

1	Einleitung	1
1.1	Problemstellung	1
1.2	Rechnergestützte Methoden zur Ermittlung molekularer Ähnlichkeiten	4
1.3	Steroide und Rezeptoren	5
1.4	Zielsetzung und Aufbau der Arbeit	7
2	Ein Hybridalgorithmus zur Überlagerung mehrerer konformativ flexibler dreidimensionaler Strukturen	9
2.1	Warum genetische Algorithmen?	9
2.2	Prinzipien des Hybridalgorithmus	10
2.3	Methodische Grundlagen des Optimierungsalgorithmus	12
2.3.1	Die genetische Information	12
2.3.2	Initialisierung der Individuen	15
	2.3.2.1 Atomzuordnungstabellen	15
	2.3.2.2 Torsionswinkel	16
2.3.3	Optimierungskriterien	17
	2.3.3.1 Größe der Substruktur	17
	2.3.3.2 Distanzmaß	17
	2.3.3.3 Stereochemiedeskriptor	19
2.3.4	Die Pareto Optimierung	19
2.3.5	Die genetische Selektion	21
2.3.6	Die genetischen Operatoren	25
	2.3.6.1 Mutation	25
	2.3.6.2 Crossover	28
2.3.7	Zwei nicht-genetische Operatoren: Creep und Crunch	32
2.3.8	Der Directed-Tweak Algorithmus	35
2.3.9	Kontrolle der Überlappung der van der Waals-Radien	38
2.4	Überlagerungskriterien	42
2.5	Pareto-Diagramme	44

2.6	Parameter der genetischen Selektion47
2.7	Vergleich zweier Selektionstypen.51
2.8	Optimierung der Einsatzwahrscheinlichkeiten der Operatoren53
2.9	Die Bedeutung des Creep- und Crunch-Operators.54
2.10	Allgemeine Optionen des Programms.56
3	Extraktion von Pharmakophorpunkten aus der MCSS Suche	59
3.1	Angiotensin-II-Antagonisten59
3.1.1	Vergleich starrer und flexibler Überlagerungen	61
3.1.2	Multiple Überlagerungen	68
3.1.3	Auswirkung der Überlagerungskriterien auf die Substrukturen	70
3.2	Cytochrom P450c17-Inhibitoren74
3.2.1	Überlagerung mit wissensbasiertem Ansatz	77
3.2.2	Multiple Überlagerungen	88
4	Untersuchung der Affinität von Steroiden zu Antikörpern und künstlichen Rezeptoren	93
4.1	Steroide, Immunglobuline und künstliche Rezeptoren93
4.2	Methodische Grundlagen98
4.2.1	Vorhersage der Bindungsmodi der Liganden am Rezeptor	98
4.2.2	Berechnung der freien Bildungsenthalpie von Komplexen	99
4.2.3	Vergleich von physikochemischen Oberflächeneigenschaften der Steroide mit Hilfe neuronaler Netze	100
4.2.3.1	<i>Neuronale Netze</i>	100
4.2.3.2	<i>Die Struktur des Kohonen neuronalen Netzes</i>	101
4.2.3.3	<i>Das Lernverfahren in Kohonen neuronalen Netzen</i>	104
4.2.3.4	<i>Vergleich von Moleküloberflächen</i>	107
4.2.3.5	<i>Berechnung des elektrostatischen Oberflächenpotentials</i>	110
4.2.3.6	<i>Berechnung der Oberflächenhydrophobizität</i>	110

4.3	Ergebnisse	112
4.3.1	Die Bindungskonfigurationen von Digoxin-verwandten Steroiden im Komplex mit dem Immunoglobulin-Rezeptor 26-10	112
4.3.2	Die Bindungskonfigurationen verschiedener Steroide im Komplex mit einem synthetischen Paracyclophanrezeptor	117
4.3.3	Die freien Bildungsenthalpien der Cyclophan-Steroid-Komplexe	126
4.3.4	Die physikochemischen Oberflächeneigenschaften der Steroide	130
5	Zusammenfassung	143
5.1	Der Hybridalgorithmus zur Überlagerung konformativ flexibler Strukturen	143
5.2	Relevante Wechselwirkungen in Steroid-Rezeptorsystemen	145
6	Literaturverzeichnis	147
A	Anhang	
A.1	Grundlagen rechnerischer Methoden	A-1
A.2	Publikationsliste	A-5
A.3	Lebenslauf	A-7

Abkürzungen, Begriffe und genetische Größen:

Abkürzungen und Begriffe:

- ESP** elektrostatsisches Potential
- GA-Experiment** entspricht einem GA-Lauf; um den gesamten Lösungsraum abzudecken müssen grundsätzlich mehrere Experimente durchgeführt werden; Angabe der Experimentanzahl in der Variable e
- Fitness**..... Gütemaß
- flexible Bindungen** .. alle Bindungen, die nicht in Ringen fixiert sind und die nicht nur einatomige Substituenten aufweisen
- Free Energy**
- Perturbation** Methode zur Berechnung der freien Energien bzw. Enthalpien
- Generation** eine Population; Angabe der Generationsanzahlen in der Variable g
- GA** Genetischer Algorithmus
- GA Lauf** einmaliges Durchlaufen aller Generationen; Angabe der Anzahl der Generationen in der Variable g
- HYP** Hydrophobizitäts-Potential
- Individuum** zwei Chromosomen: eines zur Kodierung der Atomzuordnungstabellen (Matchlisten) der Überlagerung und eines zur Kodierung der Torsionswinkel der frei drehbaren (flexiblen) Bindungen
- Matchliste** Atomzuordnungstabelle
- Matchpaar** Atompaar bestehend aus jeweils einem Atom zweier Moleküle
- Matchtupel** Atomtupel bestehend aus jeweils einem Atom aller Moleküle
- multiple Überlagerung**..... gleichzeitige Überlagerung mehrerer Moleküle
- MC** Monte Carlo Optimierungsmethode
- MCSS** größte gemeinsame Substruktur mehrerer chemischer Verbindungen (engl.: maximum common substructure)
- MD** Molekulare Dynamik (engl.: molecular dynamics)
- MLP**..... molekulares Lipophiliepotential (engl: molecular lipophilicity potential)
- MM** Molekülmechanik (engl.: molecular mechanics)
- Population**..... mehrere Individuen; Angabe der Populationsgröße in der Variable i
- RTS**..... Selektion mit eingeschränktem Wettkampf (engl.: restricted tournament selection)
- SA** Simulation eines Abkühlungsprozesses (engl.: simulated annealing)

-
- VAMP**..... Molecular Modelling Programmpaket (engl.: vectorized semiempirical molecular orbital package)
- VESPA-Ladungen** Atompartialladungen basierend auf dem elektrostatischen Oberflächenpotential (engl.: electrostatic potential charges)

Genetische Größen:

- D* Die geometrische Fitness eines Individuums, Distanzmaß
- e* Anzahl der Experimente
- g* Anzahl der Generationen
- i* Größe der Population (Anzahl der Individuen)
- N* Anzahl der Matchtupel der Atomzuordnungstabelle
- n* Anzahl der Moleküle bei multiplen Überlagerungen
- n_{tor}* Die flexiblen Bindungen aller Moleküle einer Überlagerung
- S* Die stereochemische Fitness eines Individuums,
Stereochemiedeskriptor
- S_{rts}* Größe der Subpopulation für die Selektion

Verzeichnis der chemischen Verbindungen:

- 1 Cholesterol
- 2 Digoxin
- 3 Cholsäure
- 4 Angiotensin II
- 5 Losartan
- 6 Iproniazid
- 7 Disulfiram
- 8 L-158,809
- 9 β -Östradiol
- 10 p',p'-DDT
- 11 Saxitoxin
- 12 Tetrodotoxin
- 13 L-159,894
- 14 CV-11947
- 15 Imidq
- 16 BW13
- 17 BW112
- 18 BW61
- 19 BW62
- 20 Paracyclophan
- 21 Digitoxin
- 22 Digoxigenin
- 23 Ouabain
- 24 Desoxycholsäure
- 25 Chenodesoxycholsäure
- 26 Ursodesoxycholsäure
- 27 Lithocholsäure
- 28 Hydrocortison
- 29 Cortison
- 30 Testosteron
- 31 Progesteron

1 Einleitung

1.1 Problemstellung

Molekulare Erkennung am Rezeptor ist ein Grundprinzip in der pharmazeutischen Wirkstoffforschung. Für die Suche nach neuen Leitstrukturen spielen daher molekulare Ähnlichkeiten innerhalb verschiedener Verbindungen eine große Rolle. Sie helfen vor allem bei der Vorhersage des biologischen Wirkprofils von Verbindungen. Die Grundlage dafür ist, daß die spezifische Bindung an einen Rezeptor durch das von E. Fischer postulierte „Schlüssel-Schloß-Prinzip“ bestimmt wird. Strukturell ähnliche Verbindungen sollten daher ähnliche biologische Wirkprofile aufweisen, das heißt eine ähnliche Affinität zu einem bestimmten biologischen Rezeptor aufweisen. Struktur-Aktivitäts-Analysen können, besonders wenn die Struktur des Rezeptors unbekannt ist, auf der Analyse der Ähnlichkeiten innerhalb der Liganden basieren.

Einfachste Analysen molekularer Ähnlichkeiten werden häufig durch einen Vergleich der Konstitutionen der Verbindungen durchgeführt. Die Konstitution einer Verbindung wird durch die topologische Konnektivität der Atome beschrieben. In Abbildung 1-1 sind die Kohlenstoffgerüste von Cyclopentan und n-Hexan gegeben, die deren topologische Konnektivität beschreiben.



Abb. 1-1: Die Konnektivität der Kohlenstoffgerüste von Cyclopentan und n-Hexan.

Die Bindungstasche des Rezeptors ist jedoch von räumlicher Gestalt. Molekulare Ähnlichkeiten nur in den Konstitutionen der Verbindungen zu suchen, wird somit nur selten ausreichen, um Aussagen über das Wirkprofil einer Struktur zu treffen. Die Beschreibung der Verbindungen muß daher auf deren räumliche Ausdehnung, das heißt auf deren dreidimensionale Struktur erweitert werden. Abbildung 1-2 zeigt die dreidimensionale Darstellung von Cyclopentan und n-Hexan.

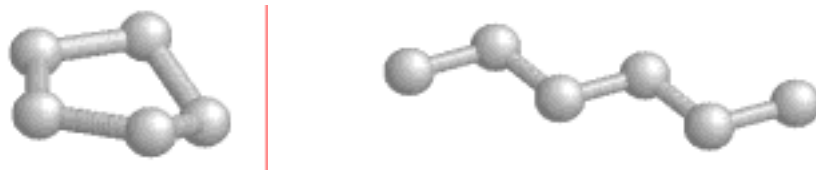


Abb. 1-2: Dreidimensionale Darstellung des Kohlenstoffgerüsts von Cyclopentan und n-Hexan.

Molekulare Ähnlichkeiten verschiedener Verbindungen im dreidimensionalen Raum lassen sich durch die Ermittlung der größten gemeinsamen Substruktur (engl.: maximum common substructure (MCSS) mehrerer dreidimensionaler Verbindungen analysieren. Je größer die ermittelte dreidimensionale Substruktur ist, desto größer ist dabei die molekulare Ähnlichkeit der untersuchten Verbindungen und desto wahrscheinlicher ist ein ähnliches biologisches Wirkprofil. Diese Vorhersagen hängen zudem nicht davon ab, ob die Struktur des Rezeptors bekannt ist. In Abbildung 1-3 ist die starre Überlagerung von Cyclopentan und n-Hexan dargestellt. Sie wurde aus je einer initialen Konformation für beide Verbindungen erhalten. Als größte gemeinsame dreidimensionale Substruktur ergeben sich drei Atome.

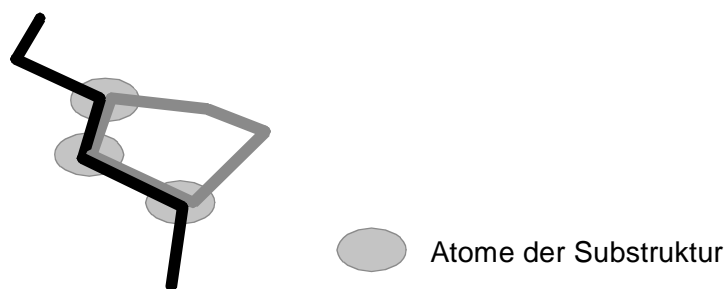


Abb. 1-3: Die starre Überlagerung der dreidimensionalen Strukturen von Cyclopentan (grau) und n-Hexan (schwarz). Die Substruktur umfaßt drei Atome und ist durch graue Ovale gekennzeichnet.

Viele biologisch aktive Verbindungen enthalten jedoch mindestens eine rotierbare Bindung. Sie können daher in vielen verschiedenen niedrig-energetischen Konformationen im dreidimensionalen Raum vorliegen. Bei Substruktursuchen mit je einer starren Struktur für jede Verbindung könnten daher leicht Ähnlichkeiten in wichtigen strukturellen Bereichen übersehen werden. Beachtet man bei der Überlagerung von Cyclopentan mit n-Hexan die konformative Flexibilität des n-Hexans, so wird die in Abbildung 1-4 dargestellte Überlagerung erhalten. Die ermittelte Substruktur umfaßt dabei bereits fünf Atome.

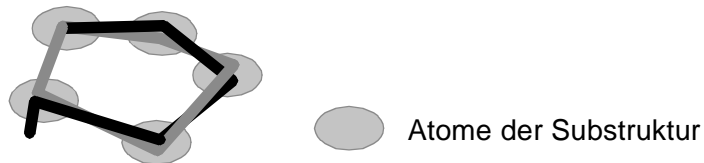


Abb. 1-4: Die flexible Überlagerung der dreidimensionalen Kohlenstoffgerüste von Cyclopentan (grau) und n-Hexan (schwarz). Die Substruktur umfaßt fünf Atome und ist durch graue Ovale gekennzeichnet.

Auf Basis dieser Überlagerungen und der daraus erfolgten Ermittlung der MCSS können für einen Satz an Verbindungen, die für die Affinität zu einem Rezeptor verantwortlichen strukturellen Bereiche herausgestellt werden. Daraus läßt sich ein gemeinsames Muster an pharmakophoren Punkten definieren oder Hinweise auf eine Leitstruktur erhalten.

Zusätzlich zu den strukturellen Ähnlichkeiten, die durch die MCSS ermittelt werden, sind auch Ähnlichkeiten der Oberflächen von Verbindungen und der darauf projizierten physikochemischen Eigenschaften für die Beurteilung der molekularen Erkennung wichtig. Bei vielen Ligand-Rezeptorsystemen spielt die Oberflächenkomplementarität eine große Rolle. Die Beschreibung der Struktur sollte somit zusätzlich auf die Oberfläche des zugrundeliegenden Gerüsts (Abbildung 1-5) erweitert werden. Die Oberflächenformen und zugehörigen physikochemischen Eigenschaften können dann ebenfalls für einen Satz von Liganden gleicher biologischer Aktivität verglichen werden.



Abb. 1-5: Die dreidimensionale van der Waals-Oberfläche von Cyclopentan.

Aus der Untersuchung der größten gemeinsamen dreidimensionalen Substruktur und der Analyse der Oberflächeneigenschaften können strukturelle und physikochemische Voraussetzungen ermittelt werden, die für ein bestimmtes biologisches Wirkprofil vorliegen müssen.

1.2 Rechnergestützte Methoden zur Ermittlung molekularer Ähnlichkeiten

Der Einsatz von Computern im Bereich der pharmazeutischen Wirkstoffforschung hilft weitgehend zur Senkung der Kosten und Zeiten, die für die Entwicklung neuer Leitstrukturen aufgebracht werden müssen. Rechnergestützte Methoden zum Auffinden der größten gemeinsamen dreidimensionalen Substruktur haben sich bereits zur Analyse molekularer Ähnlichkeit bewährt. Mit der Entwicklung von Programmen zur Berechnung dreidimensionaler Strukturen aus deren topologischer Konnektivität^{[1][2]} wurden Methoden gesucht, Ähnlichkeiten im dreidimensionalen Raum zu definieren. Die ersten Algorithmen zur Suche nach der größten gemeinsamen dreidimensionalen Substruktur basierten auf einer einzelnen, starren Konformation für jede Verbindung, ohne dabei die konformative Flexibilität in Betracht zu ziehen.^{[3]-[6]} Das Auffinden der MCSS erfolgte dabei meist durch interatomare Distanzvergleiche. Die erste detaillierte Studie an solchen distanzbasierten Methoden zur Suche nach dreidimensionalen Ähnlichkeiten wurde 1991 durch Pepperrell und Willett veröffentlicht.^[7] Daraufhin stellten Sheridan et al. 1996 eine distanzbasierte Methode vor, die zur Lösung der konformativen Flexibilität der Strukturen bereits mehrere Konformationen für eine Struktur in Betracht zog.^[8] Weitere Möglichkeiten zur Berechnung dreidimensionaler Ähnlichkeiten ergeben sich aus winkelbasierten^{[9][10]} und fragmentbasierten Ansätzen, wie z.B. bei Fisanick et al. beschrieben.^[11]

In unserem Arbeitskreis wurde von Markus Wagener eine Möglichkeit entwickelt die MCSS eines Datensatzes mittels direkter Atomüberlagerungen auf der Grundlage eines genetischen Algorithmus zu berechnen und zu optimieren.^[12] Ursprünglich war dieser Ansatz zum konstitutionellen Vergleich zweier Verbindungen gedacht. Die strukturellen Überlagerungen wurden basierend auf der Topologie eines Moleküls berechnet. Die Topologie ist dabei über die Konnektivität der Atome definiert. Daraufhin ließ sich zeigen, daß sich diese Methode auch auf die paarweise Behandlung dreidimensionaler Strukturen erweitern läßt.^[13] Mit der Einführung der Räumlichkeit der Verbindungen war wieder die Frage nach der konformativen Flexibilität gestellt. Deren Einbezug gelang durch die Möglichkeit zur Rotation der Einfachbindungen der beiden zu vergleichenden Strukturen während der Berechnung der Überlagerungen. Durch eine gegenseitige Anpassung der Konformationen während des Rechenprozesses wird gewährleistet, größere Substrukturen zu finden und die Suche mit nur einer Konformation für jede Verbindung zu starten.^[13] Damit kann auf der Basis von einer Konformation für jede Verbindung gearbeitet werden. In der vorliegenden Dissertation wird die Weiterentwicklung dieses Programms zur paarweisen Überlagerung dreidimensionaler Strukturen auf die gleichzeitige Überlagerung mehrerer Strukturen beschrieben. Der zugrundeliegende Algorithmus zur Optimierung der Überlagerungen ist ein Hybrid aus einer

genetischen und einer numerischen Optimierungsmethode. Um hoch energetische Konformationen, die bei der Rotation der Einfachbindungen entstehen können, zu vermeiden, wurde zusätzlich ein Verfahren zur Überwachung der Überlappung der van der Waals-Radien nicht gebundener Atome implementiert. Weiterhin wurden Möglichkeiten geschaffen, die es erlauben, selektives Wissen über die Bindungseigenschaften zwischen den entsprechenden Rezeptoren und den zu überlagernden Liganden einzubeziehen.

Die Untersuchung molekularer Ähnlichkeiten werden zudem auf den Vergleich der physikochemischen Oberflächeneigenschaften der Verbindungen erweitert. Diese spielen zusätzlich zu den strukturellen Ähnlichkeiten eine große Rolle bei der Betrachtung von Oberflächenkomplementaritäten zwischen Rezeptor und Ligand. Die hydrophoben oder elektrostatischen Wechselwirkungen der Oberflächen hängen nicht mehr nur von der zugrundeliegenden Gesamtstruktur ab, sondern auch von der relativen Position bestimmter funktioneller Gruppen. Für den Vergleich von Oberflächeneigenschaften und Oberflächenformen chemischer Verbindungen wurden neuronale Netze nach Kohonen eingesetzt. Diese erlauben durch die Reduktion der dreidimensionalen Oberflächenformen auf zweidimensionale Karten den direkten Vergleich der van der Waals-Oberflächen der Verbindungen und deren physikochemischer Eigenschaften.^{[14]-[18]}

1.3 Steroide und Rezeptoren

Steroide spielen in vielen medizinischen und biologischen Abläufen eine große Rolle. Das bekannteste Steroid ist das lipophile Cholesterin, **1**. Es stellt eine wichtige Komponente der Zellmembranen dar und nimmt eine Hauptfunktion in der Arteriosklerose ein. Ausgehend von dem Grundbaustein des Cholesterins, **1** startet auch die Biosynthese anderer wichtiger steroidaler Hormone, wie Glucocorticoide, Östrogene oder Androgene.^[19]

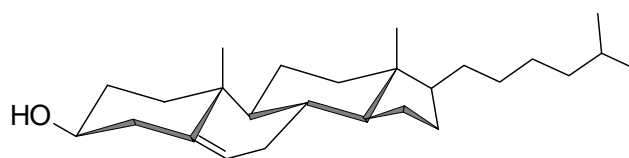


Abb. 1-6: Die Struktur des Cholesterins, **1**.

Alle Steroide haben das kondensierte Ringsystem 6-6-6-5 gemeinsam und unterscheiden sich hauptsächlich im Grad der Sättigung innerhalb der A- und B-Ringe, in der Anzahl der Sauerstoffsubstituenten sowie in der An- und Abwesenheit verschiedener Seitenketten.

Die Komplexierung von Steroiden durch Immunglobulin-Rezeptoren (Antikörper) oder künstliche Rezeptoren ist von besonderem Interesse. Es wurden bereits eine Reihe von Rezeptoren mit hohen Steroid-Affinitäten beschrieben.^{[19]-[23]} Studien über Antikörper mit hoher Affinität zu Steroiden wurden erstmals 1968 veröffentlicht.^[23] Seit dieser Zeit erfolgten intensive Untersuchungen an vielen weiteren Antikörper-Steroid-Komplexen, z.B. am DB3-Progesteron-Komplex,^[24] am 26-10- und 40-50-Digoxin-Komplex oder am 40-50-Ouabain-Komplex.^{[25]-[26]}

Grundsätzlich sind ziemlich aufwendige und kostspielige Prozesse notwendig, um Antikörper aus den entsprechenden Seren zu isolieren.^[31] Daher ist es von Vorteil, hochspezifische, künstliche Rezeptoren zu modellieren und synthetisieren, die eine hohe Affinität zu Steroiden aufweisen. Der Einsatz von künstlichen Systemen ist insbesondere für antikörper-katalysierte Reaktionen interessant.

Koga und Mitarbeiter komplexierten 1985 erstmals aromatische Steroide an künstlichen Paracyclophan-Rezeptoren.^[27] Systematische Studien wurden allerdings erst von Diederich et al. veröffentlicht, die bei verschiedenen künstlichen Rezeptoren, z.B. Paracyclophanen, eine hohe Affinität zu einer Reihe von Steroiden feststellen konnten.^{[19],[22],[23],[28]-[30]} Das allgemeine Grundgerüst des Paracyclophans ist in Abbildung 1-7 dargestellt. In das Innere des Cyclophanrings läßt sich ein Steroid einlagern.

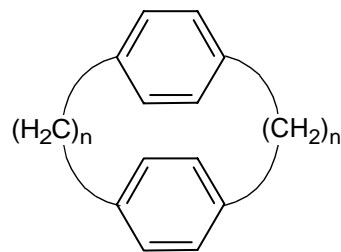


Abb. 1-7: Schematische Darstellung eines Paracyclophans.

Durch den Vergleich der Bindungseigenschaften der Steroide mit hoher Affinität zu Antikörpern und künstlichen Rezeptoren, lassen sich Gemeinsamkeiten oder Unterschiede in den Anforderungen verschiedener biologischer Systeme herausstellen.

1.4 Zielsetzung und Aufbau der Arbeit

Im Rahmen der vorliegenden Dissertation wird ein Algorithmus zur gleichzeitigen Überlagerung eines Datensatzes mit mehreren chemischen Verbindungen ähnlicher biologischer Aktivität beschrieben. Es ist ein auf den Vorarbeiten von M. Wagener basierender Algorithmus,^[12] der es erlaubt zwei chemische Verbindungen durch atomare Überlagerung miteinander zu vergleichen. Das hier implementierte Programm GAMMA (engl.: genetic algorithm for multiple molecule alignment) berechnet nun die gleichzeitige Überlagerung mehrerer chemischer Verbindungen. Zusätzlich erlaubt es die Berücksichtigung spezifischer Kenntnis über das zu optimierende Überlagerungsproblem und die entsprechende Rezeptorbindungstasche. Eine zielorientierte Konformationssuche läßt sich dabei durch die Kombination evolutionstheoretischer Ansätze mit einer numerischen Optimierungsmethode^[32] erreichen. Dieser Hybridansatz aus einem genetischen Algorithmus und der numerischen *Directed Tweak* Methode^[32] berücksichtigt die konformative Flexibilität der Verbindungen durch Rotation der Einfachbindungen während des Optimierungsprozesses. Die entstandenen Konformationen werden dabei auf eine Überlappung der van der Waals-Radien nicht gebundener Atome untersucht. Dies vermeidet es zu hoch energetische Konformationen zu erhalten, die das System nicht einnehmen kann.

Die ersten Kapitel zur Beschreibung des Hybridalgorithmus geben einen Überblick über den Ablauf des Algorithmus und der zugrundeliegenden Methoden (Kapitel 2.2 und 2.3). Anschließend werden spezifische programmtechnische Alternativen erläutert, wie z.B. der Vergleich zweier Selektionsarten und verschiedener Überlagerungskriterien sowie Parameteroptimierungen innerhalb des genetischen Algorithmus (Kapitel 2.4 bis 2.10).

Anhand der Überlagerungen von Angiotensin-II-Antagonisten und von Cytochrom P450c17-Inhibitoren lassen sich die Effektivität und die vielfältigen Möglichkeiten des Programms zeigen. Dabei werden die unterschiedlichen Ergebnisse diskutiert, die sich z.B. durch Angabe verschiedener Überlagerungskriterien oder unter Berücksichtigung von spezifischem Wissen über die Bindungseigenschaften der Rezeptortasche ergeben (Kapitel 3.1 und 3.2). Angiotensin-II- Antagonisten sind von hoher medizinischer Relevanz, da sie den letzten Schritt des Renin-Angiotensin-Systems blockieren. Sie spielen eine Schlüsselrolle in der Blutdruckregulation sowie im Elektrolyt- und Wasserhaushalt.^{[33][34]} Das P450c17-Enzym stellt das Schlüsselenzym der Androgen- und Glucocorticoidbiosynthese dar.^[35] Die Hemmung des P450c17-Enzyms ist ein vielversprechendes Konzept zur Behandlung des Prostatacarcinoms. Allerdings sind Inhibitoren dieses Enzyms mit annehmbarer Selektivität bislang noch nicht beschrieben.^{[36][37]}

Weiterhin wurden Untersuchungen zu den Grundlagen von Rezeptor-Ligand-

Wechselwirkungen durchgeführt. Diese Ergebnisse sollen vor allem zum besseren Verständnis der auftretenden Wechselwirkungen zwischen Rezeptoren und deren Liganden beitragen und dabei helfen, neue Rezeptorsysteme zu entwickeln. Eine Reihe qualitativer und quantitativer rechen-technischer Methoden dienten dabei zur Analyse der Affinität von Steroiden zu künstlichen und Immunoglobulin-Rezeptoren. Qualitativer Natur waren die Simulationen der Ligand-Rezeptor Docking-Prozesse und der Vergleich physikochemischer Oberflächeneigenschaften mit Hilfe von Kohonen neuronalen Netzen.^[38] Die Berechnungen der freien Energien der Komplexbildungsreaktionen gehörten dagegen zu den quantitativen Methoden (Kapitel 4.1 und 4.2).

Mit diesen Analysen ließen sich die Bindungsmodi des Digoxin, **2**, (Abbildung 1-8 links) und verwandter Steroide an den 26-10-Antikörper^[25] und die Bindungsmodi von Cholsäure, **3**, (Abbildung 1-8 rechts) und verwandter Steroide an ein Paracyclophan^[22] untersuchen (Kapitel 4.3). Digoxin, das sich aus dem roten Fingerhut extrahieren läßt, erlangt seine hohe medizinische Relevanz durch den Einfluß auf Herzaktivität und Affinität zu Na/K-Kanälen. Die Cholsäuren gehören zur Gruppe der Gallensäureverbindungen, deren Hauptfunktion auf dem Gebiet der Fettverdauung liegt.^[39]

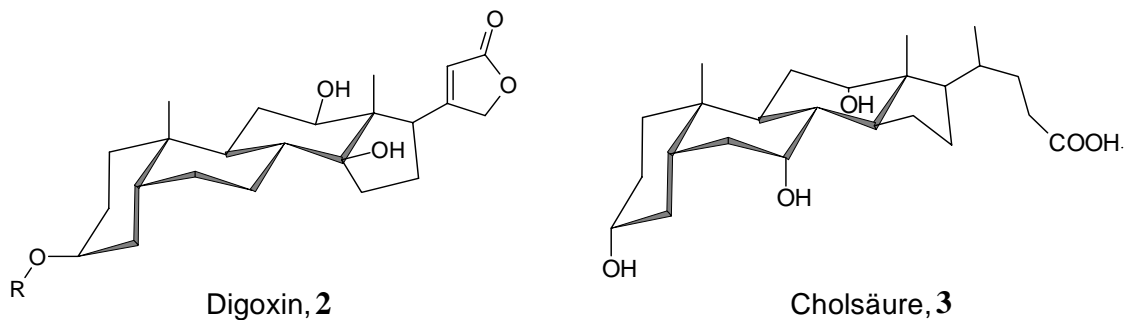


Abb. 1-8: Die Struktur von Digoxin, **2** und Cholsäure, **3**.

2 Ein Hybridalgorithmus zur Überlagerung mehrerer konformativ flexibler dreidimensionaler Strukturen

2.1 Warum genetische Algorithmen?

Die Grundlage der hier vorgestellten Hybridmethode zur gleichzeitigen Überlagerung mehrerer dreidimensionaler Strukturen unter Berücksichtigung der konformativen Flexibilität bildet ein genetischer Algorithmus (GA). Genetische Algorithmen (GA) sind robuste Optimierungsverfahren, die auf der Grundlage der natürlichen Genetik und Selektion agieren.^{[40]-[43]} Sie optimieren Lösungen effizient innerhalb großer Suchräume und lassen sich auf Probleme anwenden, die außerhalb der Möglichkeiten liegen, die mit normalen Suchalgorithmen zu erfassen sind.^{[44][45]} Ein genetischer Algorithmus imitiert die in der Natur vorgegebenen Prinzipien, über die ein vorhandenes Individuum versucht sich an eine verändernde Umwelt anzupassen. Die Individuen des genetischen Algorithmus sind dabei bestimmte Lösungen eines Problems und werden durch die genetischen Operatoren Mutation und Rekombination modifiziert. Die Selektion der Individuen, die die eigentliche Optimierung vorantreibt, orientiert sich dabei an Darwin's Theorem über das „*survival of the fittest*“.

Innerhalb der letzten Jahre hat der Einsatz genetischer Algorithmen auf Problemstellungen aus der Chemie deutlich zugenommen. Eine aktuelle Literaturrecherche im August 1999 über den *Chemical Abstract Service* ergab bei einer kombinierten Suche nach den Stichwörtern „genetic“ bzw. „evolution“ mit „algorithm“ eine Steigerung der Publikationszahl von 133 Publikationen im Jahr 1995 zu 256 im Jahr 1998 (Abbildung 2-1: graue Balken). In einer enger gefaßten Suche nach einer Kombination der oben genannten Begriffe mit den Stichwörtern „chemistry“, „bio“, „QSAR“, „structure-activity relation“, „bioactivity“ und „similar“ ließ sich für die Jahre 1995 bis 1998 eine Steigerung der Publikationen von 48 auf 104 in diesen Forschungsgebieten ermitteln (Abbildung 2-1: schwarze Balken).

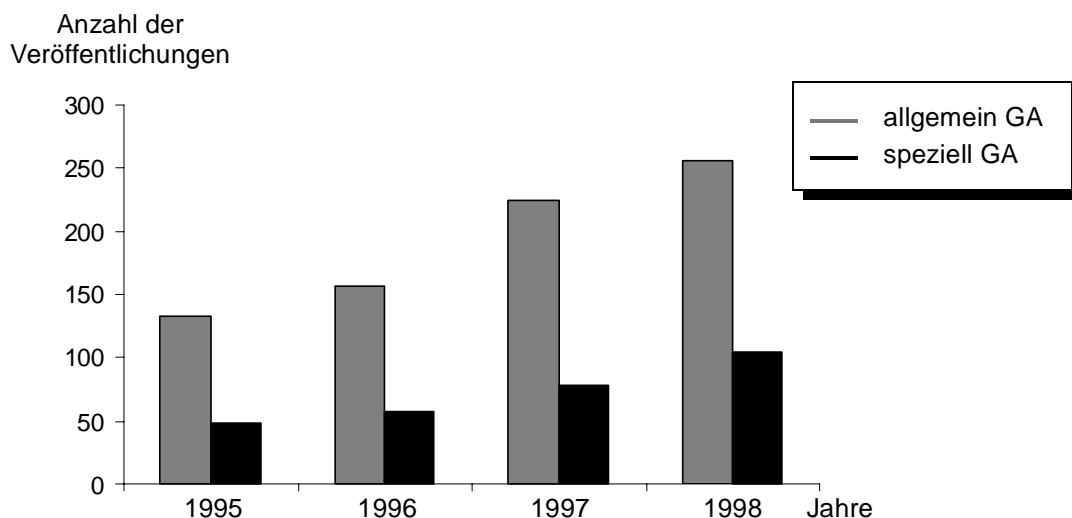


Abb. 2-1: Wissenschaftliche Veröffentlichungen (Quelle: CAS-Online) über den Einsatz genetischer und evolutionstheoretischer Algorithmen in der Chemie allgemein (grau) und in Verbindung mit den Begriffen Chemie, Biologie, QSAR und Bioaktivität (schwarz).

Genetische Optimierungsmethoden lassen sich auf vielfältige Problemstellungen anwenden, wie beispielsweise auf die Optimierung quantitativer NMR-Parameter,^[46] die Vorhersage der biologischen Aktivität chemischer Verbindungen,^{[47][48]} Konformationsanalysen^[49] und Strukturoptimierungen.^[50] Besonders im Bereich der konformativ flexiblen Suche nach strukturellen Ähnlichkeiten im dreidimensionalen Raum hat sich der Einsatz genetischer Methoden vielfältig bewährt.^{[51]-[58]}

2.2 Prinzipien des Hybridalgorithmus

Das implementierte Programm basiert auf einer bereits vorhandenen Methode zur paarweisen Überlagerung dreidimensionaler Strukturen^[12] und wurde auf die gleichzeitige Behandlung mehrerer chemischer Verbindungen erweitert. Das Ziel des Hybridalgorithmus ist die Optimierung der atomaren Überlagerung eines Satzes an dreidimensionaler chemischer Strukturen. Der Begriff „Hybrid“ wurde eingeführt, da es sich um eine Kombination einer genetischen und einer numerischen Optimierungsmethode handelt. Um einen ersten Überblick über die Prozesse zu geben, die sich während der Optimierung abspielen, ist in Abbildung 2-2 der schematische Ablauf des implementierten Hybridalgorithmus dargestellt. Die Optimierung beginnt nicht an einem einzelnen Startpunkt, sondern an einer Population verschiedener Individuen (Anfangspopulation). Jedes Individuum kann aus mehreren unabhängigen Chromosomen bestehen. Diese Chromosomen repräsentieren potentielle Teillösungen des Suchproblems. Im

vorliegenden Fall kodiert ein Chromosom die atomare Überlagerung und ein weiteres Chromosom die Änderung der Konformationen der zu vergleichenden Verbindungen. Eine Kombination der jeweiligen Teillösungen entspricht dann der Gesamtlösung eines Problems, somit einem Individuum.

Die Individuen der Anfangspopulation werden zufällig initialisiert. Darauf erfolgt die Selektion. Die Wahrscheinlichkeit für ein Individuum durch die Selektion in die Nachfolgegeneration übernommen zu werden, ist abhängig von dessen Güte (Fitness). Die Güte stellt ein Maß für die Angepaßtheit des Individuums an das zu optimierende Problem dar. Nach der Selektion wird durch die Anwendung der Operatoren und der damit verbundenen Modifizierung der Individuen eine Folgegeneration erstellt. Konsequenterweise beginnt somit eine komplette GA Optimierung mit der Initialisierung der Anfangspopulation und endet bei einem Satz optimierter Lösungen nach Durchlaufen der angegebenen Anzahl der Generationen.

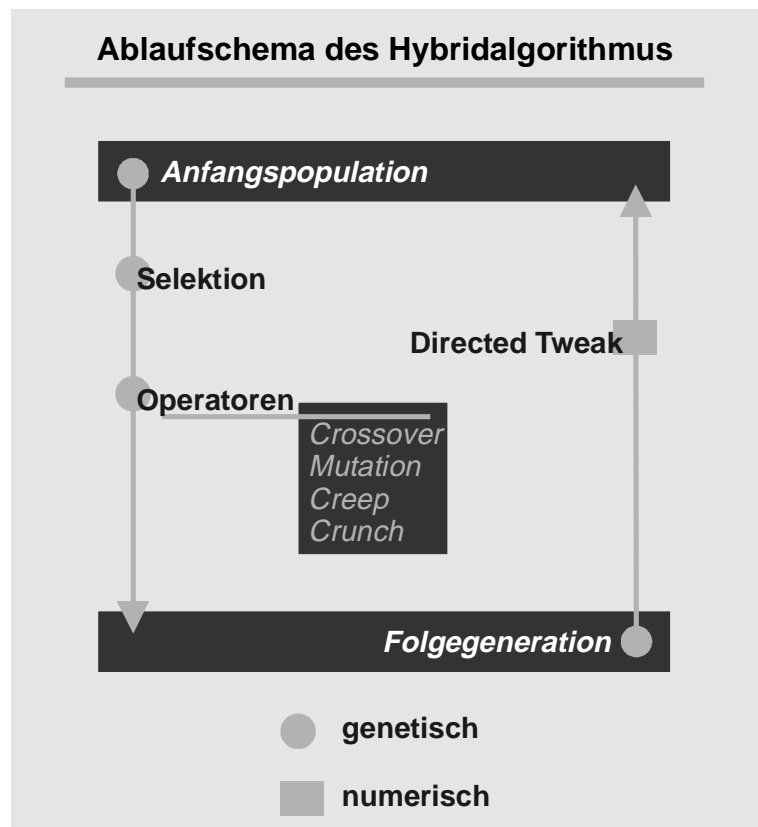


Abb. 2-2: Der schematische Ablauf des implementierten Hybridalgorithmus.

Die genetischen Operatoren *Selektion*, *Mutation* und *Rekombination* (im folgenden *Crossover* genannt), werden iterativ mit einer bestimmten, vorgegebenen Wahrscheinlichkeit auf

die Individuen der Population angewendet. In der hier vorgestellten Methode wurden zwei weitere Operatoren *Creep* und *Crunch* (engl.: creep, kriechen und crunch, zerstückeln) eingeführt, die auf das spezifische Problem der MCSS Suche zugeschnitten sind und nicht nur zufallsorientiert arbeiten.

