Gruppe VI war das trainierte Modell (I2) dagegen in der Lage, die Aktivitäten dieser Strukturen deutlich besser vorherzusagen: Die Abweichungen der Vorhersage von den experimentellen Werten nehmen ab von -0.383 auf -0.217 für **56** und noch deutlicher von 0.859 auf 0.307 für **58**. Die Verbindungen wurden im Trainingsdatensatz belassen, denn das Modell gewann so offensichtlich an Vorhersagekraft.

Gruppe X

Die Berücksichtigung dieser Strukturen im Trainingsdatensatz führte bei allen untersuchten Alignments zu einer Verschlechterung der statistischen Signifikanz (siehe Tabellen 16 und 17 auf Seite 71) der Modelle. Es konnten zwar akzeptable q^2 -Werte erreicht werden, die aber bei allen Alignments stets hinter denen der kleineren Datensätze zurückblieben.

Die Verbindungen aus Gruppe X wurden in den abschließenden Trainingsdatensatz nicht mitaufgenommen. Dafür waren nachstehende Gründe ausschlaggebend:

1. Das Alignment der Verbindungen erfolgte nicht einheitlich. Die Folge sind sehr unterschiedliche Abweichungen bei den verschiedenen Alignmentansätzen.
2. Welche Alignmentstrategie die richtige sein könnte, war aus den vorliegenden Resultaten nicht zu ersehen. Auch ein manuelles Plazieren auf entsprechenden Partialstrukturen des Referenz-Liganden erschien nicht sinnvoll, vor allem wegen des Größenunterschiedes und der daraus entstehenden vielen unterschiedlichen Möglichkeiten.
3. Der Vergleich von sehr unterschiedlich großen Strukturen in einer CoMSIA-Analyse ist aus entropischen Gründen problematisch, denn die Änderung der Entropie während des 'Andockens' eines Liganden in eine Enzymtasche ist von seiner Größe abhängig. Dieser Effekt bleibt bei einer CoMSIA-Analyse unberücksichtigt (siehe unten).

Die Entropie trägt zu einem wichtigen Teil zur Affinität eines Liganden in einer Bindungstasche bei. Der wesentliche Beitrag rührt dabei von der Änderung der Solvatation von Ligand und Protein während des 'Andockens' her,^[123] aber auch von anderen Effekten, wie beispielsweise der Einschränkung der Bewegungsfreiheit des Liganden. Es konnte gezeigt werden, daß die *CoMFA*-Deskriptoren diese Effekte nicht, oder allenfalls in zu geringem Maße berücksichtigten.^[124]

Bei der CoMSIA-Methode ist die Situation etwas anders: Hier werden entropische Effekte zumindest teilweise berücksichtigt,^[123] denn diese hängen mit hydrophoben Eigenschaften der Moleküloberfläche der Liganden zusammen.^[125,126] Die Entropieänderungen bzgl. des Proteins oder der Flexibilität des Liganden bleiben allerdings weiterhin unberücksichtigt.

Ein weiterer, möglicherweise für diese Untersuchung entscheidender Effekt führt zu einem Entropieanstieg während des 'Andockens', nämlich die Verdrängung von Wasser oder Lösungsmittelmolekülen aus der Bindungstasche. Dieser Aspekt hängt in hohem Maße von der Größe der Moleküle ab^[123] und führt möglicherweise zu Fehleinschätzungen beim Vergleich der kleineren Verbindungen aus Gruppe **X** mit Naphthylisochinolin-Alkaloiden.

4.4.5 Optimierung des CoMSIA-Modells

Gemäß den vorhergehenden Schlußfolgerungen basierend auf dem Trainingsdatensatz der 53 Biarylverbindungen und dem FLEXS^{II}-Alignment wurden nun einzelne Parameter des Modells optimiert.

4.4.5.1 Unterschiedliche Kombinationen der CoMSIA-Felder

Es wurde in der Literatur ausführlich diskutiert,^[43,127,128] inwieweit die durch die fünf CoMSIA-Deskriptoren ausgedrückten Eigenschaften unabhängig voneinander sind, denn bestimmte funktionelle Gruppen können Wechselwirkungen ausüben, die durch mehrere Felder gleichzeitig erfaßt werden. Beispielsweise können einer Hydroxyfunktion sowohl Wasserstoffbrücken-Donor- als auch -Akzeptoreigenschaften zugeordnet werden. Dies führt zu Redundanzen innerhalb der Daten und vermindert so das Verhältnis von Signal zu Rauschen^[127] und damit die statistische Signifikanz des Ergebnisses.^[43]

Welche der fünf CoMSIA-Felder für ein optimales Modell tatsächlich notwendig sind, kann man durch PLS-Analysen für alle 31 möglichen Kombinationen herausfinden. Dabei wurde auf den schnelleren SAMPLS^[53]-Algorithmus zurückgegriffen,^a die Resultate sind in Abb. 24 auf der nächsten Seite graphisch dargestellt.

^a Erläuterung auf Seite 23

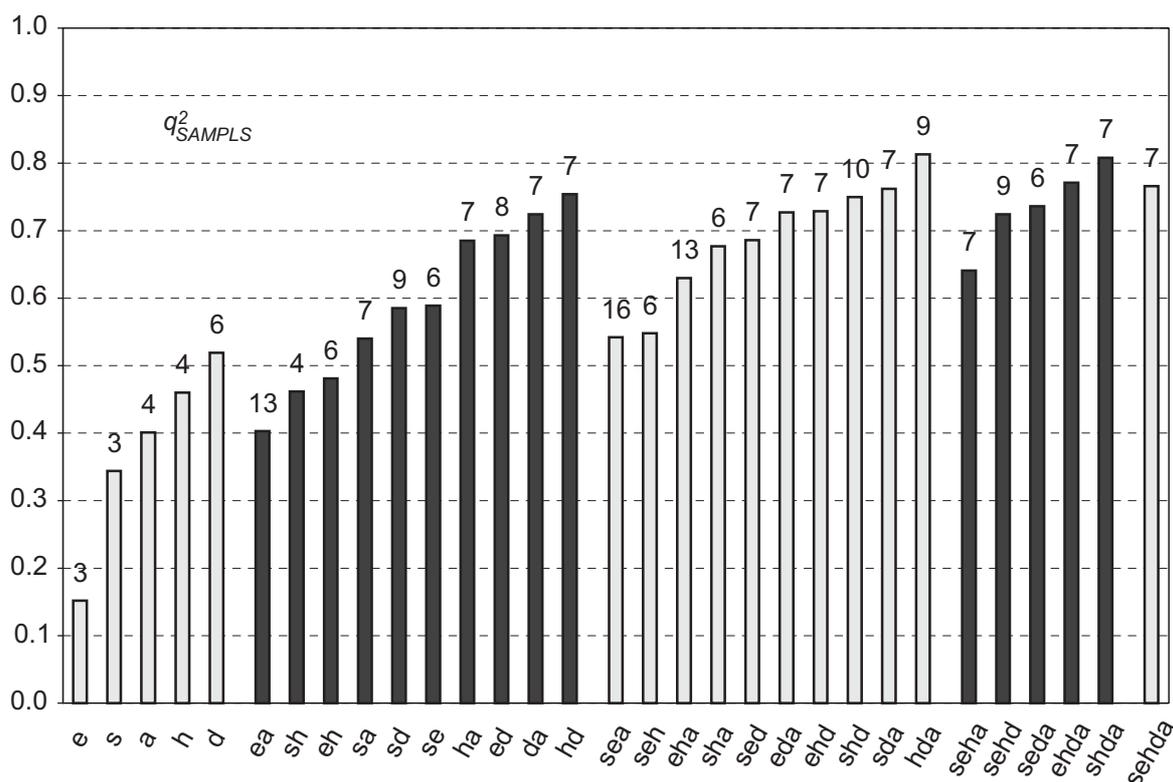


Abb. 24. Resultate der 31 möglichen CoMSIA-Feldkombinationen (s = sterisch, e = elektrostatisch, h = Hydrophobie, d/a = Wasserstoffbrücken-Donor/-Akzeptor); sortiert wurde zunächst nach der Anzahl der verwendeten Felder, dann nach q^2_{SAMPLS} , die Nummer über den Balken gibt die jeweils optimale Anzahl der Hauptkomponenten an

Die q^2_{SAMPLS} -Werte für die unterschiedlichen Feldkombinationen variierten relativ stark. Dies deutete darauf hin, daß die gegenseitigen Abhängigkeiten unter den Feldern insgesamt relativ gering sind. In vergleichbaren Untersuchungen wiesen sehr ähnliche q^2 -Werte für alle Feldkombinationen auf große Anhängigkeiten hin.^[43]

Schon die Betrachtung der einzelnen Felder für sich allein zeigte, daß Wasserstoffbrücken-Wechselwirkungen und auch die Hydrophobie signifikantere Korrelationen zeigten als die (in einer CoMFA-Analyse ausschließlich verwendeten) sterischen und elektrostatischen Funktionen.

Die optimale Zwei-Feld-Kombination ist hd ($q^2_{SAMPLS} = 0.754$, $k = 7$). Der zusätzliche Einsatz des Wasserstoffbrücken-Akzeptor-Feldes führte zu dem insgesamt besten Modell hda, $q^2_{SAMPLS} = 0.813$ ($k = 9$). Die Verwendung von statt dessen steri-

schen oder elektrostatischen Indizes führte zu jeweils schlechteren Modellen: shd ($q_{SAMPLS}^2 = 0.750$, $k = 10$) und ehd ($q_{SAMPLS}^2 = 0.729$, $k = 7$).

Dieser Befund bestätigt, daß sterische und elektrostatische Wechselwirkungen für die Aktivität der Verbindungen im Datensatz nur geringe Bedeutung haben. Möglicherweise werden entsprechende Effekte aber auch bereits über eines der drei anderen Felder (h, d, a) berücksichtigt.

Für die besten sieben Kombinationen wurde anschließend jeweils eine gewöhnliche PLS-Analyse durchgeführt, die dabei erhaltenen q^2 -Werte waren erwartungsgemäß den q_{SAMPLS}^2 -Werten sehr ähnlich (siehe Tabelle 22). Insgesamt wurden die bisherigen Überlegungen bestätigt: Die meiste Information aus dem Datensatz wird bereits durch das Wasserstoffbrücken-Donor- und das Hydrophobie-Feld ausgedrückt (hd). Dennoch wurde dieses Modell durch das Wasserstoffbrücken-Akzeptor-Feld noch einmal verbessert auf einen q^2 -Wert von schließlich 0.818 (hda, $k = 8$).

Tab. 22. CoMSIA-Modelle basierend auf den besten Feldkombinationen aus Abb. 24. Das beste der Modelle ist fett gedruckt (Abkürzungen sind ebenfalls in Abb. 24 erklärt).

	h d	s h d	e h d	h d a	s h d a	e h d a	s e h d a
q_{SAMPLS}^2	0.754	0.750	0.729	0.813	0.808	0.771	0.766
q_{LOO}^2	0.763	0.752	0.738	0.818	0.812	0.780	0.787
PRESS	0.351	0.355	0.370	0.308	0.313	0.338	0.337
k	8	7	8	8	8	8	9
r^2	0.978	0.970	0.982	0.985	0.979	0.982	0.987
s	0.107	0.124	0.097	0.101	0.104	0.096	0.084
F-Wert	243.6	208.2	298.9	276.9	257.5	304.1	357.8
Anteile							
sterisch	-	0.122	-	-	0.066	-	0.048
elektrostatisch	-	-	0.352	-	-	0.238	0.223
Hydrophobie	0.420	0.342	0.210	0.234	0.204	0.166	0.152
H-brücken-Donor	0.580	0.537	0.437	0.431	0.411	0.350	0.337
H-brücken-Akzeptor	-	-	-	0.334	0.319	0.246	0.239

Interessanterweise ist im abschließenden Modell der Informationsanteil des Wasserstoffbrücken-Akzeptor-Feldes höher als der des Hydrophobie-Feldes, obwohl bei den Einzel-Feld-Modellen das Hydrophobie-Feld noch einen besseren q^2 -Wert zeigt.

te (Abb. 24 auf Seite 78). Dies ist wiederum ein Beispiel für eine Eigenschaft der Liganden, die sich in zwei unterschiedlichen Feldern auswirkt.

4.4.5.2 Dämpfungs-Faktor

In Kapitel 1 (siehe auch Gleichung 3 auf Seite 17) wurde der Einfluß des sogenannten 'Dämpfungs'-Faktors α beschrieben. Für die bisherigen Untersuchungen wurde der empfohlene Standard-Wert von $\alpha = 0.3$ verwendet. Da dieser Wert für einzelne Substanzklassen und deren Wirkungsweise sehr spezifisch ist, muß der optimale α -Wert für jedes Problem neu gefunden werden.^[34,43,129,130]

Der berechnete Verlauf von q_{LOO}^2 unter Variation von α (in Schritten von 0.1) ist in Abb. 25 für die beiden unterschiedlichen Datensätze (jeweils FLEXS^{II}-Alignment) gezeigt.

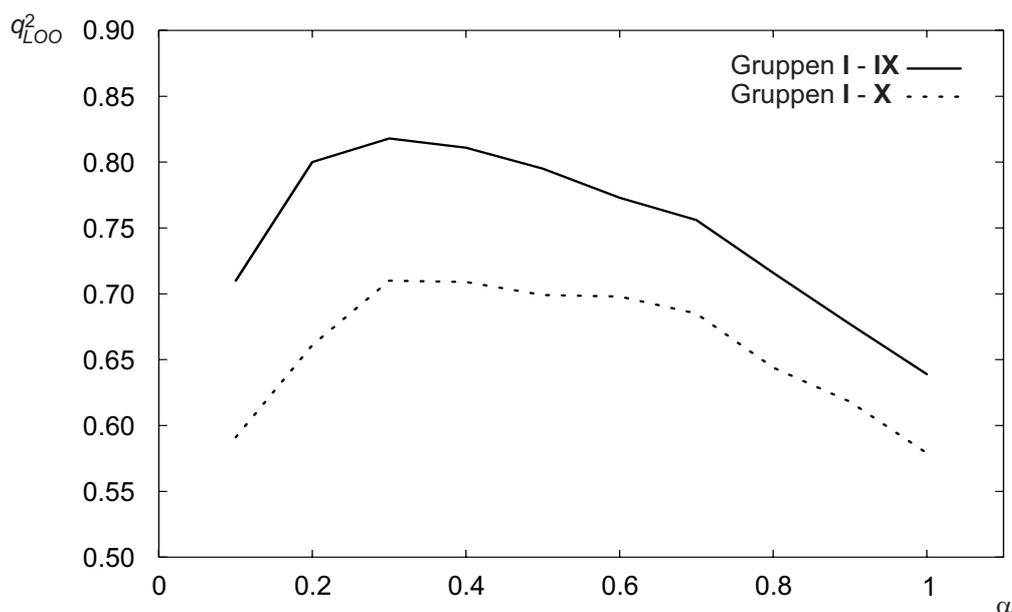


Abb. 25. Verlauf von q_{LOO}^2 für unterschiedliche Werte von α , für die Gruppen I-IX und für den kompletten Datensatz

Für kleinere Werte von α sinkt der q_{LOO}^2 -Wert relativ rasch ab, bei größeren Werten fällt die Kurve etwas weniger steil nach unten, insgesamt ist die Abhängigkeit etwas stärker als in vergleichbaren CoMSIA-Studien.^[129,130] Der für diese Untersuchung optimale Wert ist offensichtlich $\alpha = 0.3$.

4.4.5.3 Einfluß des Gitterabstandes bei der Feldberechnung

Der Abstand der Gitterpunkte hat ebenfalls Einfluß auf das QSAR-Modell. Kleinere Gitterabstände erlauben durch die Berücksichtigung von mehr Feldpunkten eine genauere Betrachtung des Systems; allerdings müssen dazu auch erheblich mehr Koeffizienten während der PLS-Analyse berechnet werden. Für das konkrete Problem bedeutete dies: Bei einer Größe der gesamten Gitterbox von etwa $20 \text{ \AA} \times 26 \text{ \AA} \times 25 \text{ \AA}$ (vgl. Tabelle 24 auf Seite 83) müssen bei einem Gitterabstand von 3 \AA insgesamt 567 ($7 \times 9 \times 9$) Punkte berücksichtigt werden, bei einem Gitterabstand von 1 \AA sind es bereits 13000 ($20 \times 26 \times 25$). Tabelle 23 zeigt, daß im konkreten Fall der Abstand der Gitterpunkte nur äußerst geringen Einfluß auf den q^2 -Wert hatte.

Tab. 23. Der Einfluß des Gitterabstandes auf die den q_{LOO}^2 -Wert, berechnet mit der SAMPLS-Methode für die Gruppen I-IX.

Gitterabstand [\AA]	1.0	1.5	2.0	2.5	3.0
q^2	0.771	0.771	0.766	0.763	0.738
k	7	7	7	7	6

Lediglich bei einer Maschenweite von 3 \AA fällt der q^2 -Wert erkennbar ab. Als Kompromiß zwischen Rechenzeit und Genauigkeit bei der Feldberechnung wurde der Standard-Wert von 2.0 \AA beibehalten.

4.4.5.4 Optimale Anzahl von Hauptkomponenten

Die Hauptkomponentenzahl k ist generell ein Maß für Komplexität eines QSAR-Modells. Zu viele Komponenten führen zu einer Über-Anpassung des Modells an die Daten, was insgesamt zu einer Verschlechterung der Vorhersagekraft führt (vgl. Abschnitt 2.5.1 auf Seite 21). Eine Reduzierung der Komponenten resultiert in robusteren QSAR-Modellen, so daß man in einigen Fällen auf Kosten des q^2 -wertes weniger Komponenten verwendet.^[26,42]

In Abb. 26 ist für einige Feldkombinationen der Verlauf von q^2 gegen die Hauptkomponentenanzahl aufgetragen.

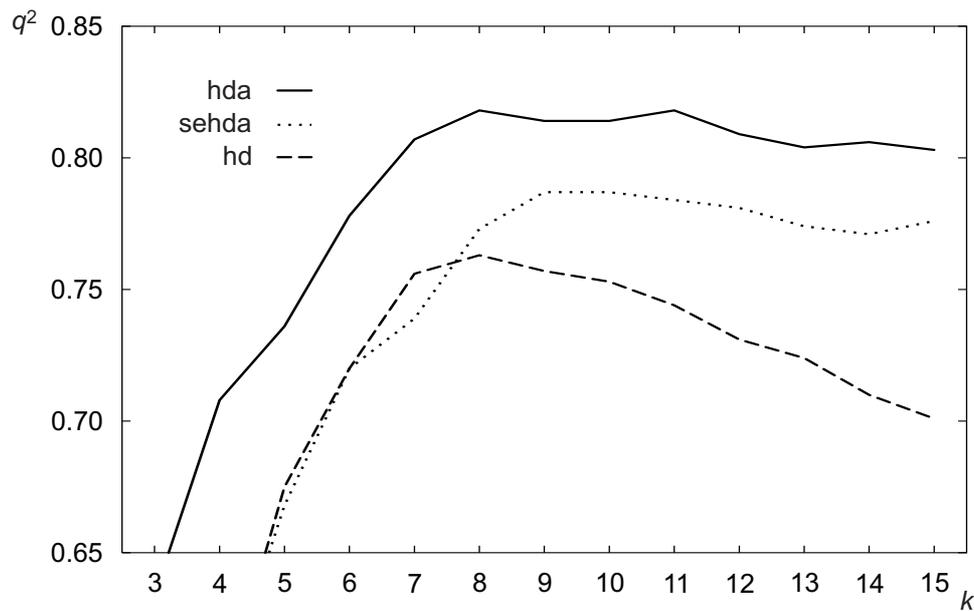


Abb. 26. Verlauf der q^2 -Werte für steigende Hauptkomponentenzahlen k ; gezeigt sind neben der Feldkombination 'hda' exemplarisch die Kombinationen 'sehda' und 'hd'

Man findet für das beste Modell 'hda' bis zu $k = 7$ einen starken Anstieg des q^2 -Wertes; für $k = 8$ ist die Zunahme etwas geringer, aber noch ausreichend, um dieses Modell als das beste anzunehmen. Höhere Komponentenzahlen führten zu keinen weiteren Verbesserungen.