Durch einen GA werden Lösungen auf eine nicht deterministische Weise erhalten. Diese Lösungen stellen nicht notwendigerweise das globale Minimum eines Suchraums dar. Um dieses Problem zu eliminieren, wird auf die Individuen jeder Folgegeneration die numerische Optimierungsmethode des *Directed-Tweaks*^[32] angewendet. Durch das *Directed-Tweak* wird der Konformationsraum der zu überlagernden Verbindungen innerhalb des Winkelraums der rotierbaren Bindungen zielgerichtet aneinander angepaßt. Die Güte der Überlagerung, die nach Anwendung der numerischen Optimierung für ein Individuum berechnet wird, stellt die Grundlage für die erneute Selektion dar. Um einen optimalen Lösungssatz zu erhalten, müssen üblicherweise mehrere GA Optimierungsläufe durchgeführt werden.

2.3 Methodische Grundlagen des Optimierungsalgorithmus

2.3.1 Die genetische Information

Bei der Entwicklung genetischer Algorithmen ist die Kodierung der möglichen Lösungen eines Problems eine grundlegende Aufgabe. Die Lösungen müssen in eine für den genetischen Algorithmus behandelbare Form transformiert werden.^{[43][51]} Die Natur speichert alle möglichen Lösungen eines Problems in Form von Chromosomen. Die Anordnung einer bestimmten Anzahl verschiedener Chromosomen charakterisiert die gesamte genetische Information eines Individuums.

Bei der Behandlung der Überlagerung konformativ flexibler dreidimensionaler Strukturen sollen durch die genetische Information die mögliche Überlagerungen mehrerer Strukturen beschrieben werden. Diese genetische Information, bzw. das Individuum, besteht dabei aus zwei unabhängigen Teilbereichen, das heißt zwei *Chromosomen*. Das erste Chromosom kodiert die Atomzuordnungstabelle oder Matchliste (siehe Abbildung 2-3 oben). Es enthält die Information über die Zuordnungen der Atome der einzelnen Moleküle zueinander. Dabei werden grundsätzlich nur nicht-Wasserstoffatome berücksichtigt. Das zweite Chromosom kodiert die durchzuführenden Konformationsänderungen in den einzelnen Molekülen. Die Änderungen der Torsionswinkel werden in Form eines *Bitstring* kodiert. Die Winkelwerte werden dabei 8-bit binär dargestellt (siehe Abbildung 2-3 unten).^{[52][53]} Beide Chromosomen, das heißt die Atomzuordnungstabelle und der Torsionswinkelbitstring bilden zusammen ein Individuum.

Mehrere dieser Individuen, also verschiedene Überlagerungen von Molekülen, stellen die Population einer Generation dar.

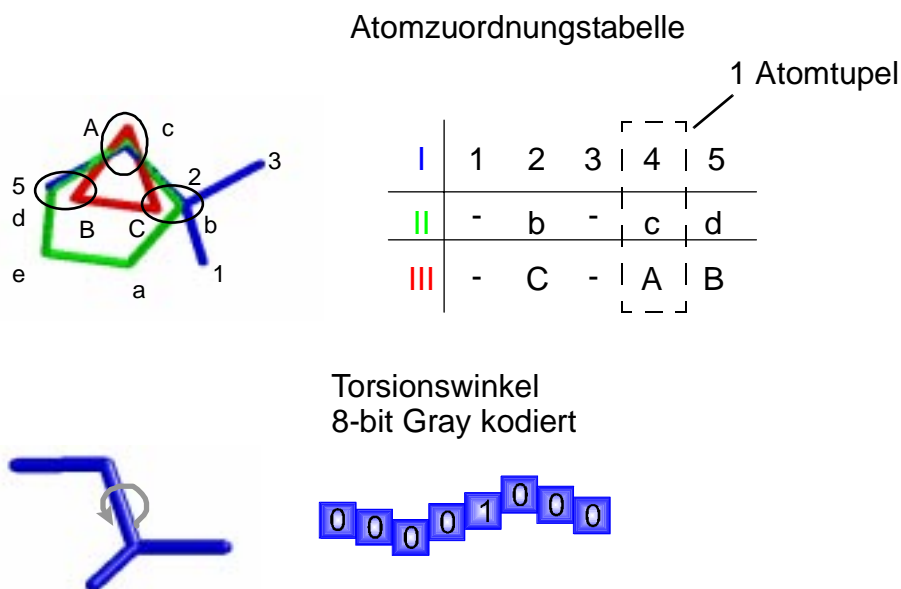


Abb. 2-3: Die genetische Information besteht aus zwei Teilbereichen: Die Atomzuordnungstabellen und die Torsionswinkel.

Die Atomzuordnungstabelle (Matchliste) wird für alle Individuen gleich lang festgelegt und zwar durch die Atomanzahl des größten Moleküls. Jedes Atom ist in der Atomzuordnungstabelle nur einmal vertreten, d.h. Doppelzuordnungen sind ausgeschlossen. In der Atomzuordnungstabelle sind die Moleküle nach Größe sortiert. Zur Größe der Substruktur tragen lediglich Zuordnungen, bei denen von jedem Molekül ein Atom beteiligt ist, bei. Eine Zuordnung dieser Art wird Match- oder Atomtupel genannt.

Bei der Torsionswinkelkodierung stößt man auf zwei grundlegende Probleme: Zum einen sollte die Verteilung der Torsionswinkel so groß sein, daß sie den gesamten Konformationsraum abdeckt. Zum anderen führt eine zu große Streuung der Torsionswinkel zu Konvergenzproblemen und damit zu hohen Rechenzeiten. Rotierbar oder flexibel ist eine Bindung dann, wenn beide gebundenen Atome mindestens einen weiteren mehratomigen Substituenten (z.B. eine Methylgruppe) aufweisen. Jeder Torsionswinkel wird in 8 bit binär kodiert. Die Torsionswinkel aller Moleküle lassen sich in einem Bitstring zusammenfassen, d.h. für n_{tor} Torsionswinkel (in Abbildung 2-3 $n_{tor} = 1$) entsteht ein $8n_{tor}$ langer Bitstring. Bei dieser Kodierung wird die *Gray-Form*^[40] der binären Darstellung gewählt. Diese bietet den Vorteil, daß sich aufeinander folgende Integerwerte nur in einem Bit unterscheiden, im Gegensatz zur normalen 2-Komplement-Darstellung der binären Kodierung von Integerwerten. Bei der 2-Komplement-

Darstellung ändern sich bei einem Schritt von einem Integerwert zum nächsten grundsätzlich mehrere Bits. In Tabelle 2-1 wird der Unterschied zwischen Gray und 2-Komplement-Kodierung der Integerzahlen 0 bis 15 aufgezeigt werden.

Tab. 2-1: Binär und Gray Kodierung vom Integer Zahlen

Integerzahl	2-Komplement-Kodierung	Gray-Kodierung
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100
9	1001	1101
10	1010	1111
11	1011	1110
12	1100	1010
13	1101	1011
14	1110	1001
15	1111	1000

Der erste Schritt der Winkelkodierung ist die Umwandlung der Winkel von -180° bis $+180^\circ$ in Integerzahlen von 0 bis 256. Der Winkel -180° entspricht dann der Zahl 0, $+180^\circ$ entspricht 255 und 0° entspricht 128. Als zweiter Schritt wird diese Integerzahl in einen 8-bit Gray kodierten String transformiert. Die kleinst mögliche Winkeländerung ergibt sich als 1.4° ($256/180^\circ$). Der Vorteil der Gray-Kodierung spiegelt sich vor allem in den Mutationen der Winkel wieder und wird in Kapitel 2.3.6.1 erklärt.

2.3.2 Initialisierung der Individuen

2.3.2.1 Atomzuordnungstabellen

Der erste Schritt bei der Initialisierung der Individuen ist die Bildung aller möglichen Atomtupel, d.h. aller Kombinationen der Atome der einzelnen Moleküle. Die Anzahl der Möglichkeiten ist dabei: $N_1 \cdot N_2 \cdot N_3$, wobei N_x die Anzahl der Nicht-Wasserstoffatome in Molekül x ist.

Tab. 2-2: Mögliche Atomkombinationen von Molekül I, II und III.

Molekül I	Molekül II	Molekül III
1	a	A
1	a	B
1	a	C
1	b	A
1	b	B
1	b	C
...

Bereits beim Bilden dieser Atomkombinationen wird das jeweilige Überlagerungskriterium berücksichtigt. Das Überlagerungskriterium kann der Atomtyp, das heißt die Ordnungszahl oder ein relatives Intervall einer bestimmten Atomeigenschaft sein. Als Atomeigenschaft kann z.B. eine physikochemische Größe, wie die Atompartialladung gewählt werden. Die endgültige Zahl der Kombinationen kann somit auch kleiner als $N_1 \cdot N_2 \cdot N_3$ werden.

Die Individuen lassen sich durch zufällige Auswahl von Atomtupel aus der vorher initialisierten Menge bilden. Die Anzahl der Individuen ist durch die vorgegebene Größe i der Population bestimmt. Treten durch die zufälligen Kombinationen der Atomtupel Mehrfachzuordnungen auf, d.h. ein oder mehrere Atome kommen in der Liste mehrfach vor, so wird das Atom durch ein noch nicht in der Liste vorhandenes oder durch eine Nullzuordnung ersetzt. Nullzuordnungen sind in Abbildung 2-3 durch einen Längsstrich gekennzeichnet.

2.3.2.2 Torsionswinkel

Jede Torsionswinkeländerung wird zu Beginn in einem 8-Bitstring initialisiert, wobei gesetzte (1) und nicht-gesetzte (0) Bits zufällig gleichverteilt werden. Für jedes Individuum entsteht so ein $8n_{tor}$ (n_{tor} ist die Anzahl der rotierbaren Bindungen aller Moleküle) langer String.

Eine Torsionswinkeländerung ruft eine Rotation der entsprechenden rotierbaren Einfachbindung hervor. Die Transformation der Koordinaten der von der Rotation betroffenen Atome, wird durch die Rotationsmatrix bestimmt. Die Rotationsmatrix für jede Rotation einer Einfachbindung berechnet sich aus den x-, y- z-Koordinaten der Atome der rotierbaren Bindung (in Abbildung 2-4 Bindung 2-4) und der Torsionswinkeländerung. Eine Koordinatentransformation verschiebt das Atom der rotierbaren Bindung, das mit dem kleineren Teil der gesamten Molekülstruktur verbunden ist, in den Koordinatenursprung (siehe Abbildung 2-4, Atom 4). Bei jeder Rotation werden jeweils die Koordinaten der Atome des kleineren Teils der Molekülstruktur abgeändert (siehe Abbildung 2-4 Atom 5 zu 5'). Die Rotation der Bindung 2-4 führt zu einer Änderung des Diederwinkels 1-2-4-5 (α) zu 1-2-4-5' (α'). Die Ausgangskordinaten der ersten Rotation sind durch die eingelesene Konformation festgelegt.

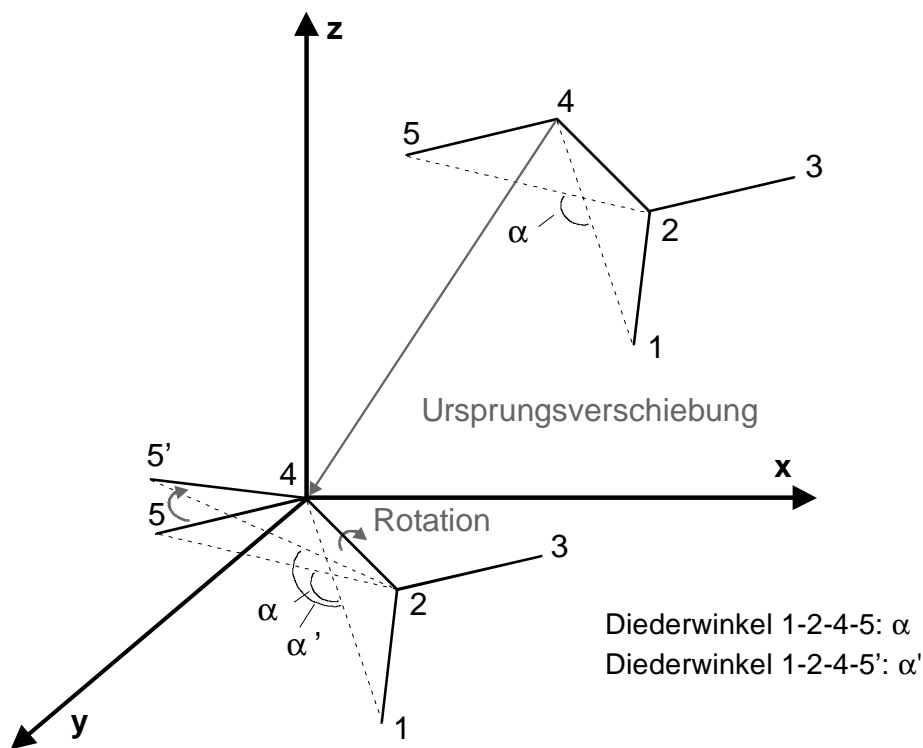


Abb. 2-4: Koordinatentransformation des Atoms 5 zu 5' durch Rotation der C-C Einfachbindung 2-4. Der Diederwinkel α ändert sich dabei zu α' .

2.3.3 Optimierungskriterien

2.3.3.1 Größe der Substruktur

Ein Ziel des Algorithmus ist die Errechnung der größten maximalen Substruktur zweier oder mehrerer Moleküle. Ein wichtiger zu optimierender Parameter ist daher die Größe der Substruktur. Diese Größe wird durch die Anzahl N der Atomtupel, die an der Überlagerung beteiligt sind, beschrieben. Bei der in Abbildung 2-5 dargestellten Überlagerung ist $N = 2$.

2.3.3.2 Distanzmaß

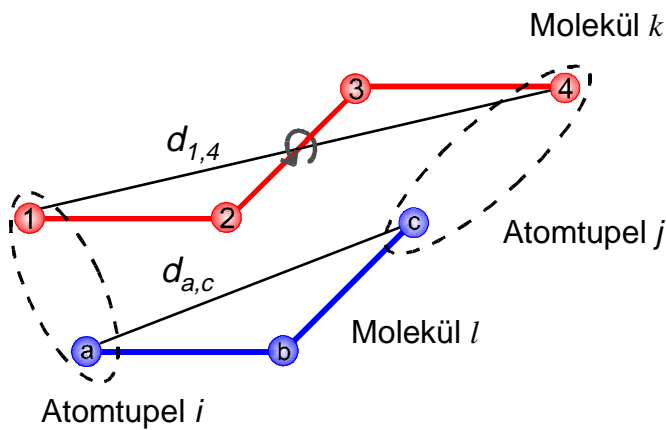
Das Distanzmaß D gibt die geometrische Güte der Überlagerung an. Er beschreibt, wie gut die Konformationen der zu vergleichenden dreidimensionalen Strukturen aneinander angepaßt sind und wird wie folgt berechnet:

$$D = \frac{1}{4} \frac{n \cdot (n - 1)}{2} \sum_{\substack{i,j \\ j \neq i}}^N \sum_{\substack{k,l \\ l \neq k}}^n (d_k(i, j) - d_l(i, j))^2 \quad (\text{Gl. 2-1})$$

mit:

i, j	=	Indizes der zu vergleichenden Atomtupel
$d_k(i, j), d_l(i, j)$	=	Atomdistanzen von Atomtupel i zu Atomtupel j in Molekül k und Molekül l
N	=	Anzahl der Atomtupel (Größe der Substruktur)
n	=	Anzahl der Moleküle

Der Distanzparameter beschreibt die Differenzen der molekül-internen Atomabstände der zu vergleichenden Atomtupel i und j . Da nur Überlagerungen gleicher Größe N zum Vergleich stehen, ist eine Normierung auf die Größe der Substruktur nicht nötig. Ein Beispiel zur Berechnung von D ist in Abbildung 2-5 dargestellt. D berechnet sich hier als Quadrat der Differenz der Atomabstände in Molekül k ($d_{1,4}$) der Atomtupel i und j und der Atomabstände in Molekül l ($d_{a,c}$) der Atomtupel i und j .



$$d_k(i, j) = d_{1,4}$$

$$d_l(i, j) = d_{a,c}$$

$$N = 2$$

$$n = 2$$

hier:

$$D = (d_{1,4} - d_{a,c})^2$$

Abb. 2-5: Das Distanzmaß D stellt die Güte der Überlagerungen dar.

Das Distanzmaß D steht im Verhältnis zu dem *rms*-Wert (root mean square error) (Gl. 2-2). Der *rms*-Wert berechnet sich für das Problem der multiplen MCSS Suche wie folgt:

$$rms = \sqrt{\frac{1}{2N \cdot n(n-1)} \left(\sum_i^N \sum_j^3 \sum_{\substack{k,l \\ l \neq k}}^n (a_{kj} - a_{lj})_i^2 \right)} \quad (\text{Gl. 2-2})$$

mit: N = Anzahl der Atomtupel (Größe der Substruktur)
 n = Anzahl der Moleküle
 a_{kj}, a_{lj} = x, y, z- Koordinaten der Atome der Moleküle k und l

Der Faktor $\frac{1}{2N \cdot n(n-1)}$ dient zur Normierung des *rms*-Wertes auf die Anzahl der Atomtupel und die Anzahl der Moleküle. Somit ist ein Vergleich von Überlagerungen, gleich welcher Größe oder Molekülanzahl, untereinander möglich.

Der *rms*-Wert ist selbst bei nur kleinen Änderungen in den Überlagerungen großen Schwankungen, unterworfen. Daher ist die Berechnung von D als Optimierungsparameter besser geeignet, als der *rms*-Wert. Zusätzlich läßt sich D schneller errechnen, da lediglich interne Distanzen berücksichtigt werden und nicht die absoluten Koordinaten der einzelnen Atome. Die Berechnung des *rms*-Wertes einer Überlagerung erfolgt einmalig am Ende jedes Programmlaufs.

2.3.3.3 Stereochemiedeskriptor

Im Falle, daß die zu überlagernden Strukturen Enantiomere sind, wären die berechneten D Parameter völlig identisch. Daher wurde ein Stereochemiedeskriptor S eingeführt. S besteht aus der Summe der Stereochemieparameter S_i über alle Atomtupel ($S = \sum S_i$). Der Stereochemiedeskriptor S_i eines Atomtupels ist wie folgt zu berechnen. Zunächst werden n ($n =$ Anzahl der Moleküle) Ebenen durch die jeweils drei Atome eines jeden Moleküls der ersten drei Atomtupel der Atomzuordnungstabelle aufgespannt. Eine Ebene ist durch drei Atome des ersten Moleküls aufgespannt, die anderen Ebenen durch die zugeordneten Atome der anderen Moleküle. In Abbildung 2-6 ist die Anzahl der Moleküle zwei. Dargestellt ist ein Matchpaar 1-a für das S_i berechnet werden soll und die sechs Atome, die zwei Ebenen aufspannen. Atom 1 gehört zum ersten Molekül und Atom a zum zweiten Molekül. Die Distanzen S_i und die Orientierungen der beiden Atome 1 und a zur entsprechenden zugehörigen Ebene in Molekül I und II sind zu berechnen. Liegen die Atome 1 und a auf verschiedenen Seiten der Ebenen, so unterscheiden sie sich in ihrer Stereochemie. Als Parameter S_i wird die größere der beiden berechneten Distanzen genommen. S kann nur dann ungleich Null sein, wenn die zugrundeliegenden Strukturen insgesamt vier oder mehr Atome haben.

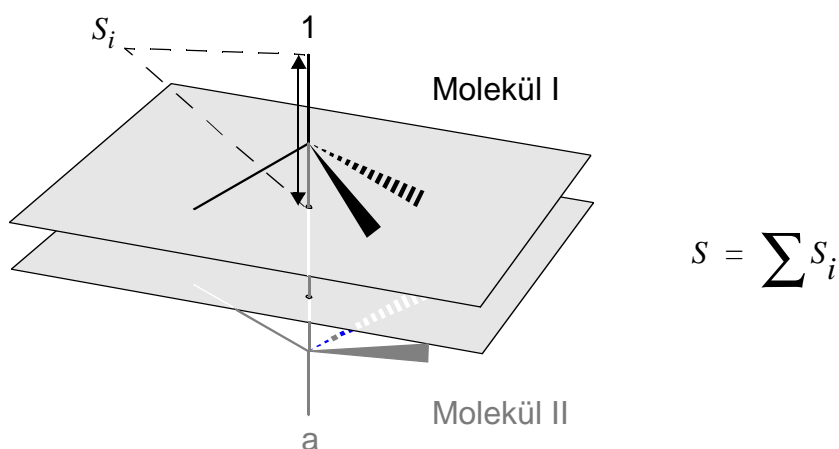


Abb. 2-6: Die Berechnung des Stereochemiedeskriptor S . Die Atome 1 und a sind Bestandteil eines Matchpaares und unterscheiden sich in ihrer Stereochemie.

2.3.4 Die Pareto Optimierung

Die Suche nach der maximalen größten Substruktur eines Datensatzes ist ein multidimensionales Optimierungsproblem mit verschiedenen Kriterien. Drei gegenläufige Parameter tragen zur Güte jeder Überlagerung bei: die Größe der Substruktur, das Distanzmaß und der Stereochemiedeskriptor. Die gesuchte Substruktur sollte einerseits so groß wie möglich sein,

andererseits sollten die Strukturen so gut wie möglich aufeinander passen. Es ist ein Optimum zu finden, das beide Kriterien in Betracht zieht. Die Optimierung muß dabei die Integrität beider Parameter respektieren. Vilfredo Pareto entwickelte ein Konzept zur Lösung multidimensionaler Optimierungen.^{[41][53]} Bei einer Optimierung nach Pareto ist ein optimierter Status dann erreicht, wenn kein Parameter mehr verbessert werden kann, ohne einen anderen zu verschlechtern.

Falls eine bestimmte Lösung einem Vektor entspricht von dem man annimmt er wäre besser als ein anderer, da er partiell kleiner ist, so ist eine mathematische Definition der Pareto Optimalität, daß Vektor \mathbf{u} partiell kleiner ist als \mathbf{v} , symbolisch $\mathbf{u} <_p \mathbf{v}$, wenn folgende Bedingungen eingehalten werden (eq 2-3):^[53]

$$(\mathbf{u} <_p \mathbf{v}) \Leftrightarrow (\forall i)(u_i \leq v_i) \wedge (\exists i)(u_i < v_i) \quad (\text{Gl. 2-3})$$

mit: u_i, v_i = Komponenten von \mathbf{u} und \mathbf{v}
 \forall, \exists = Allquantor (alle Elemente...), Existentialquantor (wenigsten ein Element...)

Unter diesen Voraussetzungen ist es möglich zu sagen, daß Vektor \mathbf{u} über Vektor \mathbf{v} dominiert. Falls Vektor \mathbf{u} weder dominierend noch nicht dominierend ist, sind beide Vektoren äquivalente Lösungen des Problems.

Die Anwendung der Pareto Optimierung auf das Problem der MCSS Suche im dreidimensionalen Raum resultiert in einer Maximierung der Größe der Substruktur unter gleichzeitiger Optimierung der geometrischen Anpassung. Dadurch erhält man nicht eine einzige möglichst perfekte Substruktur, sondern einen Satz an Substrukturen, d.h. für jede mögliche Substrukturgröße eine beste Überlagerung. Ein einmaliges Durchlaufen aller Generationen (d.h. die Durchführung eines Experiments oder Optimierungslaufs) führt zu einem Paretosatz, d.h. eine beste Lösung für jede mögliche Größe der Substruktur.

In Abbildung 2-7 ist die Anwendung der Pareto Optimierung auf die Überlagerung von Vinylcyclobutan und Propylcyclobutan dargestellt. Die grau markierten Atome schließen die Substruktur ein. Das Ergebnis der Pareto Optimierung ist ein Satz gemeinsamer Substrukturen, für die sich die geometrische Anpassung nicht weiter optimieren läßt. In den vier verschiedenen Überlagerungen von Vinylcyclobutan und Propylcyclobutan aus Abbildung 2-7 genügen drei der Pareto Optimierung.

Überlagerung IV zeigt große geometrische Distanzen in den einander zugeordneten Atomen, wogegen Überlagerung I mit der gleichen Anzahl an zugeordneten Atomen die weit bessere geometrische Anpassung bildet. Offensichtlich dominiert Überlagerung I Überlagerung IV (Abbildung 2-7). In diesem Sinne repräsentiert Überlagerung I eine Pareto optimale

Lösung, da keine andere Substruktur der gleichen Größe mit besserer geometrischer Anpassung gefunden werden kann. Überlagerungen II und III sind ebenfalls Mitglieder des Pareto-satzes. Es läßt sich keine andere Überlagerung gleicher Größe mit besserer geometrischer Anpassung finden. Zusammenfassend sind Überlagerung I, II, und III äquivalente Pareto Lösungen und keine dominiert über die andere.

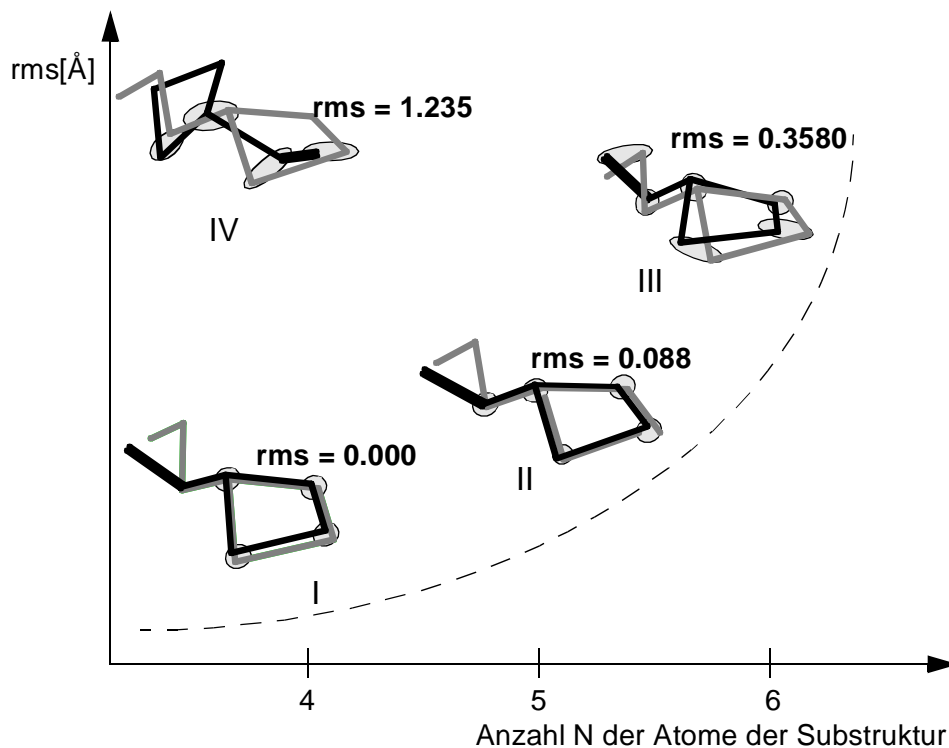


Abb. 2-7: Die Ergebnisse der Pareto Optimierung der Überlagerung von Vinylcyclobutan (schwarz) und Propylcyclobutan (grau). Der *rms*-Wert der Überlagerungen ist gegenüber der Größe *N* der Substrukturen aufgetragen. Die gestrichelte Linie markiert den sog. Paretosatz, dieser stellt alle Überlagerungen dar, die bei gleichbleibender Größe geometrisch nicht weiter aneinander angepaßt werden können.

Für jede Überlagerung kann die spezifische Paretokurve angegeben werden. Beispiele dazu sind in Kapitel 2.5 zu finden.

2.3.5 Die genetische Selektion

Durch die Selektion der Individuen wird die Optimierung vorangetrieben. Aufgrund des Evolutionsdrucks vermehren sich besser an die Optimierungsaufgabe angepaßte Individuen auf Kosten der schlechteren. Dieses Kriterium für die Angepaßtheit wird Güte oder auch Fit-

ness des Individuums genannt. Von einer Generation in die nächste werden sich nur die im Darwinschen Sinn besten Individuen in der Population durchsetzen. Durch die Selektion lassen sich Individuen mit einer Wahrscheinlichkeit abhängig von deren relativer Fitness, von einer Generation in die nächste übertragen. Die Güte für jedes zu behandelnde Problem ist eine nach eigenem Ermessen festgelegte Bedingung, z.B. der kleinste Abstand der überlagerten Atome bei Ermittlung der gemeinsamen Substruktur. Die meisten in der Literatur beschriebenen Genetischen Algorithmen behelfen sich der sogenannten Rouletterad-Selektion (engl.: *Roulette Wheel Selection*):^{[42]-[51],[52]} Dabei wird jedem Individuum (Abbildung 2-8: Individuum I, II, III und IV) ein Sektor eines Rouletterades zugewiesen. Die Größe des Sektors ist proportional zur Fitness des Individuums:

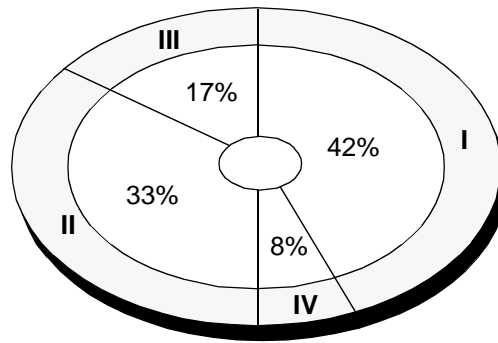


Abb. 2-8: Durch die Selektion werden die Individuen I bis IV gemäß ihrer prozentualen Anteile an der Gesamtgüte (hier als Fächer eines Rouletterades) in die nächste Generation übernommen.

Das gewichtete Rouletterad aus Abbildung 2-8 bietet ein anschauliches Bild dafür, daß manche Individuen entsprechend ihrer besseren Fitness häufiger ausgewählt werden. Die Anzahl der Fächer entspricht dabei der Anzahl der Individuen, ihre Größe dem zugeordneten prozentualen Anteil der jeweiligen Güte in Bezug auf die Güte der gesamten Population. Die Wahrscheinlichkeit p_i , daß ein Individuum i selektiert wird, berechnet sich wie folgt:^[53]

$$p_i = \frac{G_i}{\left(\sum_i G_i\right)} \quad (\text{Gl. 2-4})$$

mit: G_i = Güte oder Fitness eines Individuums i
 $\sum_i G_i$ = die Summe der Güten aller Individuen

Bei einer Drehung des Rouletterads, wird das Individuum 2 mit einer Wahrscheinlichkeit

von 0.33 selektiert, da es einen Anteil von 33% am Rouletterad besitzt. Um eine vollständige neue Population zu generieren, muß das Rad so oft gedreht werden wie Individuen in der Population enthalten sind, in diesem Beispiel vier mal.

Um vorzeitige Konvergenz zu einer nicht optimalen Lösung und die Abnahme der genetischen Vielfalt im Laufe der Generationen zu vermeiden, sollte die Anfangspopulation aus möglichst diversen, für die Lösung relevanten Mitgliedern (Individuen) bestehen. Einige in der Literatur beschriebene genetische Algorithmen nützen das Konzept der Nischen um eine vorzeitige Konvergenz zu vermeiden.^{[12][51][53]} Eine Nische ist ein lokales Minimum, das von mehreren ähnlichen Individuen besetzt wird. Mit Hilfe des „Crowding“ hat man eine Methode entwickelt, das Konzept der Nischen zu verwirklichen:^{[40][59]} ein neues Individuum, das durch die genetischen Operatoren generiert wurde, ersetzt nicht seinen direkten Vorfahren ersten Grades sondern ein Individuum aus einer zufallsmäßig ausgewählten Teilpopulation, die dem neuen Individuum möglichst ähnlich ist.

In unserem Ansatz wurde, um dem Verlust an genetischer Vielfalt vorzubeugen, der bei der Rouletterad-Selektion auftreten könnte, ein weiterer Typ der Selektion implementiert. Diese Alternative nennt sich „Selektion des eingeschränkten Wettkampfs“ (engl.: *Restricted Tournament Selection*) (RTS).^[60]

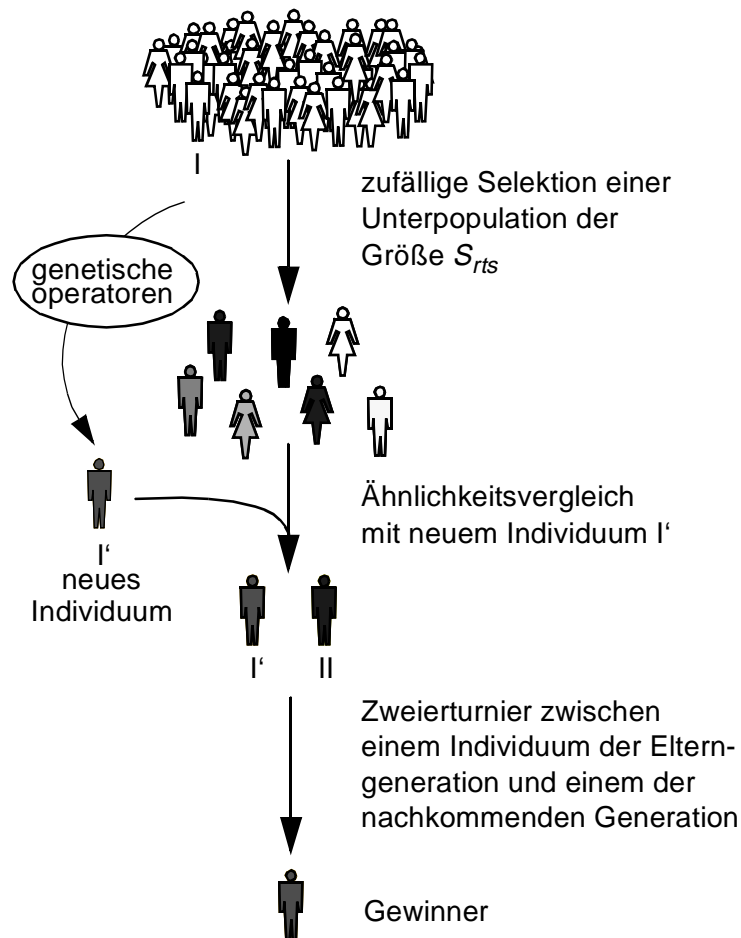


Abb. 2-9: Schematische Präsentation der „Restricted Tournament“-Selektion. Turniere werden nur zwischen ähnlichen Individuen abgehalten. Dies garantiert die Erhaltung der Vielfalt der genetischen Information.

Es hat sich herausgestellt, daß die *Restricted Tournament*-Selektion sehr hilfreich bei der Lösung mehrdimensionaler Probleme sein kann.^[60] Sie ist eine Erweiterung der Selektion der „Paarweisen Wettkämpfe“ (engl.: *Binary Tournament Selection*). Während einer *Binary Tournament*-Selektion werden Wettkämpfe um einen Platz in der neuen Generation zwischen jeweils zwei zufällig aus der Population gewählten Individuen abgehalten. In diesem Sinne beschreibt die „Selektion des eingeschränkten Wettkampfs“ oder *Restricted Tournament*-Selektion, daß die Wettkämpfe nicht zwischen irgendwelchen zufällig ausgesuchten Individuen ausgetragen werden, sondern nur zwischen ähnlichen Individuen.

Somit basiert die *Restricted Tournament*-Selektion auf dem Konzept der lokalen Konkur-

renz. Die Gewinner eines jeden Kampfes werden in die nächste Generation übernommen. In Abbildung 2-9 ist das Schema einer solchen Selektion dargestellt. Das Individuum I wird zufällig aus der Basispopulation ausgewählt und durch die Operatoren des genetischen Algorithmus in ein neues Individuum I' abgeändert. Für jedes Individuum I' wird eine kleine Teilpopulation der Größe S_{rts} ausgewählt. Das Individuum II aus der Teilpopulation, das am ähnlichsten zu I' ist wird gespeichert. I' konkurriert dann mit II um einen Platz in der neuen Generation. Diese Art der Einschränkung des *Binary Tournament* verhindert, daß ein Individuum mit einem zu diesem sehr unterschiedlichen Individuum konkurriert. Die Folge davon ist die Erhaltung der Vielfalt der genetischen Information. Ein weiterer Vorteil des beschriebenen Mechanismus des RTS ist die Möglichkeit zur kontinuierlichen Selektion; bei dieser Art der Selektion können Individuen verschiedener Generationen miteinander konkurrieren (z.B. II und I' in Abbildung 2-9).

2.3.6 Die genetischen Operatoren

2.3.6.1 Mutation

Im folgenden werden die genetischen Operatoren beschrieben. Dabei ist bei der Wirkungsweise der Operatoren zwischen den beiden Chromosomen des Individuums zu unterscheiden. Eines der Chromosomen kodiert - wie bereits in Kapitel 2.3.1 beschrieben - die Atomzuordnungen, das andere die Änderungen in den Torsionswinkeln.

Bei der Mutation der Zuordnungstabellen bzw. Matchlisten werden bei n Molekülen der Überlagerung $n-1$ Mutationsstellen zufällig ausgewählt (siehe Abbildung 2-10). Mit Ausnahme des Referenzmoleküls (größtes Molekül des Datensatzes) lassen sich die Atome aller Moleküle eines Atomtupels mutieren. Das entsprechende Atom des Atomtupels (bei der ersten Mutation das Atom des zweiten Moleküls, bei der zweiten Mutation das Atom des dritten Moleküls usw.) wird jeweils unter Beachtung folgender Grenzbedingungen mutiert: ein Atom eines Moleküls darf in der Zuordnungstabelle nicht mehrfach auftreten und die Matchkriterien müssen eingehalten werden. Das heißt, es ist ein Atom abzuändern, das noch nicht in der Liste vorhanden ist. Sind alle Atome des Moleküls bereits vergeben, so muß eine Nullzuordnung eingeführt werden (siehe Kapitel 2.3.2.1). Entspricht keines der noch nicht in der Zuordnungstabelle vorhandenen Atome dem Überlagerungskriterium des Atoms des ersten Moleküls, muß ebenfalls eine Nullzuordnung eingeführt werden. Nullzuordnungen können sich jedoch auch zufällig ergeben (1. Mutation in Abbildung 2-10: Das Atomtupel 1-a_ wird zu 1-_-).

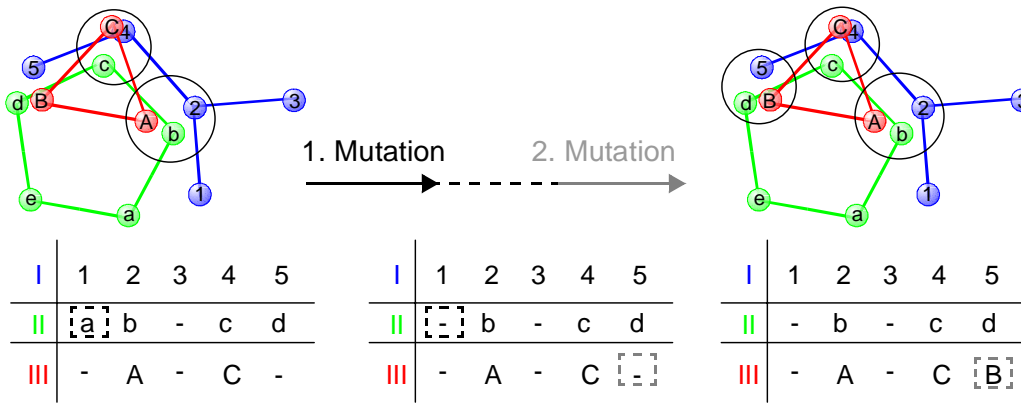


Abb. 2-10: Der Mechanismus des *Mutations*-Operators bzgl. der Atomzuordnungstabelle. Bei Überlagerung von n ($= 3$) Molekülen werden $n-1$ ($= 2$) Mutationen durchgeführt. Für jede Mutation wird zufallsmäßig ein Atomtupel, bzw. das entsprechende Atom des Atomtupels, ausgewählt (hier gekennzeichnet durch ein gestricheltes Rechteck).