4.4.6 Eigenschaften des endgültigen CoMSIA-Modells

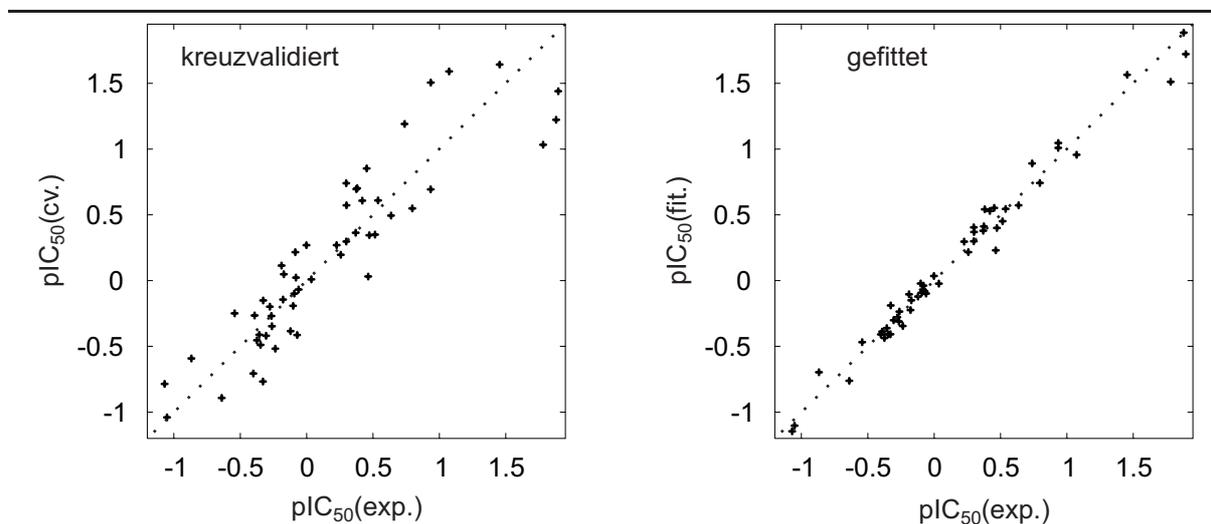
Als Zusammenfassung der Resultate des vorangegangenen Kapitels sind in Tabelle 24 noch einmal die Parameter des Modells mit den optimierten Eigenschaften zusammengefaßt.

Tab. 24. Für die Berechnung verwendete Parameter und Eigenschaften des abschließenden CoMSIA-Modells.

Parameter	
Anzahl der Verbindungen im Trainingsdatensatz	53
Koordinaten des Gitters [Å]:	
x-Richtung	-9.261100 – 10.322200
y-Richtung	-11.897900 – 13.278600
z-Richtung	-12.044300 – 12.335900
Abstand der Gitterpunkte	2 Å
k	8
Eigenschaften des optimalen CoMSIA-Modells	
q^2	0.818
q^2_{SAMPLS}	0.813
PRESS ^a	0.308
r^2	0.985
s	0.101
F	276.9
Anteile der Felder	
Hydrophobie	0.234
H-brücken-Donor	0.431
H-brücken-Akzeptor	0.334

In Tabelle 25 sind sowohl die während der Kreuzvalidierung (cv) gefundenen Abweichungen zu den experimentellen Daten wie auch die Differenz zu den gefitteten Vorhersagen des abschließenden Modells (fit) aufgeführt.

Tab. 25. Abweichungen der kreuzvalidierten Vorhersagen von den tatsächlichen biologischen Aktivitäten (cv) aller Verbindungen im Trainingsdatensatz; außerdem gezeigt ist jeweils die Differenz der mit dem besten CoMSIA-Modell gemachten Vorhersagen (fit) zu den experimentellen Werten.



Gruppe		pIC ₅₀ (exp.)	Abw.(cv)	Abw.(fit)
I	Ancistrocladisin (13)	-0.236	0.282	0.109
I	<i>cis</i> -1,2-Dihydroancistrocladisin (14)	-0.327	-0.176	-0.138
I	<i>trans</i> -1,2-Dihydroancistrocladisin (15)	-0.265	0.004	0.046
I	Ancistrobertsonin D (16)	-0.639	0.253	0.123
I	Dioncopeltin A (3)	1.880	0.657	-0.005
I	<i>N</i> -Benzyl-dioncopeltin A (17)	0.299	-0.441	-0.106
I	<i>N</i> -8- <i>O</i> -Dibenzyl-dioncopeltin A (18)	-0.001	-0.270	-0.036
I	Dioncophyllin A (19)	0.419	-0.189	-0.111
I	5'- <i>O</i> -Demethyl-dioncophyllin A (20)	0.381	-0.321	-0.161
I	Habropetalin A (11)	1.896	0.456	0.174
II	7- <i>epi</i> -Dioncopeltin A (22)	0.538	-0.072	-0.007
II	<i>N</i> -Benzyl-7- <i>epi</i> -dioncopeltin A (23)	0.257	0.061	0.041
II	7- <i>epi</i> -Dioncophyllin A (24)	0.516	0.166	0.065
II	<i>N</i> -Methyl-7- <i>epi</i> -dioncophyllin A (25)	-0.172	-0.221	-0.022
II	Dioncophyllin D (26)	0.797	0.248	0.055
III	Dioncophyllin C (4)	1.782	0.749	0.271
III	Hamatein (31)	-0.261	0.086	-0.026

(Fortsetzung auf der nächsten Seite)

Tab. 25. (Fortsetzung)

Gruppe		pIC ₅₀ (exp.)	Abw.(cv)	Abw.(fit)
III	Ancistrolikokin B (28)	0.226	-0.044	-0.070
III	Ancistrolikokin C (29)	-0.356	0.056	0.005
III	Ancistrobertsonin B (27)	-0.868	-0.276	-0.171
III	N-Methyldioncophyllin C (30)	0.453	-0.400	-0.100
III	Korupensamin B (32)	0.936	-0.569	-0.110
IV	Ancistrocladein (34)	-0.277	-0.079	0.002
IV	Ancistrocladin (35)	-0.542	-0.293	-0.074
IV	N-Methylancistrocladin (36)	-0.329	0.438	0.079
IV	Ancistrocongolin A (37)	0.300	-0.273	-0.070
IV	Ancistroealain A (38)	-0.401	0.305	0.008
IV	Ancistroealain B (39)	-0.102	0.090	-0.080
IV	Ancistrolikokin A (40)	0.464	0.433	0.234
IV	Ancistrolikokin D (41)	-0.305	0.115	-0.004
IV	Ancistrobertsonin A (12)	-1.051	-0.011	0.052
IV	Ancistrobertsonin C (33)	-0.347	0.142	0.061
IV	Korupensamin A (42)	0.738	-0.453	-0.153
V	Dioncophyllin B (43)	0.636	0.142	0.063
V	1- <i>epi</i> -Dioncophyllin B (44)	0.370	0.006	-0.009
V	Dioncophyllinol B (46)	1.074	-0.516	0.117
VI	Knipholon (48)	-0.189	-0.303	-0.085
VI	4- <i>O</i> -Demethylknipholon (50)	-0.071	0.343	0.010
VI	Knipholon-6-methylether (51)	-0.374	0.080	0.062
VI	52	-0.178	-0.034	0.046
VI	53	-0.121	0.263	0.000
VII	54	0.935	0.242	-0.074
VII	55	0.375	-0.320	-0.038
VII	56	-0.080	-0.103	-0.042
VII	9	-0.093	0.006	-0.003
VIII	57	0.473	0.128	0.073
VIII	58	0.299	0.002	0.000
VIII	59	0.036	0.027	0.059

(Fortsetzung auf der nächsten Seite)

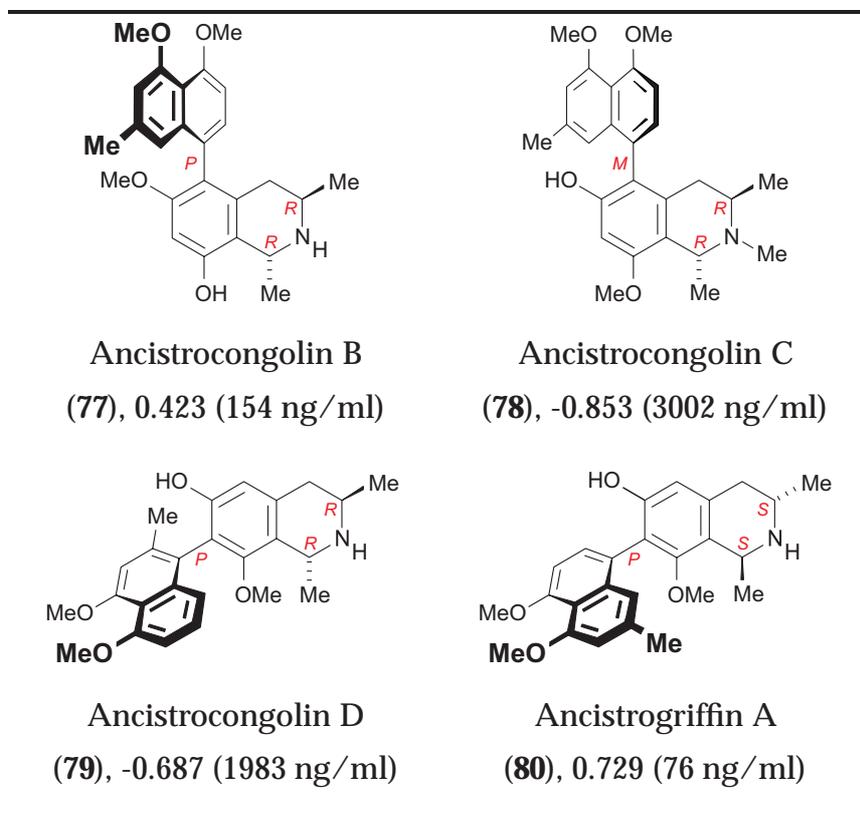
Tab. 25. (Fortsetzung)

Gruppe		pIC ₅₀ (exp.)	Abw.(cv)	Abw.(fit)
VIII	10	-0.060	0.008	0.038
IX	Dioncolacton A (60)	-0.085	-0.301	-0.016
IX	<i>N</i> -Benzylidioncolacton A (61)	-0.392	-0.127	-0.004
	62	1.454	-0.188	-0.111
	64	-1.070	-0.285	0.076

4.4.7 Evaluierung des Modells anhand eines Testdatensatzes

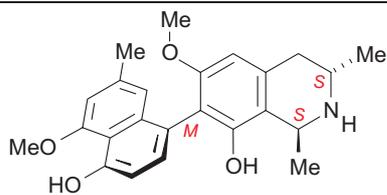
Ein wichtiger Test für ein QSAR-Modell ist die Vorhersage von Verbindungen, die nicht im Trainingsdatensatz enthalten sind. Als Testdatensatz wurden Strukturen verwendet, die während der Erarbeitung dieses Modells synthetisiert oder isoliert und anschließend getestet wurden (siehe Tabelle 26).

Tab. 26. Die Verbindungen des Testdatensatzes.

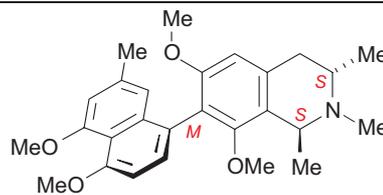


(Fortsetzung auf der nächsten Seite)

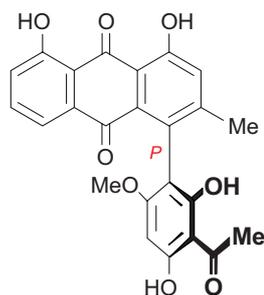
Tab. 26. (Fortsetzung)



Ancistrogriffin C
(81), -0.138 (541 ng/ml)



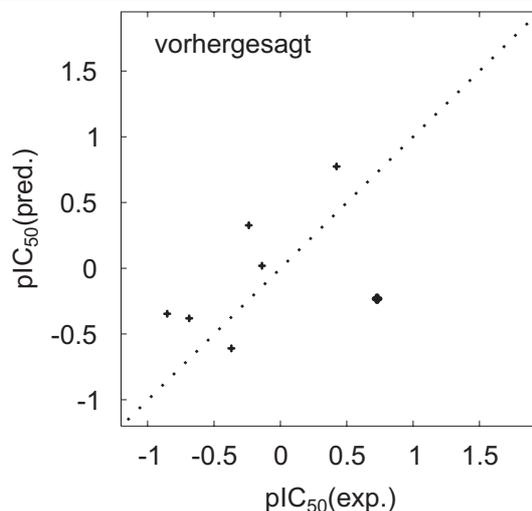
1-*epi*-Ancistrobrevin A
(82), -0.369 (1019 ng/ml)



Bulbine-Knipholon
(83), -0.238 (751 ng/ml)

Die Aktivitäten dieser Verbindungen wurden nun mit dem CoMSIA-Modell vorhergesagt, die dabei erhaltenen Abweichungen sind in Tabelle 27 auf der nächsten Seite zusammengefasst.

Tab. 27. Tatsächliche Aktivitäten und Abweichungen der Vorhersage für die Verbindungen des Testdatensatzes (Tabelle 26); der schlechte Wert für Ancistrogriffin A (**80**) wurde hervorgehoben.



Substanz	pIC ₅₀ (exp.)	Abweichung
Ancistrocongolol B (77)	0.423	-0.351
Ancistrocongolol C (78)	-0.853	-0.507
Ancistrocongolol D (79)	-0.687	-0.306
Ancistrogriffin A (80)	0.729	0.961
Ancistrogriffin C (81)	-0.138	-0.157
1- <i>epi</i> -Ancistrobrevin A (82)	-0.369	0.240
Bulbine-Knipholon (83)	-0.238	-0.565

Die so gefundene Vorhersagekraft eines Modells drückt man mit Hilfe des 'predictive r^2 ' aus.^a Auf Basis dieses Testdatensatzes ergibt sich ein $r_{pred.}^2 = 0.279$ (für 7 Verbindungen).

Dieser etwas enttäuschende Wert beruht im Wesentlichen auf den schlechten Vorhersage für Ancistrogriffin A (**80**). Obwohl diese Verbindung den relativ häufig vorkommenden 7,8'-Kupplungstyp aufweist und deshalb sehr gut durch das Modell vorhergesagt werden sollte, wurde die Aktivität um fast eine Größenordnung falsch eingeschätzt. Inzwischen wurde die Substanz ein zweites mal getestet und der Wert innerhalb einer akzeptablen Toleranz bestätigt. Für die schlechte Vorhersage konnte bisher keine Erklärung gefunden werden.

^a Erklärung siehe Seite 23

Anhand der Vorhersage von Bulbine-Knipholon (**83**) zeigte sich, daß offenbar die Gruppe VI der Phenylanthrachinone noch nicht ausreichend durch das Modell repräsentiert wurde; dies hatte sich schon in Abschnitt 4.4.4.2 (Seite 73) angedeutet. Ein möglicher Grund für die trotzdem relativ guten Vorhersagen der Verbindungen im Modell ist, daß die Aktivitätswerte aller^a Substanzen aus Gruppe VI im Trainingsdatensatz nur eine geringe Varianz aufweisen, zwischen -0.071 für 4-O-Demethylknipholon (**50**) und -0.374 für Knipholon-6-methylether (**51**). Zudem liegen diese Werte im mittleren Bereich der Varianz des gesamten Datensatzes, der von Aktivitäten zwischen 1.896 für Habropetalin A (**11**) und -1.050 für Ancistrobertsonin A (**12**) aufweist. Es besteht die Möglichkeit, daß die Phenylanthrachinone nur 'mitgeschleift' wurden und ihre tatsächliche Wirksamkeit von dem erstellten Modell nicht wiedergegeben werden kann.

Sieht man von den beiden schlechten Vorhersagen (für **80** und **83**) ab, erhält man einen sehr beachtlichen r_{pred}^2 von 0.578 für 5 Verbindungen.

4.5 Interpretation der CoMSIA-Kontur-Maps

Neben der Vorhersage der Aktivität bisher nicht getesteter Substanzen ermöglicht die CoMSIA-Methode, die Koeffizienten aus der QSAR-Gleichung dreidimensional darzustellen. Auf diese Weise können für die Aktivität wichtige funktionellen Gruppen direkt identifiziert werden.

4.5.1 Die Wasserstoffbrücken-Donor- und -Akzeptor-Felder

Das Wasserstoffbrücken-Donor-Feld hat am abschließenden CoMSIA-Modell den größten Anteil (siehe Tabelle 22 auf Seite 79). Für die Simulation von Wasserstoffbrücken-Wechselwirkungen verwendet die CoMSIA-Methode Pseudoatome, die potentielle Bindungspartner des Liganden im Rezeptor darstellen.^[131] Im Donor-Feld werden beispielsweise an einer Hydroxy-Funktion zwei Pseudoatome angehängt, die mögliche Positionen einer Wasserstoffbrückenbindung darstellen. Beispielsweise ist die Aminofunktion im Isochinolinteil (rot gepunkteter Kasten in Abb. 27) eine

^a Abgesehen von Knipholon-Anthron (**49**), das als Outlier ausgeschlossen wurde (Kapitel 4.4.1)

funktionelle Gruppe mit Einfluß auf die biologische Aktivität: Die Pseudoatome am Stickstoff liegen direkt innerhalb der als positiv ausgewiesenen cyan-gefärbten Konturen.

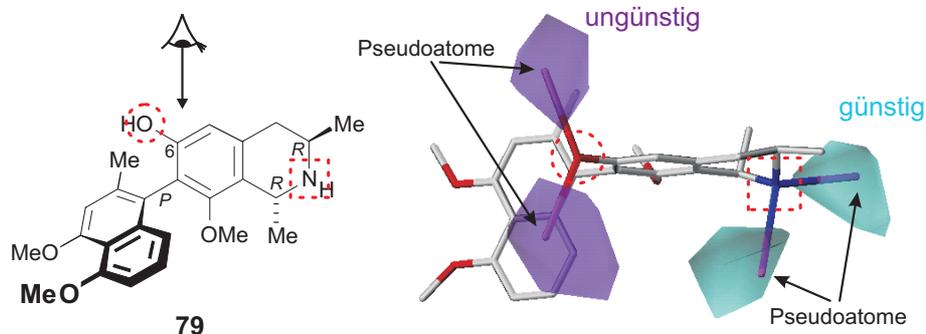


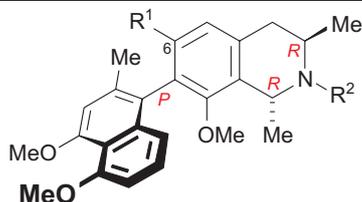
Abb. 27. Ancistrocongol D (**79**) innerhalb des Wasserstoffbrücken-Donor-Feldes; die entscheidenden Gruppen sind in beiden Darstellungen mit rot gepunkteten Markierungen versehen, Pseudoatome wurden durch Pfeile gekennzeichnet

Das Gegenteil gilt für die Hydroxyfunktion an C-6 in **79**, markiert durch einen rot gepunkteten Kreis in Abb. 27. Hier gehen die simulierten Bindungen direkt in die als negativ ausgewiesenen violetten Konturen. Wird diese Funktion (beispielsweise durch einen Methylrest) blockiert, so kann sie keine Wasserstoffbrücken-Donor-Wechselwirkung mehr ausüben. Eine solche Modifikation sollte sich entsprechend positiv (für 6-O-Methyl) oder negativ (für N-Methyl) auf die Aktivität von Ancistrocongol D (**79**) auswirken.

Diese Feststellung konnte anhand des folgenden Tests bestätigt werden: Die Berechnung der Aktivität durch das CoMSIA-Modell für **79** wurde den Vorhersagen für entsprechend modifizierte Analoga gegenübergestellt.^a Tabelle 28 faßt die Ergebnisse zusammen.

^a Die veränderten Strukturen wurden nicht neu mit FLEXS ausgerichtet, sondern lediglich das bereits vorhandene Alignment entsprechend modifiziert.

Tab. 28. Der durch das CoMSIA-Modell vorhergesagte Einfluß von Methylgruppen an den entscheidenden Wasserstoffbrücken-Donor-Funktionen in Ancistrocongol D (**79**).



79,79a-c

	R ¹	R ²	pIC ₅₀ (pred.)
Ancistrocongol D (79)	OH	H	-0.380
79a	OMe	H	0.178
79b	H	H	0.313
79c	OH	Me	-1.283

Wie vermutet, führt eine Methoxygruppe an C-6 zu einem Anstieg der berechneten Aktivität (s. Substanz **79a**). Durch die komplette Entfernung der Sauerstoff-Funktion (vgl. Struktur **79b**) verbessert sich die Aktivität nochmals. Erklärt werden kann dies über das Wasserstoffbrücken-Akzeptor-Feld. Dort findet man ebenfalls den negativen Einfluß der Hydroxyfunktion in Position 6 (siehe Abb. 28 oben), ausgewiesen durch die blauen Kontur-Plots um die entsprechenden Pseudoatome herum.