Der Mutations-Operator ist ein zufälliger Prozeß und die Änderungen haben keine bestimmte Intention, wie eine Vergrößerung der Substruktur. Bei Mutationen lassen sich die Substrukturen vergrößern oder verkleinern. Da es während der genetischen Optimierung zum irreversiblen Verlust von bestimmten Atomzuordnungen kommen kann, würden bestimmte Lösungen nicht mehr erreicht. Die Mutation ist ein Mechanismus zum Schutz dies irreversiblen Verlusts einzelner Atomzuordnungen.

Die Mutation der Chromosomen, die die Änderungen der Torsionswinkel kodieren, invertiert einzelne Bits in dem zugehörigen Bitstring des Winkels. Wie bereits in Kapitel 2.3.1 und 2.3.2 erwähnt, ist jedes Chromosom der Torsionswinkel ein $8n_{tor}$ langer Bitstring (n_{tor} = die Gesamtanzahl der rotierbaren Winkel aller Moleküle). Für jeden Torsionswinkel wird eine Mutation, wie in Abbildung 2-11 dargestellt, durchgeführt, wobei der erste Mutationspunkt zufällig gesucht wird und jeder weitere 8 Bits vom ersten entfernt ist. So erreicht man eine Gleichverteilung der Änderungen auf alle Winkel.

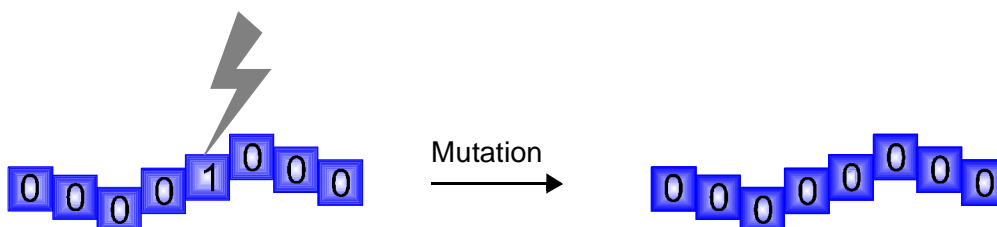


Abb. 2-11: Die Mutation der Torsionswinkel Bitstrings.

Die Integerwerte der Winkel sind, wie bereits in Kapitel 2.3.1 beschrieben, in einem 8-Bit langen String Gray kodiert. Der erste Schritt ist dabei die Umrechnung der Winkel von -180° bis $+180^\circ$ in das Intervall 0 bis 255. Die kleinste auftretende Winkeländerung ist 1.4° . Anhand der Mutation eines Winkels soll nun der Vorteil der Gray-Kodierung gegenüber der normal gebräuchlichen 2-Komplement-Kodierung der Winkel erklärt werden.

Die Mutation des Winkels 67.5° ist durch jeweilige Änderung der einzelnen Bits aus der Anfangskodierung zu erreichen (Tabelle 2-3). Die fettgedruckte Ziffer bezeichnet das jeweils abgeänderte Bit.

Tab. 2-3: Mutationen des Winkels $67,5^\circ$ in verschiedenen Kodierungen.

Mutation der Gray-Kodierung	Winkel in Grad	Mutation der 2-Komplement-Kodierung	Winkel in Grad
11101000	67.5	10110000	67.5
1110100 1	68.91	1011000 1	68.91
111010 10	71.72	101100 10	70.31
11101 100	77.34	10110 100	73.12
1110 0000	88.59	1011 1000	78.75
111 11000	66.09	101 00000	45.00
11 001000	21.09	1 0010000	22.5
10101000	111.09	11110000	157.5
01101000	-68.91	00110000	-112.5

Stellt man die Mutation des Winkels 67.5° auf einer Skala von -180° bis $+180^\circ$ graphisch dar (siehe Abbildung 2-12) lassen sich die genauen Unterschiede in den Kodierungsformen herausstellen. Ausgehend von 67.5° wird durch die Mutation der Gray-Kodierung und der 2-Komplement-Kodierung der nächst größere Winkel durch Addition von 1.4° nämlich 68.91° erreicht (Tabelle 2-1 fett markiert). Der nächst kleinere Winkel, der durch die Mutation erreicht wird, liegt bei der Gray-Kodierung nach einer Differenz von 1.4° bei 66.09° . Bei der 2-Komplement-Kodierung kommt man durch eine einfache Mutation jedoch nicht zu 66.09° sondern zu einem Wert von 45° , was einem weit größerem Schritt entspricht.

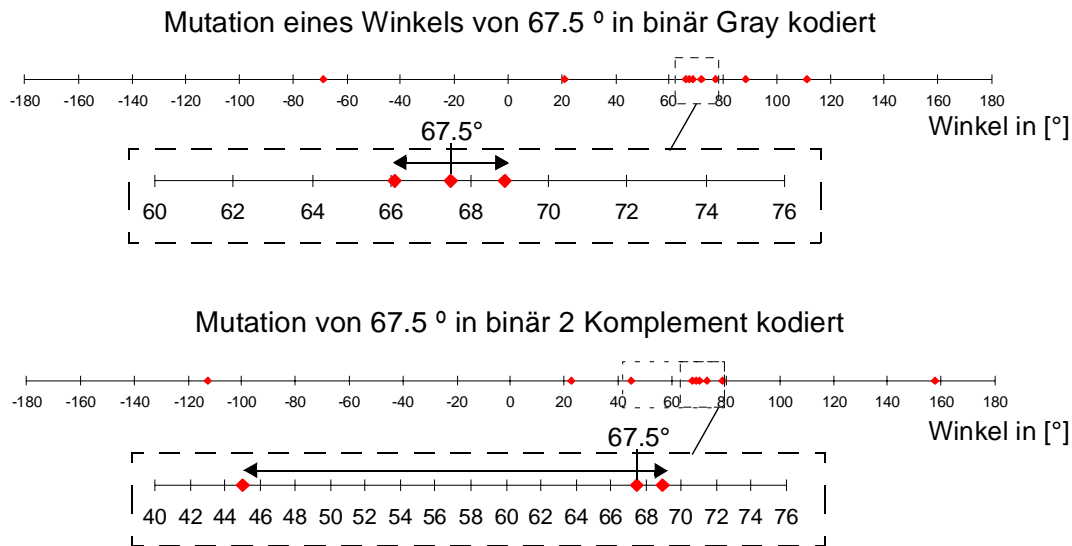


Abb. 2-12: Mutationen des Winkels 67.5° , oben: Gray kodiert, unten: 2 Komplement binär kodiert.

Bei der Gray-Kodierung von Integerzahlen kann durch eine einfache Mutation ein bestimmter Winkelwert (oder ein Minimum der Optimierung) von kleineren und größeren Werten ausgehend, in kleinen Schritten eingefasst werden. Bei der 2-Komplement-Darstellung der Winkel ist dies nicht möglich. Das zeigt sich durch die Mutation von 68.9° zu 67.5° und 66.09° zu 67.5° bei binärer Gray-Kodierung und 68.91° zu 67.5° , jedoch aber 45° zu 67.5° (als kleinster Schritt auf der Winkelskala nach unten) bei binärer 2-Komplement-Kodierung. Der nächstliegende Winkel ausgehend von 45° bei der 2-Komplement-Darstellung zu größeren Werten hin ist somit 67.5° . Bei der Gray-Kodierung erreicht man 67.5° ausgehend von 66.09° . Da es bei Optimierungen notwendig ist, ein bestimmtes Minimum (lokal oder global) von beiden Seiten (größere oder kleinere Werte) her mit kleinen Schritten einfassen zu können, ist die Gray-Kodierung für die Behandlung mit genetischen Optimierungsalgorithmen daher weit besser geeignet.

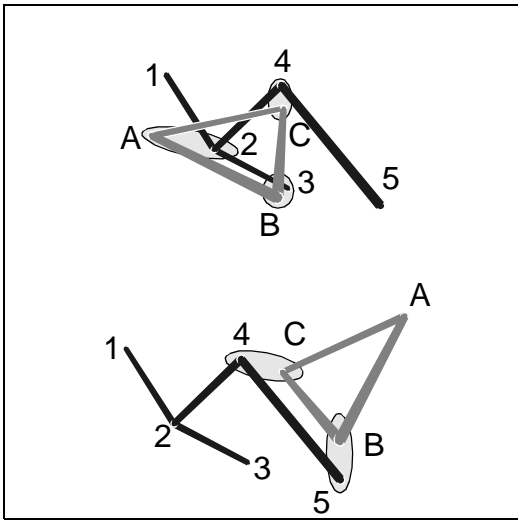
2.3.6.2 Crossover

Der Crossover-Operator tauscht zufällig gewählte Bereiche zwischen zwei Individuen aus. Auf das Problem der MCSS Suche übertragen bedeutet dies, daß Teile der Substrukturen die durch zwei Individuen definiert sind, ausgetauscht werden. Die aus der Kombination der beiden ursprünglichen Individuen (der Eltern) entstehenden Nachkommen, können dabei eine bessere Lösung zur MCSS Suche darstellen. Der Crossover-Operator ist der wichtigste Mechanismus zur Verbesserung der Individuen während der genetischen Optimierung. Im vorliegenden Fall wurde ein sogenannter Zwei-Punkt-Permutations-Crossover-Operator implementiert.^[61] Infolgedessen sind innerhalb beider Eltern-Chromosomen, zwei Schnittpunkte zufällig auszuwählen. Zwischen diesen Schnittpunkten liegt dann der zu kreuzende

Informationsstring. Dieser Teilstring muß bei beiden Elternteilen von gleicher Länge sein. In Abbildung 2-13 ist die Kreuzung zweier Chromosomen dargestellt, die die Zuordnungstabellen der Überlagerung kodieren. Dabei werden die durch Rechtecke gekennzeichneten Teile der Chromosomen (Länge = 3) gekreuzt und jede partielle Liste an das Ende des jeweilig anderen Eltern-Chromosoms gehängt. Doppelt auftretende Atomindizes (sog. Doppelreferenzen) eines Moleküls sind wie folgt aus der ursprünglichen Liste zu entfernen:

- a) Tritt ein Atom des Moleküls I in der ursprünglichen und in der an das Ende kopierten Teilliste auf, so wird das ursprüngliche Atomtupel durch das gesamte neue ersetzt (in Abbildung 2-13 sind dies z.B. die Atomtupel der Atome 3, 4 oder 5 in der ersten Zuordnungstabelle).
- b) Die nach diesem Schritt verbleibenden Doppelreferenzen in allen weiteren Molekülen werden zufällig durch andere Atome des Moleküls ersetzt, die noch nicht in der Zuordnungstabelle vorhanden sind und den Randbedingungen genügen (z.B. einem bestimmten Überlagerungskriterium). Gibt es keine übrigen Atome mehr, die diesen Ansprüchen entsprechen, so wird eine Nullzuordnung eingeführt (z.B. für die Doppelreferenz auf Atom B in der zweiten Atomzuordnungstabelle).

Diese Prozedur ermöglicht es, daß die Atomzuordnungstabellen immer eine konstante Länge aufweisen und jedes Atom nur einfach referenziert wird. Das Prinzip des Crossovers läßt sich analog auf multiple Überlagerungen übertragen, wie in Abbildung 2-14 am Beispiel der Dreifachüberlagerung von Isopentan (I), Cyclopentan (II) und Cyclopropan (III) der multiple Crossover veranschaulicht. Der zu kreuzende Teilstring ist von der Länge $d = 1$. Diese Strings werden wiederum an das jeweilige Ende des anderen Chromosoms kopiert und unter Beachtung der Schritte a) und b) entstehen die beiden Nachfolge-Chromosomen.



I	1	2	3	4	5
III	-	A	B	C	-
I	1	2	3	4	5
III	-	-	-	C	B

I	1	2	3	4	5	3	4	5
III	-	A	B	C	-	-	C	B
I	1	2	3	4	5	2	3	4
III	-	-	-	C	B	A	B	C

I	1	2	3	4	5
III	-	A	-	C	B
I	1	5	2	3	4
III	-	B	A	B	C

I	1	2	3	4	5
II	-	A	-	C	B
I	1	5	2	3	4
II	-	B	A	-	C

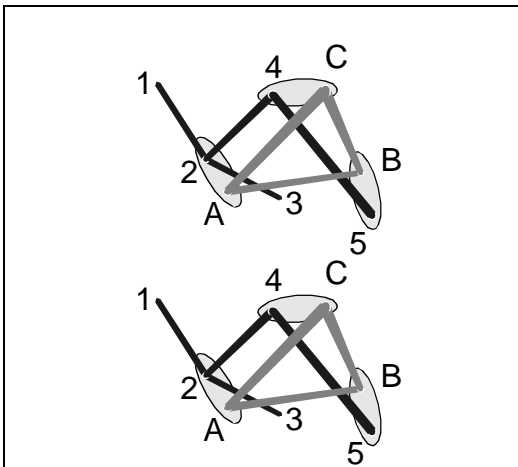


Abb. 2-13: Der Crossover-Operator am Beispiel der binären Überlagerung von Isopentan (I) und Cyclopropan (II).

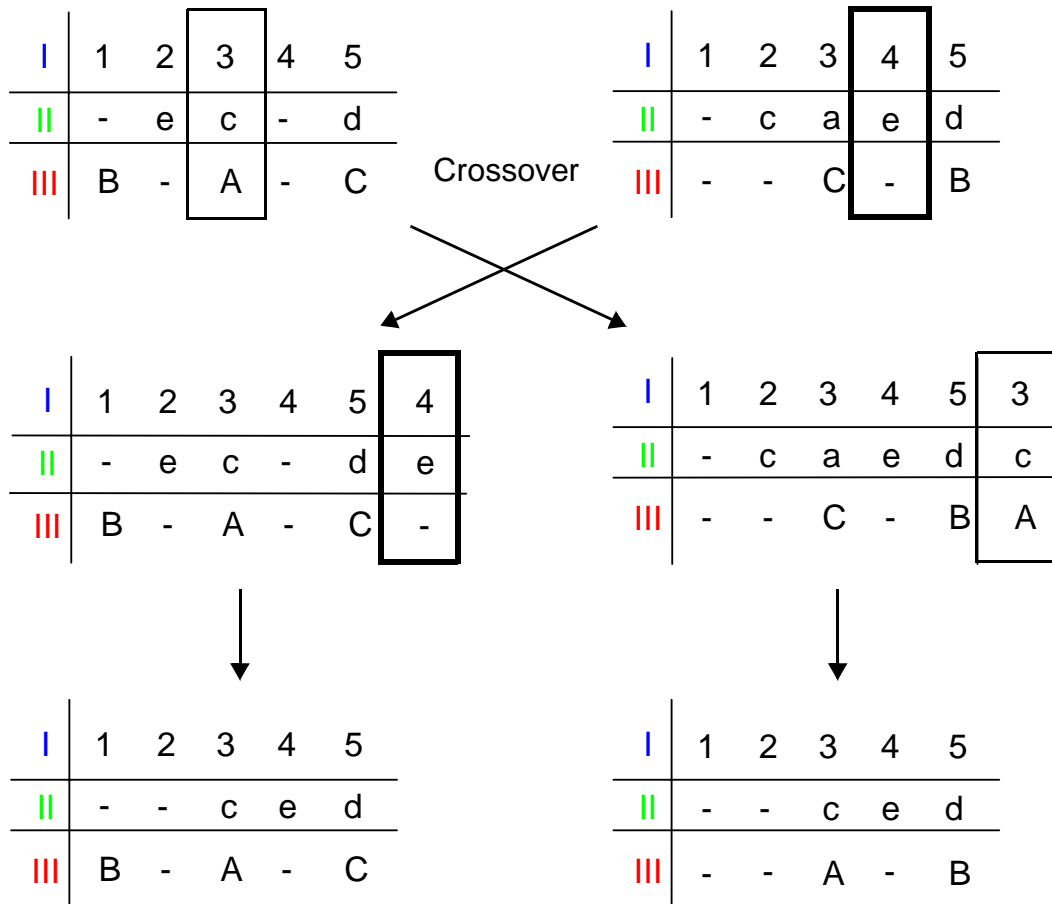


Abb. 2-14: Crossover einer Dreifachüberlagerung von Isopentan (I), Cyclopentan (II) und Cyclopropan (III).

Der Crossover-Operator wirkt ähnlich auf den zweiten Teil eines jeden Individuums, auf die Repräsentation der Torsionswinkel. Innerhalb jedes Torsionswinkel-Strings wird ein Kreuzungspunkt an der jeweils gleichen Stelle gewählt. Abbildung 2-15 stellt die Kreuzung und die Entstehung zweier neuer Torsionswinkelstrings dar. Diese beiden neuen Bitstrings führen zu einer weiteren Änderung der Konformationen der Moleküle. In einem weiteren Schritt läßt sich dann für die Konformationen durch Berechnung des *D*-Parameters erneut die geometrische Fitness bestimmen.

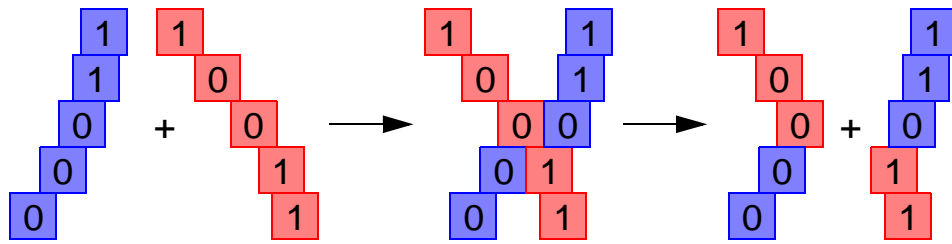


Abb. 2-15: Der Crossover-Operator bei Torsionswinkeln.

2.3.7 Zwei nicht-genetische Operatoren: Creep und Crunch

Zusätzlich zu den genetischen Operatoren Crossover und Mutation, die in einem nicht deterministischen Prozeß zur Abänderung der genetischen Information führen und auf Basis genetischer und evolutionärer Mechanismen arbeiten, wurden Operatoren entwickelt, die direkt auf das Problem der MCSS Suche zugeschnitten sind.^[62] Diese Operatoren sollten nicht stochastisch wirken, sondern vom Wissen über das spezifische Problem profitieren. Solche Operatoren werden im allgemeinen „knowledge-augmented“ (wissensangereicherte) Operatoren genannt.^[53]

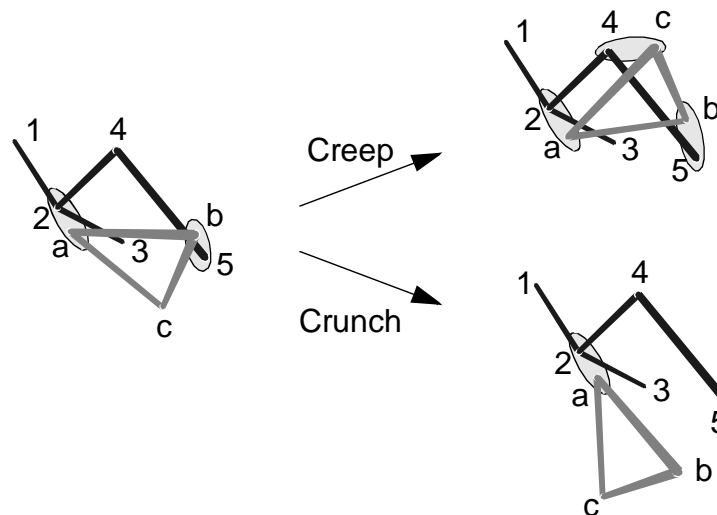


Abb. 2-16: Der Creep- und Crunch-Operator einer binären Überlagerung von Isopentan (dunkelgrau) und Cyclopropan (hellgrau).

Um die Effizienz des vorliegenden genetischen Algorithmus zu verbessern, wurden zwei wissensbasierte Operatoren, Creep und Crunch implementiert (creep (engl.): kriechen, crunch

(engl.): zerstückeln). Abbildung 2-16 und 2-17 zeigen schematisch den Creep- und Crunch-Operator für binäre und multiple Überlagerungen.

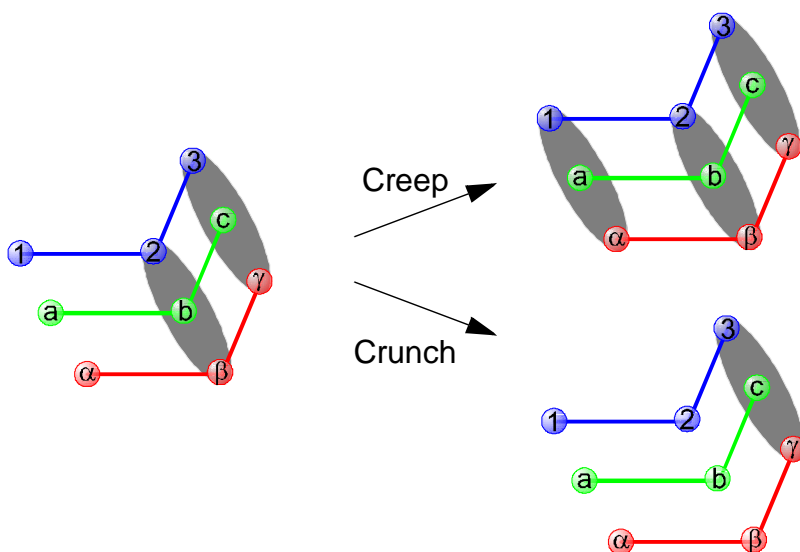
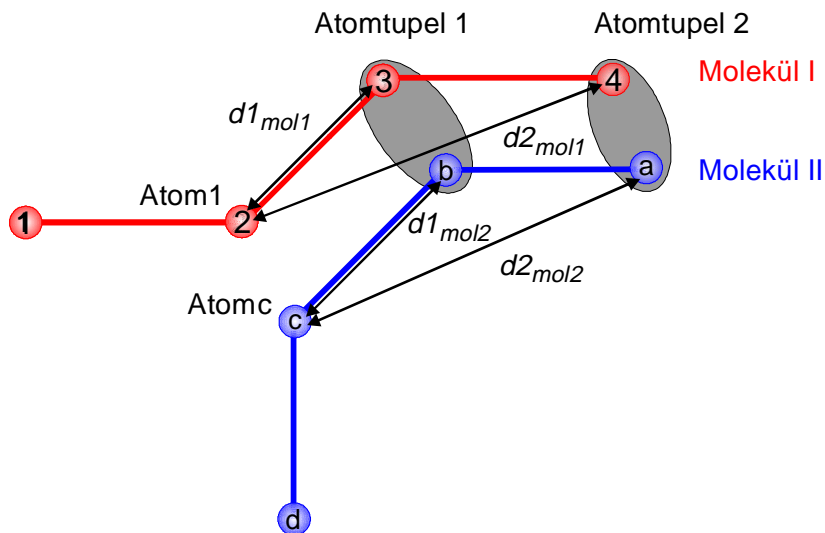


Abb. 2-17: Der Creep- und Crunch-Operator bei multipler Überlagerung.

Der Creep-Operator vergrößert dabei die bereits gefundenen Überlagerungen bzw. die Substrukturen durch Hinzufügen eines Atomtupels unter Beachtung von Restriktionen durch die räumliche Anordnung der Atome. Das neu hinzugefügte Atomtupel sollte zu einer geringen Erhöhung des D -Parameters führen (siehe Abbildung 2-16 und 2-17). Im ersten Schritt werden zwei Atomtupel zufällig ausgewählt (in Abbildung 2-18 Atomtupel 1 und Atomtupel 2, grau unterlegt). Die Distanzen des jeweiligen Atomtupels zu jedem Atom, das noch nicht an der Überlagerung beteiligt ist (in Abbildung 2-18 z.B. Atom 1 und Atom c), werden darauffolgend molekulintern berechnet. Das ergibt die Distanzen $d1_{mol1}$ und $d2_{mol1}$ für Molekül I und $d1_{mol2}$ und $d2_{mol2}$ für Molekül II. Aus diesen Distanzen sind durch Berechnung der Differenzen $d1_{mol1} - d1_{mol2}$ und $d2_{mol1} - d2_{mol2}$ Atome zu suchen, die ähnliche Distanzen zu den beiden bisherigen Atomtupel 1 und 2 aufweisen. Eines der gefundenen Atomtupel, das dieser Bedingung gerecht werden, bildet dann ein neues Atomtupel (im Beispiel der Abbildung 2-18 sind dies Atom 1 und Atom c). Auf diese Weise führt der Creep-Operator zu einer Art „hill-climbing“ Mechanismus innerhalb des Genetischen Algorithmus, da er sich wissensbasiert an ein lokales oder globales Extremum annähert.



falls

$$d1_{mol1} - d1_{mol2} \leq TOL \cdot MIN(d1_{mol1}, d1_{mol2})$$

und falls

$$d2_{mol1} - d2_{mol2} \leq TOL \cdot MIN(d2_{mol1}, d2_{mol2})$$

*dann bilden Atom1 und Atom2
ein neues Atomtupel*

mit:

$$TOL = 0.05$$

$MIN(a, b) = \text{kleinerer der beiden Werte } a \text{ und } b$

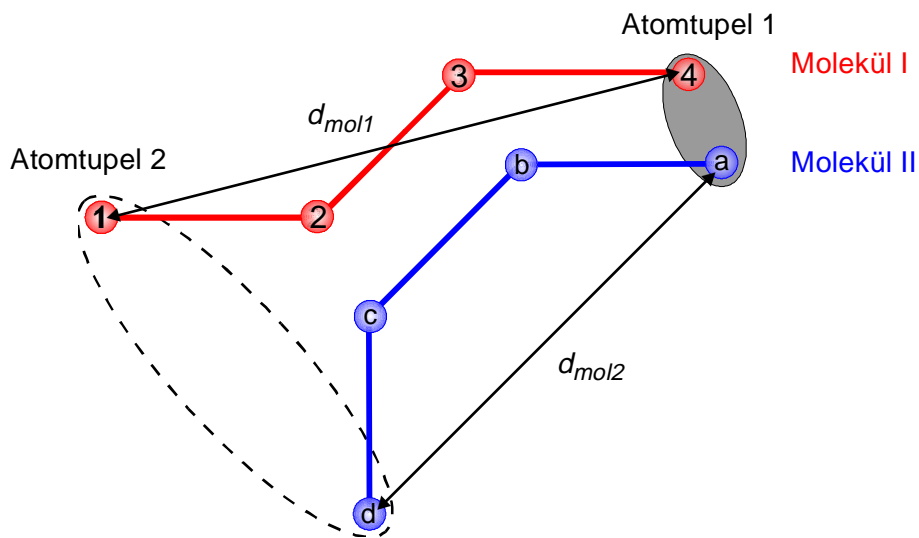
Abb. 2-18: Der Mechanismus der Suche neuer Atomtupel mit dem Creep-Operator.

Der Crunch-Operator, wie in Abbildung 2-16 für binäre und 2-17 für multiple Überlagerung dargestellt, ist der zweite nicht-genetische Operator. Er handelt als Gegenspieler (Antagonist) zum Creep-Operator, indem er die Größe der Substruktur durch Entfernung eines Atomtupels der zugrundeliegenden Überlagerung verringert.

Das Ziel der Crunch-Operators ist dabei Atomtupel zu eliminieren, die für eine schlechte geometrische Fitness der Überlagerung verantwortlich sind. Der erste Schritt ist hier die Auswahl eines Atomtupels aus der zugrundeliegenden Atomzuordnungstabelle des Individuums (in Abbildung 2-19 Atomtupel 1). Im zweiten Schritt werden die molekulinternen Distanzen zwischen diesem Atomtupel und allen anderen Atomtupel berechnet. In Abbildung 2-19 ergeben sich daraus beispielsweise die Distanzen d_{mol1} und d_{mol2} zu Atomtupel 2. Durch die Bildung der Differenz $d_{mol1} - d_{mol2}$ lassen sich Atomtupel finden, deren molekulinternen Distanzen zu anderen Atomtupel stark voneinander abweichen, d.h. die Differenz $d_{mol1} - d_{mol2}$

überschreitet einen bestimmten Toleranzwert (siehe Abbildung 2-19). Schließlich wird das Atomtupel durch eine Nullzuordnung ersetzt (hier: Matchpaar 1-d würde durch 1-₋ ersetzt), oder es werden in einem Zufallsmechanismus neue Atom der Moleküle II oder höher (bei multipler Überlagerung) gewählt, die dann mit dem entsprechenden Atom des Moleküls I (hier: Atom 1) ein neues Atomtupel bilden könnten (hier: z.B. Atom c oder b).

Der Creep-Operator vermeidet das Fallen in lokale Minima während des Optimierungsprozeß, indem er die Optimierung aus einem Minimum heraus in eine andere Richtung vorantreibt.



falls

$$d_{mol1} - d_{mol2} \geq TOL \cdot \text{MIN}(d_{mol1}, d_{mol2})$$

dann wird Atomtupel 2 gelöscht

mit:

$$TOL = 0.05$$

$\text{MIN}(a, b) = \text{kleinerer der beiden Werte } a \text{ und } b$

Abb. 2-19: Der Mechanismus der Suche von Atomtupel, die zu einem schlechten Distanzmaß führen und damit zu einer schlechten geometrischen Fitness der Überlagerung.

2.3.8 Der Directed-Tweak Algorithmus

Die „Directed Tweak“-Methode wurde erstmals von T. Hurst^[32] publiziert und hier basierend auf dem Davidon-Fletcher-Powell Optimierer implementiert.^[63] Das Motiv ist die Kombination eines nicht deterministischen genetischen Mechanismus mit einem numerischen Optimierer, um potentielle Lösungen bezüglich ihrer geometrischen Fitness weiter zu verbes-

sern. Nach dem Durchlaufen einer Generation ist die Fitness eines jeden Individuums zu bestimmen, um den Gewinner der „Restricted Tournament“-Selektion festzulegen. Vor der Bestimmung dieser Fitness wird jedoch die geometrische Anpassung jedes Individuums in Form des D -Parameters (siehe Kapitel 2.3.3.2) durch Anwendung der Directed-Tweak-Methode verbessert. Die Selektion basiert dann auf diesen verbesserten Werten. In die nächste Generation werden allerdings die originalen Individuen, vor Anwendung des Directed-Tweaks, übernommen, da sich so eine vorzeitige Konvergenz und der Verlust an genetischer Information verhindern lassen.

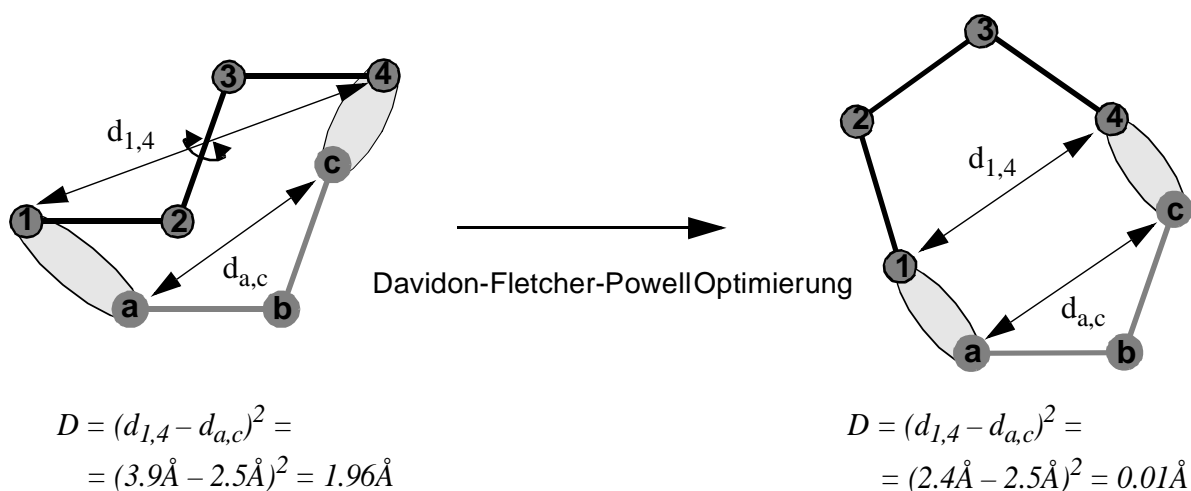


Abb. 2-20: Der Mechanismus des Directed-Tweak-Operators. Mit Hilfe eines Davidon-Fletcher-Powell Optimierers wird der geometrische Parameter D minimiert.

Die Technik basiert auf dem Davidon-Fletcher-Powell Algorithmus für multidimensionale Optimierungsprobleme, einer „numerical gradient“ Methode, die sich der ersten Ableitung der zu minimierenden Funktion bedient (die erste Ableitung ist Null bei einem Extremum der Funktion).^[63] Die Differenzen in den Konformationen werden durch Verkleinerung der quadratischen Differenzen der Distanzen einander entsprechenden Atomtupel minimiert. In Abbildung 2-20 ist dies die Minimierung des Terms $D = (d_{1,4} - d_{a,c})^2$. Das Optimieren findet im Winkelraum unter Beachten der Atomkonnektivität statt. Durch die Änderung des Diederwinkels (hier: Winkel 1-2-3-4) und erneuter Berechnung aller betroffenen Atompositionen, wie bereits in Kapitel 2.3.2.2 beschrieben, ergibt sich jeweils ein neuer D -Parameter. Somit ist D eine Funktion des Diederwinkels 1-2-3-4. Auf der Basis der, bezüglich ihrer geometrischen Anpassung

verbesserten Individuen wird die geometrische Fitness berechnet. Sie stellt die Startbedingungen für die Selektion dar. Am Ende des Durchlaufens aller Generationen präsentiert der Directed Tweak die endgültigen Überlagerungen nach der Verbesserung ihrer geometrischen Anpassung.

Als Beispiel für die Anpassung von Konformationen während eines Programmlaufs wird die Überlagerung von Methylcyclohexan und n-Butylcyclopropan (siehe Abbildung 2-21) untersucht. Der Winkel c-d-e-f von n-Butylcyclopropan wird durch Rotation der Bindung e-d an die Cyclohexankonformation angepaßt. In Abbildung 2-21 ist die Überlagerung mit der besten geometrischen Fitness, d.h. dem kleinsten *D*- bzw. *rms*-Wert dargestellt.

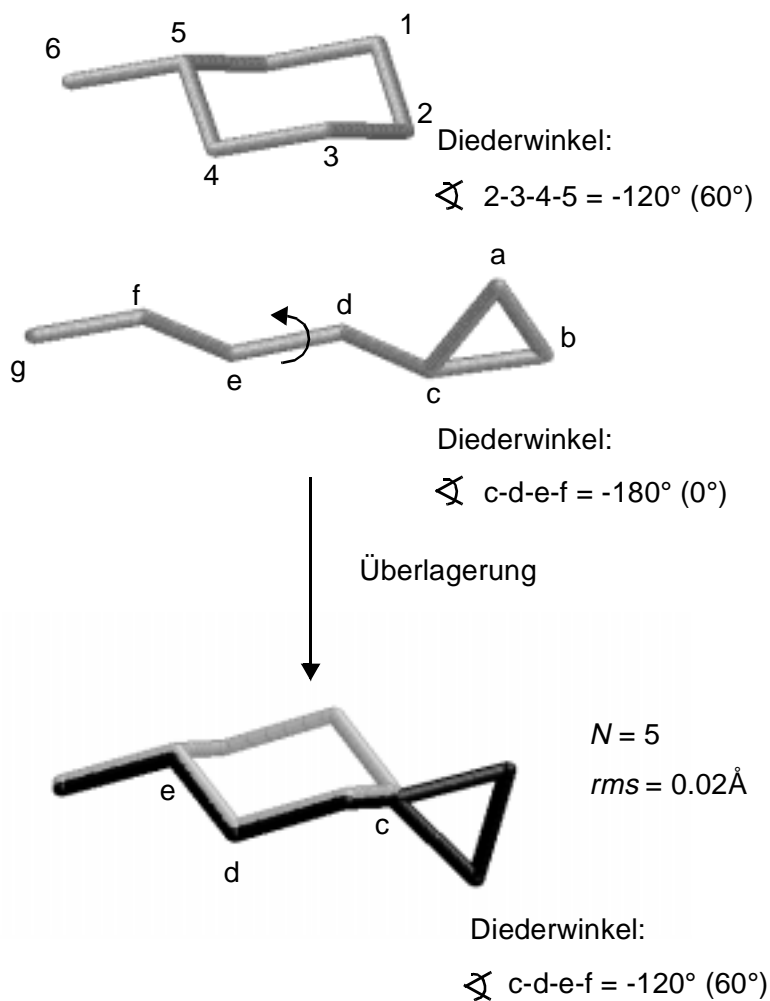


Abb. 2-21: Methylcyclohexan und n-Butylcyclopropan in den Anfangskonformationen der Überlagerung und nach der Optimierung mit der Directed-Tweak-Methode.

Idealerweise wird während der Optimierung der gesamte Konformationsraum der Bin-

dung e-d durchlaufen. Abbildung 2-22 zeigt die Verteilung des Diederwinkels c-d-e-f von n-Butylcyclopropan während eines Optimierungslaufs. Es ist deutlich zu erkennen, daß der Winkel -120° (Diederwinkel 2-3-4-5), den die Konformation des Cyclohexans als ideal vorgibt, während des Durchlaufens der Generationen am häufigsten erreicht wird. Dieser Winkel stellt somit ein Minimum für die Optimierung der geometrischen Anpassung der Konformationen dar.

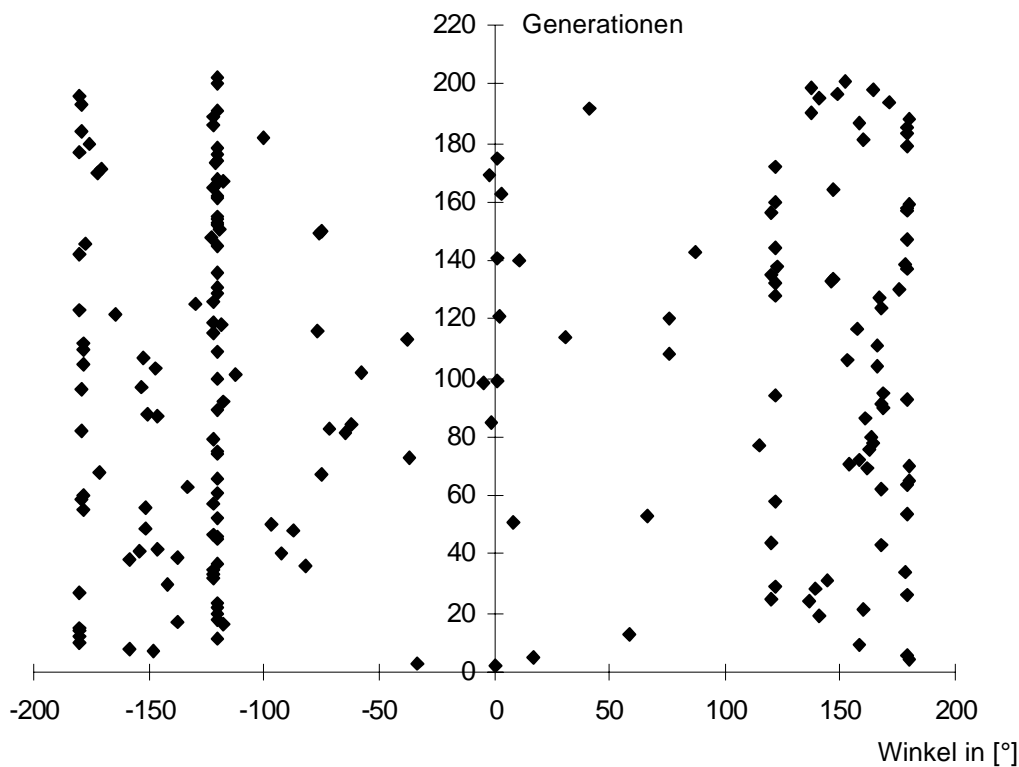


Abb. 2-22: Die Verteilung des Torsionswinkels c-d-e-f von n-Butylcyclopropan während der Optimierung der Überlagerung mit Methylcyclohexan. Es wurden 200 Generationen bei einer Populationsgröße von 50 berechnet.

2.3.9 Kontrolle der Überlappung der van der Waals-Radien

Genetische Algorithmen zur Überlagerung flexibler chemischer Verbindungen, wie z.B. von Jones et al. beschrieben,^[52] beschränken sich meist auf niedrig-energetische Konformationen. Die in Kapitel 2.3.8 beschriebene Methode der Absuche des gesamten Konformationsraums einer flexiblen Bindung zum Auffinden der besten geometrischen Anpassung zweier oder mehrerer dreidimensionaler Strukturen, endet auch in höher energetischen Konformationen. Dies verhindert den Verlust von Konformationen, die zwar im freien Zustand (Gaszustand

oder in Lösung) nicht realisiert würden, jedoch während der Bindung an einen Rezeptor, aufgrund der frei werdenden Bindungsenergie, angenommen werden könnten.

Die Überlagerung von Angiotensin-II, **4**, mit einem Angiotensin-Antagonisten Losartan, **5**, zeigt, daß bei den endgültig erhaltenen Konformationen eine Überlappung der van der Waals-Radien nicht bindender Atome auftreten kann (siehe dazu Abbildung 2-23).

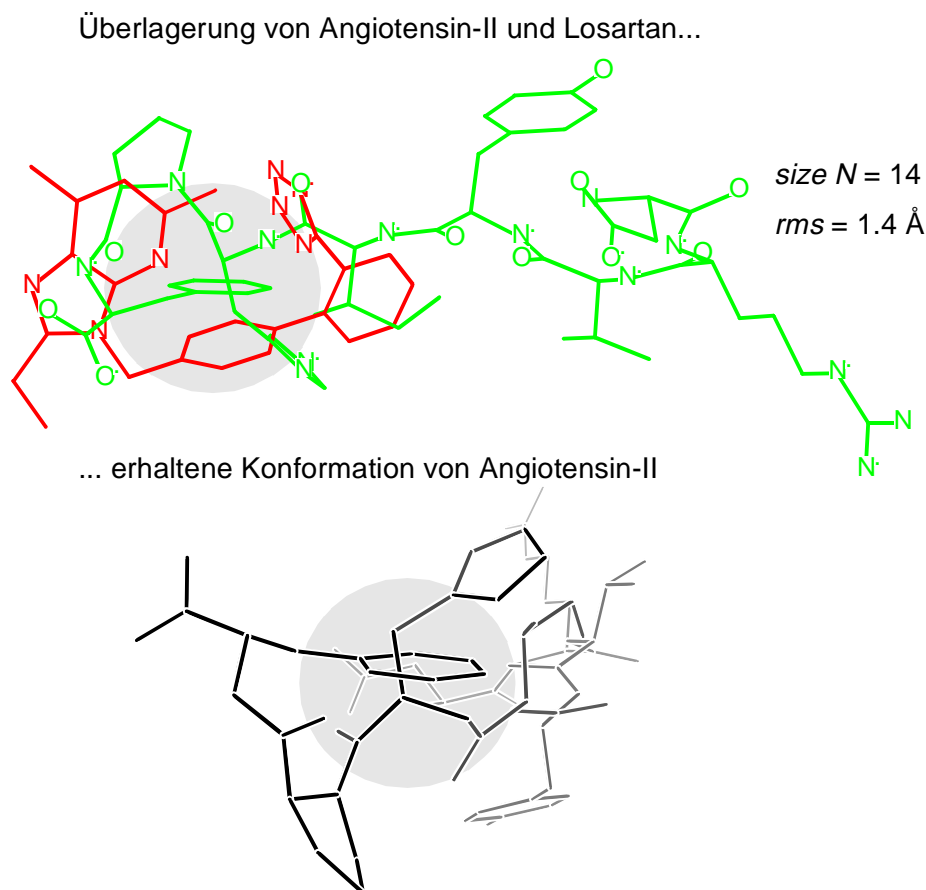
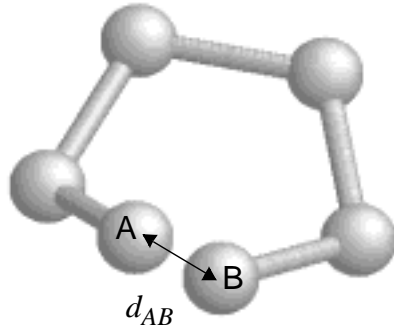


Abb. 2-23: Van der Waals-Radien-Überlappung von Angiotensin-II, **4**, bei einer nach dem Programmlauf erhaltenen Konformation.

Innerhalb des Algorithmus werden Atome als Vektoren behandelt und haben somit keine räumliche Ausdehnung. Um Konformationen, wie in Abbildung 2-23 dargestellt, zu vermeiden, ist eine Routine implementiert, die die Überlappung der van der Waals-Radien nichtgebundener Atome überprüft.

Die van der Waals-Radien der Elemente im Periodensystem wurden, wie in Abbildung 2-24 gezeigt, parametrisiert.^{[64][65]} Alle Nebengruppenelemente sowie die Hauptgruppenelemente, für die keine Angaben gefunden werden konnten, erhalten den Wert 2.0Å .

Die Berechnung der van der Waals-Radien-Überlappung wurde wie folgt implementiert: Erfolgt die Überprüfung positiv (d.h. eine Überlappung wurde gefunden) wird das Distanzmaß D mit einem entsprechenden Strafterm multipliziert. Dies führt dazu, daß das zugrundeliegende Individuum eine sehr schlechte geometrische Fitness erhält und in einem Paretowettkampf nicht dominiert.



$$\min_{AB} = 0,75 \cdot (vdW(A) + vdW(B))$$

$$fac = \frac{d_{AB}}{\min_{AB}}$$

falls:

$fac < 0.5$: Überlappung ja

$fac > 0.5$: Überlappung nein

Abb. 2-25: Die Berechnung der van der Waals-Radien-Überlappung.

2.4 Überlagerungskriterien

Für die Affinität zu einem bestimmten Rezeptor spielen zusätzlich zum Atomtyp auch dessen Bindungseigenschaften eine wichtige Rolle, z.B. das Wasserstoffbrückenbindungspotential, das elektrostatische Oberflächenpotential oder die Hydrophobizität. Bindungseigenschaften, die hauptsächlich auf Dipol-Dipol Wechselwirkungen basieren, können dabei indirekt über Variablen, wie Atompartialladung, Elektronegativität, Anzahl freier Elektronen oder Anzahl bindender Wasserstoffatome ausgedrückt werden. Als Ähnlichkeitskriterium für die Überlagerung kann in GAMMA nicht nur die Ordnungszahl gewählt werden, sondern auch folgende physikochemische Atomeigenschaften, die durch das Programmpaket PETRA^[66] berechnet werden können.

- Atompartialladung, q (q_{tot} , q_{π} oder q_{σ})
- Elektronegativität, χ
- Polarisierbarkeit, α .

Ein zu überlagerndes Atom i sollte Eigenschaftswerte haben, die innerhalb bestimmter Grenzwerte der entsprechenden Werte des Atoms j des ersten Moleküls liegen: $q_{tot}(i) = q_{tot}(j) \pm x$. Ist die Atompartialladung eines Atoms im ersten Molekül z.B. $-0.2e$, und das Intervall x ist mit 0.05 angegeben, so können nur Atome innerhalb des Bereichs $-0.25e$ bis $-0.15e$ mit diesem Atom Atomtupel bilden. Ist kein Überlagerungskriterium angegeben, so können nur Atome des gleichen Typs, d.h. mit gleicher Ordnungszahl, Atomtupel bilden.

Auch andere Kriterien, wie z.B. ob ein Atom einem aromatischen oder nicht-aromatischen Ring angehört oder die Unterscheidung zwischen Ringatomen und Nichtringatomen, können gewählt werden.

Am Beispiel der Überlagerung von Iproniazid, **6**, und Disulfiram, **7**,^[34] (siehe Abbildung 2-26) sollen die unterschiedlichen Überlagerungsergebnisse dargestellt werden, die man durch Angabe verschiedener Matchkriterien erhält. Iproniazid, **6**, ist ein Monoaminoxidase-Hemmer und wird als Antidepressivum verabreicht, während Disulfiram, **7**, bzw. Tetraethylthiuramdisulfid ein Hemmstoff der Aldehyd-Dehydrogenase ist und in der Therapie gegen chronischen Alkoholismus eingesetzt wird (führt zu Unverträglichkeitserscheinungen gegenüber alkoholischen Getränken).^[34] Diese Verbindungen wurden nicht überlagert um ein gemeinsames biologisches Wirkprofil zu ermitteln, sondern da sie sich als einfaches Beispiel, für die Demonstration der Auswirkungen der Angabe unterschiedlicher Überlagerungskriterien besonders gut eignen. Sie weisen eine geringe Atomanzahl mit unterschiedlichen Heteroatomen auf, wobei alle Atome eine relativ unterschiedliche Atompartialladung haben. In der folgenden Abbildung 2-26 werden beide Strukturformeln und

die von PETRA berechneten Atompartialladungen q_{tot} angezeigt. [67]

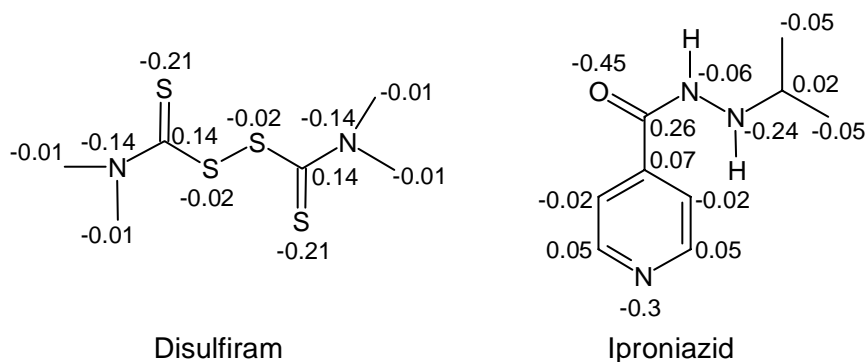


Abb. 2-26: Die Strukturen von Disulfiram, **7**, und Iproniazid, **6**. Die angegebenen Werte stellen die Partialladungen q_{tot} der Atome dar, die mit PETRA berechnet wurden. [67]

Zunächst wird eine Überlagerung durchgeführt, für die als Überlagerungskriterium nur die Identität der Atome, d.h. ihre Ordnungszahl betrachtet wird (siehe Abbildung 2-27 oben). Im zweiten Fall ist das Überlagerungskriterium die Atompartialladung q_{tot} mit einem erlaubten Bereich von $\pm 0.25e$. Dies bedeutet, daß im zweiten Molekül ein Atom nur dann mit einem Atom des ersten Moleküls überlagert werden kann, wenn dessen Atompartialladung $q_{tot}(i)$ im angegebenen Bereich von $q_{tot}(j) \pm 0.25$ bezogen auf das Atom im ersten Molekül liegt (siehe Abbildung 2-27 unten).

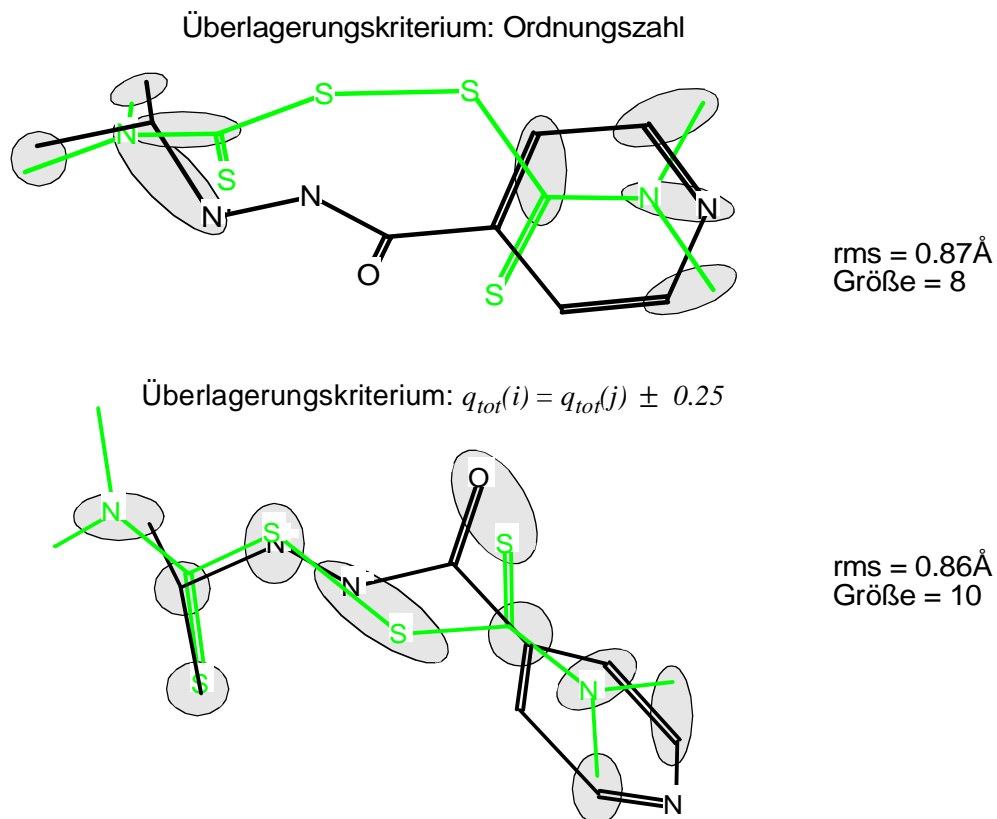


Abb. 2-27: Die unterschiedlichen Ergebnisse der Überlagerungen von Iproniazid, **6**, (schwarz) und Disulfiram, **7**, (grau) mit verschiedenen Matchkriterien. Die grauen Ovale markieren die MCSS.

In der durch graue Ovale markierten MCSS der beiden Strukturen sind eindeutige Unterschiede, hervorgerufen durch die verschiedenen Matchkriterien, zu erkennen. Der aliphatische C-S-S-C Kern von Disulfiram, **7**, ist bei der Ordnungszahlüberlagerung nicht an der Substruktur beteiligt, da kein entsprechender Partner in Iproniazid, **6**, gefunden wird (kein S-Atom enthalten). Die beiden S-Atome ($q_{tot} = -0.02e$) weisen jedoch ähnliche physikochemische Eigenschaften wie die entsprechenden N-Atome ($q_{tot} = -0.06e$ und $-0.24e$, siehe Abbildung 2-26) auf und sind in der zweiten Überlagerung als Paare an der Substruktur enthalten. Analoges gilt für die Carbonyl- (C: $0.26e$, O: $-0.45e$) und Thiocarbonylgruppe (C: $0.14e$, S: $-0.21e$), die nur in der unteren Überlagerung ein Matchpaar bilden.

2.5 Pareto-Diagramme

Als Beispiel für die Interpretation der Pareto-Diagramme soll die Überlagerung zweier Antagonisten des Angiotensin-II, **4**, am AT₁-Rezeptor, Losartan, **5**, und L-158,809, **8**,^[69] dienen (Abbildung 2-28).

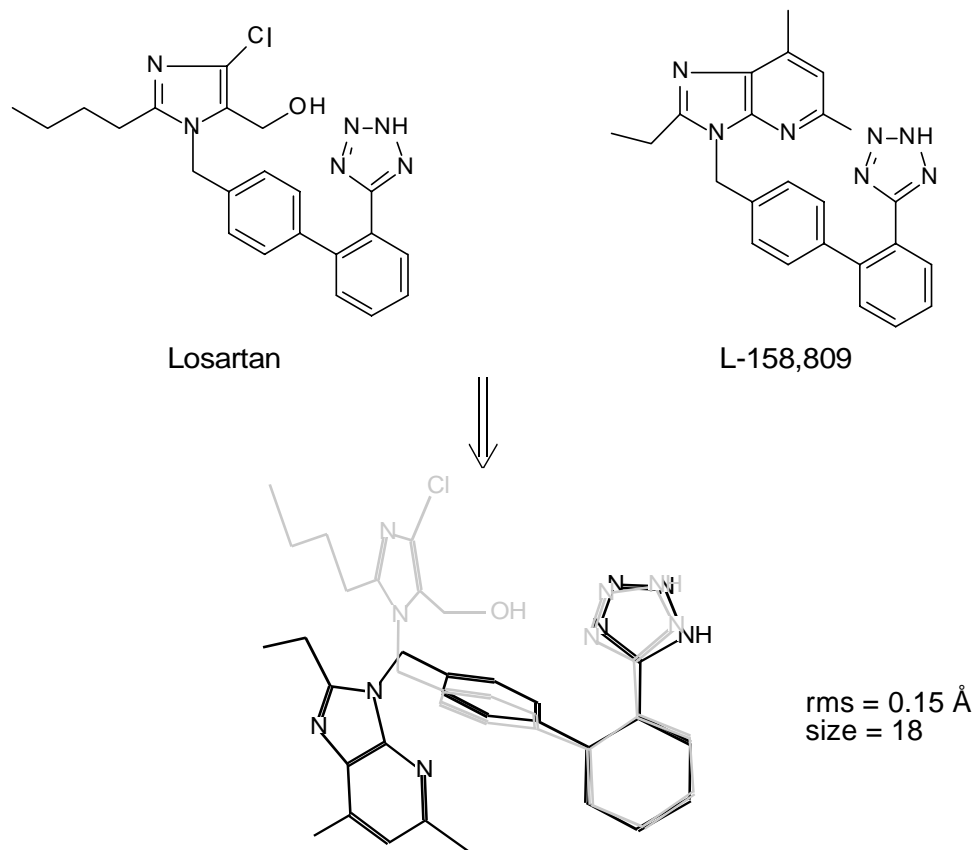
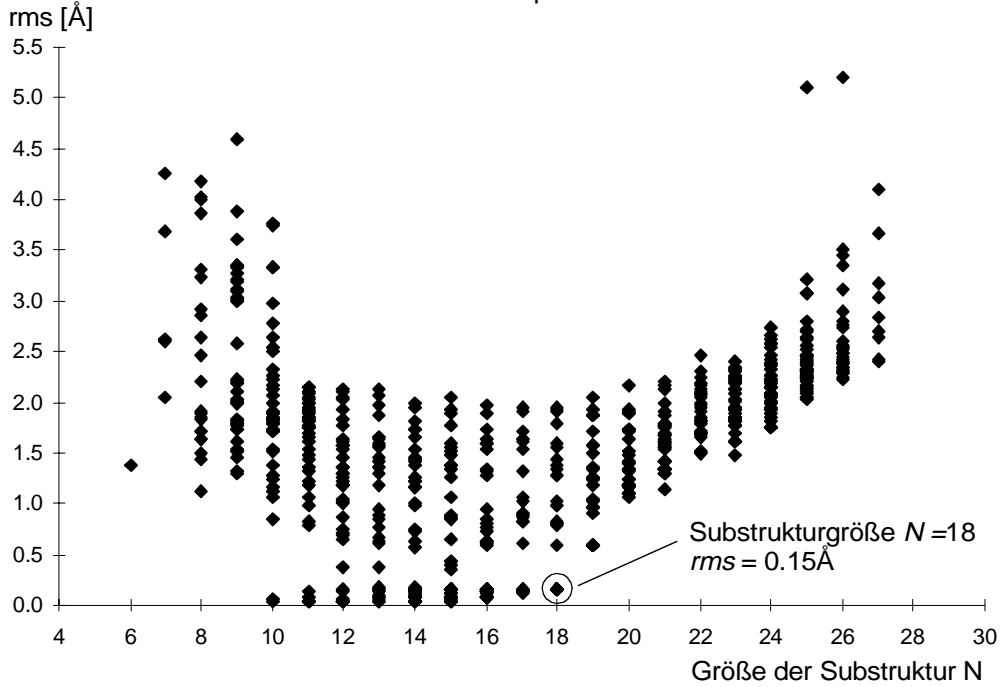


Abb. 2-28: Die Strukturformeln von Losartan, **5**, und L-158,809, **8**, und die häufigst aufgetretene Überlagerung der dreidimensionalen Strukturen von Losartan, **5**, (grau) und L-158,809, **8**, (schwarz) innerhalb 40 GA Läufen.

Die Pareto-Kurve einer Überlagerung stellt den Zusammenhang zwischen der Substrukturgrößen und dem zugehörigen *rms*-Wert der Überlagerungen innerhalb eines Experiments dar. In Abbildung 2-29 ist das Pareto-Diagramm von 40 Optimierungsläufen der Überlagerung zu sehen. Bei einer Größe von 10 bis 18 Atomen wurden in allen Experimenten die besten geometrischen Anpassungen der beiden dreidimensionalen Strukturen gefunden. Ab einer Substrukturgröße von 18 Atomen nimmt der *rms*-Wert rapide zu. In keinem Experiment wird dann noch ein *rms*-Wert unterhalb von 0.5 \AA gefunden. Im unteren Teil der Abbildung 2-29 ist ein einzelnes Experiment einer Pareto-Kurve herausgenommen, um des Verlauf der Kurven zu verdeutlichen.

Pareto-Diagramm der Überlagerung von Losartan und L-158,809
innerhalb von 40 Experimenten



Pareto-Diagramm
eines Experiments

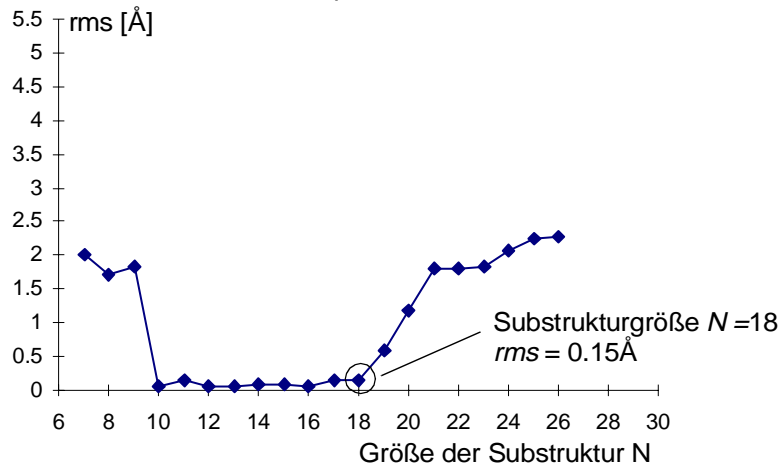


Abb. 2-29: Das Pareto-Diagramm der Überlagerung von Losartan, **5**, und L-158,809, **8**, oben: die Ergebnisse aller 40 Experimente, unten: das beste Experiment. Die beste Überlagerung mit einem rms -Wert von 0.15Å und einer Größe N von 18 Atomen ist in beiden Diagrammen hervorgehoben.

Die Häufigkeit, mit der eine bestimmte Überlagerung auftritt, ist durch die Relation der durchgeführten Experimente bzw. Optimierungsläufe zum rms -Wert einer Überlagerung mit konstanter Größe (z.B. $N=18$) zu erkennen (Abbildung 2-30). Es wird deutlich, daß die Über-

lagerung mit einer Güte von $rms = 0.15\text{\AA}$ am häufigsten (70%) während der 40 Optimierungsläufe erreicht wird.

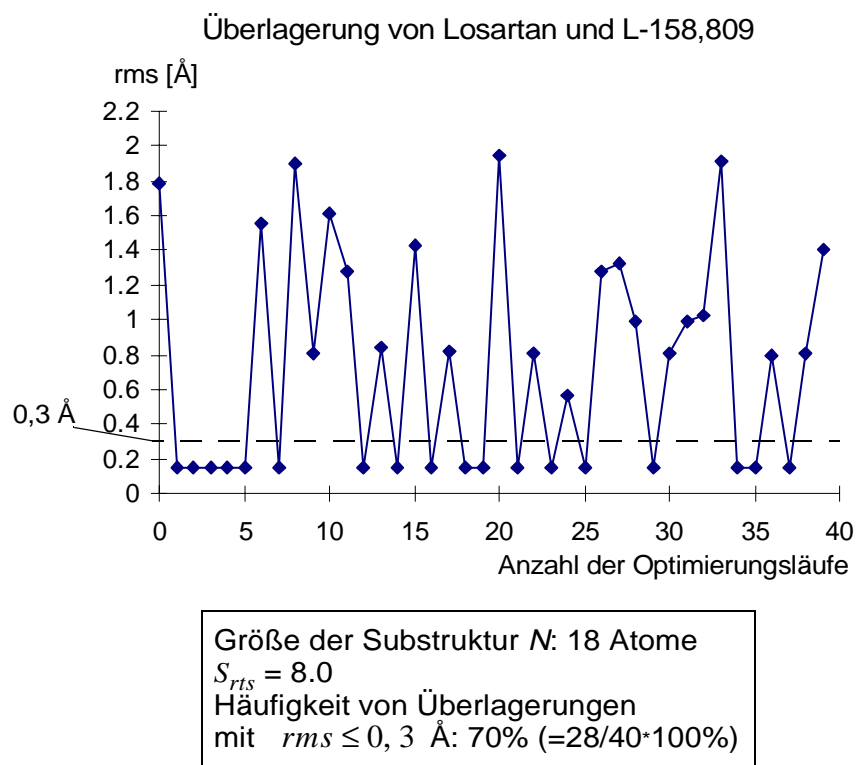


Abb. 2-30: Erhaltene Überlagerungen, bzw. geometrische Fitness in rms -Werten, innerhalb von 40 Optimierungsläufen für eine konstant gehaltene Größe der Substruktur (18 Atome) von Losartan, **5**, und L-158,809, **8**.

2.6 Parameter der genetischen Selektion

Um eine optimale Größe S_{rms} der Teilpopulation für die Ähnlichkeitsvergleiche der Selektion (siehe Kapitel 2.3.5 und Abbildung 2-9) zu ermitteln, wurden die Berechnungen der dreidimensionalen MCSS für folgende Strukturen untersucht: β -Östradiol, **9**, und das Östrogen Mimetikum p,p'-DDT, **10** (siehe: Abbildung 2-31), die beiden Neurotoxine Saxitoxin **11** und Tetrodotoxin^{*)}, **12**, (siehe Abbildung 2-32) und die beiden Angiotensin-II-Antagonisten des AT₁-Rezeptors, Losartan, **5**, und L-158,809, **8** (siehe Abbildung 2-28).

^{*)}Saxitoxin ist ein paralytisches Gift, das hauptsächlich von der kalifornischen Seemuschel *mytilus californianus* produziert wird. Tetrodotoxin ist eine toxische Substanz aus der Leber der Gattung der *Tetraodontidae*, im speziellen des Kugelfischs.^[68]

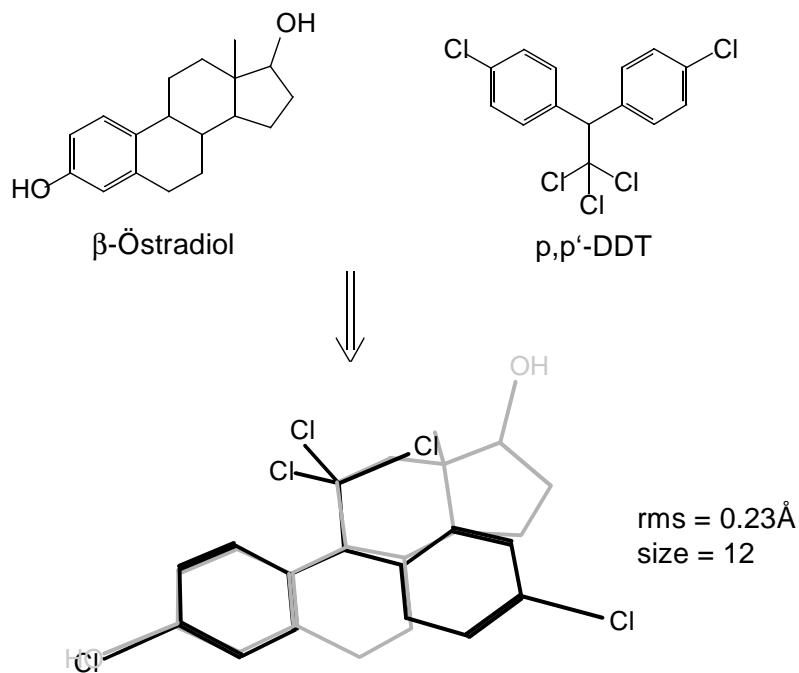


Abb. 2-31: Die Strukturformeln von β -Östradiol, **9**, und p,p'-DDT, **10**, und die am häufigsten aufgetretene Überlagerung der dreidimensionalen Strukturen von β -Östradiol (grau) und p,p'-DDT (schwarz) innerhalb von 40 GA-Läufen.

Um die optimale Größe von S_{rts} zu ermitteln (siehe Abbildung 2-9), wird die Größe dieser Subpopulation von 2 bis 50 Individuen in Schritten von 1 variiert. Die Größe i der Gesamtpopulation ist 100 Individuen. Für jede Größe von S_{rts} soll die Anzahl der gefundenen akzeptablen Lösungen miteinander verglichen werden.

Um eine Überlagerung als akzeptable Lösung des MCSS Suchproblems zu bezeichnen, wurden Bedingungen basierend auf den ermittelten Pareto-Kurven der jeweiligen Überlagerungen (siehe Abbildung 2-33) gesetzt. Es wird deutlich, daß ab einem bestimmten Schwellenwert für die Substrukturgröße (in Abbildung 2-33 gekennzeichnet durch ein gestricheltes Rechteck) der rms -Wert steil ansteigt. Die Güte jeweils dieser Substrukturgröße wurde für alle durchgeführten Optimierungsläufe untersucht, indem die Häufigkeit im Auftreten von Überlagerungen dieser Grenzgröße mit einem rms -Wert kleiner als 0.3\AA herangezogen wird.

Als „besten Läufe“ der Überlagerung werden diejenigen bezeichnet, bei denen der rms -Wert für die ermittelte Schwellenwert der Substrukturgröße am geringsten ist.

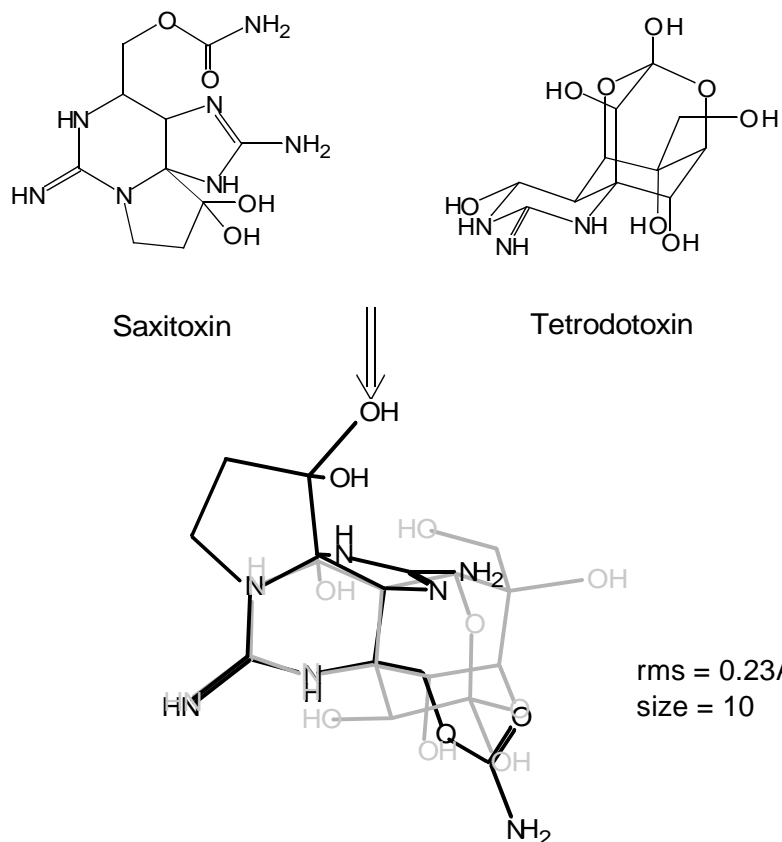


Abb. 2-32: Die Strukturformeln von Saxitoxin, **11**, und Tetrodotoxin, **12**, und die am häufigst aufgetretene Überlagerung der dreidimensionalen Strukturen von Saxitoxin, **11**, (schwarz) und Tetrodotoxin, **12**, (grau) innerhalb von 40 GA-Läufen.

Für die Überlagerung von β -Östradiol, **9**, und p,p'-DDT, **10**, ergibt sich dabei eine Substrukturgröße von 12 Atomen. Im Fall von Saxitoxin, **11**, und Tetrodotoxin, **12**, müssen die Substrukturen eine Größe von 10 Atomen aufweisen, für Losartan, **5**, und L-158,809, **8**, wird eine Substrukturgröße von 18 Atomen betrachtet. Die Häufigkeit im Auftreten der Substrukturen, die jeweils diesen Bedingungen genügen und einen rms-Wert kleiner als 0.3\AA aufweisen, wurden innerhalb von 40 GA-Läufen ermittelt (siehe Abbildung 2-34).

Für jede Größe S_{rms} der Selektion werden somit 40 GA-Optimierungsläufe für jedes der drei MCSS Suchprobleme berechnet.

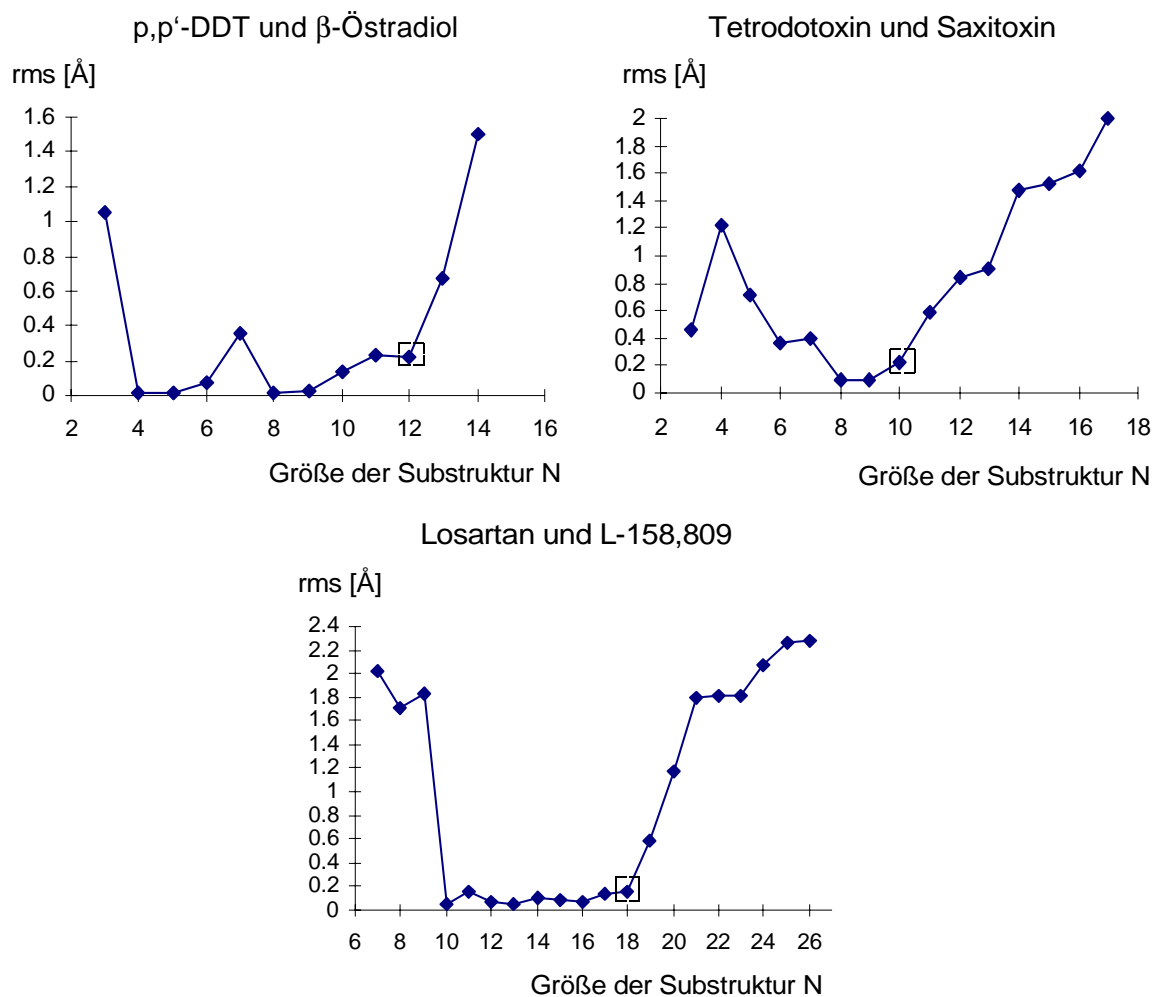


Abb. 2-33: Die Pareto-Kurven der besten Läufe der Überlagerungen von β -Östradiol, **9**, mit p,p'-DDT, **10**, Tetrodotoxin, **12**, mit Saxitoxin, **11**, und Losartan, **5**, mit L-158,809, **8**. Sie zeigen die Relation zwischen der Größe einer Substruktur und der geometrischen Fitness (*rms*-Wert) der Überlagerung. Die durch Rechtecke gekennzeichneten Werte markieren die herausgenommene Größe jeder Überlagerung.

Abbildung 2-34 zeigt die Relation zwischen der Häufigkeit der akzeptablen Lösungen und des aktuellen S_{rts} Werts innerhalb von 40 GA-Optimierungsläufen. Basierend auf diesen Ergebnissen bilden die Werte 7 und 8 für den Parameter S_{rts} bei einer Gesamtpopulation von 100 Individuen ein Optimum. Bei höheren S_{rts} -Werte erhält man zwar ebenfalls gute Lösungen, allerdings führt eine Verdopplung von S_{rts} (von 10 zu 20 Individuen) zu einer 20%igen Erhöhung der gesamten Rechenzeiten. Folgend daraus werden alle weiteren Überlagerungen mit einer Größe von maximal $S_{rts} = 1/10 \cdot i$ durchgeführt, wobei i die Gesamtpopulationsgröße darstellt.