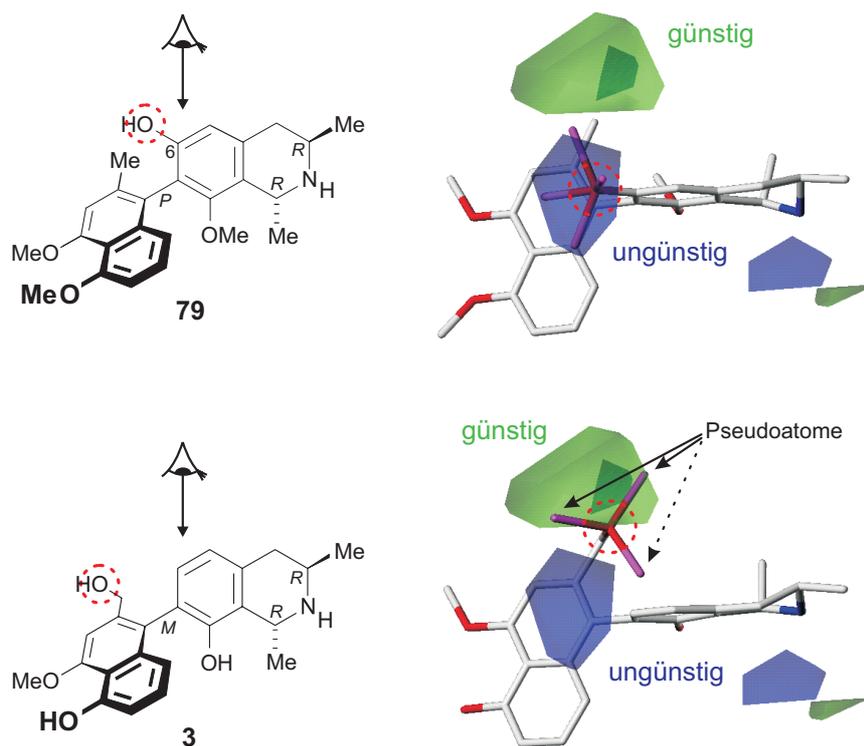
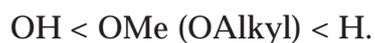


Abb. 28. Ancistrocongolins D (**79**) (oben) und Dioncopeltin A (**3**) (unten), versehen mit entsprechenden Pseudoatomen an den markierten funktionellen Gruppen, innerhalb des Wasserstoffbrücken-Akzeptor-Feldes

Im Gegensatz zu den Donor-Funktionen werden die Wasserstoffbrücken-Akzeptor-Eigenschaften von Hydroxy- oder Aminofunktionen durch Schutzgruppen zwar eingeschränkt, aber nicht vollständig unterbunden. So ist zu erklären, daß eine Verbindung mit geschützter Sauerstoff-Funktion wie **79a** (siehe Tabelle 28 auf der vorherigen Seite) schlechter eingeschätzt wird als eine 6-deoxygenierte Struktur wie **79b**. Folglich kann man für Substituenten an C-6 im Tetrahydroisochinolinring folgende Aktivitätsreihenfolge aufstellen (sortiert nach steigender Aktivität):



Die grüne Kontur (Abb. 28 unten) ist offensichtlich der Hydroxymethylgruppe in Dioncopeltin A (**3**) zuzuordnen. Die durch Pfeile markierten Pseudoatome liegen (bis auf eines, s. gestrichelter Pfeil) alle direkt in der als positiv ausgewiesenen Zone. Das dritte Pseudoatom kommt dem aromatischen Ring des Tetrahydroisochinolinteils im Molekül sehr nahe und wird deshalb für die Auswertung nicht

berücksichtigt.^a

Die entsprechende funktionelle Gruppe (2'-CH₂OH) tritt in sieben Verbindungen im Datensatz auf. Drei der vier aktivsten Substanzen gehören dazu: Neben Dioncopeltin A (**3**) sind dies Habropetalin A (**11**) (siehe Seite 35) und das Phenyltetrahydroisochinolin **62** (Seite 41). Diese Gruppe ist offensichtlich enorm von Vorteil für die Antimalaria-Aktivität.

Die übrigen vier Verbindungen, die ebenfalls diese Gruppen tragen, sind 7-*epi*-Dioncopeltin A (**22**), *N*-Benzyl-7-*epi*-dioncopeltin A (**23**), Dioncolacton A (**60**) und *N*-Benzyl-dioncolacton A (**61**). Diese zeigen aus anderen Gründen keine guten Aktivitäten: Dies sind die ungünstigere Achsenkonfiguration (**22**, **23**), eingeschränkte Beweglichkeit an der Biarylachse (**60**, **61**) und die Gegenwart von Benzylschutzgruppen (**23**, **61**).

4.5.2 Das Hydrophobie-Feld

Diese Feldfunktion unterstützt ebenfalls die Wichtigkeit der 2'-CH₂OH-Funktion in Dioncopeltin A (**3**);^b graue Kontur-Flächen sind hier einem negativem Einfluß von hydrophoben Gruppen zugeordnet, im Umkehrschluß führen in diesen Bereichen hydrophile Gruppen zu höherer Aktivität (siehe Abb. 29).

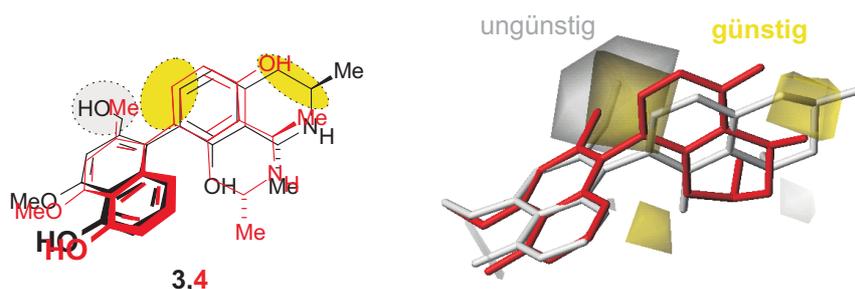


Abb. 29. Das Alignment von Dioncopeltin A (**3**) (grau) und Dioncophyllin C (**4**) (rot) innerhalb des Hydrophobie-Plots

^a Der erlaubte Grenzabstand zu anderen Atomen beträgt 1.8 Å. Bei kleinerer Distanz wäre an einer solchen Stelle die Annahme einer Ligand-Rezeptor-Wechselwirkung kaum sinnvoll.

^b Diese Abhängigkeiten zwischen Wasserstoffbrücken-Akzeptor- und Hydrophobie-Feld erklären die Tatsache, daß im abschließenden CoMSIA-Modell der Anteil des Akzeptor-Feldes größer ist als der des Hydrophobie-Feldes (s. Tabelle 24), obwohl dieses für sich alleine eine signifikantere Korrelation aufwies; siehe dazu Abschnitt 4.4.5.1 auf Seite 79.

Sehr interessant sind die beiden großen gelben Bereiche um die Tetrahydroisochinolinteile herum (Abb. 29). Sie zeigen Gebiete, die jeweils entweder von 5- oder 7-gekoppelten Naphthylisochinolin-Alkaloiden okkupiert werden. Weil aktive Substanzen beider Kupplungstypen vorliegen, werden hydrophobe Atome in beiden Bereichen als positiv eingeschätzt.

4.5.3 Elektrostatisches und sterisches Feld

Diese beiden Felder hatten bei der Analyse der einzelnen Beiträge zum CoMSIA-Modell die geringste Signifikanz gezeigt (s. Abb. 24 auf Seite 78) und waren deshalb im abschließenden Modell nicht berücksichtigt worden. Die hier diskutierten Kontur-Plots (Abb. 30) basieren auf dem Modell, das unter Verwendung aller fünf CoMSIA-Funktionen erstellt wurde (rechte Spalte in Tabelle 22 auf Seite 79).

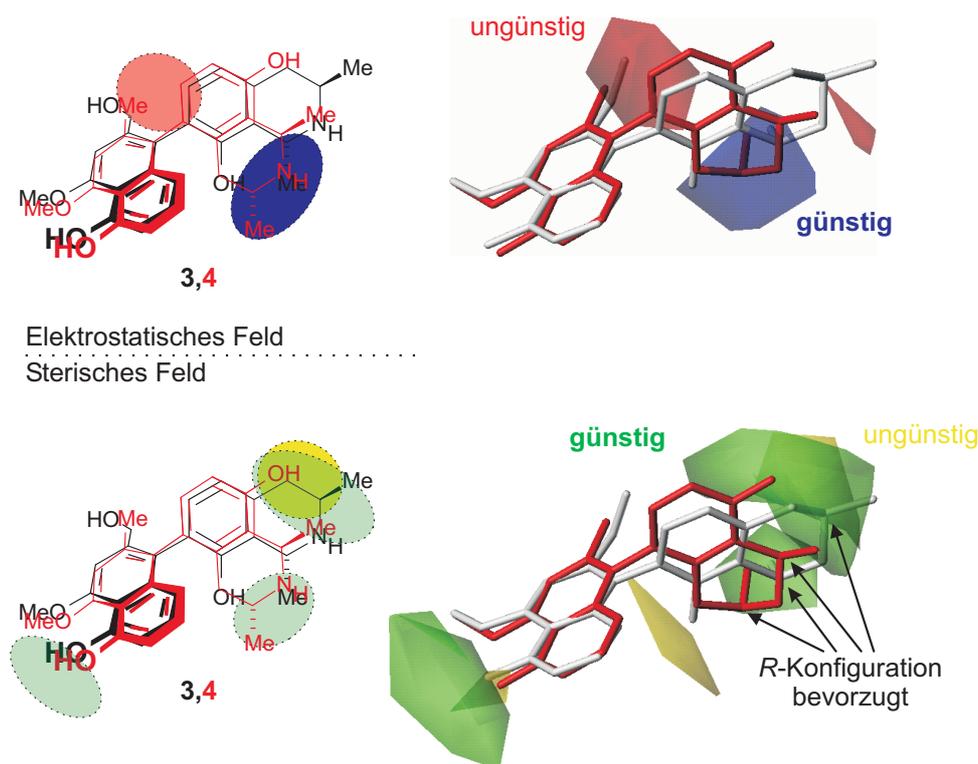


Abb. 30. Das Alignment von Dioncophyllin C (4) (rot) und Dioncopeltin A (3) (schwarz) innerhalb der Konturen des elektrostatischen (oben) und des sterischen Feldes (unten)

Das elektrostatische Feld (Abb. 30 oben) zeigt zwei bereits durch andere Felder

beschriebene Eigenschaften: Die roten Konturen (negativer Einfluß von elektronegativen Atomen) liegen im Bereich der Sauerstoff-Funktion in 6-Position, die blauen Bereiche (positiver Einfluß) sind wiederum dem Stickstoff-Atom zuzuordnen.

Das sterische Feld (Abb. 30 unten) zeigt eine sehr interessante Verteilung: Die grünen (sterische Hinderung erwünscht) und die gelben (sterische Hinderung hat negativen Einfluß) Konturen um den Isochinolinteil sind in einer Weise angeordnet, die andeutet, daß die *R*-Konfiguration an C-1 und C-3 (wie in **3** und **4**) gegenüber der *S*-Konfiguration bevorzugt ist.^a Die grüne Kontur im Bereich der 5'-Hydroxyfunktionen könnte einer Methylgruppe in diesem Bereich zugeordnet werden, allerdings hat das sterische Feld mit insgesamt weniger als 5% der Gesamt-Information im Modell (vgl. Tabelle 22 auf Seite 79) nur eine äußerst geringe Bedeutung, und man sollte deshalb diese Ergebnisse nicht überbewerten.

4.5.4 Zusammenfassung der Feldeigenschaften: Ein Pharmacophor-Modell

Die wichtigsten Resultate aus den Analysen der CoMSIA-Felder wurden in einem Pharmacophor-Modell zusammengefaßt, das in Abb. 31 schematisch dargestellt ist. Die dazu führenden Überlegungen folgen im Anschluß daran.

^a Man beachte, daß C-1 in Dioncopeltin A (**3**) mit C-3 in Dioncophyllin C (**4**) überlagert ist.

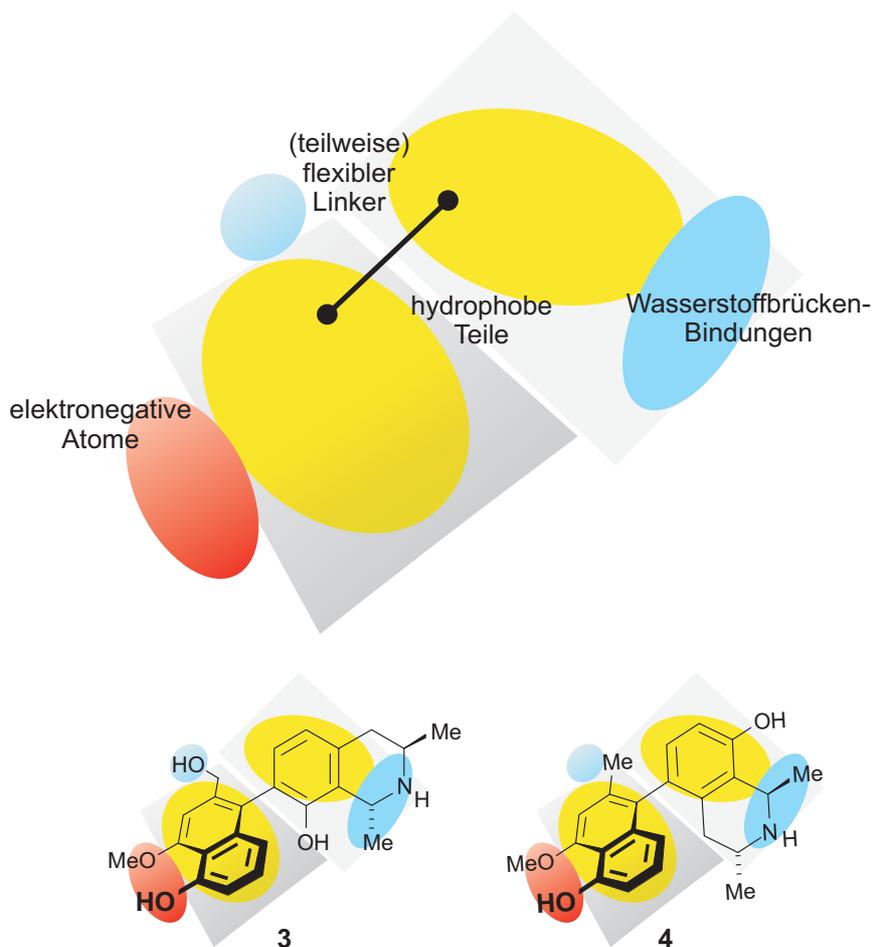


Abb. 31. Ein schematisches Pharmacophor-Modell für die Antimalaria-Aktivität von Naphthylisochinolin-Alkaloiden, beispielhaft gezeigt anhand von **3** und **4**

An dieser Stelle ist es wichtig zu bedenken, daß die CoMSIA-Methode den Einfluß bestimmter Gruppen nur dann identifizieren kann, wenn in den entsprechenden Bereichen über den gesamten Trainingsdatensatz eine gewisse Diversität vorliegt. Beispielsweise gibt es eine gute Verteilung zwischen drei verschiedenen Substituenten in Position 6 (Wasserstoff, Hydroxy- oder Methoxygruppe), so daß hier eine eindeutige Differenzierung erfolgen konnte. Elemente, die in der überwiegenden Zahl aller Verbindungen gleich sind, werden innerhalb der CoMSIA-Konturen nicht ausgewiesen.

Dieses gilt für die Naphthylhälfte mit den beiden Sauerstoffsubstituenten und dem großen aromatischen Bereich, nicht aber für die 2'-CH₂OH-Funktion.

Das erstellte Pharmacophor-Modell (Abb. 31) kommt wie folgt zustande:

1. Der zentrale Bestandteil sind zwei hydrophobe Molekülhälften (gelbe Bereiche in Abb. 31), verbunden durch einen (nicht notwendigerweise) flexiblen Linker. Über den in fast allen Verbindungen vorhandenen aromatischen Naphthyl- oder Phenylteil hinaus ist eine zweite aromatische oder zumindest hydrophobe Hälfte notwendig, erkennbar im Hydrophobie-Feld (Abb. 29 auf Seite 93).
2. Zur Konfiguration der Achse wurde bereits im früher erstellten CoMFA-Modell gezeigt^[24,25], daß Dioncopeltin A (**3**) und Dioncophyllin C (**4**) jeweils die günstigeren der beiden theoretisch denkbaren Atropdiastereomere sind. Auch an den sehr unterschiedlichen Aktivitäten von Dioncopeltin A (**3**) ($pIC_{50} = 1.880$) und seinem Atropdiastereomer *7-epi*-Dioncopeltin A (**22**) ($pIC_{50} = 0.538$) ist dies erkennbar. Die günstigste räumliche Form ist also die in Abb. 31 anhand von **3** und **4** gezeigte.
3. Wasserstoffbrücken-Wechselwirkungen (cyan-gefärbte Bereiche in Abb. 31) spielen offenbar beim Bindungsmodus der Naphthylisochinolin-Alkaloide eine wichtige Rolle. Die Bedeutung des Stickstoff-Atoms im Tetrahydroisochinolin-teil (Wasserstoffbrücken-Donor-Feld) und die Hydroxymethylfunktion in 2'-Position von Dioncopeltin A (**3**) konnten eindeutig nachgewiesen werden (Wasserstoffbrücken-Akzeptor- und Hydrophobie-Feld). Die Tatsache, daß die Aktivität bei Abschirmung durch Schutzgruppen dramatisch absinkt (vgl. Abschnitt 4.5.1 auf Seite 93), weist darauf hin, daß die Wechselwirkungen nicht nur auf die elektrostatischen oder hydrophoben Eigenschaften der Stickstoff- und Sauerstoff-Atome zurückgehen kann.
4. Für die beiden Sauerstofffunktionen in den Positionen 4' und 5' (rot unterlegter Bereich in Abb. 31) gilt ähnliches wie für die hydrophoben Bereiche, denn auch hier weist der Datensatz keine große Varianz auf.^a Möglicherweise konnte deshalb ein entscheidender Einfluß einer phenolischen OH-Gruppe im Naphthylteil, wie er bisher postuliert wurde,^[24,25] nicht gefunden werden. Die hier erzielten Ergebnisse stehen im Einklang mit experimentellen Daten,

^a Durch diese Tatsache kann erklärt werden, warum die manuelle Modifikation des FLEXS^{II}-Alignments zu FLEXS^{III} keine Verbesserung brachte (Siehe Kapitel 4.3.3.2 auf Seite 62.

denn auch Verbindungen ohne phenolische OH-Gruppe im Naphthylteil zeigen hervorragende Aktivitäten, wie zum Beispiel Habropetalin A (**11**) (siehe Seite 35) oder das Phenyltetrahydroisochinolin **62** (siehe Seite 41).

Weitere Eigenschaften sind in Abb. 31 nicht explizit dargestellt:

5. Eine Sauerstoff-Funktion (Hydroxy- oder Methoxygruppe) in Position 6 hat stark negativen Einfluß auf die Antimalaria-Aktivität. Dies ist anhand des Wasserstoffbrücken-Donor-Feldes eindeutig erkennbar (Abb. 27 auf Seite 90). Dieser Punkt widerspricht den Resultaten aus dem früheren CoMFA-Modell,^[25] das einer solchen Gruppe positiven Einfluß auf die Aktivität zuwies (vgl. auch Abb. 10 auf Seite 32).
6. Im sterischen CoMSIA-Feld gibt es Hinweise darauf, daß eine *R*-Konfiguration sowohl an C-1 als auch an C-3 vorteilhaft sein könnte.

4.6 Abschätzung der Aktivität modifizierter Strukturen und neu isolierter Naturstoffe

4.6.1 Mögliche strukturelle Vereinfachungen

Aufbauend auf das beschriebene Pharmacophor-Modell wurde versucht, eine Strategie zu entwickeln, ausgehend von den Naturstoffen Dioncophyllin C (**4**) und Dioncopeltin A (**3**) die komplizierten Strukturelemente von Naphthylisochinolin-Alkaloiden unter Erhalt der biologischen Eigenschaften zu vereinfachen. Dazu wurden die Strukturen in mehreren Schritten immer weiter vereinfacht und die Aktivitäten der entstehenden Verbindungen mit dem CoMSIA-Modell vorhergesagt.