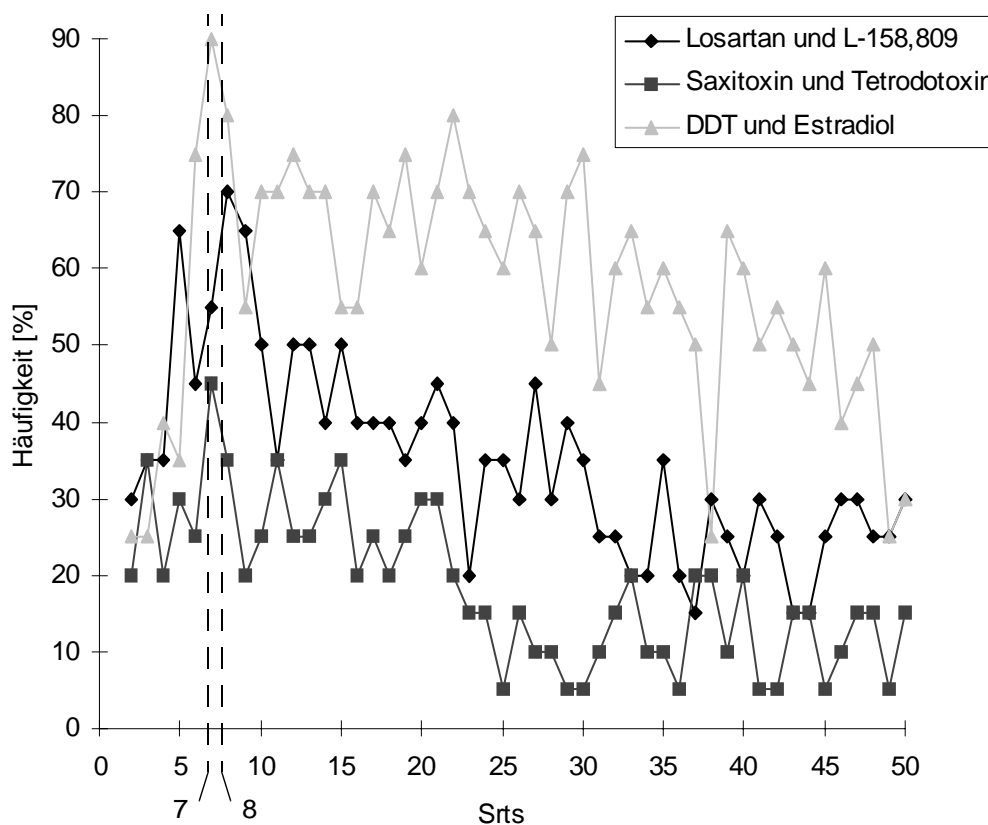


Abb. 2-34: Die Häufigkeit der aufgetretenen besten Lösungen einer jeden Überlagerung für einen bestimmten Wert der Subpopulationsgröße der Selektion S_{rts} . Untersucht wurden die Überlagerungen von β -Östradiol, 9, mit p,p'-DDT, 10, Tetrodotoxin, 12, mit Saxitoxin, 11, und Losartan, 5, mit L-158,809, 8.

2.7 Vergleich zweier Selektionstypen

Während des Durchlaufens der Generationen sollte trotz Selektion der Erhalt der genetischen Informationsvielfalt gewährleistet werden. Anhand zweier Optimierungsläufe wird gezeigt, wie sich die genetische Vielfalt im Laufe des genetischen Algorithmus ändert. Die beiden Läufe unterscheiden sich nur in der Art der genetischen Selektion, der erste wird mit der „Selektion des eingeschränkten Wettkampfs“ (engl.: *restricted tournament selection*) der zweite mit „Rouletterad-Selektion“ (engl.: *roulette wheel selection*) durchgeführt. Es wurde wiederum die Überlagerung von β -Östradiol mit dem Östrogenmimetikum p,p'-DDT (siehe Abbildung 2-31) untersucht.

Als Maß für die Überwachung der Konvergenz des Algorithmus wurde der Bias b (Gl. 2-5) berechnet. Wenn $A_{1,i}$ ein Atom von Molekül 1 (M_1) ist, und $A_{2,j}$ ein Atom von Molekül 2 (M_2) ist, so versteht man unter dem Bias b die mittlere Häufigkeit, mit der Atom $A_{2,j}$, das am

häufigsten zusammen mit $A_{1,i}$ ein Atomtupel bildet, auf das Atom $A_{1,i}$ abgebildet wird. Wenn N_1 die Anzahl der Atome in M_1 ist, $maxfreq(A_{1,i}, M_2)$ die Anzahl des am häufigsten auf das Atom $A_{1,i}$ abgebildeten Atoms aus M_2 ist und I die Größe der Population ist, so ergibt sich der Bias b als Durchschnittswert über alle Atome $A_{1,i}$ und Individuen wie folgt:

$$b = \frac{1}{N_1} \sum_i^{N_1} \frac{maxfreq(A_{1,i}, M_2)}{I} \quad (\text{Gl. 2-5})$$

mit:

N_1	=	Anzahl der Atome in Molekül 1 (M_1)
$A_{1,i}, A_{2,j}$	=	Ein Atom von M_1 , bzw. M_2
$maxfreq(A_{1,i}, M_2)$	=	Die Anzahl der Abbildungen von Atom $A_{2,j}$ in M_2 auf das Atom $A_{1,i}$ in M_1 , auf das $A_{2,j}$ am häufigsten abgebildet wird
I	=	Größe der Population

Somit meint ein Bias von 0.75, daß ein bestimmtes Atomtupel mit einer Wahrscheinlichkeit von 75% in jedem Individuum auftritt. Den höchsten Wert, den der Bias einnehmen kann ist 1.0. Ein Bias von 1.0 bedeutet, daß in jedem Individuum der Population ein bestimmtes Atom von Molekül 1 immer ein und dem selben Atom von Molekül 2 zugeordnet ist. Daher würde das gleiche Atomtupel in jedem Individuum auftreten. Je höher der Wert des Bias ist, desto geringer ist die genetische Vielfalt in der Population.

Abbildung 2-35 zeigt die Entwicklung des durchschnittlichen Bias während des Durchlaufens von 400 Generationen bei der Optimierung der Überlagerung von β -Östradiol, **9**, mit p,p'-DDT, **10**. Die graue Kurve zeigt den erhaltenen Verlauf des durchschnittlichen Bias für die *Rouletterad-Selektion* und die schwarze für die *Restricted Tournament-Selektion* (RTS).

Während des Optimierungsprozesses ist der Bias der Population, der durch die RTS erhalten wird, grundsätzlich niedriger als der Bias der Population mit *Rouletterad-Selektion*. Durch die RTS wird ein höhere Informationsvielfalt während des Durchlaufens der Generationen garantiert, und die Wahrscheinlichkeit, eine optimale Lösung zu erreichen, erhöht.

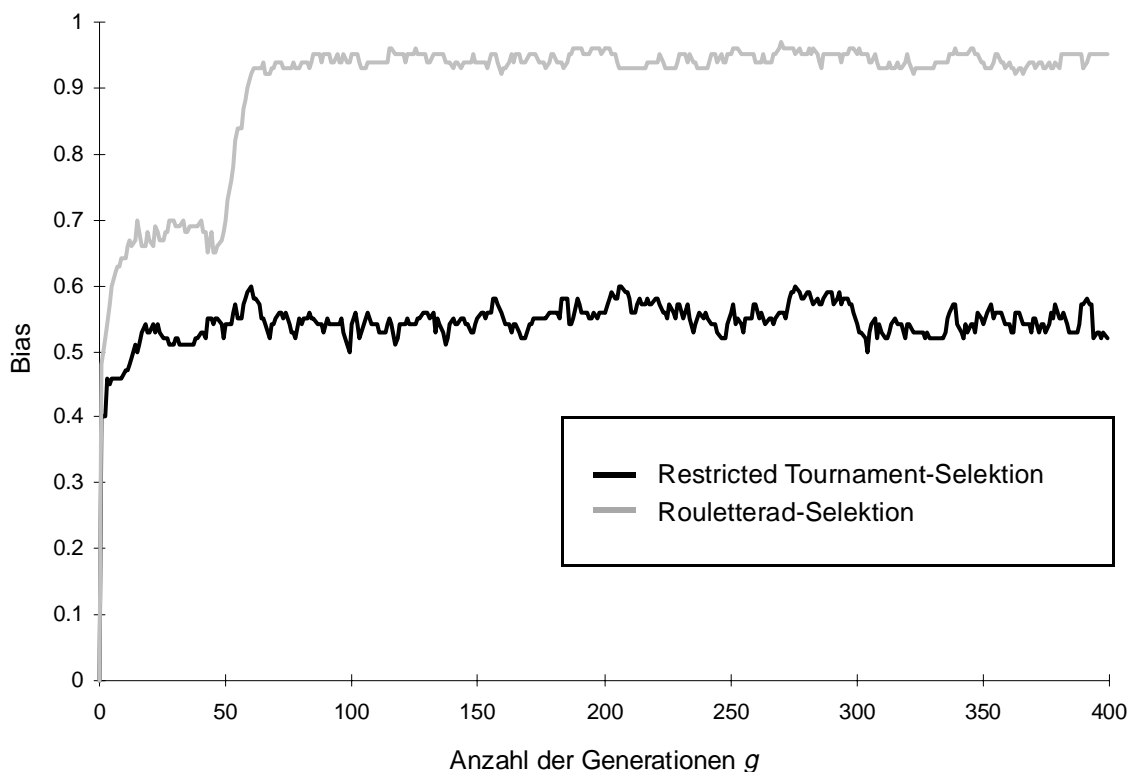


Abb. 2-35: Der durchschnittliche Bias der Überlagerung von β -Östradiol, **9**, mit p,p'-DDT, **10**, abhängig von der Anzahl der Generationen. Die graue Kurve zeigt die Optimierung mit *Rouletterad-Selektion*, die schwarze mit *Restricted Tournament-Selektion*. Die *Restricted Tournament-Selektion* bewahrt die genetische Vielfalt während der Optimierung.

2.8 Optimierung der Einsatzwahrscheinlichkeiten der Operatoren

Es hat sich gezeigt, daß die Qualität der erhaltenen Überlagerungen, binär sowie multipel, und die Geschwindigkeit der Optimierungen hauptsächlich von der Höhe der Operatoreinsatzwahrscheinlichkeiten abhängt. Ermittelt wurden die idealen Operatoreinsatzwahrscheinlichkeiten grundsätzlich durch Bestimmung der Häufigkeit, wie oft sich optimale Lösungen des MCSS-Suchproblems finden lassen (analog zu Kapitel 2.6). Für die Überlagerung von β -Östradiol, **9**, und p,p'-DDT, **10**, ist für alle Permutationen der Operatoreinsatzwahrscheinlichkeiten die Zahl ausschlaggebend, mit der innerhalb von 40 GA-Läufen bei einer Substrukturgröße von 12 Atomen ein *rms*-Wert kleiner als 0.3\AA erreicht wird (vgl. Kapitel 2.6). Durch einen Vergleich der Häufigkeiten läßt sich die beste Kombination der Operatoreinsatzwahrscheinlichkeiten ermitteln. Für das angeführte Beispiel von β -Östradiol, **9**, und p,p'-DDT, **10**, wurden folgende Werte erhalten: $p_{\text{mut}}/p_{\text{cross}}$: 0.4/0.6, $p_{\text{tormut}}/p_{\text{torcross}}$: 0.7/0.4 und $p_{\text{creep}}/p_{\text{crunch}}$: 0.5/0.3.

Im allgemeinen lassen sich für die Einsatzwahrscheinlichkeiten der Operatoren die in Tabelle 2-4 aufgelisteten Erfahrungswerte erhalten. Die genauen Einsatzwahrscheinlichkeiten hängen v.a. von der Anzahl der zu überlagernden Strukturen und der Zahl der rotierbaren Bindungen ab.

Tab. 2-4: Die Einsatzwahrscheinlichkeiten der Operatoren.

Operator	Zahlenwert
$P_{mutation}$	$\geq 0,4$
$P_{crossover}$	$\geq 0,6$
$P_{torcross}^*$	$0,5 \pm 0,1$
P_{tormut}^\dagger	$0,6 \pm 0,1$
P_{creep}	$\geq 0,5$
P_{crunch}	$< 0,5$

*. Crossover-Operator der Torsionswinkel

†. Mutations-Operator der Torsionswinkel

2.9 Die Bedeutung des Creep- und Crunch-Operators

Um die Bedeutung des Creep- und Crunch-Operators zu verdeutlichen, wurde wiederum die Überlagerung von β -Östradiol, **9**, mit p,p'-DDT, **10**, untersucht. 40 GA-Läufe wurden mit einer p_{creep}/p_{crunch} -Kombination von 0.0/0.0 durchgeführt, weitere 40 mit einem Verhältnis von p_{creep}/p_{crunch} von 0.5/0.3. Die Einsatzwahrscheinlichkeiten der anderen Operatoren blieben dabei konstant auf den in Kapitel 2.8 ermittelten Werten von p_{mut}/p_{cross} : 0.4/0.6, $p_{tormut} / p_{torcross}$: 0.7/0.4 (vgl. Tabelle 2-4). Der Vergleich der beiden Optimierungsläufe ohne Creep-, und Crunch-Operator und mit Creep- und Crunch-Operator ist in Abbildung 2-36 dargestellt. Im oberen Diagramm ist das Verhältnis des *rms*-Werts zur Größe der Substruktur aller 40 GA Läufe zu sehen, wogegen im unteren Diagramm der durchschnittliche *rms*-Wert der 40 Läufe für jede Substrukturgröße ermittelt wurde.

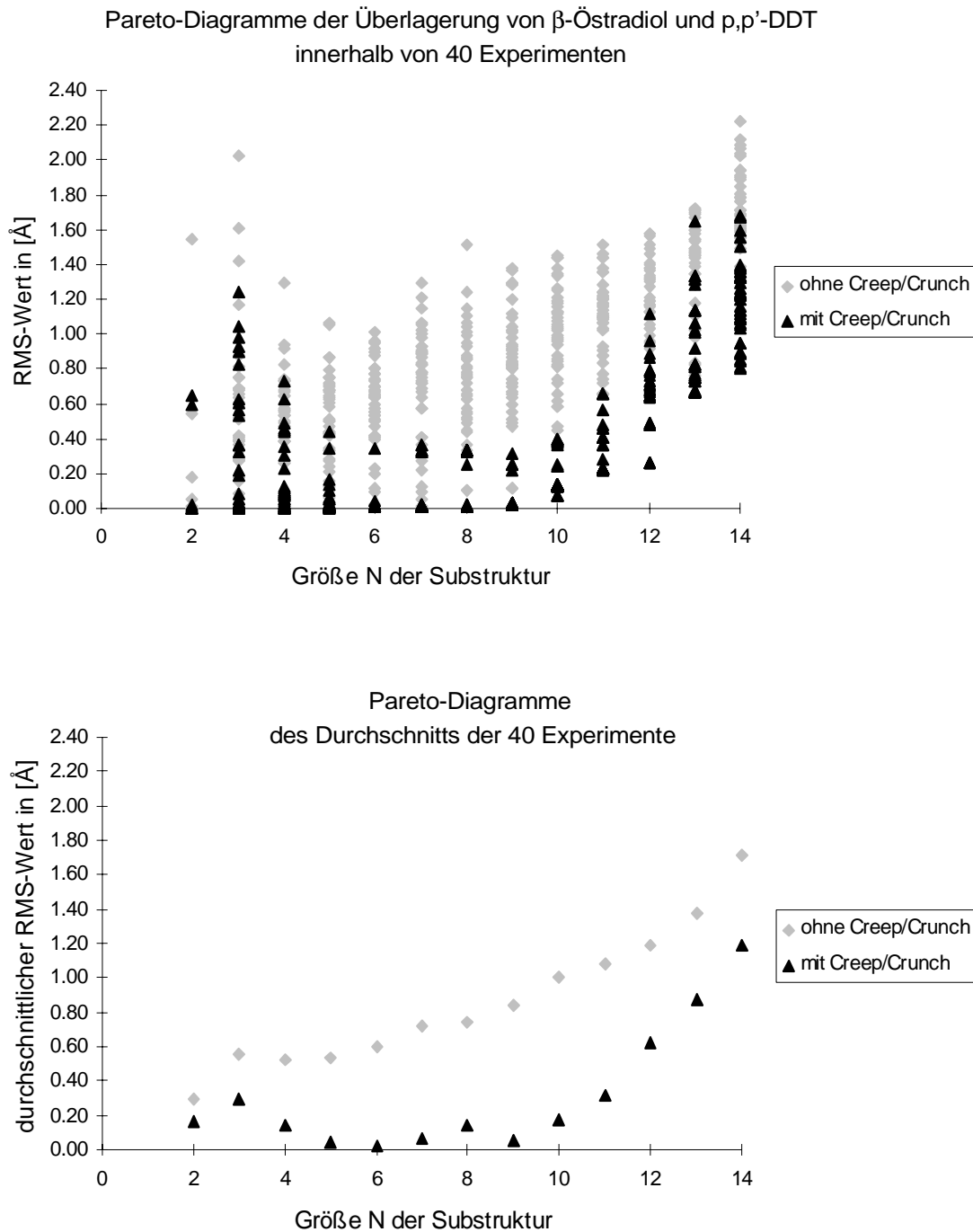


Abb. 2-36: Der Vergleich der GA-Optimierungsläufe unter Einsatz des Creep- und Crunch-Operators und ohne Creep und Crunch. Untersucht wurde die Überlagerung von β -Östradiol, **9**, und p,p'-DDT, **10**. Die Pareto-Diagramme zeigen die Relation des *rms*-Werts zur Größe der Substruktur, oben: innerhalb von 40 GA-Läufen, unten: als Durchschnitt der 40 GA-Läufe.

In beiden Diagrammen ist deutlich zu erkennen, daß die Einführung des Creep- und Crunch-Operators zu weit besseren Lösungen des Überlagerungsproblems führt. Die Substrukturen aller Größen weisen bei Einsatz des Creep- und Crunch-Operators bessere geometrische Anpassungen der Konformationen auf (niedrigere *rms*-Werte) als jene, die ohne Creep und Crunch berechnet wurden.

2.10 Allgemeine Optionen des Programms

Soll der Optimierungsalgorithmus ideale Lösungen liefern, so müssen v.a. die Startbedingung und die davon abhängige Initialisierung der Individuen variiert werden. Innerhalb des genetischen Algorithmus können folgende Parameter optional gesetzt werden (Tabelle 2-5).

Tab. 2-5: Wichtige Kontrollparameter des genetischen Algorithmus.

Kontrollparameter	Beschreibung	Erfahrungswert
i	Anzahl der Individuen einer Population	maximal: (Produkt der Atomanzahl aller Moleküle) / Molekülanzahl ²
g	Anzahl der Generationen eines Optimierungslaufs	$4 i$
e	Anzahl der Optimierungsläufe oder Experimente	25 - 50
S_{rts}	Größe der Subpopulation der Selektion	maximal $1/10 i$
$P_{Operatorname}$	Einsatzwahrscheinlichkeit der Operatoren	0.0 - 1.0 (siehe Tabelle 2-4)
$F = molx, moly...$	Liste von Molekülen, die während der Überlagerung als flexibel behandelt werden	optional, problemabhängig
$R = molx, moly...$	Liste von Molekülen, die während der Überlagerung als starr behandelt werden	optional, problemabhängig
A_{molx}	Atomindizes in Molekül x , die ausschließlich an der Überlagerung beteiligt sind, d.h. nur diese werden Teil der Substruktur	optional, problemabhängig
B_{molx}	Bindungsindizes in Molekül x aller Bindungen, die ausschließlich als rotierbar betrachtet werden (automatische Erkennung flexibler Bindungen ist ausgeschaltet)	optional, problemabhängig
F_{molx}	Atomindizes in Molekül x , die unbedingt an der Überlagerung teilhaben sollen (Fixpunkte)	optional, problemabhängig
$m_{atoprop} = x$	Physikochemische Atomeigenschaft $atoprop$ als Überlagerungskriterium, x ist ein Toleranzintervall, z.B. $q_{tot}(i) = q_{tot}(j) \pm x$.	optional, default ist die Ordnungszahl
C	Van der Waals-Radien-Überlappung wird geprüft	optional, problemabhängig

Die Berechnung von mehr als einem Optimierungslauf (Angabe der Anzahl der Optimierungsläufe bzw. Experimente durch die Option *-e*) ist nötig, da es sich bei genetischen Algorithmen um nicht-deterministische Prozesse handelt, die nicht notwendigerweise durch einen Optimierungslauf auf die beste Lösung führen. Es muß für jedes MCSS-Suchproblem eigens entschieden werden, wie viele Läufe durchgeführt werden, so daß mit einer gewissen Wahrscheinlichkeit eine beste Lösung erreicht wird (vgl. Kapitel 2.2). Das wiederholte Auftreten einer oder mehrerer Lösungen innerhalb der Experimente gibt einen Hinweis auf die maximale Anzahl der durchzuführenden Läufe bzw. Experimente. Bei multiplen Überlagerungen muß aufgrund auftretender Konvergenzprobleme bei der Konformationssuche ein Molekül als starr (Option *-R*) vorgegeben werden. Diese Verbindung bildet das Template für die Anpassung der Konformationen aller übrigen Moleküle. Ohne Vorgabe eines Templates konnte teilweise ein Hin- und Heroszillieren der Konformationen festgestellt werden, ohne daß ein Minimum gefunden wurde.

3 Extraktion von Pharmakophorpunkten aus der MCSS Suche

3.1 Angiotensin-II-Antagonisten

Angiotensin-II-Antagonisten sind von großer medizinischer Relevanz, da sie den letzten Schritt des Renin-Angiotensin-Systems (RAS), nämlich die Bindung von Angiotensin-II (AngII), **4**, an den AT₁- oder AT₂-Rezeptor blockieren. Man konnte zeigen, daß das RAS eine Schlüsselrolle in der Blutdruckregulation und des Elektrolyt- und Wasserhaushalts spielt.^[33] Die bisherige Behandlung von Bluthochdruck oder Herzinsuffizienz basierte auf der Blockierung der Umwandlung von AngI zu AngII durch das Angiotensin-Konversionsenzym (engl.: angiotensin converting enzyme, ACE). ACE ist jedoch eine unspezifische Protease und zudem verantwortlich für den Abbau von Bradykinin sowie anderer Peptide, wie z.B. Enkephalin.^[34] Daraus folgt, daß ACE-Hemmer zu verschiedenen Nebeneffekten führen, die hauptsächlich dem gehemmten Abbau des Bradykinins zugeschrieben werden.^[34] Der Versuch neue Methoden zu finden, das RAS zu blockieren, führte zu der Idee, den letzten Schritt in der Reaktionskaskade, die Bindung von Angiotensin-II, **4**, zum AT₁-oder AT₂-Rezeptor, zu verhindern. Takeda Chemical Industries^[70] bauten auf dem Konzept, daß nicht-peptidische AngII-Antagonisten nicht die Nachteile peptidischer AngII-Antagonisten, insbesondere die schnelle Spaltung nach oraler Aufnahme, aufweisen würden. An den von Takeda Chemical Industries entwickelten Verbindungen wurden durch DuPont weitere Untersuchungen durchgeführt, die zur Entdeckung von Losartan, **5**, (DuP 753, Abbildung 3-1) führten. Dies stellt einen Prototyp für eine neue Klasse potenter und oral verfügbarer nicht-peptidischer AngII-Rezeptor-Antagonisten dar.^[69]

Um molekulare Ähnlichkeiten zwischen Angiotensin-II-Antagonisten zu untersuchen und daraus Schlüsse auf Pharmakophorpunkte zu ziehen, wurden einige hoch aktive Angiotensin Antagonisten **5**, **8**, **13** und **14** (Abbildung 3-1 und Tabelle 3-1) mit der Leitstruktur Losartan, **5**, überlagert. Die Strukturen, sowie die Nomenklatur der AngII-Rezeptor-Antagonisten wurden der Publikation von Wexler et. al.^[69] entnommen.

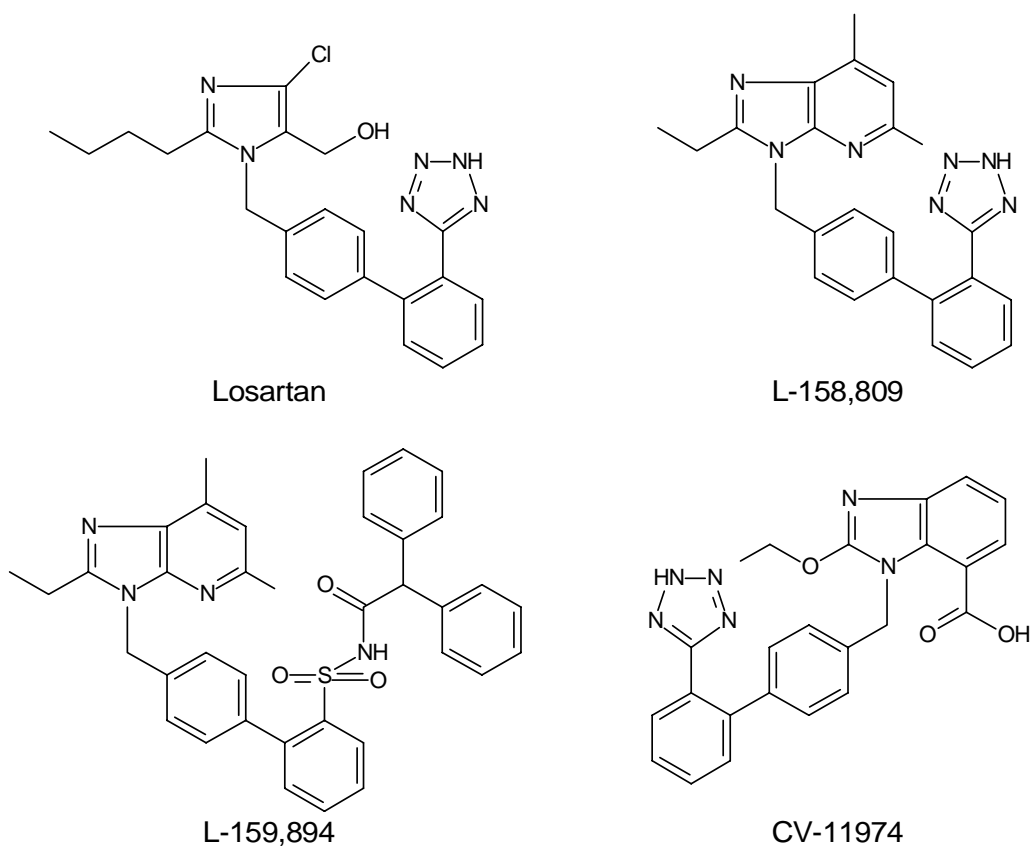


Abb. 3-1: Die Strukturformeln der Angiotensin-II Antagonisten Losartan, **5**, L-158,809, **8**, L-159,894, **13** und CV-11974, **14**.

Tab. 3-1: Bindungsaffinitäten in IC_{50} (Angiotensin-II, 25 μ M) der Angiotensin-II Antagonisten bezüglich des AT_1 -Rezeptors^[69]

Verbindung	IC_{50} [nM]
Losartan, 5	19
L-158,809, 8	0.3
L-159,894, 13	0.05
CV-11974, 14	28

Die dreidimensionalen Strukturen der Antagonisten und die des Octapeptids Angiotensin-II, **4**, (Asp-Arg-Val-Tyr-Ile-His-Pro-Phe) wurden mit Hilfe des 3D-Strukturgenerators CORINA^{[1],[2]} berechnen.

Da die in Kapitel 2.3.4 beschriebene Pareto-Optimierung bei der Berechnung der besten

Überlagerung integriert ist, erhält man nicht eine Überlagerung pro genetischem Experiment, sondern einen Paretosatz von Überlagerungen mit steigender Substrukturgröße und jeweils optimierter geometrischer Fitness. Jede hier gezeigte Überlagerung wurde dem besten Paretosatz entnommen (vgl. Kapitel 2.6). Die Operatoreinsatzwahrscheinlichkeiten während der genetischen Algorithmus-Läufe entsprechen den in Tabelle 2-4 angegebenen Werten. Alle weiteren Parameter, die in den folgenden Überlagerungen gesetzt wurden, sind in Tabelle 3-2 gegeben.

Tab. 3-2: Allgemeine Parameter der GA Experimente zur Überlagerung der Angiotensin-II Antagonisten

Parameter	Zahlenwert
i	100
g	1000 (AngII und Losartan) 400 (alle übrigen)
e	40
S_{rts}	16 (AngII und Losartan) 8
$m_{atoprop}$	Ordnungszahl

3.1.1 Vergleich starrer und flexibler Überlagerungen

Im folgenden wird zunächst besonders auf die Effizienz der Hybridmethode, d.h. der Kombination aus einem genetischen Algorithmus und einer numerischen Optimierungsmethode zur Anpassung der Konformationen eingegangen. Dazu werden die Ergebnisse der Überlagerungen für folgende Verbindungen verglichen: das Oktapeptid Angiotensin-II, **4**, (Asp-Arg-Val-Tyr-Ile-His-Pro-Phe) mit Losartan, **5**, (Abbildung 3-2), Losartan, **5**, mit L-158,809, **8**, (Abbildung 3-5), Losartan, **5**, mit L-159,894, **13**, (Abbildung 3-6) und Losartan, **5**, mit CV-11974, **14**, (Abbildung 3-7). Alle Verbindungen werden einerseits mit dem reinen genetischen Algorithmus und andererseits im Rahmen der Hybridmethode durch Rotation der Einfachbindungen^[12] flexibel überlagert.

Die erste Überlagerung von Losartan, **5**, und AngII, **4**, (oberer Bereich von Abbildung 3-2) wurde mit starren Konformationen für beide Moleküle erhalten. Während der Berechnung der zweiten Überlagerung (unterer Bereich von Abbildung 3-2) wird nur die Konformation des Oktapeptids starr gehalten und die Rotation der flexiblen Bindungen in der Losartanstruktur, **5**,

erlaubt. Dies führt zu einer konformativen Anpassung von Losartan, **5**, an AngII, **4**. In diesem Beispiel agiert AngII, **4**, somit als *Templat* für die Anpassung der Konformationen während der Überlagerungen.

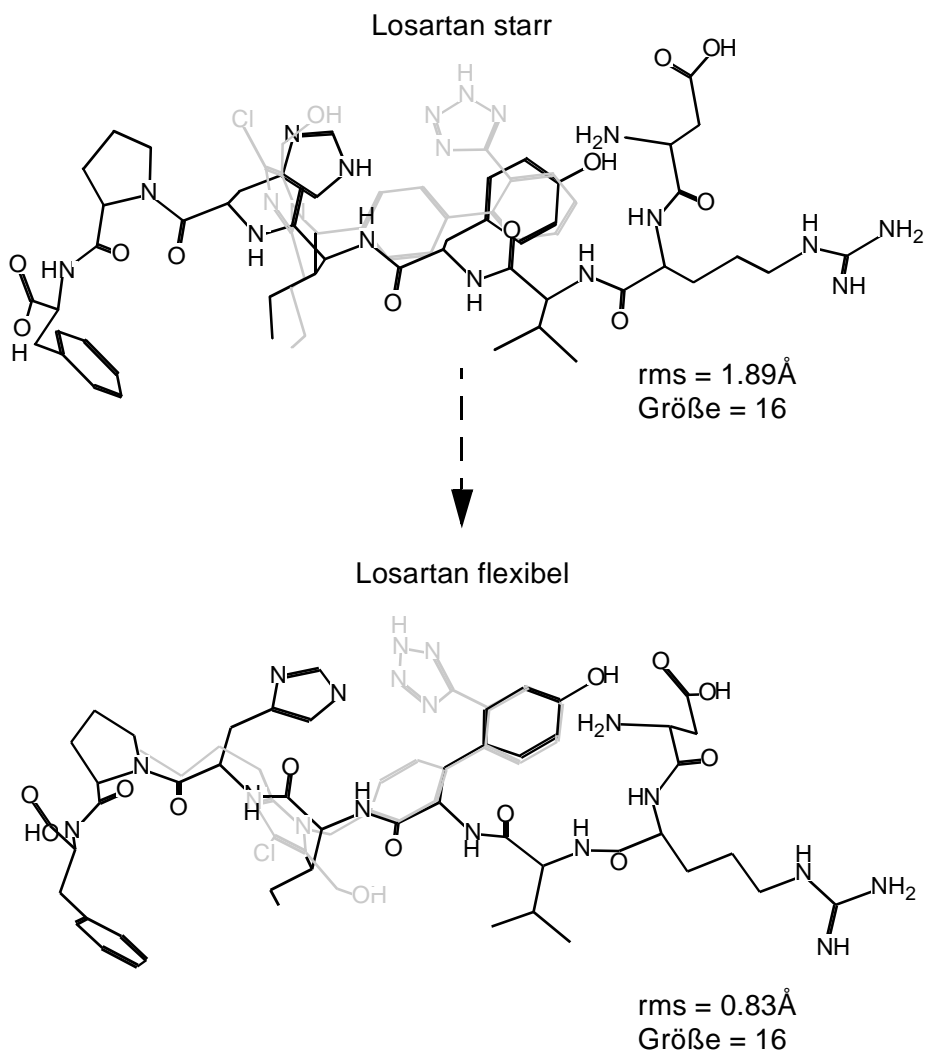


Abb. 3-2: Der Vergleich einer starren und flexiblen Überlagerung von AngII, **4**, (schwarz) mit Losartan, **5**, (grau). In der oberen Überlagerung wurden beide Strukturen starr gehalten, in der unteren war die Rotation der flexiblen Bindungen von Losartan, **5**, erlaubt.

Der Vergleich beider Überlagerungen basiert auf Substrukturen gleicher Größe. Die Pareto-Kurven der besten von 40 genetischen Experimenten zur starren und flexiblen Überlagerung von Losartan, **5**, mit AngII, **4**, sind in Abbildung 3-3 dargestellt. Die besten Experimente sind jeweils, die mit dem niedrigsten *rms*-Wert für die Substrukturgröße, ab der die Pareto-Kurven steil ansteigen (vgl. Kapitel 2.6).

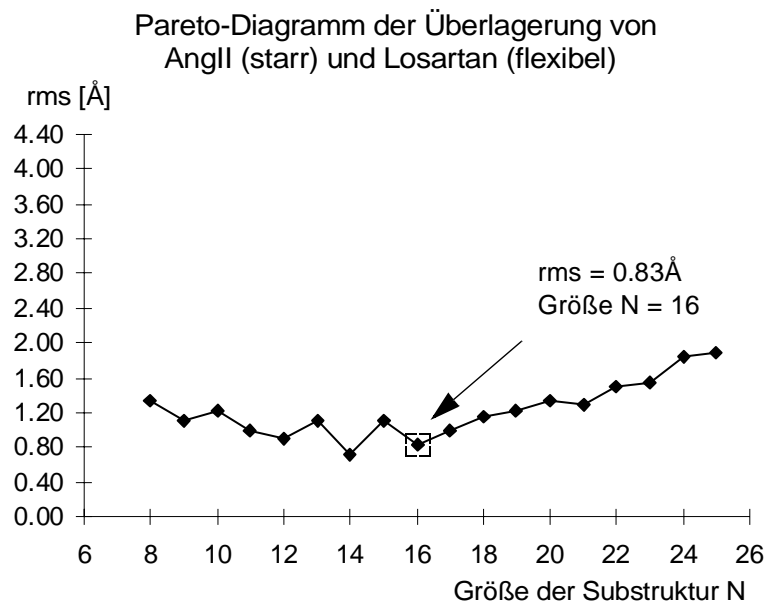
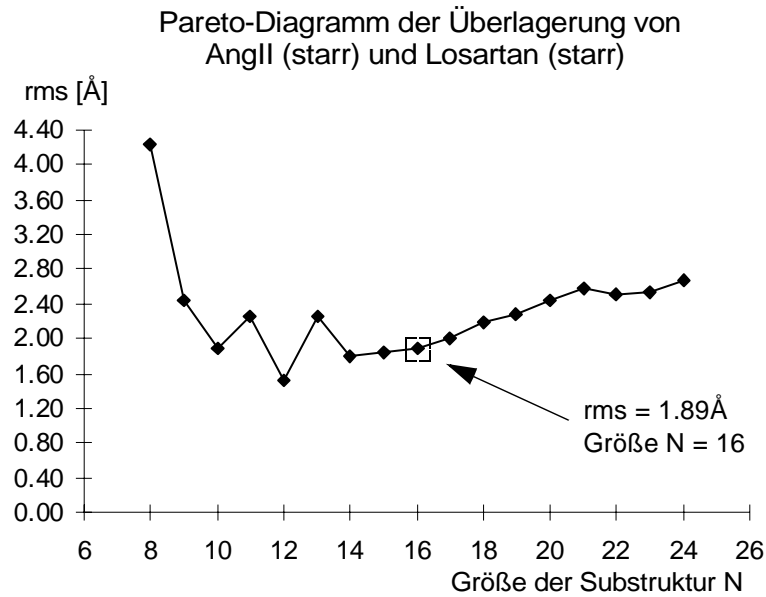


Abb. 3-3: Die Pareto-Diagramm der starren (oben) und flexiblen (unten) Überlagerung von Losartan, 5, und AngII, 4, des jeweils besten erreichten Paretosatzes innerhalb von 40 Experimenten. Die Rechtecke markieren die Datenpunkte der in Abbildung 3-2 gezeigten Überlagerungen.

Bei der flexiblen Überlagerung lassen sich gegenüber der starren Überlagerung für jede Substrukturgröße kleinere *rms*-Wert erhalten. Bei der starren Überlagerung ist zusätzlich der Anstieg des *rms*-Wertes und die damit verbundene Verschlechterung der geometrischen Anpassung weit weniger steil als bei der flexiblen. Die Substrukturgröße $N = 16$ bildet ein Minimum in der Berechnung der flexiblen Überlagerung. Sie ist daher die größte Substruktur

mit noch akzeptabler geometrischer Abweichung ($rms < 1.0\text{\AA}$). Für alle nachfolgenden Substrukturgrößen steigt der rms -Wert rapide an. Die verbesserte geometrische Anpassung der Konformationen nach der flexiblen Überlagerung gegenüber der starren ist deutlich an dem niedrigeren rms -Wert, 0.83\AA gegenüber 1.89\AA , zu erkennen.