Dabei wurden folgende Überlegungen verfolgt:

1. Schritt: Die Vereinfachung des Naphthylteils hin zu einem Phenylring mit nur einer phenolischen Sauerstoff-Funktion hat laut CoMSIA-Modell weder für Dioncopeltin A (**3**) noch für Dioncophyllin C (**4**) großen Einfluß auf die

Aktivität (Abb. 32); für die Synthese ermöglicht diese Veränderung allerdings erhebliche Vereinfachungen, da keine Axialchiralität mehr vorliegt.

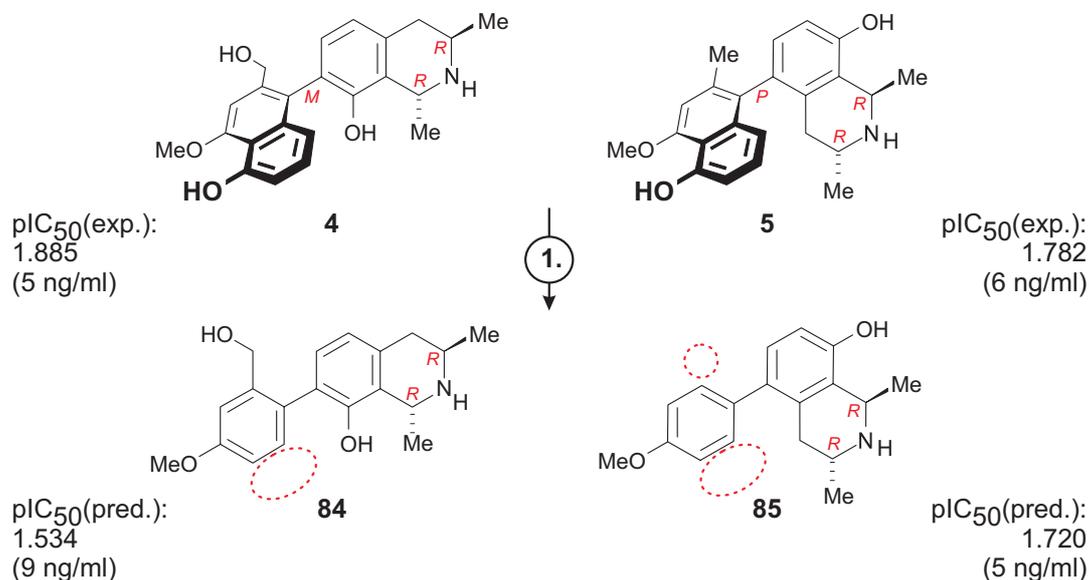


Abb. 32. Die Reduzierung des Naphthylrestes auf einen Phenylrest und der Einfluß auf die berechnete Aktivität

Verbindung **84** wurde in leicht modifizierter Form bereits synthetisiert^[105] und getestet (vgl. **62** auf Seite 41, pIC₅₀ = 1.454). Der für **84** vorhergesagte pIC₅₀-Wert lag bei 1.534. Die gleiche Veränderung (bei zusätzlicher Entfernung der Methylgruppe an C-2') führte für Dioncophyllin C (**4**) zu praktisch keiner Abnahme der Aktivität (Abb. 32 rechts); das entstandene Phenyltetrahydroisochinolin **85** war in entschützter Form ebenfalls bereits bekannt^[106] (vgl. **63** auf Seite 41), war aber aufgrund der freien OH-Gruppen chemisch sehr labil. Die Methoxygruppe an 4'-Position sollte die Empfindlichkeit der Substanz gegenüber Oxidation verringern und so zu deren Stabilität beitragen.^a

- Schritt: Der Aufbau der beiden Stereozentren im Tetrahydroisochinolinring kann durch die Plazierung der beiden Methylgruppen an dasselbe C-Atom umgangen werden. Zudem ermöglicht diese Modifikation für Wasserstoffbrückenbindungen einen leichteren Zugang zum Stickstoff-Atom (siehe **86** und **87** in Abb. 33).

^a Vgl. hierzu die Diskussion der Outlier, besonders Verbindung **63**, in Abschnitt 4.4.1.2 auf Seite 66

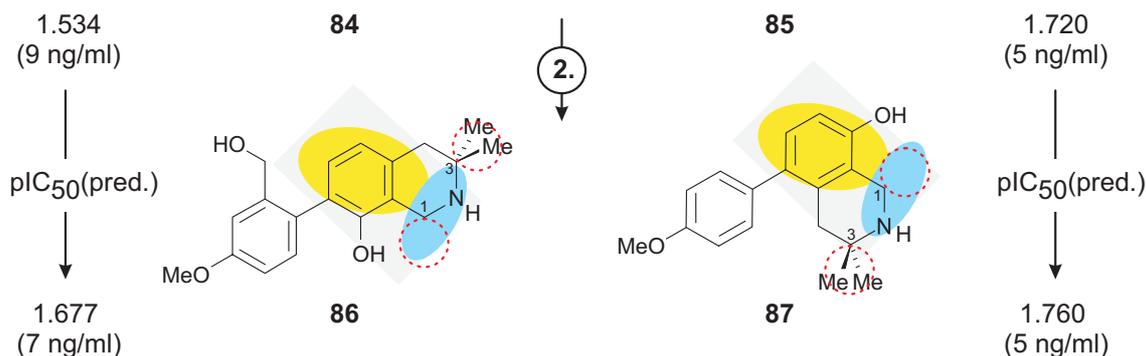


Abb. 33. Die Positionierung beider Methylgruppen im Tetrahydroisochinolinring am gleichen C-Atom

Aufgrund des vom Modell für entsprechende Wasserstoffbrückenbindungen als günstig ausgewiesenen Bereichs (cyan in Abb. 33) sollte C-3 besser geeignet sein als C-1; würden beide Methylgruppen statt dessen an C-1 positioniert, so lägen sie direkt in dieser Zone (siehe rot gepunktete Kreise in Abb. 33). Das Modell schreibt in dieser Art modifizierten Verbindungen sogar leicht verbesserte Aktivitäten zu. Der Syntheseweg zu derartigen Verbindungen ist durch die Bereitstellung eines entsprechend modifizierten Isochinolin-Bausteines^[109] zum Teil bereits erschlossen (siehe unten).

3. Als weitere Vereinfachung ist denkbar, die Hydroxyfunktion an C-8 ebenfalls zu entfernen (siehe **88** und **89** in Abb. 34).

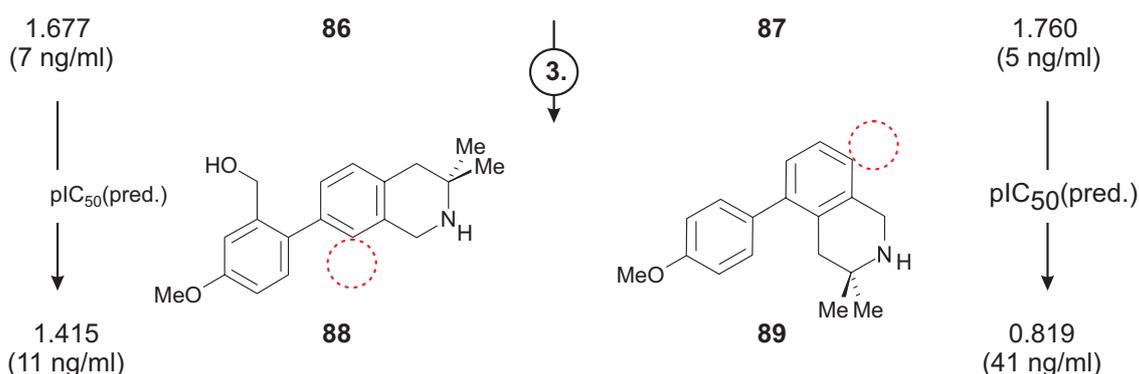


Abb. 34. Die Entfernung der Hydroxyfunktion an C-8 und die unterschiedlichen Auswirkungen für beide Kupplungstypen

Diese Funktion liegt für in 5- und in 7-Position gekuppelte Verbindungen in unterschiedlichen Raumbereichen, weshalb laut Modell unterschiedliche Aus-

wirkungen zu erwarten sind: Während die Aktivität für **88** nur leicht zurückgehen sollte [$\Delta pIC_{50}(\text{pred.}) = -0.262$], bricht sie laut Vorhersage für **89** deutlich ein [$\Delta pIC_{50}(\text{pred.}) = -0.941$] (Abb. 34).

4. (a) Ein interessanter, aber einschneidender Versuch wäre es, den aliphatischen Teil des Tetrahydroisochinolinrings durch ein primäres Amin zu ersetzen (s. **90** in Abb. 35).

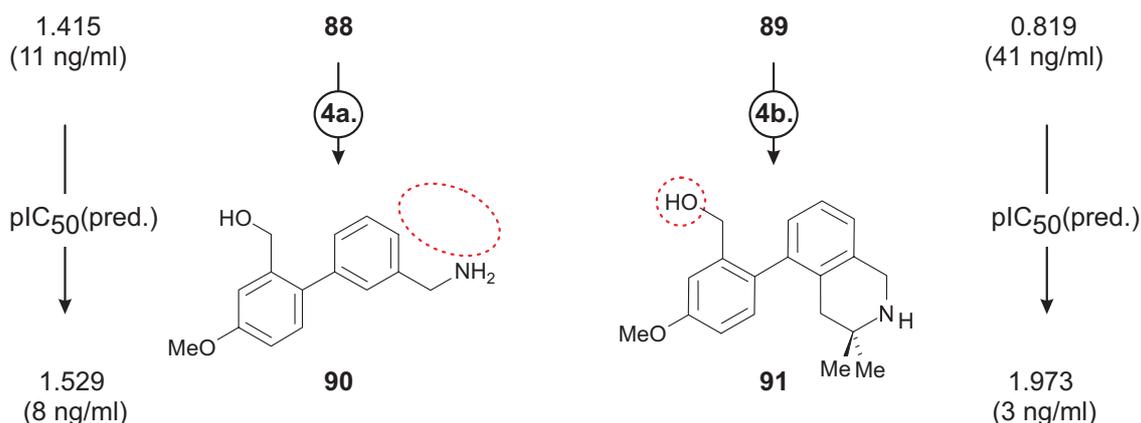


Abb. 35. Die Reduzierung des aliphatischen Teils des Tetrahydroisochinolinringes auf eine Aminfunktion (links) und die Einführung der vielversprechenden 2'- CH_2OH -Funktion in ein 5-gekuppeltes Phenyltetrahydroisochinolin

Das Modell sagt für die so modifizierte Verbindung **90** sogar eine Verbesserung des pIC_{50} -Wertes voraus, allerdings muß dies wegen der sehr drastische Veränderung des Grundgerüsts vorsichtig betrachtet werden.

- (b) Sehr aussichtsreich ist die Einführung einer Hydroxymethylgruppe in einer 5-gekuppelten Verbindung, wie **91** (Abb. 35). Das Modell sagt für diese Modifikation eine aufgrund der Erkenntnisse aus dem Hydrophobie-Plot zu erwartende (vgl. Abschnitt 4.5.2 auf Seite 94) drastische Verbesserung voraus.

Da die Entfernung der Hydroxyfunktion in **87** zu einer drastischen Verschlechterung geführt hatte (s. Abb. 34), wurde neben **90** und **91** auch die Aktivität für **92** (siehe Abb. 36) berechnet. Diese Struktur zeigte mit $pIC_{50} = 2.473$ (1 ng/ml) die bisher beste berechnete Aktivität und ist damit ein aussichtsreiches Syntheseziel.

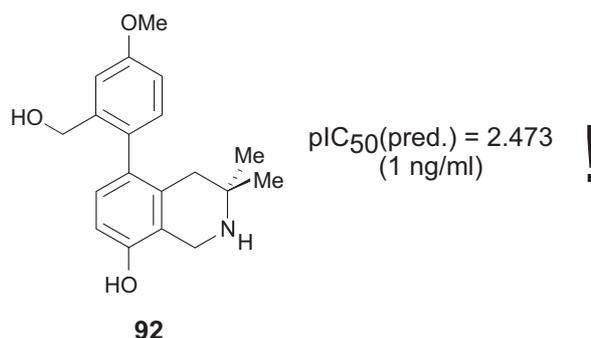


Abb. 36. Die Verbindung mit der bisher besten vorhergesagten Aktivität

4.6.2 Weitere im Isochinolinteil vereinfachte Strukturen

Wie bereits angedeutet, konnten von Neumann^[109] verschiedene Kupplungsbausteine synthetisiert werden, die beide Methylgruppen des Isochinolinteils an C-3 tragen (siehe Abb. 37). Die Einführung des für die Biarylkupplung nötigen Halogenatoms gelang in Position 5 ebenfalls bereits. Auch erste Kupplungsversuche wurden inzwischen durchgeführt.

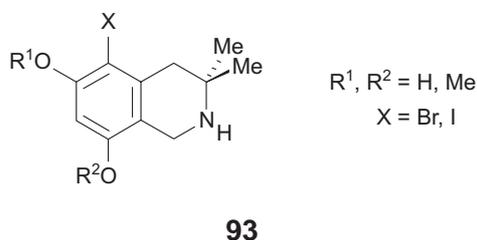
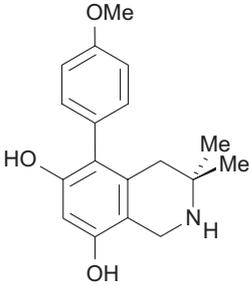
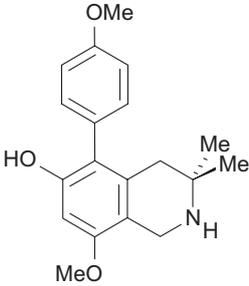
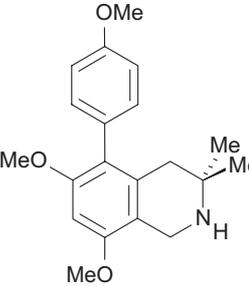
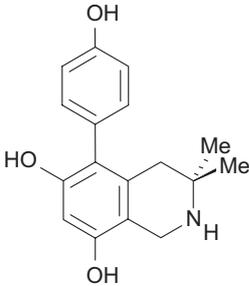
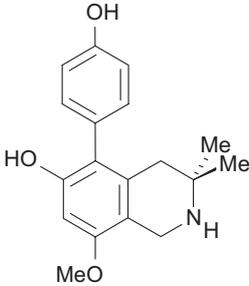
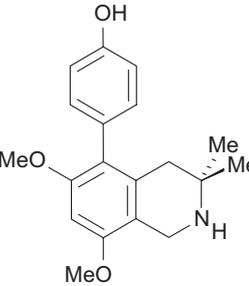
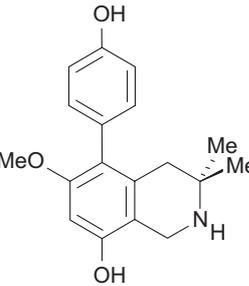
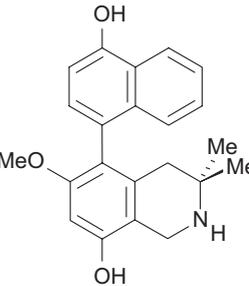
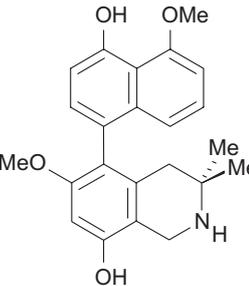


Abb. 37. Von Neumann in unserer Arbeitsgruppe synthetisierte Kupplungsbausteine

Über die im vorhergehenden Abschnitt hinaus gemachten Vorschläge wurde für eine detaillierte Abschätzung, welche Kupplungspartner für **93** am besten geeignet sein sollten, die Aktivität verschiedener Zielmoleküle vorausgesagt. Die Resultate sind in Tabelle 29 auf der nächsten Seite zusammengefaßt.

Tab. 29. Vorhersagen der Aktivität von neuen, noch nicht bekannten Verbindungen, die sich ausgehend von modifizierten Isochinolinbausteinen (siehe Abb. 37) synthetisieren lassen sollten.

 <p>94 0.999 (30 ng/ml)</p>	 <p>95 0.137 (229 ng/ml)</p>	 <p>96 0.642 (75 ng/ml)</p>
 <p>97 1.171 (19 ng/ml)</p>	 <p>98 0.359 (131 ng/ml)</p>	 <p>99 0.814 (48 ng/ml)</p>
 <p>100 1.743 (5 ng/ml)</p>	 <p>101 1.614 (9 ng/ml)</p>	 <p>102 1.699 (8 ng/ml)</p>

Insgesamt zeigten die vorhergesagten Aktivitäten zu erwartende Trends. Die Demethylierung der Methoxygruppe in 4'-Position erhöht den pIC_{50} -Wert bei allen drei Verbindungen um etwa 0.2 (siehe **97**, **98** und **99** in Tabelle 29). Die Methylierung der Hydroxyfunktion an C-6 erhöht die Aktivität (erwartungsgemäß) drastisch auf

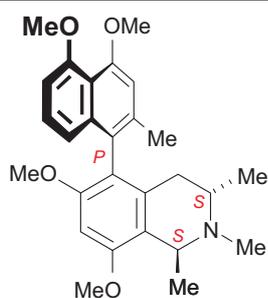
1.743 (5 ng/ml) für **100**. Da diese Verbindung (wie auch **94** und **97**) wahrscheinlich sehr oxidationsempfindlich sein wird, könnte man die Möglichkeit zur Einebnung der beiden Biarylhälften einschränken^a durch den Schritt von Phenol hin zu Naphthol, was aber zu einem Rückschritt in der Aktivität führt. Verbindung **102** ähnelt bereits sehr der Struktur von Dioncophyllin C (**4**), und zeigt deshalb laut Modell entsprechend gute Aktivität.

Insgesamt wurden keine herausragenden Aktivitäten vorhergesagt. Für bessere Resultate sollte vor allem versucht werden, die Sauerstoff-Funktion an C-6 zu entfernen. Dies hatten ja bereits die Ergebnisse des vorhergehenden Kapitels gezeigt.

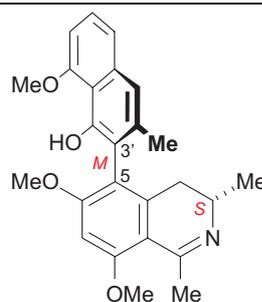
4.6.3 Neu isolierte Verbindungen

Nach der Fertigstellung des CoMSIA-Modells wurden in unserem Arbeitskreis^[132,133] weitere Naphthylisochinolin-Alkaloide isoliert. Für bisher nicht auf Antimalaria-Aktivität getestete Verbindungen sind die Vorhersagen des Modells in Tabelle 30 zusammengefaßt.

Tab. 30. Neue in unserem Arbeitskreis isolierte Verbindungen, deren Aktivität noch nicht getestet wurde.



Ancistrotaninanin E^[132]
(**103**), -0.303 (875 ng/ml)

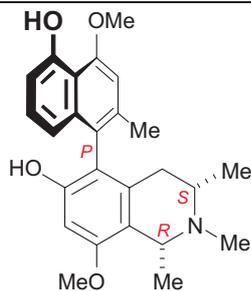


Ancistrotaninanin A^[132]
(**104**), 0.212 (249 ng/ml)

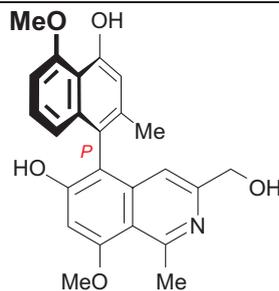
(Fortsetzung auf nächster Seite)

^a Der größere Naphthylrest verhindert die für eine Oxidation zum Diphenochinon erforderliche Einebnung der beiden durch die Achse verbundenen Aromaten und erhöht so die Stabilität der entsprechenden Verbindung.