Der C-Terminus von AngII, **4**, wird für die AT_1 -Rezeptor Affinität des Peptids verantwortlich gemacht. Die AngII-Antagonisten mimetieren daher genau diesen Bereich des Octapeptids.^[69] Die von Wexler et. al. publizierten, für die Affinität relevanten Substrukturen von AngII, **4**, und dessen Antagonisten^[69] umfassen die Pro-His-Ile-Tyr-Reste des Peptids, somit also den C-Terminus. Nur die hier ermittelte flexible Überlagerung von AngII, **4**, mit **5** (unterer Bereich von Abbildung 3-2) führt zu einer Substruktur, die sich in Richtung C-Terminus des Peptids verschoben hat. Sind, wie in diesem Beispiel, pharmakophore Teilbereiche einer der zu vergleichenden Strukturen bekannt, können diese als Grundlage für die Substruktur angegeben werden. Im folgenden wird getestet, in wie weit sich die jeweils andere Struktur auf diese pharmakophoren Bereiche anpaßt. In Abbildung 3-4 ist die Überlagerung von AngII mit **5** dargestellt, wobei der C-Terminus, durch Angabe der entsprechenden Atomindizes, als ausschließlicher Teil der Substruktur vorgegeben ist.

Überlagerung von AngII (C-Terminus, starr) und Losartan (flexibel)

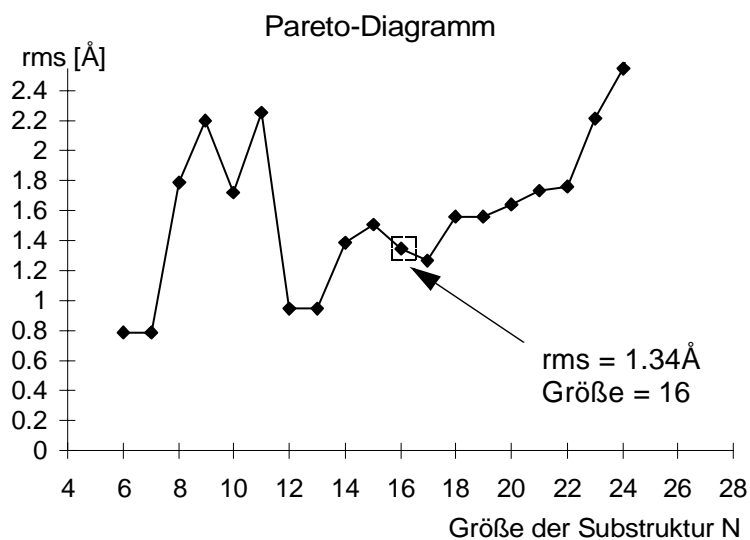
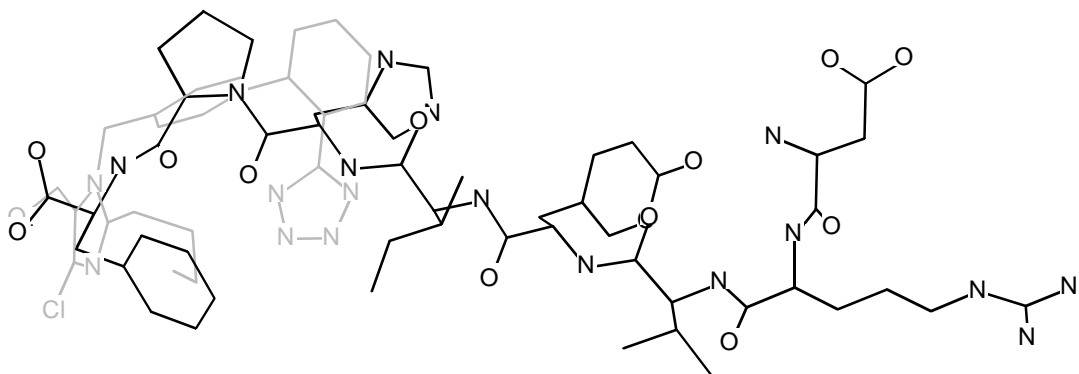


Abb. 3-4: Die Überlagerung von AngII, **4**, (starr) mit Losartan, **5** (flexibel). Der C-Terminus von AngII, **4**, ist als Teil der Substruktur vorgegeben. Das Pareto-Diagramm ist das des besten Experiments innerhalb von 40 Experimenten.

Bei gleicher Substrukturgröße (16 Atome) ist der *rms*-Wert höher ($rms = 1.34\text{\AA}$ gegenüber $rms = 0.83\text{\AA}$) als in der flexiblen Überlagerung ohne Angabe von Atomindizes (Abbildung 3-2). Allerdings stellt die Größe von 16 Atomen auch kein Minimum des Pareto-Diagramms dar. Bessere Überlagerungen basieren auf den Substrukturgrößen 12, 13 ($rms = 0.94\text{\AA}$) und 17 Atome ($rms = 1.27\text{\AA}$).

Die Überlagerungen von **5** mit **13** (Abbildung 3-5), mit **8** (Abbildung 3-5) und mit **14** (Abbildung 3-6) wurden ebenfalls einerseits mit starren Konformationen (linke Seiten der Abbildungen 3-5 bis 3-7), andererseits durch Berücksichtigung der konformativen Flexibilität der Strukturen (rechte Seiten der Abbildungen 3-5 bis 3-7) erstellt. Die flexiblen Überlagerungen passen bei beiden Molekülen die Torsionswinkel innerhalb der Substrukturen aneinander

an, so daß die geometrische Abweichung der Konformationen in diesem Bereich minimiert wird.

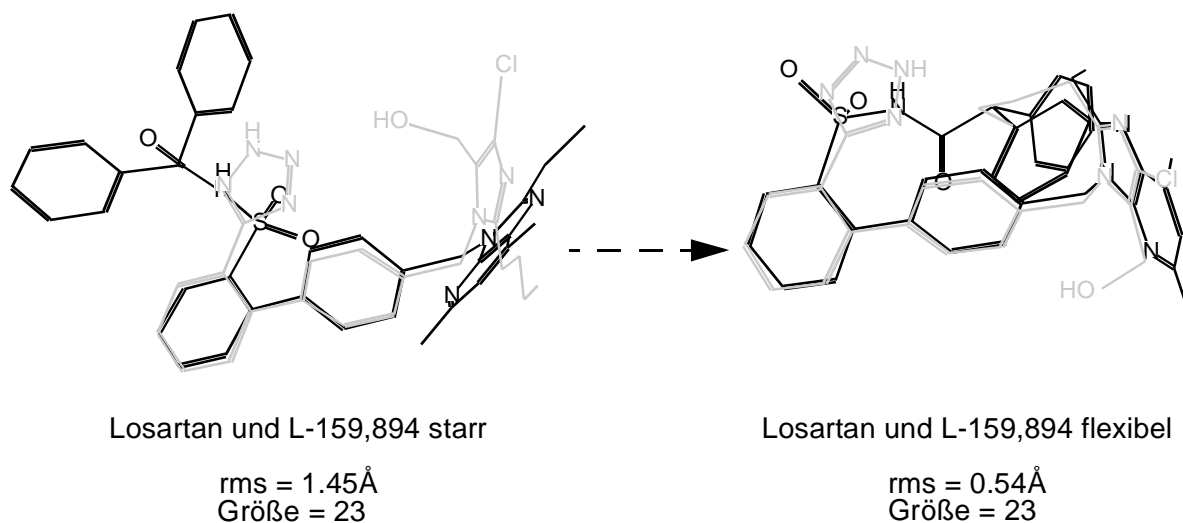


Abb. 3-5: Die starre und flexible Überlagerung von Losartan, **5**, (grau) und L-159,894, **13**, (schwarz). In der linken Überlagerung wurden beide Strukturen als starr behandelt, in der rechten flexibel.

Wenn beide Strukturen als flexibel behandelt werden, beträgt die mittlere Rechenzeit eines GA Experiments 7min 20s auf einer SGI Origin 200 (single R10000, 180 MHz), während sie sich bei starrer Überlagerung auf 43s beschränkt. Die Berücksichtigung der konformationellen Flexibilität beider Strukturen führt zu einem ungefähr zehnfachen Anstieg der Rechenzeiten.

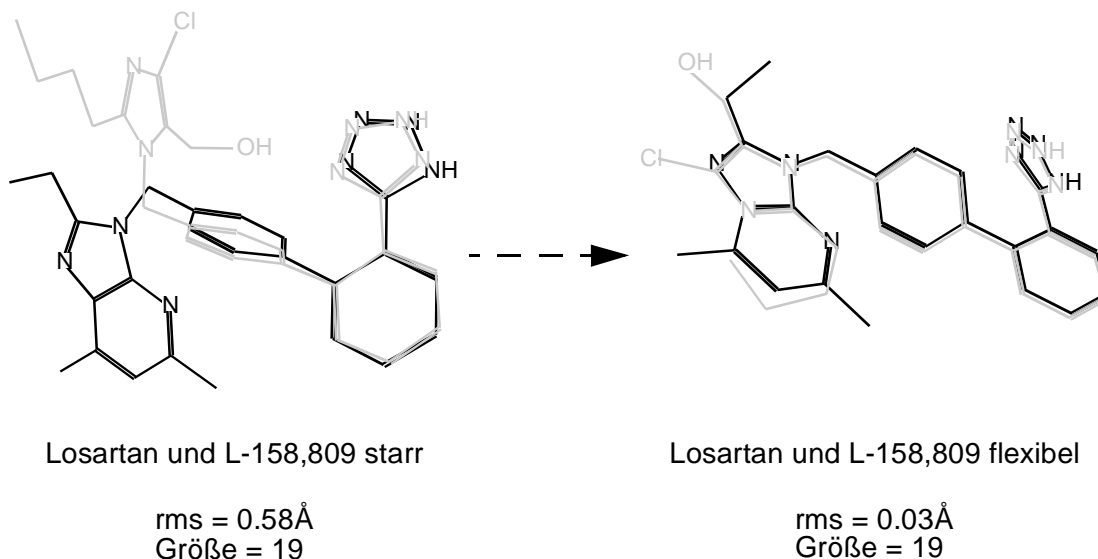


Abb. 3-6: Die starre und flexible Überlagerung von Losartan, **5**, (grau) und L-158,809, **8**, (schwarz). In der linken Überlagerung wurden beide Strukturen als starr behandelt, in der rechten flexibel.

Ein Vergleich der starren und flexiblen Überlagerungen (Abbildungen 3-5 bis 3-7) zeigt in allen drei Beispielen die deutliche Verbesserung in der geometrischen Anpassung der Konformationen, was sich in niedrigeren *rms*-Werten bei gleicher Substrukturgröße ausdrückt.

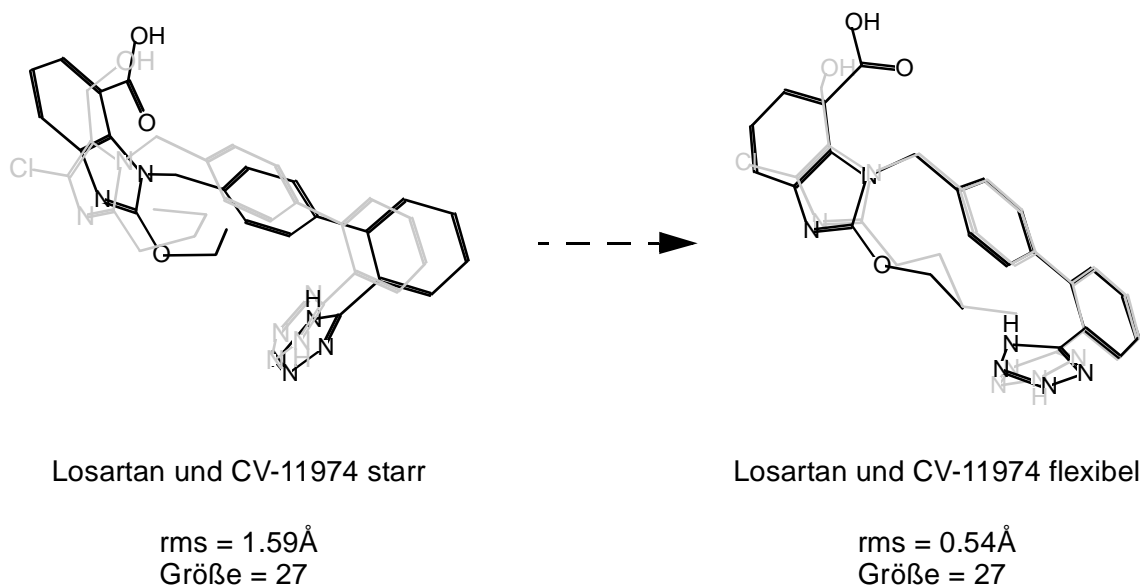


Abb. 3-7: Die starre und flexible Überlagerung von Losartan, **5**, (grau) und CV-11974, **14** (schwarz). In der linken Überlagerung wurden beide Strukturen als starr behandelt, in der rechten flexibel.

In den Substrukturen der flexiblen Überlagerungen sind, zusätzlich zu den beiden aroma-

tischen Ringen, der Imidazolring und die Alkylseitenketten des Imidazolrings mit entsprechenden Regionen von **8**, **13** oder **14** überlagert. Diese Ergebnisse entsprechen den von Lin et al.^[71] publizierten Struktur-Wirkungs-Beziehungen (engl.: structure activity relationships, SAR) der AngII-Antagonisten, die vor allem lipophile (Benzolring und Alkylseitenkette) und basische Bereiche (Imidazolring) der Verbindungen als verantwortlich für eine hohe Affinität der AngII-Antagonisten zum AT₁-Rezeptor betonen.

Die hier vorgestellte Hybridmethode zum Auffinden der größten gemeinsamen Substruktur eines Satz von Verbindungen ist zudem nicht auf den Erhalt von niedrigenergetischen Verbindungen beschränkt. Daher ist die Wahrscheinlichkeit, eine Konformation zu finden, die die Verbindung innerhalb der Rezeptorbindungstasche aufweist, erhöht.

3.1.2 Multiple Überlagerungen

Zusätzlich zum Vergleich zweier dreidimensionaler Strukturen ist es möglich, Sätze von Verbindungen zu analysieren, die die gleiche biologische Aktivität aufweisen. Um dies zu zeigen, wurden die dreidimensionalen Strukturen der vier Angiotensin-II-Antagonisten **5**, **8**, **13** und **14** (Abbildung 3-1) in einem simultanem Prozeß überlagert. Die in Abbildung 3-8 gezeigte Überlagerung ist die innerhalb von 40 GA Läufen am häufigsten erreichte.

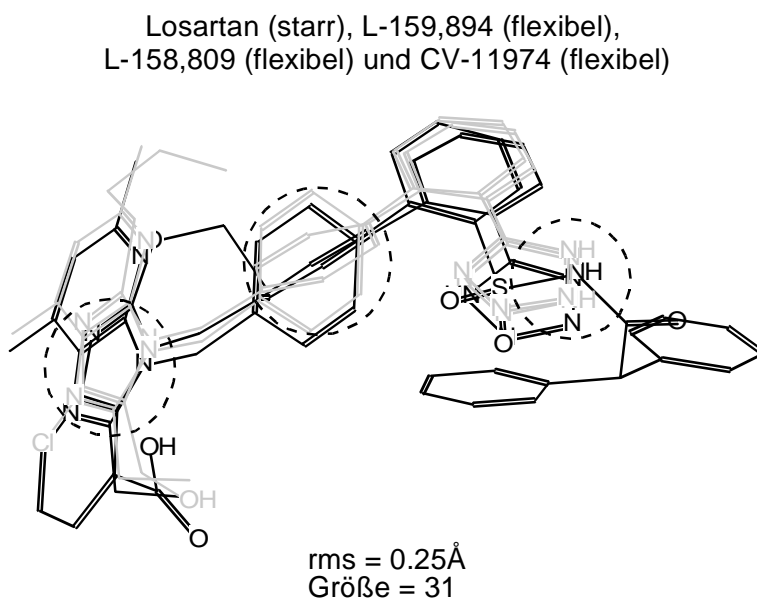


Abb. 3-8: Das Ergebnis der gleichzeitigen Überlagerung von Losartan, **5** (grau), L-159,894, **13** (schwarz), L-158,809, **8** (grau) und CV-11974, **14** (schwarz). Die gestrichelten Kreise markieren pharmakophore Bereiche.

Wiederum enthält die 3D-MCSS beide Benzolringe und den Imidazolring und umfaßt somit die hydrophoben und basischen Bereiche der Strukturen, die für die Bindung zum AT₁-Rezeptor eine wichtige Rolle spielen und damit pharmakophore Punkte bilden.^[71] Jones et. al.^[52] publizierte eine äquivalenten Überlagerung eines Satzes von Angiotensin-II-Antagonisten mit zwei der hier ebenfalls im Datensatz enthaltenen Strukturen (**5** und **8**). Auch Jones et. al. finden nach Analyse der Überlagerungen als Pharmakophorpunkte ein basisches Stickstoffatom und die Benzolringe. In Abbildung 3-8 sind beide pharmakophoren Strukturbereiche und der Tetrazolring durch gestrichelte Kreise markiert. Die Imidazolringe aller Verbindungen, die das basische Stickstoffatom enthalten, die Benzolringe und der Tetrazolring sind in der gezeigten Überlagerung an der Substruktur beteiligt.

3.1.3 Auswirkung der Überlagerungskriterien auf die Substrukturen

Um festzulegen, ob nicht nur strukturelle Ähnlichkeiten, sondern auch Ähnlichkeiten in den physikochemischen Eigenschaften der Angiotensin-II-Antagonisten vorliegen, werden die folgenden Überlagerungen mit dem Überlagerungskriterium eines bestimmten Ladungsintervalls durchgeführt.

In Abbildung 3-9 sind die Werte der partiellen Atomladungen q_{tot} der einzelnen Angiotensin-II-Antagonisten **5**, **8**, **13** und **14** dargestellt.

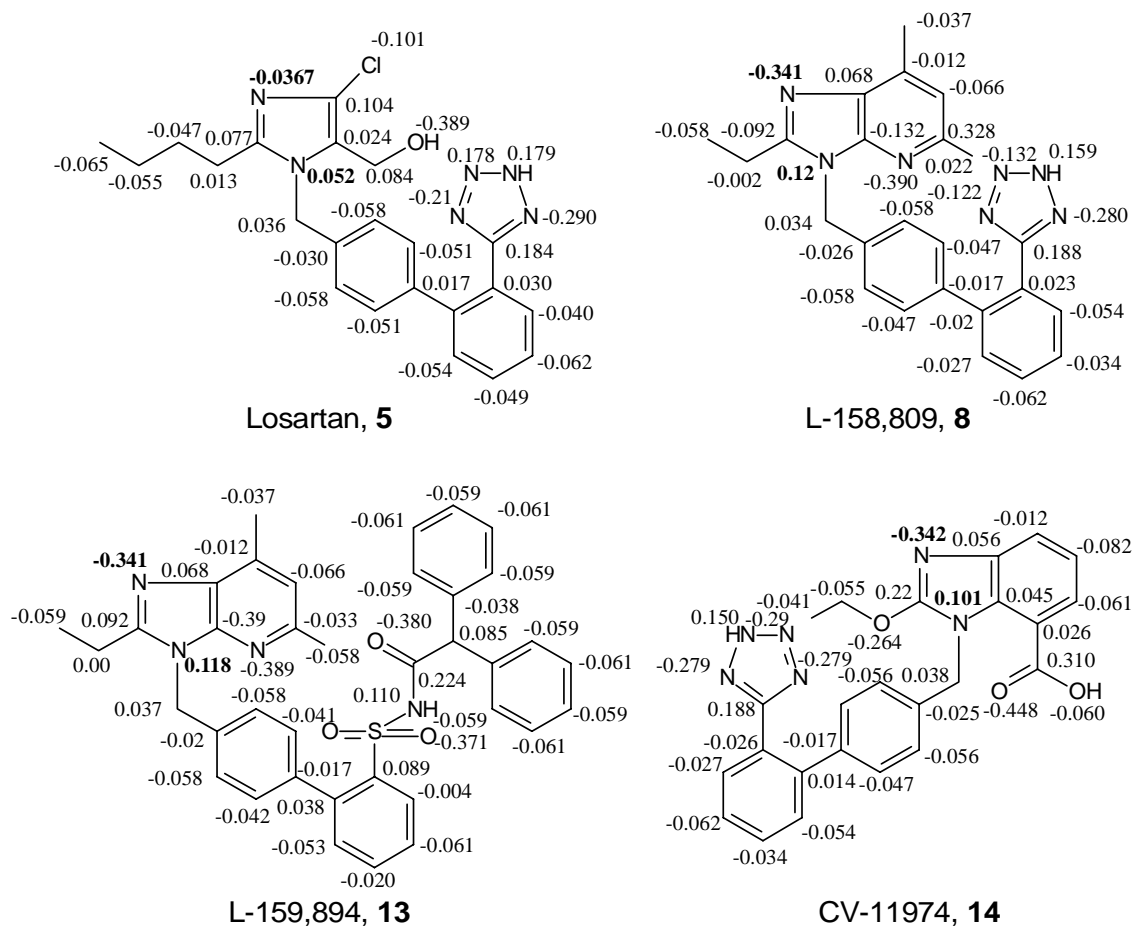
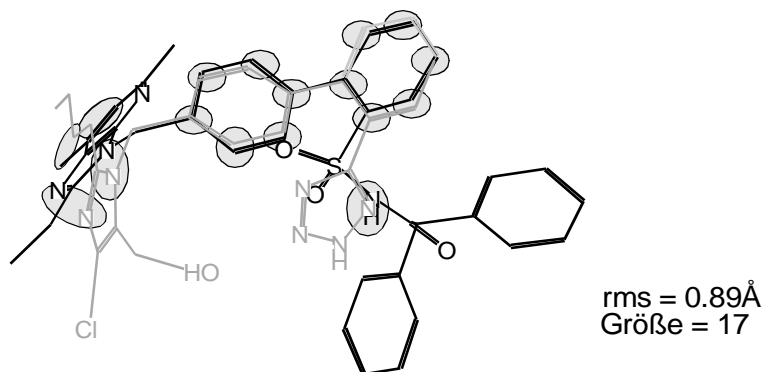


Abb. 3-9: Die Strukturen und partiellen Atomladungen q_{tot} der Angiotensin-II-Antagonisten **5**, **8**, **13** und **14**.

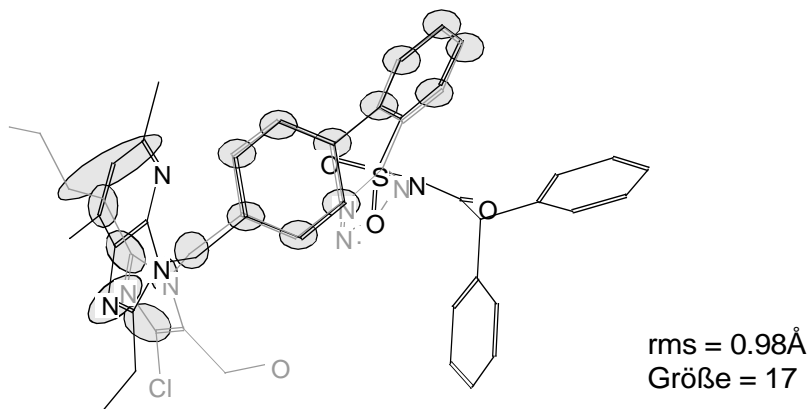
Anhand der Überlagerung von Losartan, **5**, und L-159,894, **13**, lassen sich die Auswirkungen anderer Überlagerungskriterien als den Ordnungszahlen auf die Substrukturen zeigen.

Überlagerung von Losartan mit L-159,894

Überlagerungskriterium: Ordnungszahl



Überlagerungskriterium: $q_{tot}(i) = q_{tot}(j) \pm 0.01$



Überlagerungskriterium: $q_{tot}(i) = q_{tot}(j) \pm 0.001$

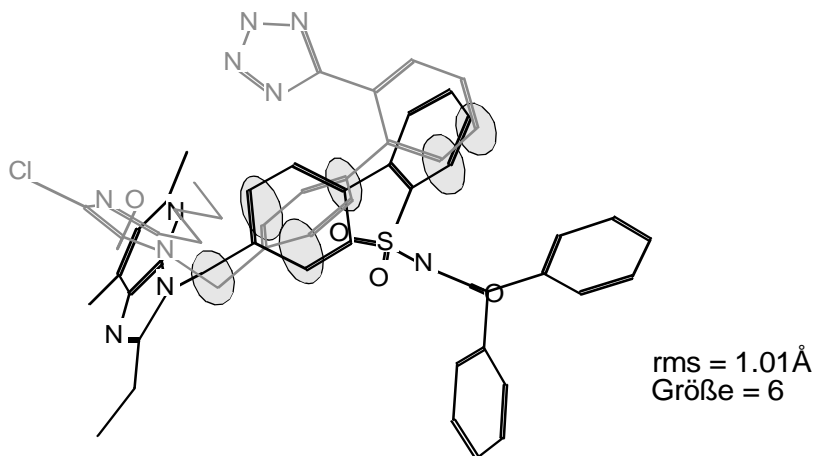


Abb. 3-10: Verschiedene Substrukturen der Überlagerung von Losartan, **5**, mit L-159,894, **13**, die durch Angabe unterschiedlicher Matchkriterien erhalten werden. Die Atome der MCSS sind durch graue Ovale markiert.

In Abbildung 3-10 sind einerseits die Überlagerung von Losartan, **5**, und L-159,894, **13**, mit der Ordnungszahl als Überlagerungskriterium (Abbildung 3-10 oben) und andererseits mit der Atompartialladungen $q_{tot}(i) = q_{tot}(j) \pm x$ als Überlagerungskriterium gezeigt, wobei x im ersten Fall 0.01e (Abbildung 3-10 Mitte) ist und im zweiten Fall 0.001e (Abbildung 3-10 unten) beträgt.

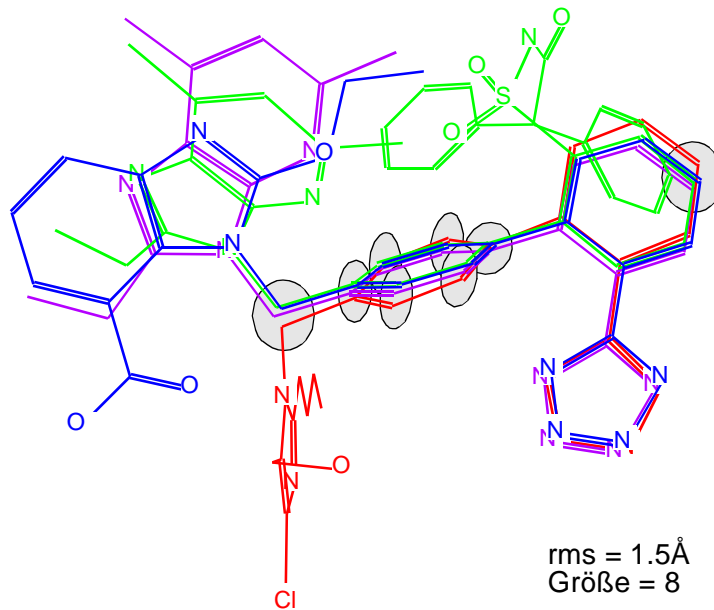
Wie bereits in Kapitel 3.1.1 beschrieben, spielen für die Affinität der Angiotensin-II-Antagonisten hauptsächlich die beiden hydrophoben Benzolringe, der Imidazolring bzw. das basische Stickstoffatom des Imidazolrings und die Alkylseitenketten am Imidazolring eine große Rolle.

Die Überlagerung von **5** und **13** mit dem Überlagerungskriterium $q_{tot}(i) = q_{tot}(j) \pm 0.01$ ist noch sehr ähnlich zur Überlagerung mit der Ordnungszahl als Überlagerungskriterium und umfaßt große Bereiche beider Strukturen (siehe Abbildung 3-10, oben und Mitte). Das Überlagerungskriterium von $q_{tot}(i) = q_{tot}(j) \pm 0.001$, wurde so eng gewählt, daß sich daraus nur Atomtupel bilden lassen, die sehr gering abweichende bzw. identische q_{tot} -Werte aufweisen. Dabei lassen sich besonders hohe physikochemische Ähnlichkeiten ermitteln, die hier in den Atomen des einen Phenylenrings und der Brücke zum Imidazolring liegen.

Die multiple Überlagerung von **5**, **8**, **13** und **14** aus Kapitel 3.1.2 wurde ebenfalls mit dem Überlagerungskriterium $q_{tot}(i) = q_{tot}(j) \pm 0.01$ (Abbildung 3-11 a) und $q_{tot}(i) = q_{tot}(j) \pm 0.001$ (Abbildung 3-11 b) wiederholt, um physikochemische Ähnlichkeiten in allen vier Verbindungen herauszustellen. Es ist die jeweils größte gemeinsame Substruktur für beide Bedingungen repräsentiert. Die höchsten Ähnlichkeiten bezüglich der Partialladungen q_{tot} ergeben sich wiederum innerhalb der Benzolringe. Bei einem q_{tot} -Intervall von 0.01 ist die Orientierung der Verbindungen mit Ausnahme von **5** ähnlich zu der in Abbildung 3-8, bei der das Überlagerungskriterium, die Ordnungszahl war. Werden die Werte der partiellen Atomladungen (Abbildung 3-9) der Stickstoffatome der Imidazolringe betrachtet, so ist zu erkennen, daß die jeweiligen Ladungen in den Verbindungen **8**, **13** und **14** sehr ähnlichen sind, in Verbindung **5** jedoch von diesen abweichen. Der Grund für die abweichenden q_{tot} -Werte der Imidazolstickstoffatome von **5** liegt an dem stark elektronegativen Chlorsubstituenten des Imidazolrings und dessen elektronenziehenden Eigenschaften.

Losartan (starr), L-159,894 (flexibel),
L-158,809 (flexibel) und CV-11974 (flexibel)

a) Überlagerungskriterium: $q_{tot} = q_{tot}(x) \pm 0.01$



b) Überlagerungskriterium: $q_{tot} = q_{tot}(x) \pm 0.001$

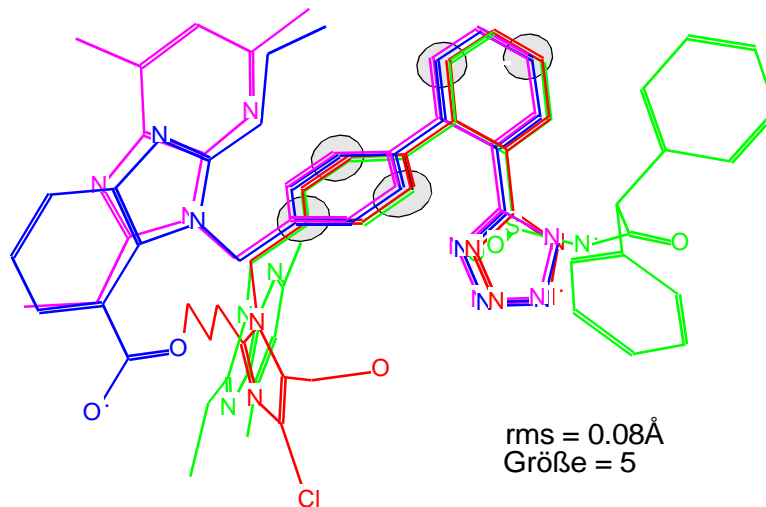


Abb. 3-11: Das Ergebnis der gleichzeitigen Überlagerung von Losartan, **5** (rot), L-159,894, **13** (blau), L-158,809, **8** (violett) und CV-11974, **14** (grün), mit einem Überlagerungskriterium von a) $q_{tot}(i) = q_{tot}(j) \pm 0.01$ und b) $q_{tot}(i) = q_{tot}(j) \pm 0.001$. Die Atome der MCSS sind durch graue Ovale markiert.

3.2 Cytochrom P450c17-Inhibitoren

Der Metabolismus von Fremdstoffen (z.B. Arzneimittel oder Giftstoffe) läuft im Organismus in zwei verschiedenen Phasen ab. Die Reaktionen der Phase I führen geeignete funktionelle Gruppen in das Molekül ein, die anschließend als Ausgangspunkt für die Reaktion der Phase II dient. Während der Phase-II-Reaktion konjugiert ein endogenes (körpereigenes) Molekül an den Fremdstoff. Die meisten dieser endogenen Verbindungen sind hydrophil und bilden durch die Konjugation zum Fremdstoff eine wasserlösliche Substanz.^[39]

Phase-I-Reaktionen können Oxidationen, Hydrolysen oder Demethylierungen sein. Oxidationen spielen jedoch mit Abstand die größte Rolle. Hierfür sind mikrosomale Monooxygenasen verantwortlich. Der zentrale Baustein der Monooxygenasen ist das Cytochrom P450 (membranständiges Hämoprotein) oder eines seiner Isoenzyme. Während der Oxidation durchläuft das zentrale Eisenion des Cytochrom P-450 im Häm eine Änderung von Fe^{3+} nach Fe^{2+} und zurück nach Fe^{3+} .^[39]

Das hier behandelte P450c17-Enzym (17- α -Hydroxylase/C17-20 Lyase) stellt das Schlüsselenzym der Androgen- und Glucocorticoidbiosynthese dar.^[35] Mit den oben beschriebenen Cytochrom P450 Isoenzymen hat P450c17 vor allem das Häm als prosthetische Gruppe gemeinsam. Es ist bekannt, daß die an das Enzym konjugierenden Substanzen eine Koordination zum zentralen Eisenatom am einen Ende des Grundgerüsts, sowie eine Wasserstoffbrücke am anderen Ende aufweisen.^[35] Daraus bedingt sich Anwesenheit eines Atoms mit freiem Elektronenpaar (z.B. Stickstoff) und mindestens eines Wasserstoffbrückenakzeptors oder -donors. Die Hemmung der 17- α -Hydroxylase/C17-20 Lyase stellt ein vielversprechendes Konzept zur Behandlung des Prostatacarcinoms dar, doch Inhibitoren dieses Enzyms mit annehmbarer Selektivität sind bislang nicht beschrieben.^{[36][37]}

P450c17-Inhibitoren können in drei verschiedene Strukturklassen bezüglich des Steroidgrundgerüsts (Abbildung 3-12) unterteilt werden: in A/B/C/D-Ringsysteme (Tabelle 3-3), A/B-Ringsysteme (Tabelle 3-4) und A/C-Ringsysteme (Tabelle 3-5). In Tabelle 3-3 bis 3-5 sind die chemischen Strukturen, die jeweiligen internen Bezeichnungen und die IC_{50} Human-Werte (Progesterone, 25 μM)^{[36][37]} der Verbindungen angegeben. Imidq, **15**, ist dabei der aktivste Steroidligand mit einem starren Kohlenstoffgerüst. BW13, **16**, und BW112, **17**, sind beides konformationell flexible Liganden, wobei **16** der aktivste 3,4-Dihydronaphthalinligand und **17** der aktivste Biphenylligand ist.

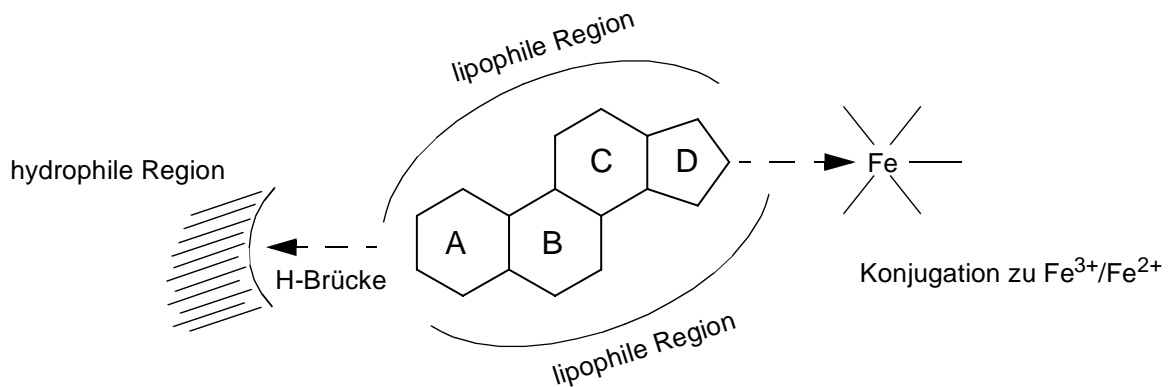


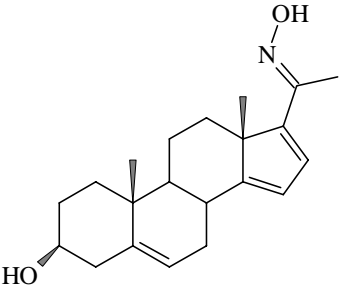
Abb. 3-12: Das Steroidgrundgerüst und die mutmaßlichen Koordinationsstellen der einzelnen Bereiche innerhalb des PC450c17-Rezeptors.

Alle in Tabelle 3-3 bis 3-5 aufgelisteten Verbindungen koordinieren an einem Ende über ein basisches Stickstoffatom an das zentrale Eisenatom im P450c17 und binden am anderen Ende über einen Wasserstoffbrückenakzeptor (OH-, OMe- oder NR₂-Gruppe) an eine basische Gruppe.

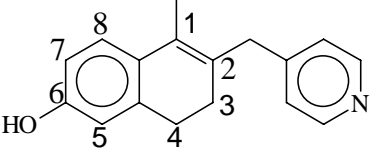
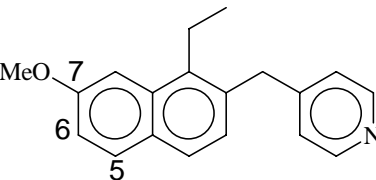
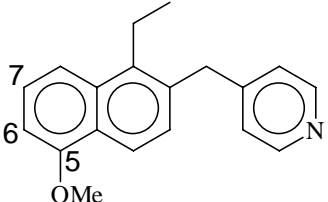
Tab. 3-3: P450c17-Inhibitoren mit A/B/C/D-Ringsystem mit den zugehörigen IC₅₀ Human-Werten (Progesteron, 25 µM).

Verbindung	Bezeichnung	IC ₅₀ [nM]
	Imidq, 15	0.040
	MH61 (Z-Isomer, E-Isomer ist inaktiv)	0.077

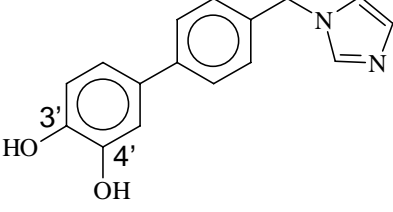
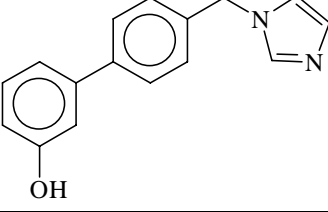
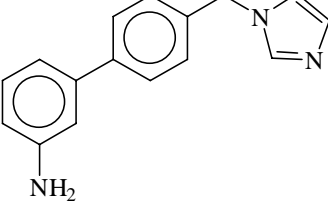
Tab. 3-3: P450c17-Inhibitoren mit A/B/C/D-Ringsystem mit den zugehörigen IC_{50} Human-Werten (Progesteron, 25 μ M).

Verbindung	Bezeichnung	IC_{50} [nM]
	MH55	0.17

Tab. 3-4: P450c17-Inhibitoren mit A/B-Ringsystem mit den zugehörigen IC_{50} Human-Werte (Progesteron, 25 μ M).

Verbindung	Bezeichnung	IC_{50} [nM]
	BW13, 16	0.036
	BW61, 18	0.074
	BW62, 19	0.085

Tab. 3-5: P450c17-Inhibitoren mit A/C-Ringsystem mit den zugehörigen IC_{50} Human-Werten (Progesteron, 25 μ M).