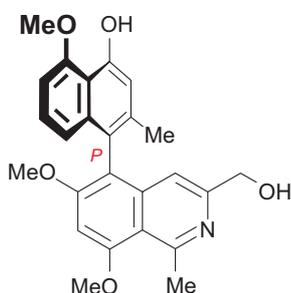
Tab. 30. (Fortsetzung)



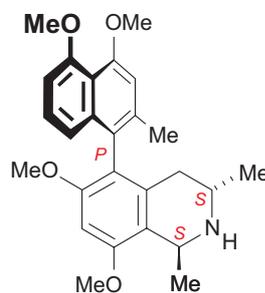
5'-O-Demethylancistroclin^[132]
(105), -0.793 (2532 ng/ml)



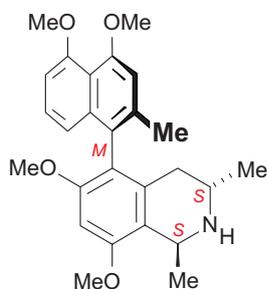
106^[132]
-0.663 (1865 ng/ml)



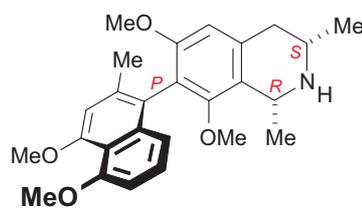
107^[132]
-0.211 (681 ng/ml)



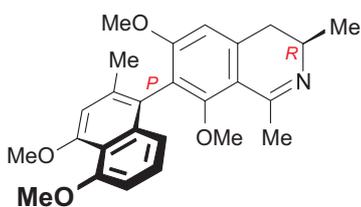
6-O-Methylancistrocladin^[133]
(108), -0.014 (435 ng/ml)



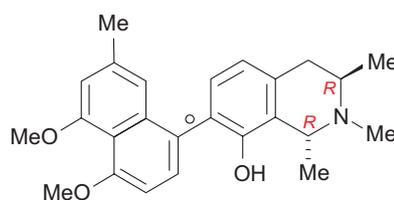
6-O-Methylhamatin^[133]
(109), -0.215 (691 ng/ml)



cis-Dihydroancistrocladisins^[133]
(110), -0.317 (874 ng/ml)



3-epi-Ancistrocladisins^[133]
(111), -0.332 (901 ng/ml)



5'-O-Methyl-N-Methyldioncophyllin D^[133]
(112), -0.242 (684 ng/ml)

Die Untersuchung zur Aufklärung der Achsenkonfiguration der 5,3'-gekuppelten Verbindung Ancistrotanzanin A (**104**) wurde erst kürzlich abgeschlossen.^[134] Für die entsprechende, *P*-konfigurierte Struktur war zuvor ein pIC₅₀-Wert von -0.515 (1326 ng/ml) berechnet worden.

Insgesamt wurden für die neuen Verbindungen keine herausragenden Aktivitäten gefunden. Die Hydroxymethylgruppe in Position 3 im Isochinolinteil von **106** und **107** (Tabelle 30 auf Seite 104) ist ein bisher unbekanntes Strukturelement und liegt deshalb im Trainingsdatensatz nicht vor. Ob und in welcher Weise diese Funktion die Antimalaria-Aktivität beeinflusst, können erst die biologischen Tests zeigen.

4.7 Vergleich mit Chloroquin (1) und ähnlichen Wirkstoffen

Nach der Identifizierung der entscheidenden funktionellen Gruppen und deren relativer Lage im Raum erschien es interessant zu prüfen, ob und wie weit die Aktivitäten anderer Malaria-Wirkstoffe mit dem Pharmacophor-Modell in Einklang gebracht werden können. Dies könnte auf Analogien im Wirkmechanismus hinweisen.

Die Wirkungsweise von Chloroquin (**1**) wurde in Abschnitt 3.1.2 auf Seite 26 beschrieben; entsprechende Versuche mit freiem Häm haben Hinweise ergeben, daß Naphthylisochinolin-Alkaloide eine ähnliche Affinität zu Häm aufweisen,^[13] wengleich die Wechselwirkung des untersuchten Dioncopeltin A (**3**) im Vergleich zu **1** relativ schwach ausfiel.^[13] Die Befunde deuteten auf Analogien der Wirkungsweise der beiden Substanzklassen hin, allerdings ist das komplette Wirkspektrum von Naphthylisochinolin-Alkaloiden wahrscheinlich komplexer. Dafür spricht auch, daß deren Antimalaria-Aktivität bereits exoerythrocytisch nachgewiesen werden konnte.^[10]

Dennoch erschien es lohnend zu überprüfen, ob die Aktivität von Chloroquin (**1**) und ähnlichen Substanzen mit den Ergebnissen des CoMSIA-Modells vergleichbar sind.

Chloroquin (**1**) wurde bei der Durchführung aller derjenigen Tests, auf deren Basis diese Arbeit erstellt wurde, als Vergleichssubstanz ebenfalls vermessen. Der dabei erzielte IC_{50} -Wert liegt im Normalfall bei etwa 50 ng/ml (entspricht einem pIC_{50} -Wert von 0.806) für den Chloroquin-resistenten K1-Stamm, die Werte für den nicht-resistenten NF54-Stamm liegen bei etwa 0.5 ng/ml ($pIC_{50} = 2.806$). Für die Naphthylisochinolin-Alkaloide zeigten die Testwerte gegen den Chloroquin-sensitiven NF-54-Stamm zwar moderat abweichende Werte, nicht aber grundsätzlich verschiedene Tendenzen.

Als weitere, möglicherweise vergleichbare Strukturen wurden die bereits erwähnten (Kapitel 3.1.3 auf Seite 27) Verbindungen Chinin (**5**), Mefloquin (**2**) und Primaquin (**6**) untersucht. Die Wirkungsweisen für dieser Verbindungen sind ebenfalls nicht vollständig geklärt. Wahrscheinlich greifen sie wie Chloroquin (**1**) im erythrocytischen Entwicklungsstadium der Plasmodien ein, im Detail werden aber andere Mechanismen diskutiert.^[14]

Die Strukturen von **1**, **2**, **5** und **6** und wurden mit FLEXS und Dioncopeltin A (**3**) als Referenz-Liganden ausgerichtet. Chloroquin (**1**) und Primaquin (**6**) werden therapeutisch als Racemat eingesetzt, im Gegensatz zu den Naphthylisochinolin-Alkaloiden hat aber die Konfiguration der Stereoelemente weder bei **1** noch bei **6** bedeutenden Einfluß auf die räumliche Form des Moleküls. In Abb. 38 ist deshalb für beide Strukturen nur jeweils das Alignment des *R*-Isomers abgebildet, berechnet wurden jeweils beide möglichen Enantiomere

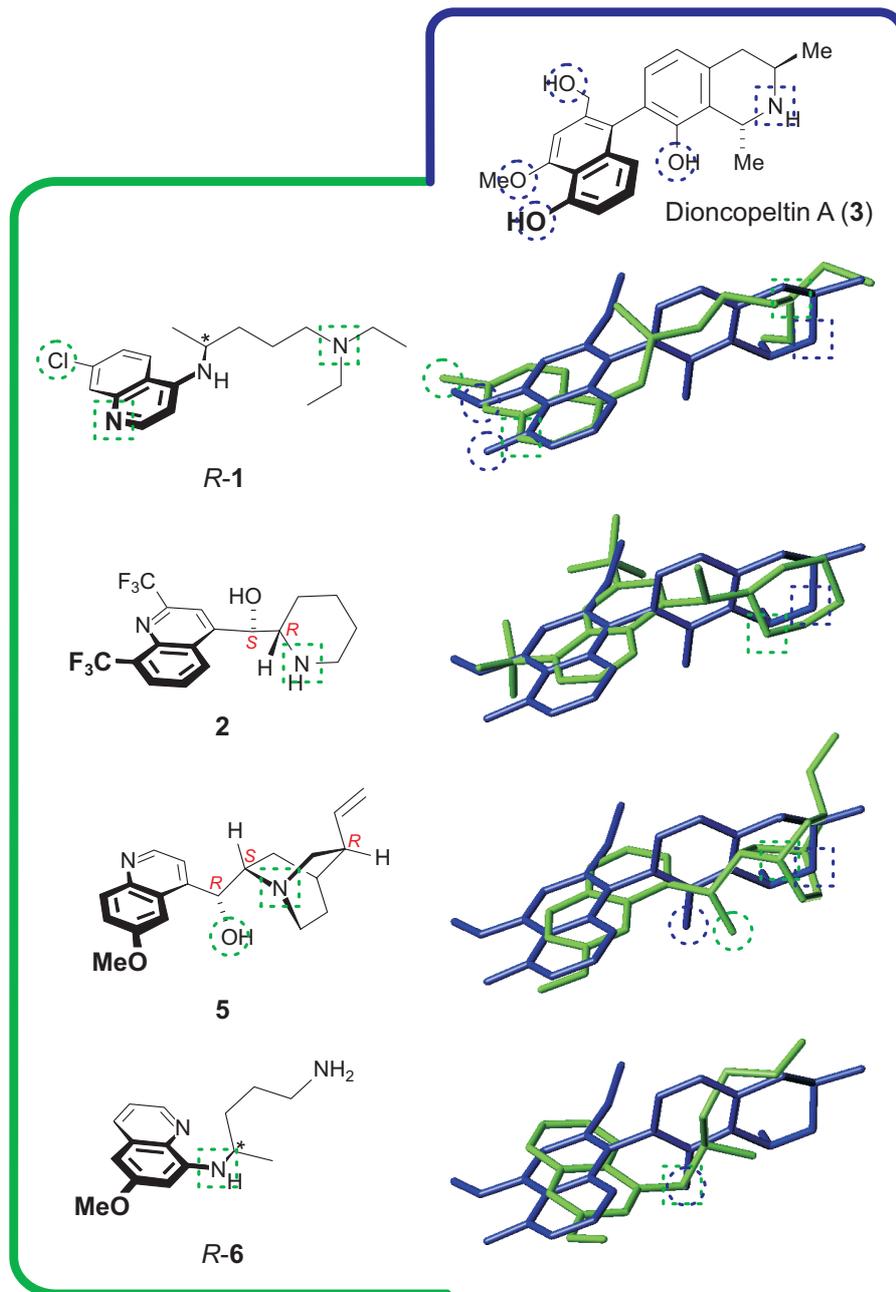


Abb. 38. Die mit FLEXS berechneten Alignments von Chloroquin (1), Mefloquin (2), Chinin (5) und Primaquin (6) mit Dioncopeltin A (3); für die beiden therapeutisch als Racemat eingesetzten Verbindungen Chloroquin (1) und Primaquin (6) ist jeweils das *R*-Isomer gezeigt

Das Alignment von Chloroquin (1) erschien schlüssig; das Chlor-Atom und die Stickstoff-Funktion des Chinolinrings lagen im Bereich der elektronegativen phenolischen Sauerstoff-Funktionen von Dioncopeltin A (3), und der Stickstoff am Ende der Alkylkette von 1 wurde im Bereich des Isochinolin-Stickstoff-Atoms von

Dioncopeltin A (**3**) plaziert, so daß die beiden Ethyl-Reste in Chloroquin (**1**) mit den beiden Methylgruppen des Tetrahydroisochinolinringes in Dioncopeltin A (**3**) überlagert wurden (s. Abb. 38 auf der vorherigen Seite).

Die Struktur von Mefloquin (**2**) hingegen schien nicht ganz so gut auf Dioncopeltin A (**3**) zu passen. Dennoch konnte der Chinolinteil (von **2**) mit dem Naphthylteil (von **3**) und der Piperidinring in **2** mit dem Tetrahydroisochinolinring in **3** überlagert werden. In ähnlicher Weise wurde auch Chinin (**5**) ausgerichtet, zusätzlich wurde hier auch die Hydroxyfunktion von Chinin (**5**) mit der 8-OH-Gruppe in **3** überlagert. Das Primaquin-Alignment zeigte kaum Gemeinsamkeiten dieser Verbindung mit Dioncopeltin A (**3**), selbst die Aminofunktionen wurden hier nicht mehr entsprechend ausgerichtet.

Die durch das Modell vorhergesagten Aktivitäten für die entsprechenden Substanzen erbrachten keine auffallend hohen Werte für alle vier untersuchten Strukturen (siehe Tabelle 31).

Tab. 31. Vorhersagen des CoMSIA-Modells für eingesetzte Antimalaria-Wirkstoffe.

Substanz	pIC ₅₀ (pred.)	IC ₅₀ [ng/ml]		
		pred.	exp.	(Schwankung)
<i>R</i> -Chloroquin (1)	0.105	251.01	48	(19-66) ^α
<i>S</i> -Chloroquin (1)	0.130	237.16	48	(19-66) ^α
Mefloquin (2)	-0.346	838.40	1.9	(0.8 - 3.7)
<i>R</i> -Primaquin (6)	0.816	39.62	β	
<i>S</i> -Primaquin (6)	0.488	84.35	β	
Chinin (5)	-0.042	357.77	50	(26-70)

^α Racemat.

^β Im entsprechenden System nicht getestet.

Konkrete Hinweise auf eventuell vorhandene Gemeinsamkeiten der Wirkmechanismen dieser Verbindungen mit Naphthylisochinolin-Alkaloiden konnten hier nicht gefunden werden, obwohl durch das Alignments zumindest einige strukturelle Gemeinsamkeiten aufgezeigt wurden. Dagegen spricht ebenfalls, daß Chloroquin (**1**) sehr unterschiedliche Aktivitätswerte gegen den K1-Stamm und gegen den NF54-Stamm zeigt (vgl. hierzu S. 33). Naphthylisochinolin-Alkaloide hingegen zeigen in diesen beiden Testsystemen jeweils ähnliche Aktivitäten.

Kapitel 5

Zusammenfassung

Bei der Suche nach neuen Leitstrukturen gegen verschiedenste Erreger spielen derzeit Methoden wie die kombinatorische Synthese oder das High-Throughput-Screening entscheidende Rollen.^[135] Leider steht aber den so täglich gefundenen Tausenden von neuen 'hits' noch ein weiter Weg bis hin zum fertigen Medikament bevor. Naturstoffchemiker gehen mit der gezielten Isolierung wirksamer Substanzen aus natürlichen ethnopharmakologischen Quellen einen scheinbar konträren, mit Sicherheit aber viel selektiveren Weg. Bei der Optimierung einer einmal gefundenen Leitstruktur leisten heutzutage – unabhängig vom Ursprung der Substanzen – computerchemischen Methoden einen unabdingbaren Beitrag.^[136]

Ein interessantes Strukturprinzip im Kampf gegen Malaria weisen die Naphthylisoquinolin-Alkaloide^[137,138] auf, deren Potential gegen diese Tropenkrankheit bereits Mitte der 90er Jahre festgestellt wurde.^[4] In Abb. 39 sind exemplarisch einige Vertreter dieser Naturstoffklasse gezeigt.

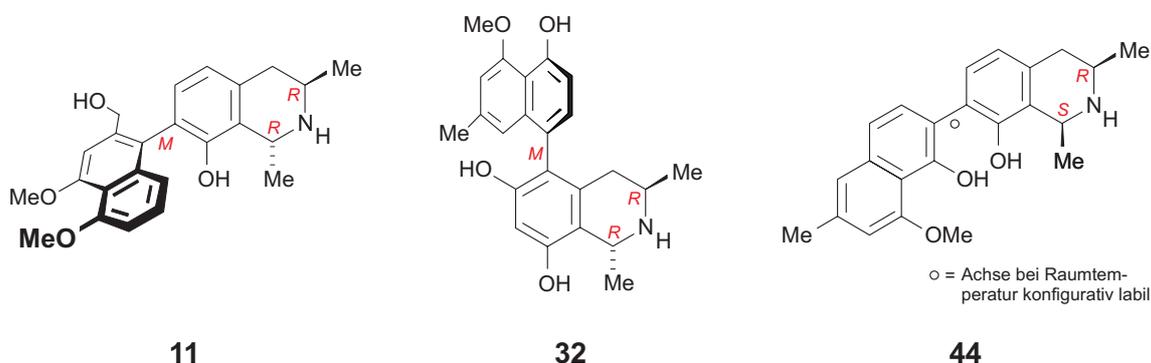


Abb. 39. Habropetalin A (11), Korupensamin B (32) and 1-*epi*-Dioncophyllin B (44)

Schon die gezeigte kleine Auswahl verdeutlicht die strukturelle Vielfalt dieser Verbindungen. Die Kupplungspositionen an die beiden Hälften der in den meisten Fäl-

len stereogenen Biarylachse ist entscheidend für die räumliche Form der Moleküle.

In langjähriger Isolier- und Synthese-Arbeit konnte in unserer Arbeitsgruppe ein strukturell diverser Datensatz von mehr als 70 Verbindungen erarbeitet werden, die auf ihre biologische Wirksamkeit gegen *Plasmodium falciparum*, den Erreger der Malaria tropica, getestet wurden. Auf Basis dieses wertvollen Datensatzes sollte in dieser Arbeit mit Hilfe der CoMSIA-Methode^[34] ein QSAR-Modell erstellt werden, das in der Lage ist, die strukturellen Eigenschaften von Naphthylisochinolin-Alkaloiden mit ihrer Antimalaria-Aktivität zu korrelieren.

Dabei wurden im einzelnen folgende Ergebnisse erzielt:

1. Der wichtigste Schritt während einer CoMSIA-Analyse – das molekulare Alignment (s. Abb. 40) – wurde weitestgehend automatisiert. Die zu diesem Zweck verwendeten Programme FLEXS^[31,30] und GASP^[27,28] zeigten allerdings Defizite bei der Behandlung stereogener Biarylachsen: Der FLEXS-Algorithmus kann axialchirale Verbindungen aus methodischen Gründen nicht korrekt behandeln, durch die Effizienz des Programms konnte dieser Nachteil jedoch leicht umgangen werden. GASP dagegen erlaubt zwar prinzipiell derartige Eigenschaften zu berücksichtigen, bei der genauen Analyse der Resultate wurden allerdings gravierende Mängel gefunden, die schließlich dazu führten, daß die vielversprechenden Möglichkeiten des Programms nicht genutzt werden konnten.

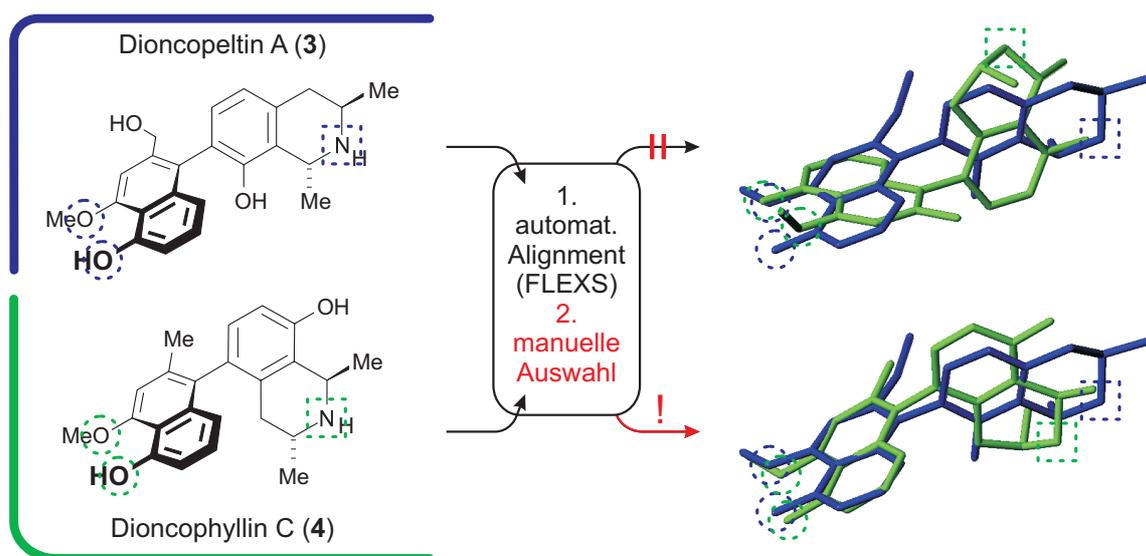


Abb. 40. Die manuelle Auswahl der Alignments unter Berücksichtigung des Stickstoff-Atoms, gezeigt für Dioncopeltin A (3) und Dioncophyllin C (4)

Die Erfahrungen mit Antimalaria-Aktivitäten von Naphthylisochinolin-Alkaloiden in unserem Arbeitskreis hat gezeigt, daß eine Abschirmung des Stickstoffatoms im Isochinolinteil zu einem drastischen Rückgang der Aktivität führt, gleiches gilt für die Veränderung der elektronischen Eigenschaften der Aminfunktion. Eine Analyse des FLEXS-Alignments zeigte jedoch, daß diese funktionelle Gruppe beim automatisierten Ausrichten der Strukturen nicht immer berücksichtigt wurde. Durch manuelle Auswahl von Plazierungen mit korrekt ausgerichtetem Stickstoff (s. Abb. 40) konnte das zunächst automatisch erzeugte Alignment weiter verbessert werden.

2. Mit unterschiedlichem Erfolg wurde versucht, neben den Naphthylisochinolin-Alkaloiden weitere Substanzklassen (s. Abb. 41) in das Modell mit aufzunehmen.

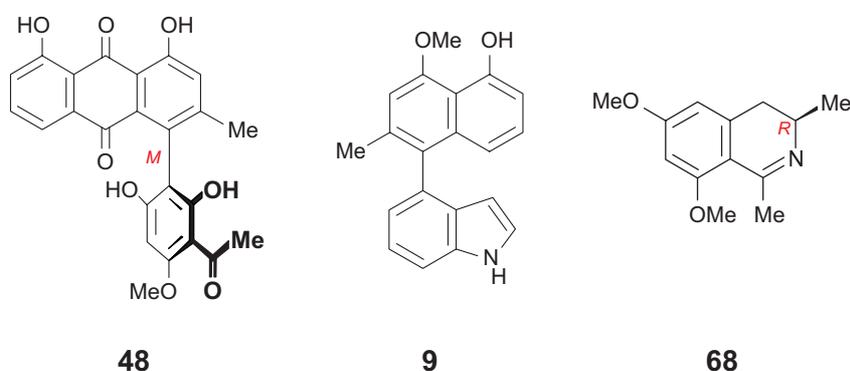


Abb. 41. Vertreter anderer ebenfalls im ursprünglichen Datensatz enthaltener Substanzklassen

Knipholon (**48**) und verschiedene seiner Derivate, sowie einige Naphthylindole (z. B. **9**) wurden in den Trainingsdatensatz aufgenommen. Die Eigenschaften der Phenylnaphthylisochinone können allerdings aufgrund des grundsätzlich anderen Strukturprinzips möglicherweise noch nicht ausreichend repräsentiert werden; die Naphthylindole hingegen ermöglichten ein an struktureller Diversität und Qualität verbessertes Modell

Die monomeren, synthetischen Vorläufer von Naphthylisochinolin-Alkaloiden (z. B. **68**, s. Abb. 41) wurden nicht in den abschließenden Trainingsdatensatz aufgenommen. Die Strukturen waren nur schwer eindeutig zu alignen; darüber hinaus sprachen auch entropische Gründe gegen einen direkten Vergleich mit Naphthylisochinolin-Alkaloiden.

3. Für den Ausschluß von fünf Outliern (8.6% der Verbindungen) aus dem Trainingsdatensatz konnten jeweils schlüssige Gründe formuliert werden. Das noch nicht optimierte Modell zeigte anschließend bereits einen beachtlichen q_{LOO}^2 -Wert von 0.787 (9 Hauptkomponenten).

4. Die Untersuchung der Abhängigkeiten zwischen den einzelnen CoMSIA-Feldern verdeutlichte, daß die Berücksichtigung von lediglich den Wasserstoffbrücken- und den hydrophoben Wechselwirkungen zu einem optimalen Modell führt. Der (zusätzliche) Einsatz der sterischen oder elektrostatischen CoMSIA-Funktionen führte jeweils zu schlechteren q^2 -Werten. Das abschließende Modell basierte auf 53 Verbindungen und zeigte einen beachtlichen q_{LOO}^2 -Wert von 0.818 (8 Hauptkomponenten). Mittlere q_n^2 -Werte zwischen 0.757 (10 Gruppen) und 0.651 (3 Gruppen) für die Gruppen-Kreuzvalidierungen zeigten, daß keine zu großen strukturellen Redundanzen im Datensatz vorhanden waren.

5. Die Qualität der möglichen Vorhersagen wurde anhand eines Testdatensatzes überprüft. Die Aktivität von Verbindungen mit den häufigsten Kupplungstypen 7,1' und 7,8' sowie 5,1' und 5,8' konnte, abgesehen von einer Ausnahme, mit einem r_{pred}^2 -Wert von 0.578 berechnet werden. Inwieweit dies auch für den 7,6'-Kupplungstyp gilt, konnte mangels Testverbindungen bisher nicht ermittelt werden.

6. Die Kontur-Darstellungen der fünf CoMSIA-Felder wurden graphisch ausgewertet. Abb. 42 zeigt Ancistrocongol D (**79**) innerhalb des Wasserstoffbrücken-Donor-Feldes, das zum abschließenden CoMSIA-Modell den größten Informationsbeitrag leistete.

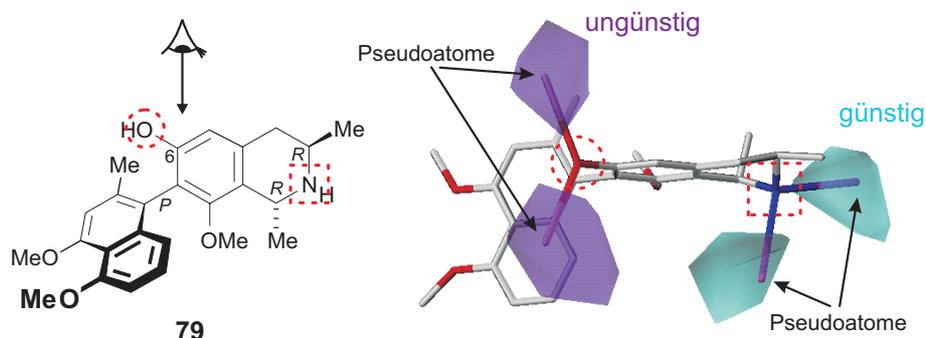


Abb. 42. Ancistrocongol D (**79**) mit entsprechender Markierung der entscheidenden funktionellen Gruppen innerhalb des Wasserstoffbrücken-Donor-Feldes

Die entscheidenden, in diesem Plot ausgewiesenen Gruppen sind die Sauerstoff-Funktion an C-6 mit negativem und die Aminfunktion im Isochinolinteil mit positivem Einfluß auf die Aktivität. Die als wichtig identifizierten Eigenschaften aus *allen* CoMSIA-Kontur-Plots wurden in einem Pharmacophor-Modell (Abb. 43) zusammengefaßt.

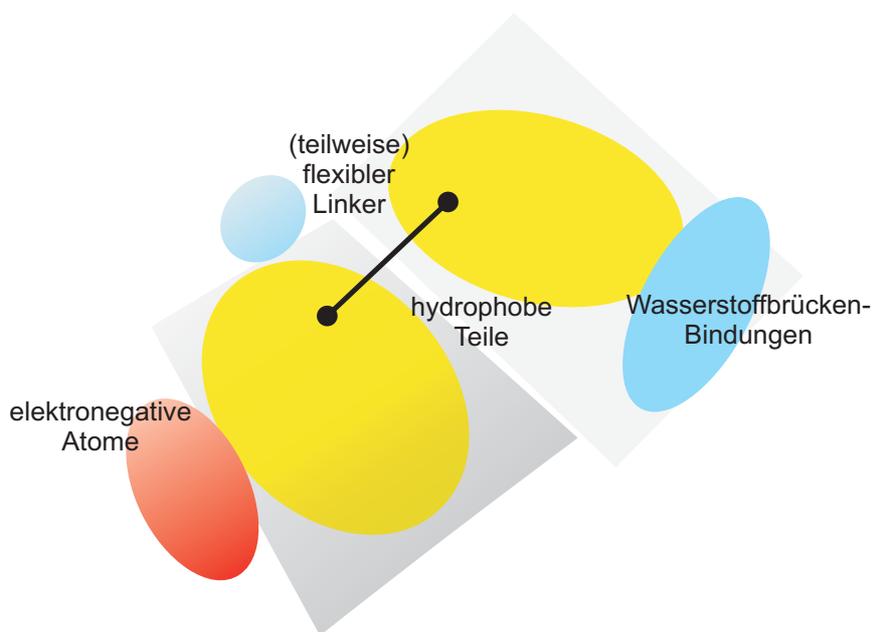


Abb. 43. Ein Pharmacophor-Modell für die Antimalaria-Aktivität von Naphthylisochinolin-Alkaloiden

Das Modell basiert auf zwei, durch einen (nicht notwendigerweise) flexiblen Linker verbundenen, hydrophoben Hälften. Entscheidend sind die cyan gefärbten Zonen (Abb. 43), in denen Wasserstoffbrücken-Donor oder -Akzeptor Funktionen vorhanden sein müssen. Das rote Gebiet sollte von elektronegativen Atomen besetzt werden

7. Aufbauend auf dieses Modell wurden neue Strukturvorschläge gemacht, wobei folgende Aspekte beachtet wurden:

- Verzicht auf stereogene Elemente;
- Verbesserung der Zugänglichkeit des Stickstoffatoms;
- Einführung einer 2'-CH₂OH-Funktion in 5-gekuppelte Strukturen.

Die aussichtsreichsten Strukturen sind in Abb. 44 gezeigt:

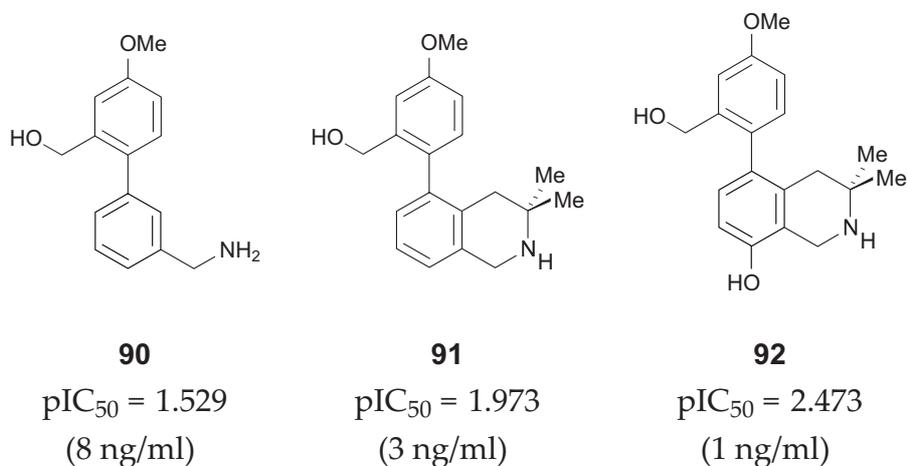


Abb. 44. Von Naphthylisochinolin-Alkaloiden abgeleitete vereinfachte Strukturen mit aussichtsreichen Vorhersagen

8. Die Strukturen der gängigen Antimalaria-Wirkstoffe Chloroquin (1), Mefloquin (2), Chinin (5) und Primaquin (6) wurden auf Gemeinsamkeiten mit den Eigenschaften des CoMSIA-Modells verglichen. Aufgrund einiger struktureller Übereinstimmungen mit Naphthylisochinolin-Alkaloiden konnten plausible Alignments erstellt werden. Die Vorhersagen für diese Verbindungen wiesen aber keine auffallend hohen Werte auf, so daß insgesamt keine Hinweise auf gleiche biologische Targets gefunden werden konnten.

Zusammenfassend konnte mit dieser Arbeit auf breiter Basis von insgesamt 53 Strukturen ein QSAR-Modell erstellt werden, das die Antimalaria-Aktivität von Naphthylisochinolin-Alkaloiden modellieren kann. Darüber hinaus wurden die für die Antimalaria-Aktivität (in positiver und negativer Weise) wichtigen funktionellen Gruppen identifiziert. Das erstellte Pharmacophor-Modell wird einen entscheidenden Beitrag zur gezielten Steuerung zukünftiger Synthesen wirksamerer Naphthylisochinolin-Analoga leisten.

Ausblick

Das während dieser Arbeit aufgestellte Alignment kann in Zukunft verwendet werden, auch andere biologische Eigenschaften der Naphthylisochinolin-Alkaloide, wie beispielsweise die Aktivität gegen die Schlafkrankheit (*Trypanosoma brucei rhodesiense*), mit Hilfe der CoMSIA-Methode genauer zu untersuchen.

In Bezug auf die Antimalaria-Aktivität wäre es über diese Arbeit hinaus in Zukunft wünschenswert, für einzelne Strukturklassen, wie etwa die 7,6'-gekoppelten Naphthylisochinolin-Alkaloide, eine breitere Datenbasis zu erlangen, damit auch die Eigenschaften derartiger Strukturen besser im Modell verankert werden können.

Auch der Einsatz verschiedener translations- und rotationsinvarianter Strukturdeskriptoren^[139] (TRI-Deskriptoren), wie er in Zusammenarbeit mit der Arbeitsgruppe Baumann derzeit durchgeführt wird, verspricht weitere Einblicke in die Zusammenhänge von Struktur und Wirkung von Naphthylisochinolin-Alkaloiden zu geben.

Kapitel 6

Summary

Today's strategy of modern drug discovery basically relies on techniques that are able to process many substances in a short time, e.g. combinatorial synthesis and high throughput screening.^[135] Nevertheless, the optimization of thousands of hits found every day to yield applicable new drugs is still a time-consuming and expensive development process. Natural product chemists use an apparently contradictory, but certainly more selective way towards active substances, by identifying and isolating new compounds from natural, ethno pharmacological origins. The optimization of a once found lead nowadays is increasingly guided by a computer-aided, virtual route^[136] – regardless whether the substance originates from natural sources or from a high-throughput assay.

Naphthylisoquinoline alkaloids^[137,138] form a promising lead structure against malaria parasites. Isolated from tropical lianas, their potential against this disease has already been discovered in the mid-90s.^[4] Some representatives of this class of natural products are exemplarily shown in Fig. 45.

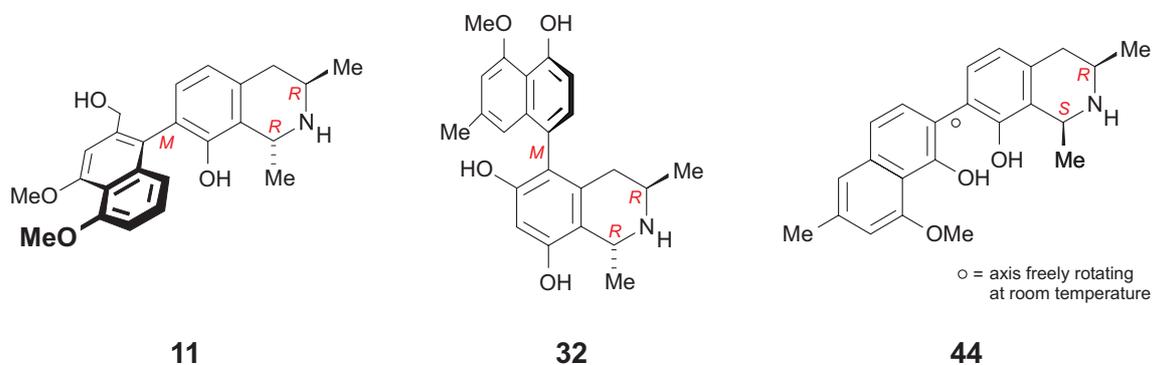


Fig. 45. Habropetaline A (**11**), korupensamine B, (**32**) and 1-*epi*-dioncophylline B (**44**)

Even this small selection already points out the structural diversity of this intriguing class of compounds. The mostly stereogenic biaryl axis can be coupled in different positions at the two aryl moieties, which is crucial for the three-dimensional shape of the molecules.

During many years of isolation and chemical synthesis our group established a diverse dataset of more than 70 substances, all of which have been tested for their biological activity against *Plasmodium falciparum*, the exciter of Malaria tropica. On the basis of this highly valuable dataset the main goal of this work was to establish a QSAR model (using the CoMSIA method^[34]), which is able to correlate the structural properties of naphthylisoquinoline alkaloids with their antimalarial activity.

In detail, the following results were achieved:

1. The most important step of a CoMSIA analysis – the molecular alignment (s. Fig. 46) – has been automated as far as possible. However, the programs used for this purpose, FLEXS^[31,30] and GASP,^[27,28] revealed several flaws in processing stereogenic biaryl axes: The FLEXS algorithm by itself can not treat axial chirality correctly, although the efficiency of the program allowed to bypass this shortcoming easily. GASP in principle is able to deal with such properties, however, a careful examination of the results revealed serious flaws; unfortunately, the promising possibilities of the program could therefore not be fully used.

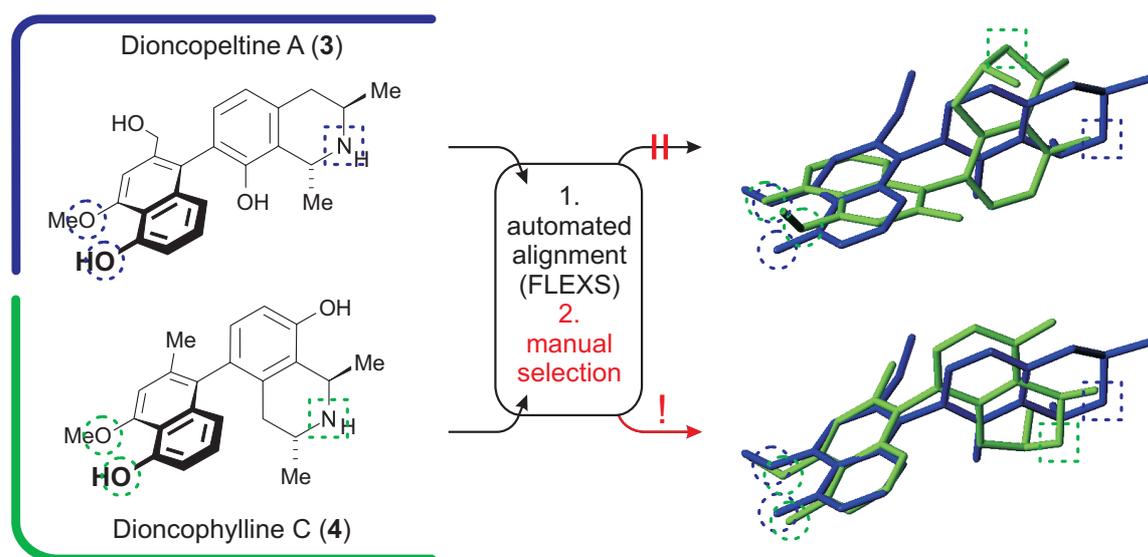


Fig. 46. Manual selection of the alignments (with respect to the nitrogen atom) shown for dioncopeltine A (3) and dioncophylline C (4)

According to our experience, an *N*-alkylation of the isoquinoline part of naphthylisoquinoline alkaloids usually decreases the *in vitro* antimalarial activity dramatically. The same is the case when altering the electronic properties of the amino function. A detailed investigation of the automated FLEXS alignment however showed that this important functional group had not been aligned in every case. By manually selecting only placements with a correctly aligned nitrogen function (see Fig. 46), the results were further improved.

2. The attempt to treat other classes of compounds (see Fig. 47) together with naphthylisoquinoline alkaloids within one single model was only partially successful.

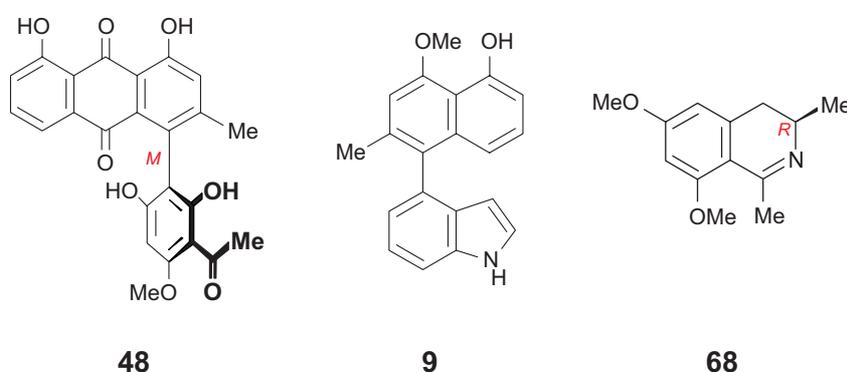


Fig. 47. Examples of other classes of compounds, also included in the initial dataset.

Knipholone (**48**) and several of its derivatives, as well as naphthylindenes (e. g. **9**) were included into the final training set. The phenylanthraquinones however are, due to quite different structural properties, not well represented by the model. On the other hand, the naphthylindenes led to an improved model in diversity, as well as in statistical quality.

The monomeric, synthetic precursors of naphthylisoquinoline alkaloids (e. g. **68**, see Fig. 47) were not included into the final training set. The fact that no unequivocal alignment was found for this kind of structures was mainly the reason for that; additionally entropical considerations disfavored the set up of a direct correlation with the activities of naphthylisoquinoline alkaloids.

3. From the training set of in total 58 compounds five outlier (8.6% of all substances) had to be eliminated. After an examination of these structures conclusive reasons for the deviating activities for all of these substances were found. After that the model

showed already a considerable q_{LOO}^2 value of 0.787 (9 components).

4. The necessity of each of the five CoMSIA fields for an optimal model was investigated thoroughly. It was shown that the best combination is to use only the hydrogen-bonding and the hydrophobic functions. The additional use of either steric or electrostatic CoMSIA fields led to worse q^2 values. The final model was based on 53 structures and showed a remarkable q_{LOO}^2 value of 0.818 (8 components). Mean q_n^2 values between 0.757 (10 groups) and 0.651 (3 groups) proved that there were no large structural redundancies in the dataset.