Verbindung	Bezeichnung	IC_{50} [nM]
	BW112, 17	0.087
	BW95	0.13
	BW99	0.21

3.2.1 Überlagerung mit wissensbasiertem Ansatz

Für die aktivsten Verbindungen jeder Strukturklasse (**15**, **16** und **17**) sollen im folgenden die strukturellen Homologien ermittelt werden. Imidq, **15**, bildet aufgrund seines relativ starren Steroidgrundgerüsts ein Templat für die Überlagerungen mit BW13, **16**, und BW112, **17**. Die Konformation von Imidq, **15**, wurde somit bei jeder Überlagerung konstant gehalten, die Konformationen aller mit **15** überlagerten Verbindungen wurden als flexibel behandelt. War an der Überlagerung Imidq, **15**, nicht beteiligt so wurden die Konformationen beider Moleküle flexibel behandelt.

Aufgrund der unterschiedlichen Grundgerüstgrößen der drei Verbindungsgruppen führt eine konventionelle, starre Überlagerung nicht Ergebnissen, die kein Wissen über die Struktur und Anforderungen des spezifischen Rezeptors berücksichtigen. Die in Abbildung 3-13 gezeigten Überlagerungen von **15** und **16**, einer A/B/C/D-Ring und einer A/B-Ringverbin-

ding, wurden auf drei verschiedene Weisen erhalten. Das Überlagerungskriterium war bei allen drei Überlagerungen die Ordnungszahl, es wurden die üblichen, in Tabelle 3-2 beschriebenen, genetischen Parameter gesetzt und das beste Ergebnis der jeweiligen 40 Experimente repräsentiert. Die Unterschiede liegen in den vorgegebenen Bedingungen für die zu erhaltenen Substrukturen. In der Berechnung der ersten Überlagerung war die Wahrscheinlichkeit für alle Atome, in der gemeinsamen Substruktur enthalten zu sein, identisch (Abbildung 3-13, oben). In der zweiten Überlagerung wurde die Überlagerung des Atompaares **15**: N²/**16**: N¹ erzwungen (Abbildung 3-13 mitte). In der dritten Überlagerung wurde zusätzlich die Überlagerung des Atompaares **15**: O¹/**16**: O¹ verlangt (Abbildung 3-13 unten). Das Wissen über spezielle Anlagerungsformen der Liganden innerhalb der Rezeptorbindungstasche wird somit über erzwungene Matchpaare in die Berechnung der strukturellen Ähnlichkeiten eingebracht (Abbildung 3-13 Mitte und unten). Die gegebenen Bedingungen führen im vorliegenden Beispiel zu einer Verschlechterung der gesamten geometrischen Anpassung, denn es ergeben sich *rms*-Werte von 0.91Å und 1.14Å gegenüber 0.50Å. Diese Verschlechterung liegt vor allem an der unterschiedlichen Größe der beiden Strukturen. Ohne Vorgaben, wird sich die kleinere Struktur grundsätzlich an einen Bereich der größeren anpassen (Abbildung 3-13 oben). In diesem Fall liegt BW13, **16**, genau auf dem A/B/C-Ring Bereich von Imidq, **15**. Dies führt zu niedrigeren *rms*-Werten, vernachlässigt aber, daß im vorderen und hinteren Bereich der steroidal Struktur, genauer im C/D-Ring Bereich, ebenso Ähnlichkeiten zu finden sein müßten.

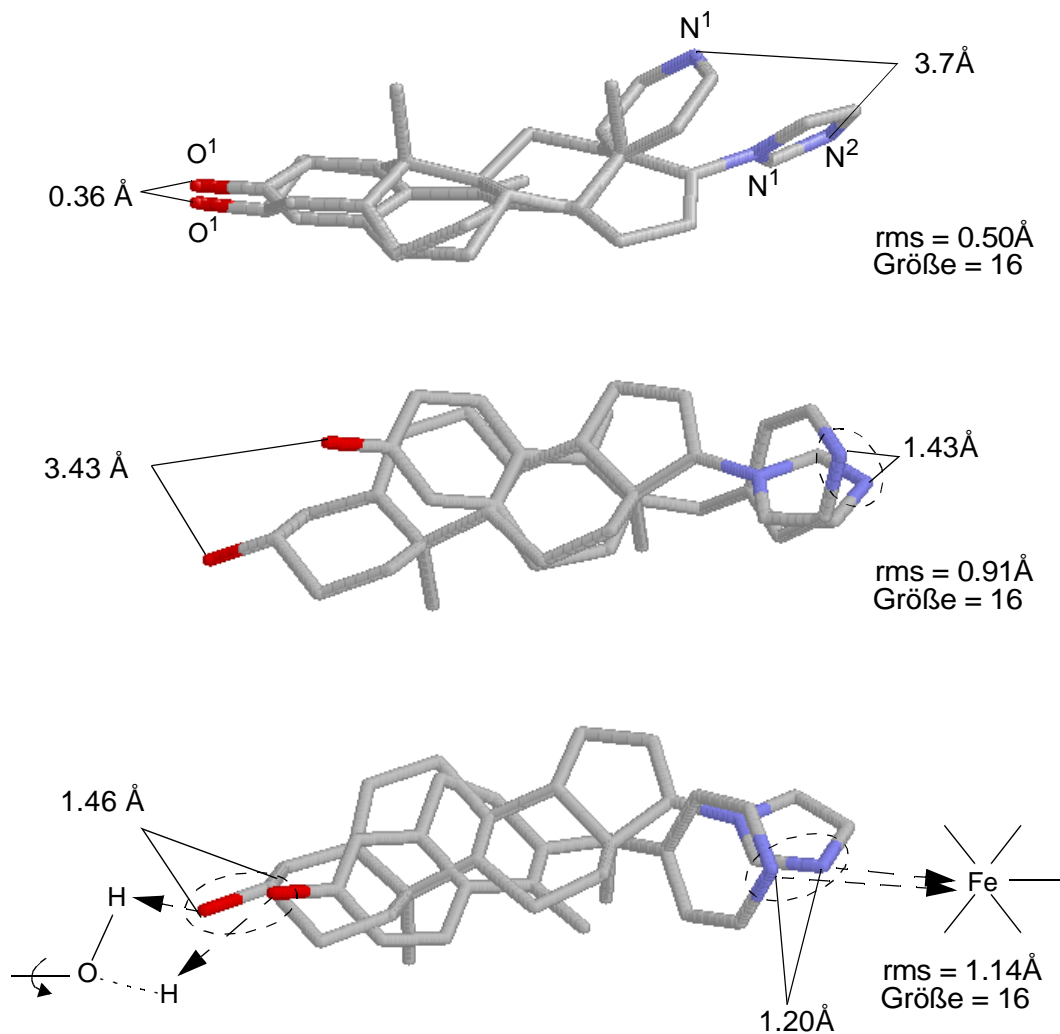
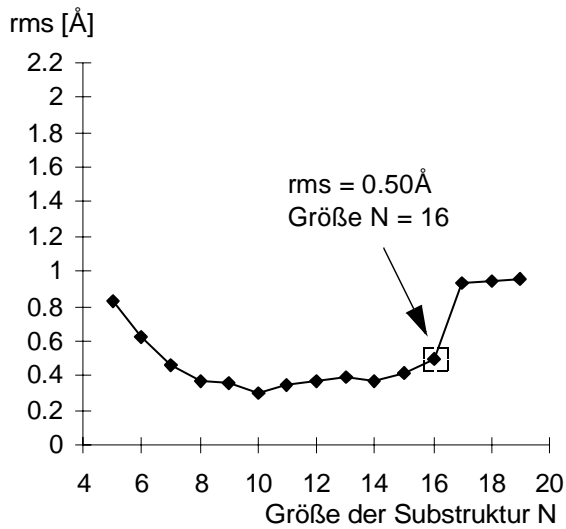


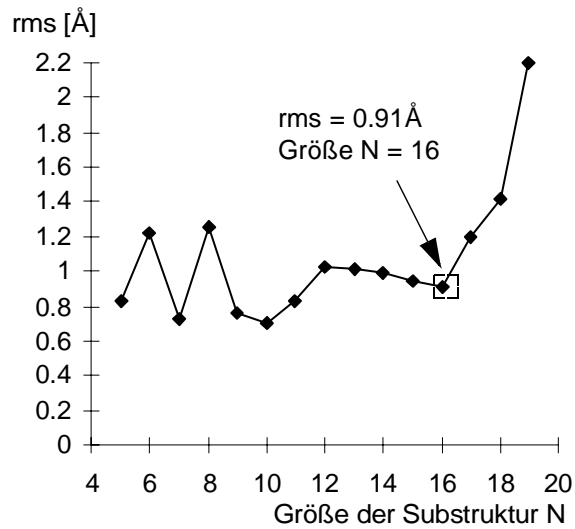
Abb. 3-13: Die Überlagerung von **15** (starr) mit **16** (flexibel). Die vorgegebenen Matchpaare sind durch gestrichelte Ovale gekennzeichnet. In der oberen Überlagerung sind keine Einschränkungen gegeben, in der mittleren ist das Matchpaar **15: N²/16: N¹** vorgeschrieben, in der unteren die Matchpaare **15: N²/16: N²** und **15: O¹/16: O¹**.

Die Pareto-Diagramme der besten GA-Läufe aller drei Überlagerungen sind in Abbildung 3-14 dargestellt.

Pareto-Diagramm der Überlagerung von Imidq (starr) und BW13 (flexibel):
keine Vorgaben



Pareto-Diagramm der Überlagerung von Imidq (starr) und BW13 (flexibel):
Vorgabe: **15**: N^2 /**16**: N^1



Pareto-Diagramm der Überlagerung von Imidq (starr) und BW13 (flexibel):
Vorgabe: **15**: N^2 /**16**: N^1 , **15**: O^1 /**16**: O^1

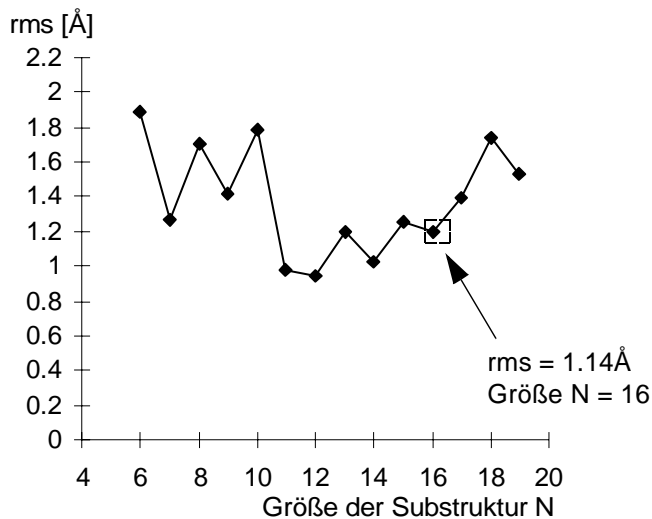


Abb. 3-14: Die besten Pareto-Kurven von jeweils 40 GA-Experimenten zur Überlagerung von Imidq, **15**, und BW13, **16**, unter Vorgabe verschiedener Matchpaare.

Zum Vergleich der Überlagerungen wurde eine Substrukturgröße von 16 herausgenommen. Diese Größe stellt jeweils die Substrukturgröße dar, die vor einem weiteren steilen Anstieg der Pareto-Kurve zu liegen kommt. Die Pareto-Kurven zeigen weiterhin deutlich die für alle Substrukturgrößen niedrigeren *rms*-Werte für den Fall, daß keine Restriktionen für die

zu erhaltenden Substrukturen angegeben werden.

Die Vorgabe bedingter Atome für die zu erhaltende Substruktur eines Datensatzes von Verbindungen, erlaubt die Beurteilung der Anpassung der Strukturen unter Einbeziehung von bereits vorhandenem Wissen über pharmakophore Bereiche dieser Verbindungen. Jeweils das Stickstoff- und Sauerstoffatom in **15** und **16** koordinieren an analoge Stellen im Rezeptor und dürfen somit bei der Beurteilung der strukturellen Ähnlichkeit nicht weit voneinander zu liegen kommen. In der untersten Überlagerung von Abbildung 3-13 ist der jeweilige Abstand nur 1.46Å (N-N) und 1.20Å (O-O). Die beiden zusammengehörenden Atompaare weisen eindeutig in die gleiche Richtung. Daher ist eine Koordination über identische Positionen zum Rezeptor wahrscheinlich. Die untere Überlagerung zeigt zwar eine schlechtere atomare Anpassung der Gesamtstruktur, trotzdem liegt der hydrophobe Bereich des Steroidkohlenstoffgerüsts insgesamt gut übereinander. Auf kurze Abstände der Stickstoffatome ist dabei mehr Wert zu legen, als auf die Abstände der Sauerstoffatome. Die Stickstoffatome koordinieren an das fixierte Eisenatom im Enzym. Der Position der Sauerstoffatome am Ligand kann der Rezeptor durch geringe konformative Änderungen, derjenigen Teile des Rezeptors, die zu diesen beiden Gruppen H-Brückenbindungen ausbilden, entgegenkommen (Abbildung 3-13).

Die Überlagerung von Imidq, **15**, (A/B/C/D-Ring) mit BW112, **17**, (A/C-Ring), die sich nicht stark in der Größe der Kohlenstoffgerüsts unterscheiden, ist in Abbildung 3-15 dargestellt. **17** verfügt über zwei Hydroxygruppen (3'-OH und 4'-OH) am A-Ring, im Gegensatz zu der A/B-Ringverbindung **16**, das nur eine Hydroxygruppe in 6-Position aufweist. Um Analogien zur Überlagerung von **15** und **16** zu ziehen, wurde zunächst nur die 3'-OH Gruppe von **17** berücksichtigt (Abbildung 3-15 oben), anschließend aber auch eine Überlagerung unter Berücksichtigung der 4'-OH Gruppe durchgeführt (Abbildung 3-15 unten).

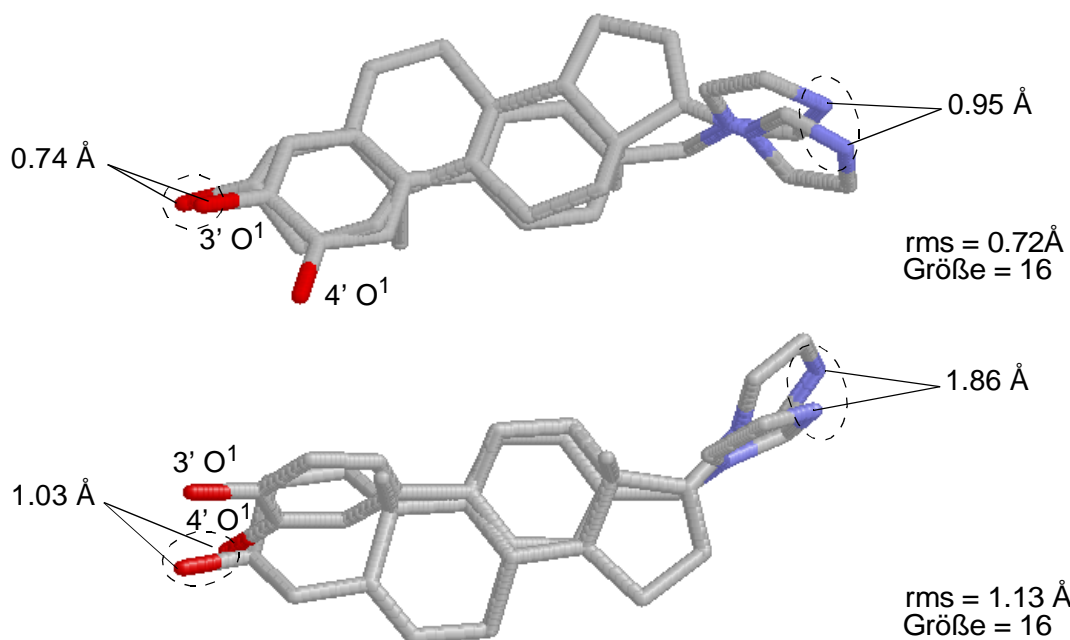


Abb. 3-15: Die Überlagerung von **15** (starr) mit **17** (flexibel). Die vorgegebenen Matchpaare sind durch gestrichelte Ovale gekennzeichnet. Oben: **15**: O¹/**17**: 3'O¹ und **15**: N²/**17**: N², unten: **15**: O¹/**17**: 4'O¹ und **15**: N²/**17**: N².

Deutlich zu erkennen ist die bessere konformationelle Anpassung des Kohlenstoffgerüsts von **17** und **15** im Vergleich zu **16** und **15** (Abbildung 3-13, unten), obwohl folgende Matchpaare vorgegeben waren **15**: O¹/**17**: 3'O¹ und **15**: N²/**17**: N². Die Distanz der einander zugeordneten Sauerstoffatome $d(\text{O-O})$ beträgt 0.74Å und die der Stickstoffatome $d(\text{N-N})$ 0.95Å. Somit liegen die wichtigen pharmakophoren Punkte deutlich näher aneinander als in der analogen Überlagerung von **16** mit **15** (Abbildung 3-13, $d(\text{O-O})$: 1.46Å und $d(\text{N-N})$: 1.20Å). Vergleicht man die Überlagerung unter unterschiedlicher Angabe der Matchpaare **15**: O¹/**17**: 3'O¹ und **15**: O¹/**17**: 4'O¹, so ist ersichtlich, daß die Angabe von **15**: O¹/**17**: 3'O¹ zu besserer konformationeller Anpassung führt ($rms = 0.72\text{Å}$ gegenüber 1.13Å). Auch die Abstände der Sauerstoffatome ($d(\text{O-O})$: 0.74Å bzw. 1.03Å) und der Stickstoffatome ($d(\text{N-N})$: 0.95Å bzw. 1.86Å) sind bei Angabe des 3'-Sauerstoffatoms von **17** kürzer. Weiterhin dreht sich die Ebene des A-Biphenylrings von **17** aus der Ebene des A-Rings des starren Steroidgerüsts von **15** heraus, so daß sie senkrecht aufeinander zu liegen kommen.

Um weitere A/B-Ringsysteme, d.h. 2,3-Dihydronaphthaline zu untersuchen, wurde das starre Steroidgerüst von **15** zusätzlich mit den beiden 7-, bzw. 5-OMe substituierten Verbindungen **18** und **19** überlagert (Abbildung 3-16).

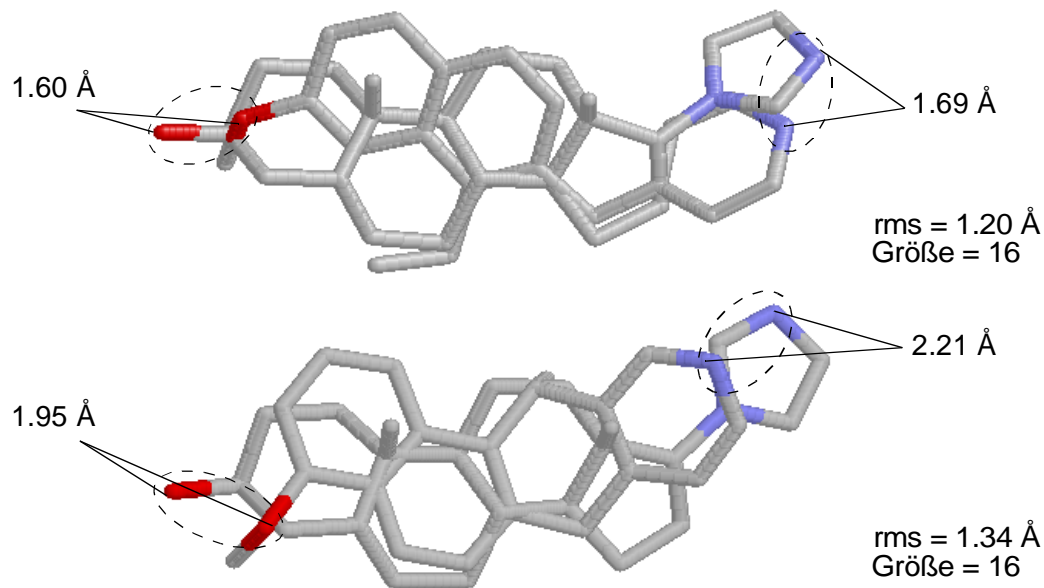


Abb. 3-16: Die Überlagerung von **15** (starr) mit **18** (flexibel, oben) bzw. **19** (flexibel, unten). Die vorgegebenen Matchpaare sind durch gestrichelte Ovale gekennzeichnet. Oben: die Matchpaare **15**: N¹/**18**: N² und **15**: O¹/**18**: O¹, unten: die Matchpaare **15**: N¹/**19**: N² und **15**: O¹/**19**: O¹.

Als Ergebnis der beiden Überlagerungen von **18** und **19** mit **15** ($rms = 1.2\text{\AA}$ bzw. 1.34\AA) ergibt sich eine schlechtere geometrische Anpassung im Vergleich zur Überlagerung mit dem entsprechenden 6-OH substituierten Dihydronaphthalin, **16** (Abbildung 3-13, $rms = 1.14\text{\AA}$). Auch die Abstände der zugeordneten Stickstoffpaare (**18**: 1.69\AA , **19**: 2.21\AA) und Sauerstoffpaare (**18**: 1.60\AA , **19**: 1.95\AA) sind relativ hoch.

Weiterhin (Abbildung 3-17) werden die Überlagerungen von BW112, **17**, (A/C-Ring System) mit BW13, **16**, (A/B-Ring System) analysiert. Diese unterscheiden sich wiederum stärker in der Größe ihrer Kohlenstoffgerüste.

Die erste Überlagerung (Abbildung 3-17, oben) setzt als konstante Matchpaare **16**: N¹/**17**: N² und **16**: O¹/**17**: 3'O¹ voraus, während bei der zweiten Überlagerung (Abbildung 3-17, unten) **16**: N¹/**17**: N² und **16**: O¹/**17**: 4'O¹ gegeben ist.

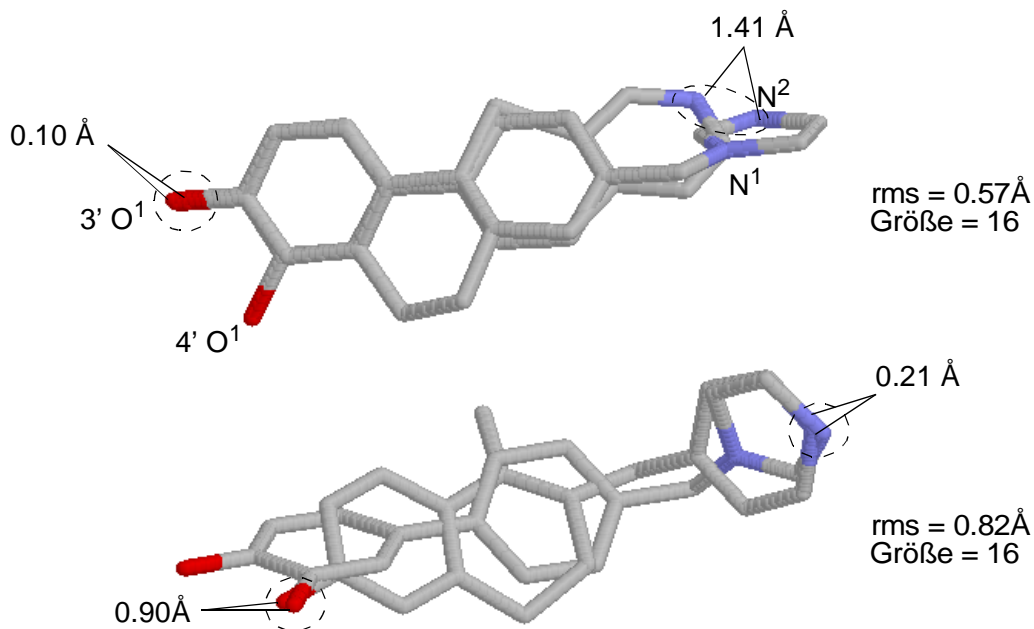


Abb. 3-17: Überlagerung von **16** (flexibel) mit **17** (flexibel). Die vorgegebenen Matchpaare sind durch gestrichelte Ovale gekennzeichnet. oben: **16**: N¹/**17**: N² und **16**: O¹/**17**: 3'O¹, unten: **16**: N¹/**17**: N² und **16**: O¹/**17**: 4'O¹.

Die erste Variante (Abbildung 3-17 oben) ergibt durch die partielle Überlagerung des Biphenyl-C-Rings von BW112, **17**, durch die 1-Methyl und 2-Brücken-Methylengruppe des 3,4-Dihydronaphthalin, **16**, zwar eine sehr gute Übereinstimmung ($rms = 0.57\text{Å}$) der Grundgerüste und der Sauerstoffatome ($d(\text{O-O})$: $0,10\text{ Å}$), liefert aber einen großen Abstand der Stickstoffatome ($d(\text{N-N})$: $1,41\text{ Å}$). Eine bessere Deckung der Stickstoffatome ($d(\text{N-N})$: $0,21\text{Å}$, $d(\text{O-O})$: $0,9\text{ Å}$) in der zweiten Variante (Abbildung 3-17 unten) ist jedoch mit einer starken Verschiebung der Orientierung ($rms = 0.82\text{Å}$) der Gerüste beider Strukturen verbunden. Diese muß jedoch kein Ausschlußgrund für einen ähnlichen Anlagerungsmodus darstellen, da der hydrophobe Bereich beider Verbindungen dennoch übereinander zu liegen kommt.

Die Überlagerung der an 7- bzw. 5-Position substituierten 3,4-Dihydronaphthalin A/B-Ringsysteme **18** und **19** mit dem Biphenylsystem **17** sind in Abbildung 3-18 und 3-19 dargestellt.

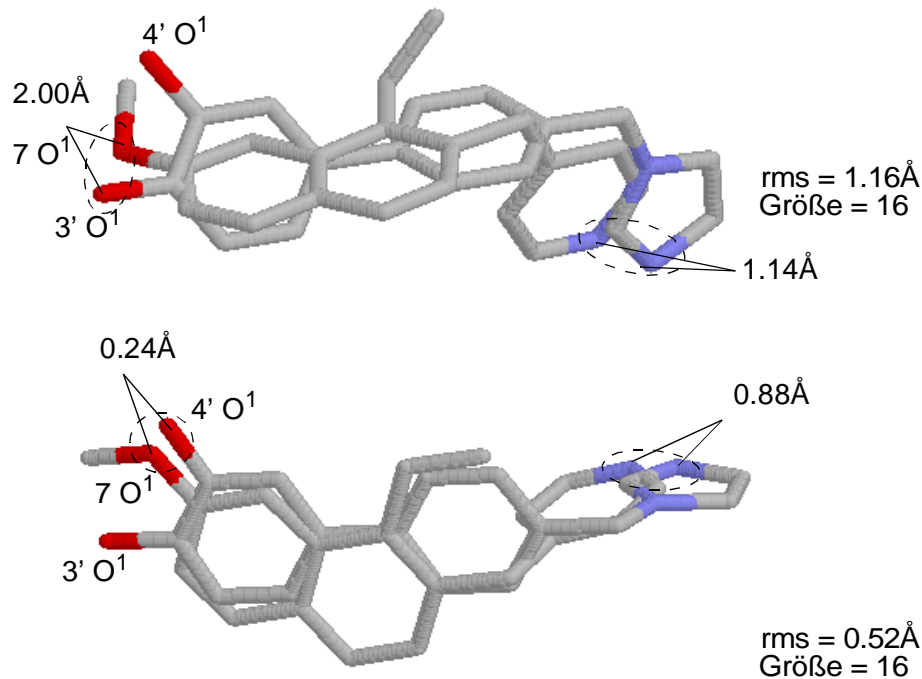


Abb. 3-18: Überlagerung von **18** (flexibel) mit **17** (flexibel). Die vorgegebenen Matchpaare sind durch gestrichelte Ovale gekennzeichnet. Oben: **18**: N¹/17: N² und **18**: O¹/17: 3'O¹, unten: **18**: N¹/17: N² und **18**: O¹/17: 4'O¹.

Verbindung **18** wurde mit **17** zweifach überlagert, wobei das Matchpaar **18**: N¹/17: N² in beiden Überlagerungen vorgegeben wurde, und die Unterschiede in der Angabe von **18**: O¹/17: 3'O¹ (Abbildung 3-18 oben) beziehungsweise **18**: O¹/17: 4'O¹ (Abbildung 3-18 unten) lagen. Die strukturelle Übereinstimmung ist bei Angabe des Matchpaars **18**: O¹/17: 4'O¹ besser als bei Angabe von **18**: O¹/17: 3'O¹ ($rms = 0.52\text{Å}$ und $rms = 1.16\text{Å}$) mit zusätzlich kürzeren Abständen der zugeordneten Atompaaire ($d(\text{N-N})$: 0.88Å und $d(\text{O-O})$: 0.24Å gegenüber $d(\text{N-N})$: 1.14Å und $d(\text{O-O})$: 2.0Å).

Im Gegensatz dazu war für die Überlagerung von **16** und **17** die strukturelle Übereinstimmung bei Angabe des **17**: 3'O¹-Sauerstoffatoms höher (Abbildung 3-17 oben, $rms = 0.57\text{Å}$).

Die Überlagerungen von **19** mit **17** durch Angabe der Matchpaare **19**: N¹/17: N² und **19**: O¹/17: 3'O¹ (Abbildung 3-19 oben) beziehungsweise **19**: N¹/17: N² und **19**: O¹/17: 4'O¹ (Abbildung 3-19 unten) ergibt analoge Substrukturen wie in der Überlagerung von **18** mit **17**.

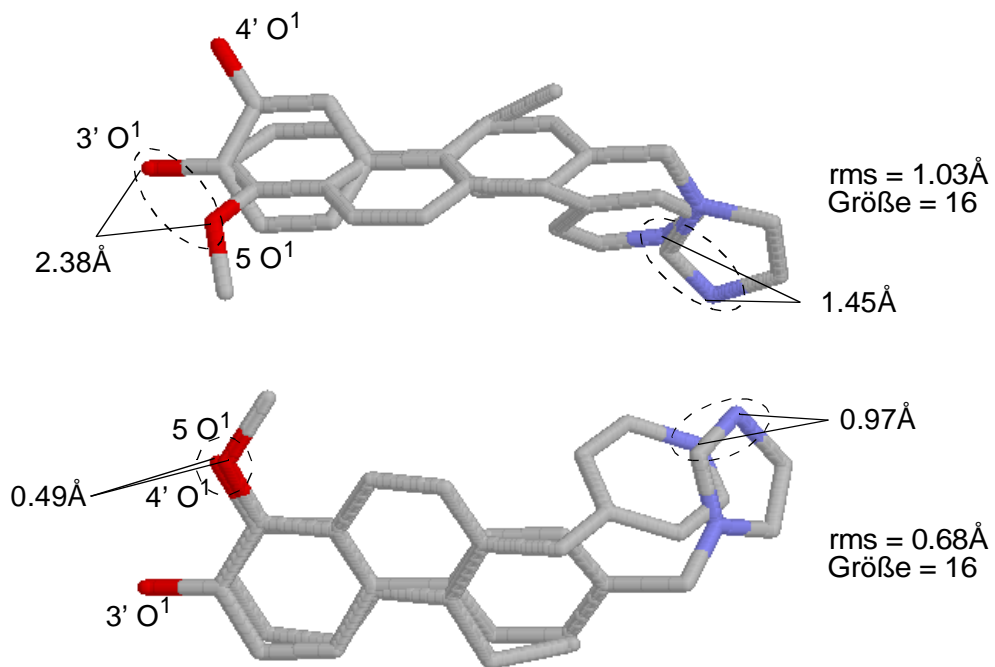


Abb. 3-19: Überlagerung von **19** (flexibel) mit **17** (flexibel). Die vorgegebenen Matchpaare sind durch gestrichelte Ovale gekennzeichnet. Oben: **19**: N¹/**17**: N² und **19**: O¹/**17**: 3'O¹, unten: **19**: N¹/**17**: N² und **19**: O¹/**17**: 4'O¹.

Die geometrische Anpassung der Überlagerung mit den vorgegebenen Matchpaaren **19**: O¹/**17**: 3'O¹ ist erheblich schlechter ($rms = 1.03\text{Å}$) als die geometrische Anpassung der beiden Verbindungen unter Angabe des Matchpaares **19**: O¹/**17**: 4'O ($rms = 0.68\text{Å}$). Wiederum sind auch die einzelnen Abstände der einander von vorneherein zugeordneten Atome bei Angabe des **17**: 4'O kleiner ($d(\text{O-O})$: 2.38Å und 0.49Å , $d(\text{N-N})$: 1.03Å und 0.68Å).

In Tabelle 3-6 sind zusammenfassend die Ergebnisse der paarweisen Überlagerungen von **15**, **16**, **17**, **18** und **19** angegeben. Zur Beurteilung der Güte der Überlagerung werden der rms -Wert bei konstanter Substrukturgröße (16 Atome) und die jeweiligen Abstände der zugeordneten Stickstoff- und Sauerstoffatome in Betracht gezogen.

Tab. 3-6: Die *rms*-Werte der berechneten Überlagerungen einiger PC450c17-Inhibitoren und die Abstände der N-N- und O-O-Paare.

Überlagerung	konstante Matchpaare	rms/[Å]	d(N-N)/[Å]	d(O-O)/[Å]
15 mit 16	N ² -N ¹ , O ¹ -O ¹	1.14	1.20	1.46
15 mit 17	N ² -N ² , O ¹ -3'O ¹	0.72	0.95	0.74
15 mit 17	N ² -N ² , O ¹ -4'O ¹	1.13	1.86	1.03
15 mit 18	N ² -N ¹ , O ¹ -O ¹	1.20	1.60	1.69
15 mit 19	N ² -N ¹ , O ¹ -O ¹	1.34	2.21	1.95
16 mit 17	N ¹ -N ² , O ¹ -3'O ¹	0.57	1.41	0.10
16 mit 17	N ¹ -N ² , O ¹ -4'O ¹	0.82	0.21	0.90
18 mit 17	N ¹ -N ² , O ¹ -3'O ¹	1.16	1.14	2.00
18 mit 17	N ¹ -N ² , O ¹ -4'O ¹	0.52	0.88	0.24
19 mit 17	N ¹ -N ² , O ¹ -3'O ¹	1.03	1.54	2.38
19 mit 17	N ¹ -N ² , O ¹ -4'O ¹	0.68	0.97	0.49

Die Überlagerung von **15**, dem steroidalen Templat, mit der 6-substituierten 3,4-Dihydronaphthalin-Verbindung, **16**, ($rms = 0.72\text{Å}$) ist schlechter als die Überlagerung von **15** mit dem Biphenylsystem, **17**, unter Angabe des 3'-Sauerstoffatoms von **17** ($rms = 0.57\text{Å}$), allerdings besser als die Überlagerung von **15** mit **17** bei Angabe des 4'-Sauerstoffatoms ($rms = 1.13\text{Å}$). Im Gegensatz dazu ist bei der Überlagerung von **17** mit den beiden 7- und 5-substituierten 3,4-Dihydronaphthalin-Verbindungen (A/B-Ringsysteme) **18** und **19** eine deutlich bessere geometrische Anpassung ($rms = 0.52\text{Å}$ und 0.68Å) bei Angabe des 4'-Sauerstoffatoms vorzufinden. Die Überlagerungen von **15** mit den beiden 7- bzw. 5-substituierten 3,4-Dihydronaphthalinverbindungen, **18** bzw. **19**, liefern jeweils ähnlich schlechte geometrische Anpassung, wie das 6-substituierte 3,4-Dihydronaphthalin, **16**, bei der Überlagerung mit **18** bzw. **19**. Diese Ergebnisse drücken sich nicht nur in den *rms*-Werten als Güte aus, sondern zusätzlich in den einzelnen Abständen der von vorneherein einander zugeordneten Atompaare. Die Auswertung der Abstände der pharmakophoren Punkte in den Überlagerungen unterstreichen daher die Ergebnisse der Auswertung der *rms*-Werte.

Zusammenfassend zeigt die Analyse der Überlagerung der drei aktivsten Verbindungen **15** (A/B/C/D-Ringsystem), **16** (A/B-Ringsystem) und **17** (A/C-Ringsystem), daß für die

Gruppe der Biphenyl-Verbindungen (A/C-Ringsysteme) von einer eher steroidal Anlage unter Koordination des 3'-Wasserstoffbrückenakzeptors ausgegangen werden kann. Die Überlagerung des Steroids **15** mit der Biphenylverbindung **17** liefert besser Ergebnisse als mit allen drei Dihydronaphthalinverbindungen **16**, **17** und **18**. Die geometrische Anpassung der weniger aktiven Verbindungen **18** und **19** an das starre Gerüst des hochaktiven Steroids **15** ist dabei schlechter als für die höher aktive Verbindung **16**. Für die am wenigsten aktive Verbindung **19** ist sie noch schlechter als für Verbindung **18**.

Auf die Koordination des 3'-Wasserstoffbrückenakzeptors in **17** weisen die Überlagerungen mit **15** und **16** hin. Die 5- und 7-substituierten Dihydronaphthalinverbindungen **17** und **18**, die eine schlechtere Affinität zum Rezeptor aufweisen als die 6-substituierte Verbindung **16**, liefern eine besser Überlagerung mit **17** bei Angabe der 4'-OH Gruppe von **17**. Da womöglich aber die 3'-OH Gruppe von **17** für die Koordination im Rezeptor verantwortlich ist und die geometrische Anpassung der 5- und 7- substituierten Dihydronaphthalinverbindungen unter diesen Voraussetzungen nicht gut möglich ist, könnte das auch einen Grund für deren schlechtere Affinität darstellen.

Die hier mit GAMMA berechneten Überlagerung unter Einbezug von spezifischen Wissen über die Rezeptorbindungstasche haben gezeigt, daß Hinweise über Anlagerungsmodi auch von konformativ flexiblen Verbindungen erhalten werden können. Aufgrund von Überlagerungen von Verbindungen ohne bekanntem biologischem Wirkprofil mit hier gezeigten aktiven Verbindungen, sollten Aufschlüsse über deren mögliche Affinität zu dem P450c17-Enzym erhalten werden. Drei verschiedene Größen, die aus den Überlagerungen erhalten werden, lassen sich dabei analysieren, erstens die *rms*-Werte der Überlagerungen, zweitens die Abstände der an das Eisenatom koordinierenden funktionellen Gruppen in den Liganden und drittens die Abstände der an den Wasserstoffbrückendonator des Enzyms koordinierende funktionellen Gruppen.

3.2.2 Multiple Überlagerungen

Eine simultane Überlagerung der vier hoch aktiven PC450c17-Inhibitoren **15**, **16**, **17** und **18** ist in Abbildung 3-20 dargestellt. Das Überlagerungskriterium war die Ordnungszahl. Als feste Matchpaare wurden **15**: O¹/**16**: O¹/**17**: 3'O¹/**18**: O¹ und **15**: N²/**16**: N¹/**17**: N²/**18**: N¹ vorgegeben, somit war das 3'-Sauerstoffatom von BW112 **17** an der Überlagerung beteiligt.