5. The quality of the predictions of new structures was checked using a test set of seven compounds. The activity of substances of the most frequently occurring coupling types (7,1' and 7,8' as well as 5,1' and 5,8') could be predicted satisfactorily with an r_{pred}^2 value of 0.578, with only one exception. Due to the lack of test compounds it could not be determined to what extent predictions for 7,6'-coupled compounds are possible.

6. The CoMSIA contour plots of every single CoMSIA field have been graphically analyzed. Fig. 48 exemplarily shows Ancistrocongolins D (79) within the H-bond donor field, which contributed the largest deal of information to the final CoMSIA model.

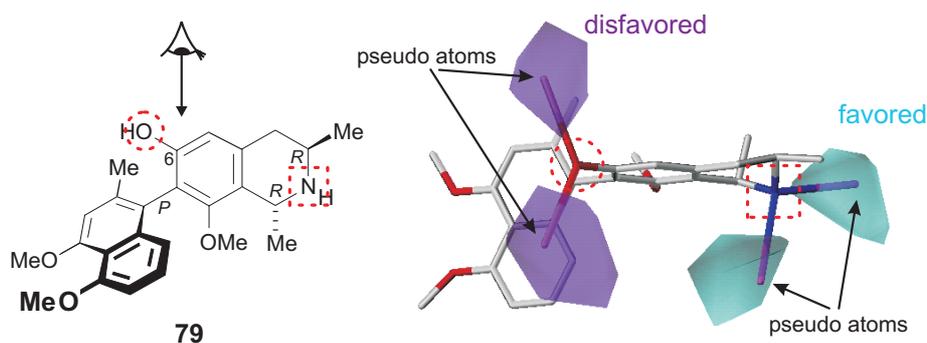


Fig. 48. Ancistrocongolins D (79) within the H-bond donor field; the crucial functional groups are highlighted by dotted circles

The most important groups outlined in this special field are the oxygen function at C-6 with a negative, and the amino function in the isoquinoline part with a positive influence on the activity.

The results from *all* CoMSIA fields were summarized in a pharmacophore model, shown in Fig. 49.

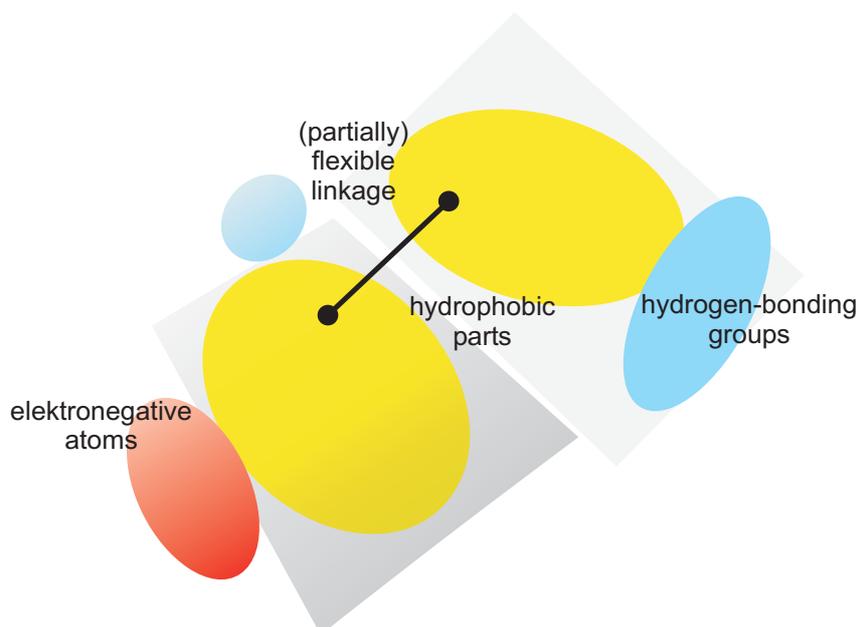


Fig. 49. A pharmacophore model for the antimalarial activity of naphthylisoquinoline alkaloids

The model is based on two hydrophobic moieties, combined through a (not necessarily) flexible linker. The cyan-colored zones (Fig. 49) indicate areas where hydrogen bonding groups should be available. The red region should be occupied with electronegative atoms.

7. Based on this model suggestions for modified, new structures were made; the following aspects were considered:

- No usage of stereogenic elements;
- Improvement of the accessibility of the nitrogen atom;
- Introduction of a 2'-CH₂OH function into structures coupled in the 5-position of the tetrahydroisoquinoline ring (see Fig. 50).

The most promising structures are shown in Fig. 50:

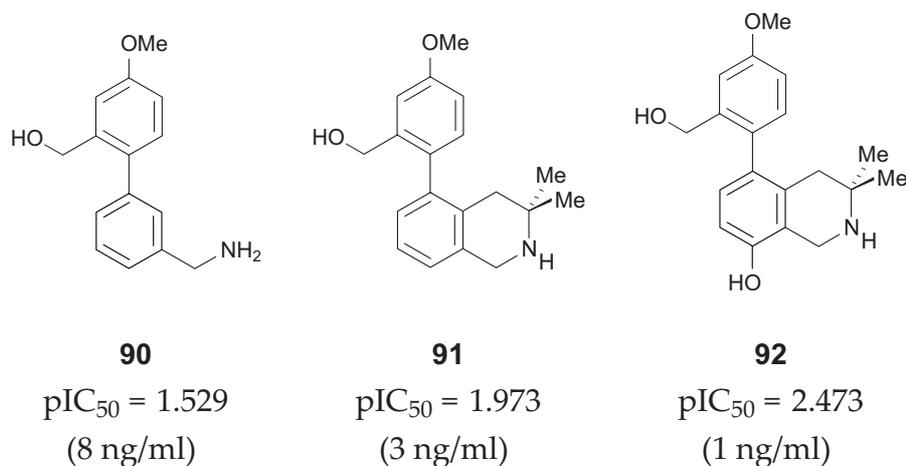


Fig. 50. Simplified analogs of naphthylisoquinoline alkaloids with promising predictions

8. The structures of the known antimalarial agents chloroquine (1), mefloquine (2), quinine (5), and primaquine (6) were examined for their accordance to the established CoMSIA model. Due to some structural similarities to naphthylisoquinoline alkaloids the alignment was quite straightforward. The predictions of the activities however did not show outstanding activities so that finally no indication for similar biological targets was found.

In conclusion in this work a QSAR model was created on a broad structural basis of 53 structures to model the antimalarial activity of naphthylisoquinoline alkaloids. Additionally, the important functional groups (with positive or negative influence on activity) were identified. The resulting pharmacophore model is supposed give an important contribution to the directed design and synthesis of even more active naphthylisoquinoline analogs.

Outlook

The alignment set up during this thesis can be used in future work to also examine other biological properties of naphthylisoquinoline alkaloids with the CoMSIA method, like e.g. the activity against *Trypanosoma brucei rhodesiense* (sleeping disease).

Concerning the antimalarial activity it would be desirable for future work to have a greater structural basis for individual compound classes, e.g. the 7,6'-coupled naphthylisoquinoline alkaloids, to also give the model the ability to describe the properties of this kind of compounds in a more comprehensive way.

The application of translationally and rotationally invariant molecular descriptors^[139] (TRI descriptors) is also promising to give further insight into the structure-activity relationship of naphthylisoquinoline alkaloids. In cooperation with the group of Dr. Baumann this work is already in progress.

Literatur

- [1] *The World Health Report 2001*, World Health Organisation, Genf, **2001**.
- [2] Basiert auf Informationen der Internetseite des 'Center for Disease Control and Prevention (CDC)', <http://www.cdc.gov/travel/reference.htm>, und der online-Version der WHO-Broschüre 'International Travel and Health': <http://www.who.int/ith/index.html>.
- [3] G. Bringmann, D. Feineis, *Act. Chim. Thérapeut.* **2000**, *26*, 151–171.
- [4] Y. F. Hallock, K. P. Manfredi, J. W. Blunt, J. H. Cardellina II, M. Schäffer, K.-P. Gulden, G. Bringmann, A. Y. Lee, J. Clardy, G. François, M. R. Boyd, *J. Org. Chem.* **1994**, *59*, 6349–6355.
- [5] G. Bringmann, M. Rübenacker, R. Weirich, L. Aké Assi, *Phytochemistry* **1992**, *31*, 4019–4024.
- [6] G. Bringmann, M. Rübenacker, P. Vogt, H. Busse, L. Aké Assi, K. Peters, H. von Schnering, *Phytochemistry* **1991**, *30*, 1691–1696.
- [7] G. Bringmann, A. Hamm, C. Günther, M. Michel, R. Brun, V. Mudogo, *J. Nat. Prod.* **2000**, *63*, 1465–1470.
- [8] G. François, G. Timperman, W. Eling, L. Aké Assi, J. Holenz, G. Bringmann, *Antimicrob. Agents Chemother.* **1997**, *41*, 2533–2539.
- [9] G. François, T. Steenackers, G. Timperman, L. Aké Assi, R. D. Haller, S. Bär, M. A. Isahakia, S. A. Robertson, C. Zhao, N. J. D. Souza, J. Holenz, G. Bringmann, *Int. J. Parasitol.* **1997**, *27*, 29–32.
- [10] G. François, G. Timperman, T. Steenackers, L. Aké Assi, J. Holenz, G. Bringmann, *Parasitol. Res.* **1997**, *83*, 673–679.
- [11] G. François, B. Chimanuka, G. Timperman, J. Holenz, J. Plaizier-Vercammen, L. Aké Assi, G. Bringmann, *Parasitol. Res.* **1999**, *85*, 935–941.
- [12] B. Chimanuka, G. François, G. Timperman, Y. V. Heyden, J. Holenz, J. Plaizier-Vercammen, G. Bringmann, *Parasitol. Res.* **2001**, *87*, 795–803.
- [13] K. Becker, persönliche Mitteilung.

- [14] M. Foley, L. Tilley, *Int. J. Parasitol.* **1997**, *27*, 231–240.
- [15] G. Bringmann, K. Messer, M. Wohlfarth, J. Kraus, K. Dumbuya, M. Rückert, *Anal. Chem.* **1999**, *71*, 2678–2686.
- [16] G. Bringmann, M. Wohlfarth, H. Rischer, M. Heubes, W. Saeb, S. Diem, M. Herderich, J. Schlauer, *Anal. Chem.* **2001**, *73*, 2571–2577.
- [17] G. Bringmann, G. Lang in: *Progr. Molec. Subcell. Biol.* (Hrsg. W. E. G. Müller), Springer Verlag, Berlin, S. im Druck.
- [18] G. Bringmann, S. Busemann in: *Natural Product Analysis* (Hrsg. P. Schreier, M. Herderich, H. Humpf, W. Schwab), Vieweg, Wiesbaden, **1998**, S. 195–212.
- [19] G. Bringmann, J. Mühlbacher, C. Repges, J. Fleischhauer, *J. Comp. Chem.* **2001**, *22*, 1273–1278.
- [20] G. Bringmann, S. Tasler, R. Pfeifer, M. Breuning, *J. Organomet. Chem.* **2002**, *661*, 49–65.
- [21] G. Bringmann, M. Breuning, R. Pfeifer, W. A. Schenk, K. Kamikawa, M. Uemura, *J. Organomet. Chem.* **2002**, *661*, 31–47.
- [22] G. Bringmann, S. Tasler, *Tetrahedron* **2001**, *57*, 331–343.
- [23] G. Bringmann, D. Menche, *Acc. Chem. Res.* **2001**, *34*, 615–624.
- [24] D. Vitt, Dissertation, Universität Würzburg **1998**.
- [25] S. Schmitt, Diplomarbeit, Universität Würzburg **1997**.
- [26] R. D. Cramer, D. E. Patterson, J. D. Bunce, *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
- [27] G. Jones, P. Willet, R. C. Glen, *J. Comput.-Aided Mol. Design* **1995**, *9*, 532–549.
- [28] G. Jones, P. Willet, R. C. Glen, *J. Mol. Biol.* **1995**, *245*, 43–53.
- [29] G. Bringmann, G. François, L. Aké Assi, J. Schlauer, *Chimia* **1998**, *52*, 18–28.
- [30] C. Lemmen, T. Lengauer, *J. Comput.-Aided Mol. Design* **1997**, *11*, 357–368.
- [31] C. Lemmen, T. Lengauer, G. Klebe, *J. Med. Chem.* **1998**, *41*, 4502–4520.

- [32] M. Rarey, M. Wefing, T. Lengauer, *J. Comput.-Aided Mol. Design* **1996**, *10*, 41–54.
- [33] M. Rarey, B. Kramer, T. Lengauer, G. Klebe, *J. Mol. Biol.* **1996**, *261*, 470–489.
- [34] G. Klebe, U. Abraham, T. Mietzner, *J. Med. Chem.* **1994**, *37*, 4130–4146.
- [35] G. Klebe, T. Mietzner, F. Weber, *J. Comput.-Aided Mol. Design* **1994**, *8*, 751–778.
- [36] G. Klebe, T. Mietzner, F. Weber, *J. Comput.-Aided Mol. Design* **1999**, *13*, 35–49.
- [37] G. Klebe, T. Mietzner, *J. Comput.-Aided Mol. Design* **1994**, *8*, 583–606.
- [38] *Sybyl 6.7.1*, Tripos, Inc., 1699 St. Hanley Road, Suite 303, St. Louis, MO, 63144.
- [39] J. Sadowski, J. Gasteiger, *Chem. Reviews* **1993**, *93*, 2567–2581.
- [40] J. Sadowski, J. Gasteiger, G. Klebe, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1000–1008.
- [41] J. Hoflack, P. J. D. Clercq, *Tetrahedron* **1988**, *44*, 6667–6676.
- [42] R. D. Cramer, J. D. Bunce, D. E. Patterson, *QSAR* **1988**, *7*, 18–25.
- [43] M. Böhm, J. Stürzebecher, G. Klebe, *J. Med. Chem.* **1999**, *42*, 458–477.
- [44] J. Gasteiger, M. Marsili, *Tetrahedron* **1980**, *36*, 3219–3222.
- [45] H. Diaz-Arauzo, K. Kohler, T. Hagen, J. Cook, *Life Sciences* **1991**, *49*, 207–216.
- [46] Y. Martin, M. Bures, E. Dahaner, J. DeLazzer, I. Lico, P. Pavlik, *J. Comp.-Aid. Mol. Des.* **1993**, *7*, 83–102.
- [47] V. N. Viswanadhan, A. K. Ghose, G. R. Revankar, R. K. Robins, *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163–172.
- [48] H. Wold in: *Research papers in statistics* (Hrsg. F. David), Wiley, New York, **1966**, S. 411–444.
- [49] H. Wold in: *Quantitative sociology* (Hrsg. H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, V. Capecchi), Academic Press, New York, **1975**, S. 307–357.
- [50] S. Wold, M. Martens, H. Wold in: *Matrix Pencils* (Hrsg. A. Ruhe, B. Kågström), Springer Verlag, Heidelberg, **1983**, S. 286–293.

- [51] U. Thibaut, G. Folkers, G. Klebe, H. Kubinyi, A. Merz, D. Rogan in: *3D QSAR in drug design: Theory, Methods and Applications* (Hrsg. H. Kubinyi), ESCOM, Leiden, The Netherlands, **1993**, S. 711–716.
- [52] H. Kubinyi, U. Abraham in: *3D QSAR in drug design: Theory, Methods and Applications* (Hrsg. H. Kubinyi), ESCOM, Leiden, The Netherlands, **1993**, S. 717–728.
- [53] B. L. Bush, R. B. Nachbar, *J. Comput.-Aided Mol. Design* **1993**, *7*, 587–619.
- [54] F. Lindgren, S. Rännar, *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 105–113.
- [55] T. L. Richie, A. Saul, *Nature* **2002**, *415*, 694–701.
- [56] A. M. Tomas, A. M. van der Wel, A. W. Thomas, C. J. Janse, A. Waters, *Parasitol. Today* **1998**, *14*, 245–249.
- [57] N. A. Campbell, *Biologie*, Spektrum Akademischer Verlag, Berlin, **1997**.
- [58] M. Balter, *Science* **2000**, *290*, 39.
- [59] J. Taverne, *Parasitol. Today* **1999**, *15*, 180–181.
- [60] F. M. Steinberg, *Brit. Med. J.* **2001**, *322*, 676.
- [61] P. Loria, S. Miller, M. Foley, L. Tilley, *Biochem. J.* **1999**, *339*, 363–370.
- [62] A. Yayon, Z. I. Cabantchik, H. Ginsberg, *Europ. Molec. Bio. Org. J.* **1984**, *3*, 2695–2700.
- [63] S. L. C. O. Kayser, A. F. Kiderlein, *Studies in Natural Products Chemistry* **2002**, 779–848.
- [64] R. F. Loeb, W. M. Clarke, G. R. Coateney, L. Coggeshall, E. K. Marshall, S. C. Marvel, O. R. McCoy, J. J. Saper, W. H. Serbell, J. A. Shannon, G. A. Carden, *J. Am. Med. Assoc.* **1946**, *130*, 1069–1070.
- [65] M. P. Pollack, Kenya bans use of chloroquine. *ProMed-mail* (19. July 1999), <http://www.promedmail.org>.
- [66] H. Jomaa, J. Wiesner, S. Sanderbrand, B. Altincicek, C. Weidemeyer, M. Hintz, I. Turbachova, M. Eberl, J. Zeidler, H. K. Lichtenthaler, D. Soldati, E. Beck, *Science* **1999**, *285*, 1573–1576.

- [67] J. Wiesner, P. Wissner, H. M. Dahse, H. Jomaa, M. Schlitzer, *Bioorg. Med. Chem.* **2001**, *9*, 785–792.
- [68] [Http://www.jomaa-pharmaka.de](http://www.jomaa-pharmaka.de).
- [69] L. Florens, M. P. Washburn, J. D. Raine, R. M. Anthony, M. Grainger, J. D. Haynes, J. K. Moch, N. Muster, J. B. Sacci, D. L. Tabb, A. A. Witney, D. Wolters, Y. Wu, M. J. Garnder, A. A. Holder, R. E. Sinden, J. R. Yates, D. J. Carucci, *Nature* **2002**, *419*, 520–526.
- [70] E. Lasonder, Y. Ishihama, J. S. Andersen, A. M. W. Vermunt, A. Pain, R. W. Sauerwein, W. M. C. Eling, N. Hall, A. P. Waters, H. G. Stunnenberg, M. Mann, *Nature* **2002**, *419*, 537–542.
- [71] K. S. Aultman, M. Gottlieb, M. Giovanni, A. S. Fauci, *Science* **2002**, *298*, 13.
- [72] C. M. Morel, Y. T. Touré, B. Dobrokhotov, A. M. J. Oduola, *Science* **2002**, *298*, 79.
- [73] M. J. Gardner, N. Hall, E. Fung, O. White, M. Berriman, R. W. Hyman, J. M. Carlton, A. Pain, K. E. Nelson, S. Bowman, I. T. Paulsen, K. James, J. A. Eisen, K. Rutherford, S. L. Salzberg, A. Craig, S. Kyes, M.-S. Chan, V. Nene, S. J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Perteau, J. Allen, J. Selengut, D. Haft, M. W. Mather, A. B. Vaidya, D. M. A. Martin, A. H. Fairlamb, M. J. Fraunholz, D. S. Roos, S. A. Ralph, G. I. McFadden, L. M. Cummings, G. M. Subramanian, C. Mungall, J. C. Venter, D. J. Carucci, S. L. Hoffman, C. Newbold, R. W. Davis, C. M. Fraser, B. Barrell, *Nature* **2002**, *419*, 498–511.
- [74] R. A. Holt, G. M. Subramanian, A. Halpern, G. G. Sutton, R. Charlab, D. R. Nusskern, P. Wincker, A. G. Clark, J. M. C. Ribeiro, R. Wides, S. L. Salzberg, B. Loftus, M. Yandell, W. H. Majoros, D. B. Rusch, Z. Lai, C. L. Kraft, J. F. Abril, V. Anthonard, P. Arensburger, P. W. Atkinson, H. Baden, V. de Berardinis, D. Baldwin, V. Benes, J. Biedler, C. Blass, R. Bolanos, D. Boscus, M. Barnstead, S. Cai, A. Center, K. Chatuverdi, G. K. C., M. A. Chrystal, M. Clamp, A. Cravchik, V. Curwen, A. Dana, A. Delcher, I. Dew, C. A. Evans, M. Flanigan, A. Grundschober-Freimoser, L. Friedli, Z. Gu, P. Guan, R. Guigo, M. E. Hillenmeyer, S. L. Hladun, J. R. Hogan, Y. S. Hong, J. Hoover, O. Jaillon, Z. Ke, C. Kodira, E. Kokoza, A. Koutsos, I. Letunic, A. Levitsky, Y. Liang, Jhy-Jhu Lin, N. F. Lobo, J. R. Lopez, J. A. Malek, T. C. McIntosh, S. Meister, J. Miller, C. Mobarry, E. Mongin,

- S. D. Murphy, D. A. O'Brochta, C. Pfannkoch, R. Qi, M. A. Regier, K. Remington, H. Shao, M. V. Sharakhova, C. D. Sitter, J. Shetty, T. J. Smith, R. Strong, J. Sun, D. Thomasova, L. Q. Ton, P. Topalis, Z. Tu, M. F. Unger, B. Walenz, A. Wang, J. Wang, M. Wang, X. Wang, K. J. Woodford, J. R. Wortman, M. Wu, A. Yao, E. M. Zdobnov, H. Zhang, Q. Zhao, S. Zhao, S. C. Zhu, I. Zhimulev, M. Coluzzi, A. della Torre, C. W. Roth, C. Louis, F. Kalush, R. J. Mural, E. W. Myers, M. D. Adams, H. O. Smith, S. Broder, M. J. Gardner, C. M. Fraser, E. Birney, P. Bork and P. T. Brey, J. C. Venter, J. Weissenbach, F. C. Kafatos and F. H. Collins, S. L. Hoffman, *Science* **2002**, *297*, 129–149.
- [75] G. Taubes, *Science* **2000**, *290*, 434–437.
- [76] J. Holenz, Dissertation, Universität Würzburg **1997**.
- [77] W. Saeb, Dissertation, Universität Würzburg **2001**.
- [78] T. Ponnudurai, A. D. Leeuwenberg, J. H. Meuwissen, *Trop. Geogr. Med.* **1981**, *33*, 50–54.
- [79] I. J. Udeinya, P. M. Graves, R. Carter, M. Aikawa, L. H. Miller, *Exp. Parasitol.* **1983**, *56*, 207–214.
- [80] G. Bringmann, H. Reuscher, *Tetrahedron Lett.* **1989**, *30*, 5249–5252.
- [81] G. Bringmann, H. Reuscher, *Angew. Chem.* **1989**, *101*, 1725–1726; *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 1672–1673.
- [82] G. Bringmann, F. Teltschik, M. Michel, S. Busemann, M. Rückert, R. Haller, S. Bär, A. Robertson, R. Kaminsky, *Phytochemistry* **1999**, *52*, 321–332.
- [83] G. Bringmann, W. Saeb, M. Rübenacker, *Tetrahedron* **1999**, *55*, 423–432.
- [84] G. Bringmann, J. R. Jansen, H. Reuscher, M. Rübenacker, K. Peters, H. G. von Schnering, *Tetrahedron Lett.* **1990**, *31*, 643–646.
- [85] G. Bringmann, W. Saeb, R. God, M. Schäffer, G. François, K. Peters, E.-M. Peters, P. Proksch, K. Hostettmann, L. Aké Assi, *Phytochemistry* **1998**, *49*, 1667–1673.
- [86] G. Bringmann, M. Rückert, K. Messer, O. Schupp, A. M. Louis, *J. Chromatogr. A* **1999**, *837*, 267–272.

- [87] G. Bringmann, M. Rübenacker, W. Koch, D. Koppler, T. Ortmann, M. Schäffer, L. Aké Assi, *Phytochemistry* **1994**, *36*, 1057–1061.
- [88] G. Bringmann, D. Lisch, H. Reuscher, L. Aké Assi, K. Günther, *Phytochemistry* **1991**, *30*, 1307–1310.
- [89] G. Bringmann, C. Günther, W. Saeb, J. Mies, R. Brun, L. Aké Assi, *Phytochemistry* **2000**, *54*, 337–346.
- [90] G. Bringmann, C. Günther, W. Saeb, J. Mies, A. Wickramasinghe, V. Mudogo, R. Brun, *J. Nat. Prod.* **2000**, *63*, 1333–1337.
- [91] G. Bringmann, J. Holenz, R. Weirich, M. Rübenacker, C. Funke, M. R. Boyd, R. Gulakowski, G. François, *Tetrahedron* **1998**, *54*, 497–512.
- [92] G. Bringmann, R. Götz, P. A. Keller, R. Walter, P. Henschel, M. Schäffer, M. Stäblein, T. R. Kelly, M. R. Boyd, *Heterocycles* **1994**, *39*, 503–512.
- [93] G. Bringmann, F. Teltschik, M. Schäffer, R. Haller, S. Bär, M. A. Robertson, M. Isahakia, *Phytochemistry* **1998**, *47*, 31–35.
- [94] J. Fleischhauer, A. Koslowski, B. Kramer, E. Zobel, G. Bringmann, K.-P. Gulden, T. Ortmann, B. Peter, *Z. Naturforsch.* **1993**, *48b*, 140–148.
- [95] G. Bringmann, J. R. Jansen, H.-P. Rink, *Angew. Chem.* **1986**, *98*, 917–919; *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 913–915.
- [96] G. Bringmann, K. Messer, R. Brun, V. Mudogo, *J. Nat. Prod.* **2002**, *65*, 1096–1101.
- [97] G. Bringmann, M. Rübenacker, T. Geuder, L. Aké Assi, *Phytochemistry* **1991**, *30*, 3845–3847.
- [98] G. Bringmann, J. Kraus, D. Menche, K. Messer, *Tetrahedron* **1999**, *55*, 7563–7572.
- [99] G. Bringmann, D. Menche, *Angew. Chem.* **2001**, *40*, 1733–1736; *Angew. Chem. Int. Ed.* **2001**, *40*, 1687–1690.
- [100] E. Dagne, W. Steglich, *Phytochemistry* **1984**, *23*, 1729–1731.
- [101] E. Dagne, A. Yenesew, *Phytochemistry* **1993**, *34*, 1440–1441.
- [102] M. Bezabih, B. M. Abegaz, *Phytochemistry* **1998**, *48*, 1071–1073.

- [103] M. Bezabih, S. Motlhagodi, B. M. Abegaz, *Phytochemistry* **1997**, *46*, 1063–1067.
- [104] D. Menche, Dissertation, Universität Würzburg **2002**.
- [105] M. Welter, Diplomarbeit, Universität Würzburg **1998**.
- [106] U. Möhler, Diplomarbeit, Universität Würzburg **1998**.
- [107] G. Bringmann, C. Günther, *Synlett* **1999**, 216–218.
- [108] G. Bringmann, R. Weirich, H. Reuscher, J. R. Jansen, L. Kinzinger, T. Ortman, *Liebigs Ann. Chem.* **1993**, 877–888.
- [109] S. Neumann, Dissertation, Universität Würzburg, vorraussichtlich **2003**.
- [110] M. Wenzel, Dissertation, Universität Würzburg **1999**.
- [111] G. Bringmann, M. Wenzel, T. R. Kelly, M. R. Boyd, R. J. Gulakowski, R. Kaminsky, *Tetrahedron* **1999**, *55*, 1731–1740.
- [112] G. Bringmann, R. Götz, S. Harmsen, J. Holenz, R. Walter, *Liebigs Ann.* **1996**, 2045–2058.
- [113] Siehe GASP-Produktbeschreibung:
<http://www.tripos.com/sciTech/inSilicoDisc/pharmaPerception/gasp.html>.
- [114] G. R. Marshall, C. D. Barry, H. E. Bosshard, R. A. Dammkoehler, D. A. Dunn in: *Computer Assisted Drug Design, American Chemical Society Symposium Series* (Hrsg. E. C. Olson, R. E. Christofferson), American Chemical Society, Washington DC, Bd. 112, **1979**, S. 205–226.
- [115] *FLEXS Version 1.8, Sybyl 6.7.1*, Tripos, Inc., 1699 St. Hanley Road, Suite 303, St. Louis, MO, 63144.
- [116] G. Bringmann, K. Messer, K. Wolf, J. Mühlbacher, M. Grüne, R. Brun, A. M. Louis, *Phytochemistry* **2002**, *60*, 389–397.
- [117] G. Bringmann, D. Menche, M. Bezabih, B. M. Abegaz, R. Kaminsky, *Planta Med.* **1999**, *65*, 757–758.
- [118] G. Bringmann, D. Vitt, *J. Org. Chem.* **1995**, *60*, 7674–7681.

- [119] *GASP 2.0 / Sybyl 6.7.1*, Tripos, Inc., 1699 St. Hanley Road, Suite 303, St. Louis, MO, 63144.
- [120] *VAMP 7.0*, G. Rauhut, J. Chandrasekhar, A. Alex, B. Beck, W. Sauer, T. Clark, Oxford Molecular Ltd., The Medawar Centre, Oxford Science Park, Sandford-on-Thames, Oxford, OX4 4GA, England (inzwischen Pharmacoepia Inc., Box 5350, Princeton, NJ 08543-5350, USA).
- [121] E. Dagne, A. Yenesew, *Phytochemistry* **1993**, *34*, 1440–1441.
- [122] H. Kubinyi in: *The Encyclopedia of Computational Chemistry* (Hrsg. P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. S. III, P. R. Schreiner), John Wiley & Sons, Chichester, **1998**.
- [123] G. Klebe, *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 87–104.
- [124] G. Klebe, U. Abraham, *J. Med. Chem.* **1993**, *36*, 70–80.
- [125] H. J. Böhm, G. Klebe, *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2588–2614.
- [126] G. Klebe in: *3D QSAR in drug design* (Hrsg. H. Kubinyi), ESCOM, Leiden, The Netherlands, **1993**, S. 173–199.
- [127] U. Norinder, *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 25–39.
- [128] K. H. Kim, G. Greco, E. Novellino, *Perspect. Drug Discovery Des.* **1998**, *12*, 257–315.
- [129] T. J. Hou, Z. M. Li, Z. Li, J. Liu, X. J. Xu, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1002–1009.
- [130] T. Hou, Y. Li, N. Liao, X. Xu, *J. Mol. Model.* **2000**, *6*, 438–445.
- [131] G. Klebe, U. Abraham, *J. Comput.-Aided Mol. Design* **1999**, *13*, 1–10.
- [132] M. Dreyer, Dissertation, Universität Würzburg, vorraussichtlich **2004**.
- [133] S. Specker, Diplomarbeit, Universität Würzburg **2002**.
- [134] M. Reichert, Diplomarbeit, Universität Würzburg, vorraussichtlich **2003**.
- [135] P. Seneci, S. Miertus, *Mol. Divers.* **2000**, *5*, 75–89.

- [136] H. J. Böhm, G. Klebe, H. Kubinyi, *Wirkstoffdesign*, Spektrum Akademischer Verlag, Heidelberg, **1996**.
- [137] G. Bringmann in: *Guidelines and issue for the discovery and drug development against Tropical Diseases* (Hrsg. H. Vial, A. Fairlamb, R. Ridley), World Health Organisation, Geneva, **2001**, im Druck.
- [138] G. Bringmann, F. Pokorny in: *The Alkaloids* (Hrsg. G. A. Cordell), Academic Press, New York, Bd. 46, **1995**, S. 127–271.
- [139] K. Baumann, *J. Chem. Inf. Comput. Sci.* **2002**, 42, 26–35.
- [140] Der Autor dankt Herrn Christian Lemmen (damals tätig bei der GMD/SCAI in Sankt Augustin) für die zur Verfügung gestellten Programme FLEXV und FLEXS (in verschiedenen Beta-Versionen).
- [141] *GAUSSIAN 98 (Revision A.2)*, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, Q. C. P. Y. Ayala, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, B. G. J. P. M. W. Gill, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, Tripos, Inc., Pittsburgh PA, **1998**.

Anhang A

Verwendete Abkürzungen

3D	Dreidimensional
AM1	Austin Method 1
B3LYP	Becke 3 Lee Yang Parr
Bn	Benzyl
CD	Circular-Dichroismus
CoMFA	Comparative Molecular Field Analysis
CoMSIA	Comparative Molecular Similarity Indices Analysis
CPU	Central Processing Unit
cv	kreuzvalidiert (crossvalidated)
DISCO	DIStance COmparisons
exp.	Experimentell
fit	angepasst (fitted)
GASP	Genetic Algorithm Similarity Program
H	Wasserstoff
HIV	Human Immunodeficiency Virus
HPLC	High Performance Liquid Chromatography
IC ₅₀	Inhibition Concentration 50
MD	Molecular Dynamics
NMR	Nuclear Magnetic Resonance
pIC ₅₀	$-\log(\text{IC}_{50})$
PLS	Partial Least Squares
PM3	Parameterization Model 3

pred.	vorhergesagt (predicted)
QSAR	Quantitativ Structure-Activity Relationship
RMS	Root Mean Square
SAMPLS	SAMple-distance Partial Least Squares
Tos	p-Toluolsulfonyl
UV	Ultra-Violett

Anhang B

Technische Details

B.1 Soft- und Hardware

Für die unterschiedlichen in dieser Arbeit durchgeführten Verfahren wurde eine Reihe von Programmpaketen verwendet. Die CoMSIA-Studien wurden mit dem Molecular-Modelling-Paket SYBYL 6.7^[38] durchgeführt. Das Alignment wurde mit Hilfe der Programme GASP 2.0^[119] und FLEXS^[115] erstellt. Zu Beginn dieser Arbeit wurde mit Beta-Versionen von FLEXS^[140] (ab Version 1.6) gearbeitet, zur Visualisierung wurde neben SYBYL auch das Programm FLEXV^[140] eingesetzt. Die Konformere für geschlossene Ringsysteme wurden mit Hilfe von CORINA^[39,40] erstellt. Semiempirische Berechnungen wurden sowohl mit Hilfe des Gaussian 98-Programmpakets^[141] als auch mit VAMP5.5^[120] durchgeführt.

Alle Programme wurden unter dem Betriebssystem IRIX6.5.10 eingesetzt, als Workstations standen mehrere Rechner des Typs SGI-OCTANE (R10000) und SGI-O2 (R5000) zur Verfügung. In der Beta-Testphase von FLEXS konnte auch eine Linux-Version des Programms getestet werden, lauffähig auf IBM-Kompatiblen PCs unter dem Betriebssystem Linux.

B.2 FLEXS-Script

Das in Abb. 51 gezeigte, selbst erstellte Script wurde für die Automatisierung des Alignment mit FLEXS verwendet. In der Datei "Liganden" werden die einzelnen, zu berechnenden Liganden abgelegt.

REF_LIG	Auswahl des Menüs für Referenz-Liganden;
READ K_09_Dioncopeltin_A.mol2	Einlesen von Dioncopeltin A (3) als Referenz;
END	Menü für Referenz-Liganden beenden;
FOR_EACH \$0 in "Liganden"	Beginn der Schleife über alle Test-Liganden;
TEST_LIG	Auswahl des Menüs für Test-Liganden;
READ \$0	Einlesen des Test-Liganden;
SELBAS a f	Selektion des Basisfragments;
END	Menü für Referenz-Liganden beenden;
SUPERPOS	Auswahl des Alignment-Menüs;
PLACEBAS 3	Plazieren des Basisfragments;
COMPLEX all	Positionierung der übrigen Fragmente;
OPTIMIZE all	Optimierung des Overlaps;
SORT E_TOTAL	Sortieren nach ansteigender Energie;
CLUSTER	Clustering der Lösungen;
END	Alignment-Menü beenden;
TEST_LIG	Auswahl des Menüs für Test-Liganden;
WRITE \$0 n all	Schreiben der verschiedenen Test-Ligand-Files;
END	Menü für Test-Liganden beenden;
END_FOR	Ende der Schleife über die Test-Liganden;
END	Beenden des Alignment Vorganges;
y	Beenden des Programmes;

Abb. 51. Das für das FLEXS-Alignment verwendete Script

danke...

Keine wissenschaftliche Arbeit ohne die Hilfen, Anregungen und sonstige Unterstützung anderer. Mein Dank gilt allen die zum Gelingen dieser Arbeit beigetragen haben.

Mein erster persönlicher Dank ist dennoch ein eher 'unwissenschaftlicher' – er gilt meinen Eltern und meiner Schwester, auf die ich mich zu jeder Zeit und ohne Einschränkung verlassen konnte und kann.

Diese Arbeit wäre nicht möglich gewesen ohne Generationen von Doktoranden und Diplomanden, die an der Erstellung des Datensatzes mitgewirkt haben. Es wäre (sogar hier) zu wenig Platz um alle zu nennen. Den Herren R. Kaminsky und R. Brun am Schweizerischen Tropeninstitut in Basel danke ich für die Durchführung der biologischen Tests.

Dr. Anthony Hopfinger und Nik Stiefl danke ich für interessante Diskussionen und Anregungen, Dr. Knut Baumann zusätzlich für viel Arbeit beim Korrekturlesen des 'QSAR-Papers'.

Während meiner Zeit im 'Rechnerraum' durfte ich mit vielen Kollegen zusammenarbeiten; dies waren zunächst Ralf Stowasser, Daniel Vitt und Stefan Busemann - temporär auch mit Virima Mudogo, G. Prasuna und Johan Faber. Besonders ans Herz gewachsen sind mir Tanja Schulz, Jürgen Kraus, Kristina Wolf und schließlich Jörg Mühlbacher; von jedem Einzelnen konnte ich – auf ganz individuelle Weise – so einiges lernen.

Von den 'Kochern' in unserem Arbeitskreis möchte ich zunächst und vor allen Wael Saeb nennen, der mich in so manches Geheimnis des praktischen Arbeitens eingeweiht hat (auch wenn ich es selbst seit dem Hauptstudium nicht mehr versuchen konnte). Darüber hinaus danke ich Andreas Wuzik und Robert Pfeifer für fruchtbare Kooperationen.

Vor allem (aber nicht nur) für seelische Unterstützung - besonders gegen Ende dieser Arbeit – möchte ich Doris Feineis danken, ohne sie wäre alles im AK nicht nur ein bißchen schwieriger gewesen.

Für das Korrekturlesen dieser Arbeit danke ich vor allem Philipp Schmidt, aber auch Thomas Gottwald, Marco Greb, Robert Pfeifer, Kristina Wolf und Doris Feineis.

Ein großes Dankeschön gilt den Doktoranden der Arbeitsgruppe Hartung: Philipp Schmidt, Marco Greb, Thomas Gottwald und Kristina Spehar haben meine Zeit ein bißchen kürzer erscheinen lassen. Darüber hinaus sei auch den Herren C. Lyle, R. Reger und J. Panse herzlichst für so manche unterhaltsame Stunde gedankt.

Am Ende folgt nicht 'das Letzte', sondern der rote Faden aus Vergangenheit in die Zukunft; ich schulde Hede für so manches Dank, ganz besonders aber für ihre unglaublichen Energie und ihren Sinn für Humor, an denen ich die letzten Jahren teilhaben durfte.

Lebenslauf

Persönliche Daten	Geburtsdatum:	22. Juli 1973
	Geburtsort:	Würzburg
	Familienstand	Ledig
	Staatsangehörigkeit:	Deutsch

Schulbildung

07.1979 – 06.1983	Grundschule in Dettelbach
07.1983 – 06.1992	Allgemeine Hochschulreife am Armin-Knab-Gymnasium in Kitzingen

Hochschulstudium

11.1992 – 03.1998	Studium der Chemie (Diplom)
10.1994	Vordiplomprüfung
05.1997	Diplomprüfung
07.1997-03.1998	Diplomarbeit am Institut für Org. Chemie, Universität Würzburg im Arbeitskreis von Prof. Dr. Gerhard Bringmann: "Dichtefunktional und Hybrid- ab-initio Studien an Übergangsmetallkomplexen und zu stereoselektiven Radikalcyclisierungen"
seit 05.1998	Doktorarbeit am Institut für Org. Chemie, Universität Würzburg im Arbeitskreis von Prof. Dr. Gerhard Bringmann: "3D-QSAR Untersuchungen an antimalaria-aktiven Naphtylisochinolin-Alkaloiden"
seit 05.1998	Wissenschaftlicher Mitarbeiter am Institut für Organische Chemie

Erklärung

Hiermit erkläre ich ehrenwörtlich, daß ich die Dissertation “3D-QSAR-Untersuchungen an antimalaria-aktiven Naphthylisochinolin-Alkaloiden” selbständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe. Ich erkläre außerdem, daß diese Dissertation weder in gleicher oder anderer Form bereits in einem anderen Prüfungsverfahren vorgelegen hat. Ich habe früher außer den mit dem Zulassungsgesuch urkundlich vorgelegten Graden keine weiteren akademischen Grade erworben oder zu erwerben versucht.

Würzburg, den