Imidq (starr), BW112 (flexibel),
 BW13 (flexibel) und BW61 (flexibel)
 3'-O¹ von **17** vorgegeben

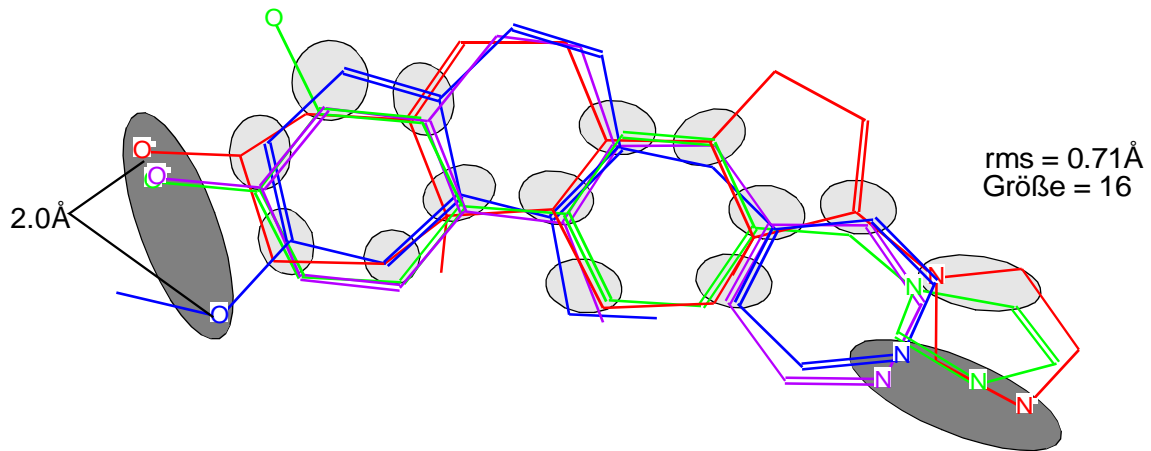


Abb. 3-20: Das Ergebnis der gleichzeitigen Überlagerung von Imidq, **15** (rot), BW112, **17** (grün), BW13, **16** (violett) und BW61, **18** (blau). Die Ovale markieren die Substruktur. Die vorgegebenen Matchpaare **15**: O¹/**16**: O¹/**17**: 3'O¹/**18**: O¹ und **15**: N²/**17**: N²/**16**: N¹/**18**: N¹ sind dunkel unterlegt.

In einer weiteren Überlagerung wurden als feste Matchpaare **15**: O¹/**16**: O¹/**17**: 4'O¹/**18**: O¹ und **15**: N²/**17**: N²/**16**: N¹/**18**: N¹ gesetzt (Abbildung 3-21), somit das 4'-Sauerstoffatom von BW112 **17**.

Imidq (starr), BW112 (flexibel),
 BW13 (flexibel) und BW61 (flexibel)
 4'-O¹ von **17** vorgegeben

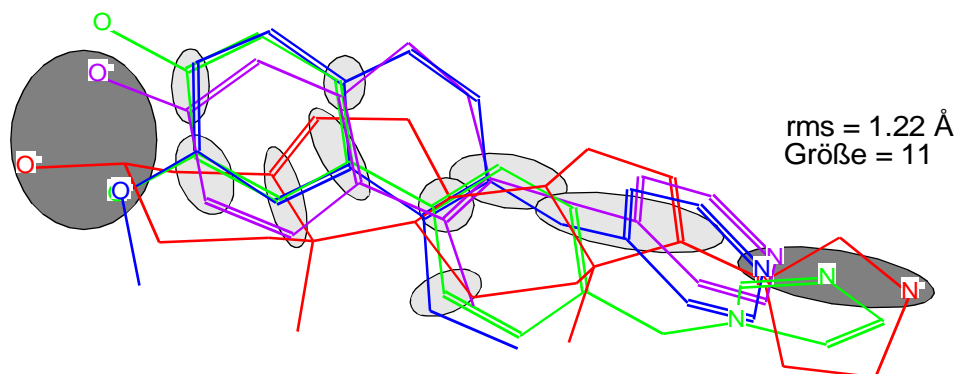


Abb. 3-21: Das Ergebnis der gleichzeitigen Überlagerung von Imidq, **15** (rot), BW112, **17** (grün), BW13, **16** (violett) und BW61, **18** (blau). Die Ovale markieren die Substruktur. Die vorgegebenen Matchpaare **15**: O¹/**16**: O¹/**17**: 4'O¹/**18**: O¹ und **15**: N²/**17**: N²/**16**: N¹/**18**: N¹ sind dunkel unterlegt.

Durch einen Vergleich von Abbildung 3-20 und Abbildung 3-21 ist deutlich zu erkennen, daß eine bessere Überlagerung erreicht wird, wenn das 3'-Sauerstoffatom von BW112, **17**, ($rms = 0.71\text{\AA}$ bzw. 1.22\AA) angegeben wird. Die Distanz des Sauerstoffatoms in der OMe-Gruppe von **18** beträgt zu den gut überlagerten Sauerstoffatomen der Verbindungen **15**, **16** und **17** ungefähr 2.0\AA bei Angabe des 3'-Sauerstoffatoms von **17**. (Abbildung 3-20). Die in der paarweisen Überlagerung besser auf das 17-4'-Sauerstoffatom angepaßte Verbindung **18** paßt sich auch hier am schlechtesten in die multiple Überlagerung ein. Es wurde wiederum die Substrukturgröße 16 herausgenommen, obwohl im Pareto-Diagramm in Abbildung 3-22 (oben) kaum ein Anstieg des rms -Wertes für höhere Substrukturgrößen zu verzeichnen ist. Bei Angabe des 4'-Sauerstoffatoms von BW112, **17**, paßt sich vor allem die Verbindung **18** gut an die Biphenylverbindung an, die gesamte Überlagerung ist jedoch bei einer Substrukturgröße von 11 Atomen mit einer schlechten geometrischen Güte von $rms = 1.22\text{\AA}$ verbunden. Aus dem Pareto-Diagramm in Abbildung 3-22 (unten) geht weiterhin hervor, daß größere Substrukturen als 11 Atome nicht ermittelt werden konnten. Für Substrukturen größer als 11 Atome wurde somit kein Minimum bei der Optimierung der gleichzeitigen Überlagerung der vier Verbindungen gefunden.

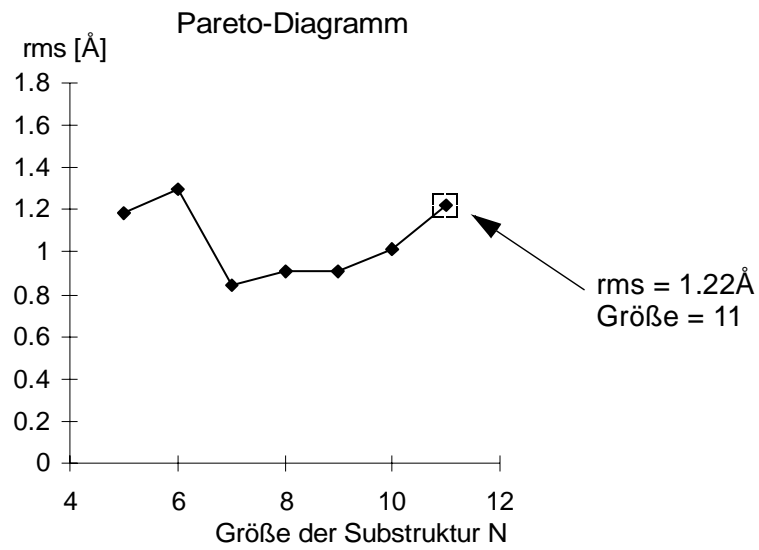
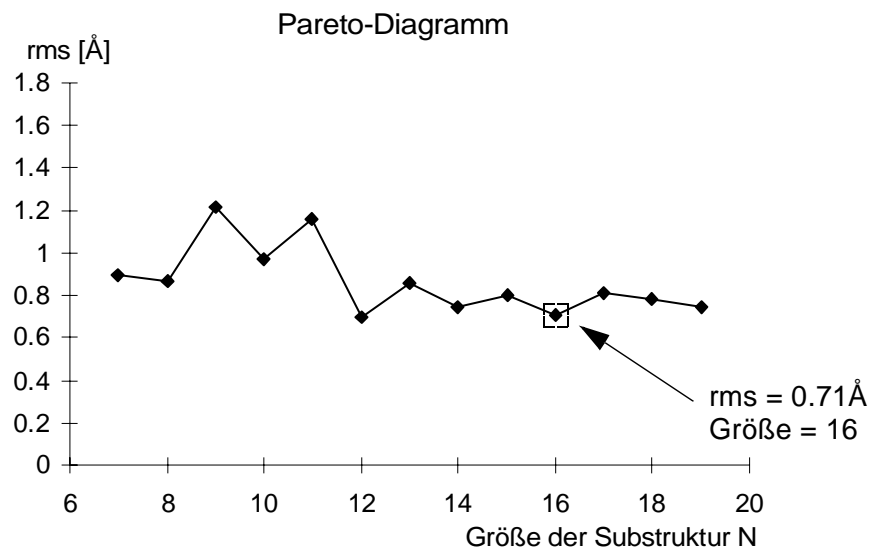


Abb. 3-22: Gemittelte Pareto-Diagramme der Überlagerung von Imidq, 15, BW112, 17, BW13, 16 und BW61, 18. Oben: Angabe des 3'O von BW112, 17, unten: Angabe des 4'O von BW112, 17.

4 Untersuchung der Affinität von Steroiden zu Antikörpern und künstlichen Rezeptoren

4.1 Steroide, Immunoglobuline und künstliche Rezeptoren

Im folgenden sollen die Unterschiede der Bindungen von Steroiden zu Immunoglobulinrezeptoren (Antikörper) und zu künstlichen Rezeptoren, wie z.B. Cyclophanen (Abbildung 1-7) untersucht werden. Dafür werden die physikochemischen Eigenschaften der Steroid-Ligand-Komplexe sowie die Steroidliganden selbst mit Hilfe von rechnergestützten Methoden verglichen. Das Ziel ist die Affinität der an Immunoglobulinrezeptoren aktiven Verbindungen zu künstlichen Rezeptoren zu untersuchen um Gemeinsamkeiten oder Unterschiede in den Bindungseigenschaften herauszustellen.

Immunoglobuline bestehen generell aus zwei Arten von Peptidketten, einer leichten (engl.: light, L) und einer schweren (engl.: heavy, H) Kette (Abbildung 4-1). Jede L-Kette ist über eine Disulfidbrücke an eine H-Kette gebunden. Die H-Ketten selbst sind untereinander über mindestens eine Disulfidbrücke gekoppelt. Jede H-Kette besteht dabei aus einer variablen Region (V_H) und drei konstanten Regionen (C_{H1} , C_{H2} und C_{H3}), jede L-Kette aus einer variablen (V_L) und einer konstanten (C_L) Region. In den variablen Regionen sind bedeutende Unterschiede in der Aminosäuresequenz der verschiedensten Antikörpern zu finden. Diese beschränken sich grundsätzlich auf den aminoterminalen Abschnitt des Polypeptids. Die Konstanten Regionen dagegen sind in den Aminosäuresequenzen der meisten Antikörper zu finden. Die variablen Regionen der H- und L-Ketten (V_H und V_L) sind grundsätzlich von gleicher Länge, die konstanten Regionen der H-Ketten (C_{H1} , C_{H2} und C_{H3}), sind zusammen dagegen dreimal so lang, wie die der L-Ketten. Drei Segmente der L- und H-Kette weisen eine weit größere Variabilität auf, als andere Reste in diesen variablen Regionen. Diese hypervariablen Regionen bilden die Antigenbindungsstellen, d.h. die Spezifität der Antikörper wird durch die Eigenschaften dieser Aminosäuren bestimmt.^[72]

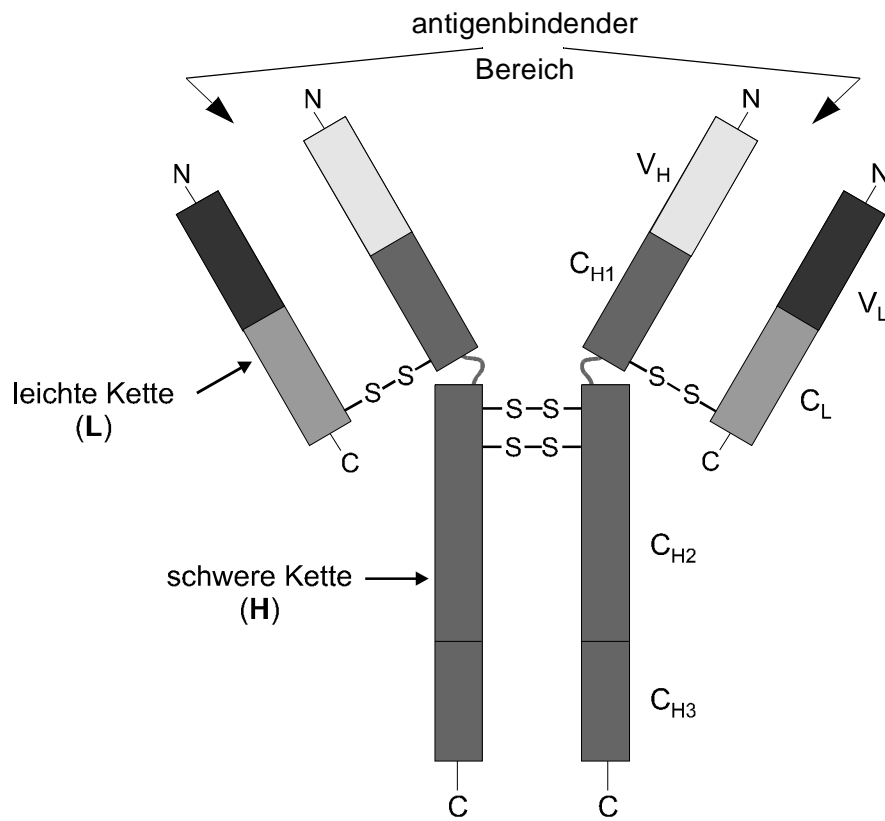


Abb. 4-1: Der allgemeine strukturelle Aufbau eines Immunglobulins.

Anhand der Kristallstruktur des Anti-Digoxin-26-10-Murin monoklonalen Antikörpers (26-10-Fab, antigenbindendes Fragment, Brookhaven Protein Database: 1IGJ)^[20], können die leichten und schweren variablen Regionen der Immunglobuline verdeutlicht werden (Abbildung 4-2).

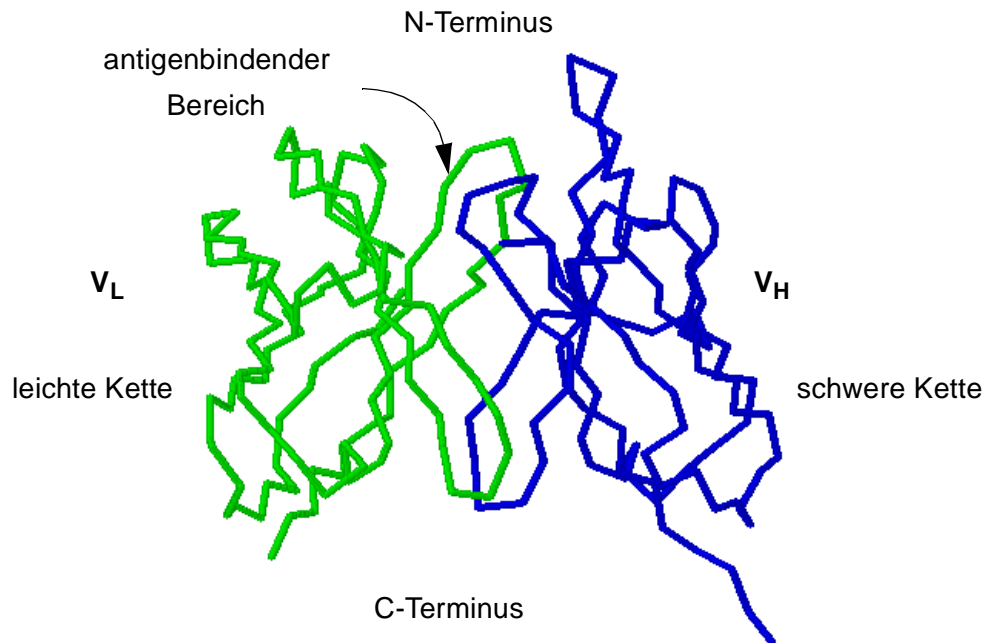


Abb. 4-2: Der antigenbindende Bereich des 26-10-Fab mit den leichten (grau) und schweren (schwarz) hypervariablen Regionen V_H und V_L , dazwischen liegt der antigenbindende Bereich.

Als künstlicher Rezeptor wurde das von Diederich et al. bereits systematisch untersuchte Paracyclophan, **20**,^[19] herausgenommen. In einer Reihe von Studien^{[22],[23][28]-[30]} wurden Cyclophanrezeptoren auf ihre Fähigkeit, Steroide zu komplexieren, untersucht. Das Ergebnis dabei war, daß vor allem Cyclophanrezeptoren, die Naphthylphenylmethyleinheiten aufweisen, besonders gut zur Komplexierung von Steroiden in wässrigem Medium geeignet sind.

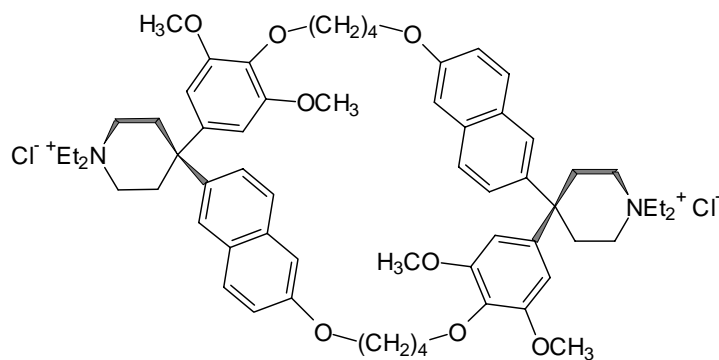


Abb. 4-3: Die Struktur des Paracyclophans, **20**.

In den folgenden Kapitel wurden vorerst die Komplexe des Anti-Digoxin-26-10-Murin monoklonalen Antikörpers^{[20],[21]} mit verschiedensten zu Digoxin, **2**, verwandten Strukturen,^[24] wie Digitoxin, **21**, Digoxigenin, **22**, und Ouabain, **23**, (Abbildung 4-4) untersucht. Dar-

auffolgend wurden die Bindungsaffinitäten der Gallensäuren Cholsäure, **3**, Desoxycholsäure, **24**, Chenodesoxycholsäure, **25**, Ursodesoxycholsäure, **26**, und Lithocholsäure, **27**,^[22] (Abbildung 4-5) und anderer bekannter Steroide wie Hydrocortison, **28**, Cortison, **29**, Testosteron, **30** und Progesteron, **31** (Abbildung 4-6) zu dem oben erwähnten Paracyclophan, **20** analysiert. Um einen Vergleich zwischen den Bindungsarten zu ermöglichen, wurden Steroide mit hoher Affinität zu dem 26-10-Fab, wie z.B. Digoxin, **2**, und Ouabain, **23** ebenfalls in die Bindungstasche des künstlichen Cyclophanrezeptors, **20**, eingelagert.

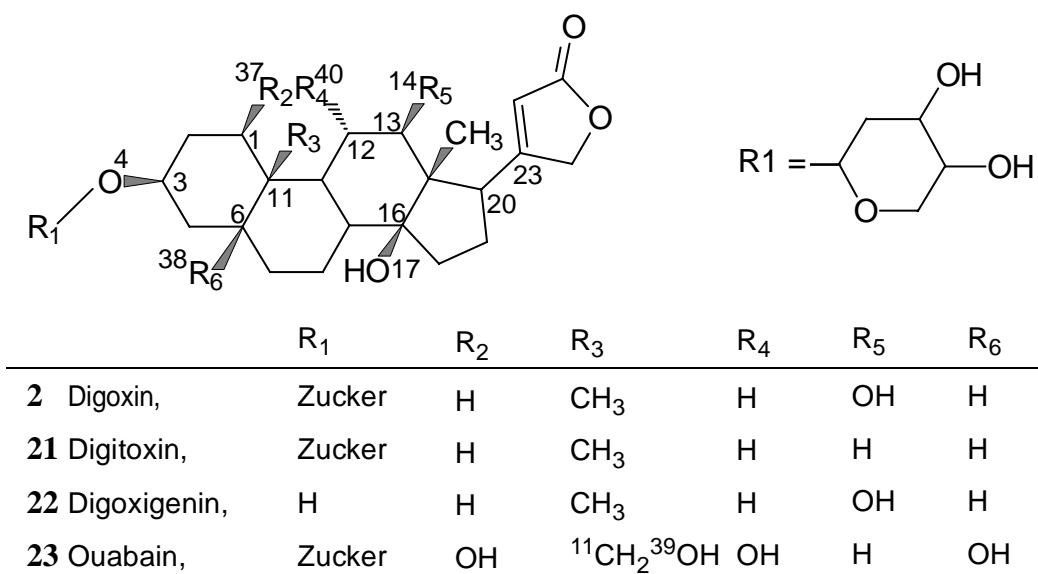
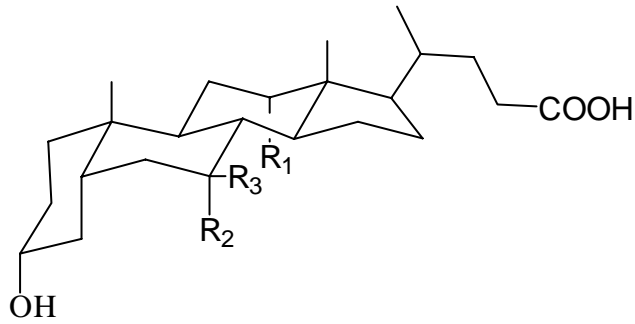
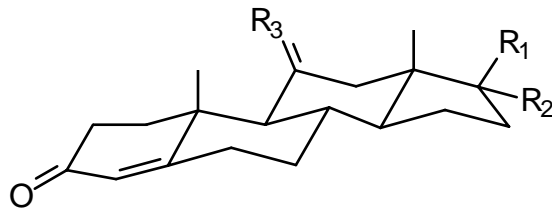


Abb. 4-4: Die zu Digoxin, **2**, verwandten Strukturen von Digitoxin, **21**, Digoxigenin, **22** und Ouabain, **23**.



	R ₁	R ₂	R ₃
3 Cholsäure	OH	OH	H
24 Desoxycholsäure	OH	H	H
25 Chenodesoxycholsäure	H	OH	H
26 Ursodesoxycholsäure	H	H	OH
27 Lithocholsäure	H	H	OH

Abb. 4-5: Die Strukturen von Cholsäure, **3**, Desoxycholsäure, **24**, Chenodesoxycholsäure, **25**, Ursodesoxycholsäure, **26**, und Lithocholsäure, **27**.



	R ₁	R ₂	R ₃
3 Hydrocortison	COCH ₂ OH	OH	αH, βOH
24 Cortison	COCH ₂ OH	OH	O
25 Testosteron	OH	H	H,H
26 Progesteron	COCH ₃	H	H,H

Abb. 4-6: Die Strukturen von Hydrocortison, **28**, Cortison, **29**, Testosteron, **30** und Progesteron, **31**.

4.2 Methodische Grundlagen

4.2.1 Vorhersage der Bindungsmodi der Liganden am Rezeptor

Die behandelten Steroidstrukturen **2** und **21** bis **31** wurden innerhalb des Molecular Modelling Programmpaketes MacroModel^[73] mit Hilfe eines MM3 Kraftfelds^[74] minimiert. Die Peptidsequenz des 26-10-Antikörpers wurde mit bekannten homologen Antikörperstrukturen überlagert um die Bereiche der schweren (H) und leichten (L) Seitenkette des Antikörpers zu ermitteln, die bei Antikörperpeptiden die Rezeptorbindungstaschen einschließen. Diese Sequenzüberlagerung wurde mit Hilfe des Immunoglobulin Domain Modelling Programm AbM^[75] durchgeführt.

Die Struktur des Paracyclophans, **20**, wurde mit Hilfe von kontinuierlichem *Simulated Annealing* basierend auf *Molecular Dynamics* (MD) mit einem MM3 Kraftfeld^[74] (siehe Anhang A.1.1) erhalten (siehe Anhang A.1.1 bis A.1.2 und A.1.4). Der lineare und kontinuierliche Abkühlungsprozeß lief von 300K bis 50K. Zuerst wurde für die Dauer von 10ps der Gleichgewichtszustand bei 300K eingestellt (Datensammlung). In den darauffolgenden 100ps wurde im Rahmen einer dynamischen *Annealing Simulation* langsam von 300K auf 50K abgekühlt (Abkühlungsprozeß).

Um sich innerhalb eines gemeinsamen Koordinatensystems zu bewegen, wurden die Steroidstrukturen mit Hilfe des Molecular Modelling Softwarepaketes InsightII/Discover^[76] manuell in die Rezeptorbindungstasche des 26-10-Fab bzw. des Paracyclophans, **20** gesetzt und die Koordinaten erneut abgespeichert. Alle Datenanalysen und Visualisierungen wurden ebenfalls innerhalb InsightII/Discover vorgenommen.

Docking Experimente um Anlagerungsmodi der Steroide **2** und **21** bis **23** innerhalb des 26-10-Fab und **2** und **23** bis **27** innerhalb des Paracyclophans, **20**, zu erhalten, wurden mit Hilfe des Programms AutoDock^[77] durchgeführt. AutoDock ist ein automatisches Dockingprogramm, das basierend auf einer *Monte Carlo Simulated Annealing* (siehe Anhang A.1.3 und A.1.4) Technik den Konformationsraum der Steroide innerhalb der Rezeptorbindungstasche des Antikörpers bzw. des Paracyclophans, **20**, absucht. Die Energieberechnungen ergeben sich dabei aus gitterbasierten molekularen Anziehungspotentialen. PM3 VESPA Ladungen,^[78] die innerhalb des Programmpaketes VAMP^[78] berechnet wurden, bilden die Basis für die Berechnung des elektrostatischen Gitters durch AutoDock. Innerhalb jedes AutoDock Laufs zur Berechnung eines Anlagerungsmodus wurden 100 Simulationen durchgeführt, wobei jede Simulation aus 150 Zyklen, mit einem Maximum von je 8,000 angenommenen und abgelehnten Schritten besteht. Die initiale Kühltemperatur war 500K und der Temperaturreduktionsfak-

tor betrug 0.85 pro Zyklus.

4.2.2 Berechnung der freien Bildungsenthalpie von Komplexen

Die freien Komplexbildungsenthalpien wurden mit dem modifizierten Amber4.1 Programmpaket berechnet.^[79] Im allgemeinen ist der beste Weg zur Berechnung der freien Bindungsenergien, die *Free Energy Perturbation* (siehe Anhang A.1.5).^{[80]-[82]} Die Rechenzeit, die für diese Simulation benötigt wird, ist jedoch aufgrund der aufeinanderfolgenden vielen kleinen Zustandsänderungen essentiell hoch. Åqvist et al. entwickelte daher eine Methode zur Abschätzung der linearen Wechselwirkungsenergien (engl.: *Linear Interaction Energy Approximation*) um das Problem der Berechnung der freien Bildungsenthalpien eines Komplexbildungsprozesses zu vereinfachen.^{[82][83]} Die freie Bildungsenthalpie wird, basierend auf einem thermodynamischen Kreislauf der Komplexbildung (siehe Anhang A.1.5), wie folgt ausgedrückt (Gl. 4-1):

$$\Delta G_{\text{Bildung}} = \frac{1}{2} (\langle V^{\text{elec}}_{\text{Antikörper}} \rangle - \langle V^{\text{elec}}_{\text{Wasser}} \rangle) + \alpha (\langle V^{\text{vdw}}_{\text{Antikörper}} \rangle - \langle V^{\text{vdw}}_{\text{Wasser}} \rangle) \quad (\text{Gl. 4-1})$$

mit: ΔV^{elec} = elektrostatisches Potential aus einer MD Simulation
 ΔV^{vdw} = van der Waals-Potential aus einer MD Simulation

Der Faktor $\frac{1}{2}$ basiert auf der *Linear Response Theorie*, während α ein empirischer Parameter ist, der für jedes spezifische Rezeptorsystem neu zu bestimmen ist. Die Methode wurde für eine Reihe von Inhibitoren des Endothiaspenin, das zur Familie der aspartischen Proteinasen gehört, getestet.^[39] Zwischen den berechneten und den experimentellen Werten von $\Delta G_{\text{Bildung}}$ wurden Differenzen von weniger als 1 kcal/mol ermittelt. Die Möglichkeit zur Simulation der freien Bildungsenthalpie durch die *Linear Interaction Energy Approximation* wurde hier innerhalb des Amber4.1 Programmpakets^[79] mit Hilfe des AMBER94 Kraftfeldes^[84] implementiert.

Die Atomladungen der Liganden und des Cyclophanrezeptors wurden durch Hartree-Fock Rechnungen des elektrostatischen Potentials mit einem 6-31G* Basissatz innerhalb des Gaussian94^[85] Modelling Programmpaketes berechnet.

Die Liganden und die Ligand-Rezeptor-Komplexe wurden in eine rechteckige Lösungsmittelbox mit 800 Wassermolekülen für den reinen Liganden und 1300 Wassermolekülen für die Ligand-Rezeptor-Komplexe, eingebettet. Die Berechnung der Lösungsmittelbox basierte auf Jørgensens TIP3P Annäherung (siehe Anhang A.1.6).^[86]

Nach anfänglicher Energieminimierung, wurden Molecular Dynamics Simulationen, bestehend aus 50ps Gleichgewichtszustandseinstellung und 80ps Datenaufnahme für jeden Liganden, vorerst nur gelöst in Wasser, dann gebunden im solvatisierten Komplex, durchgeführt. Die Simulationen, sowohl für den Liganden, als auch für den Ligand-Rezeptor-Komplex wurden innerhalb der TIP3P Lösungsmittelbox bei konstantem Druck von 1atm und konstanter Temperatur von 300K durchgeführt. Bei allen Simulationen wurden Wechselwirkungen zwischen nicht-bindenden Atomen nur bis zu einem interatomaren Abstand von 10Å berücksichtigt. Die nichtbindende Atompaarliste wurde jeweils nach 12 Schritten aktualisiert. Die Zeitspanne für jede Simulation betrug 2fs. Um alle Bindungslängen auf die Werte des Gleichgewichtszustands zu beschränken wurde die SHAKE Prozedur angewendet (siehe Anhang A.1.7).^{[87][88]}

4.2.3 Vergleich von physikochemischen Oberflächeneigenschaften der Steroide mit Hilfe neuronaler Netze

4.2.3.1 Neuronale Netze

Künstliche neuronale Netze sind informationsverarbeitende Systeme, die aus einer großen Anzahl einfacher Einheiten, den Zellen oder *Neuronen* bestehen, die in definierter Weise miteinander verbunden sind. Wird eine Neuron aktiviert, sendet es über gerichtete Verbindungen (engl.: links) Informationen zur Aktivierung weiterer Neuronen. Dabei liegt oft eine Schichtung der Neuronen vor. Wie die Neuronen entlang ihres Verbindungsnetzes miteinander kommunizieren, d.h. wie sie sich gegenseitig beeinflussen, wird mit Hilfe einer Propagierungsfunktion festgelegt.^[89]

Es existieren verschiedene Typen von neuronalen Netzen, wie z.B. das Kohonen-, das Hopfield-, das Backpropagation- oder Counterpropagation-Netz. Sie unterscheiden sich in ihrem Aufbau (Schichtung, Verknüpfung und Gestalt der Neuronen), sowie im Lernverfahren.

Ein wesentliches Element der neuronalen Netze ist ihre Lernfähigkeit von nichtlinearen Zusammenhängen. Lernfähigkeit bedeutet hier, nach der Präsentation bekannter Daten, die Klassifizierung dieser Daten und sinnvolle Reduktion komplexer Information vornehmen zu können, sowie durch Generalisierung und Abstraktion selbständig Schlüsse ziehen zu können.^{[38][90][91]}

Die *Lernregel* ist ein Algorithmus, gemäß dem das neuronale Netz lernt, aus einer vorgegebenen Eingabe (*Trainingsdatensatz*), die dem Netz mehrfach präsentiert wird, eine Ausgabe zu erzeugen.

Grundsätzlich werden zwei Lernverfahren unterschieden:^[38]

- *überwachtes* (engl.: supervised) Lernen: Dem Netz wird neben dem Eingabemuster der Trainingsdaten auch das zu erwartende Ausgabemuster bekanntgegeben, d.h. die bekannten Ausgabemuster werden beim Training verwendet. Das Netz sollte nach dem Training dazu in der Lage sein, die Generalisierung auf neue Eingabemuster durchzuführen, indem es ein zu diesem passendes Ausgabemuster ermittelt. Ein solches Lernverfahren stellt z.B. der Backpropagation-Algorithmus dar.^[38]
- *nicht überwachtes* (engl.: unsupervised) Lernen: Dem Netz werden ausschließlich Eingabemuster präsentiert. Das Ziel des Lernverfahrens ist es, anhand der Eingabemuster die Trainingsdaten in Klassen zu unterteilen, bzw. Ähnlichkeiten zwischen den Trainingsdaten zu erkennen. Das bekannteste Beispiel unüberwachten Lernens sind die selbstorganisierenden Karten von Kohonen.

Neuronale Netze sind vielseitig verwendbar. Sie werden zum Auffinden und Bilden von Klassen, zur Vorhersage und Simulation von Werten, sowie zur Reduktion der Komplexität von Eingabedaten eingesetzt. Im Bereich der Chemie lassen sich mit Hilfe neuronaler Netzen funktionelle Gruppen erkennen, Spektren simulieren oder physikochemische Oberflächeneigenschaften, wie z.B. elektrostatische Oberflächenpotentiale, Hydrophobizität oder Wasserstoffbrückenpotentiale von Molekülen auf zweidimensionale Flächen projizieren.^[38]

4.2.3.2 Die Struktur des Kohonen neuronalen Netzes

Teuvo Kohonen entwickelte das Konzept der selbstorganisierenden Karten (*SOM*, engl.: *self organizing feature maps*, Kohonen feature maps).^[14] Bei selbstorganisierenden Karten handelt es sich um einschichtige neuronale Netze mit einer Schicht aktiver Neuronen. Sie werden mit unüberwachten Lernverfahren trainiert.

Die besondere Eigenschaft dieser Netze ist die Fähigkeit, einen mehrdimensionalen Informationsraum in eine zweidimensionale Karte bzw. Neuronenschicht topologieerhaltend abzubilden. Die Fähigkeit des Kohonennetzes, Ähnlichkeiten zwischen Datenpunkten zu erkennen und Nachbarschaftsbeziehungen aufzustellen, ermöglicht weiterhin die Anwendung auf diverse Klassifizierungsprobleme. Punkte, die im mehrdimensionalen Raum benachbart sind, werden dabei in benachbarte Punkte der zwei-dimensionalen Karte projiziert.

Abbildung 4-7 zeigt die Architektur eines Kohonen neuronalen Netzes: Jede Säule repräsentiert eine Neuron, und jeder Würfel innerhalb einer Säule ein bestimmtes Gewicht eines Neurons.^[38] Jedes Neuron hat damit die gleiche Anzahl, m , an Gewichten, w_{ji} , wie die

Dimension (m) des Inputvektors vorgibt.

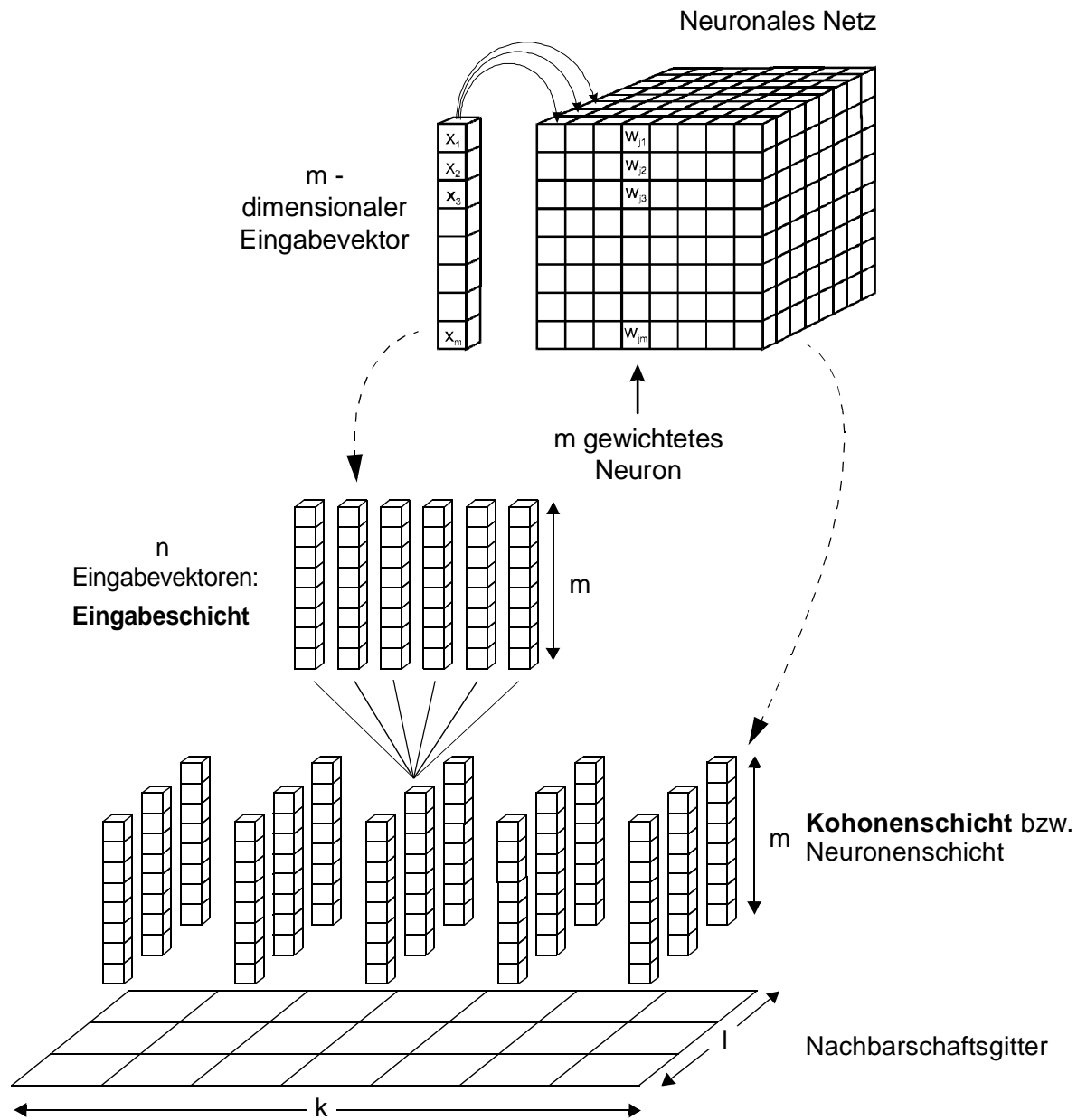


Abb. 4-7: Struktur eines Kohonennetzes. Die Eingabeschicht besteht aus beliebig vielen Eingabedatenpunkten der Länge m , die Kohonenschicht besteht aus $k \times l$ Neuronen der Länge m .

Ein Kohonennetz besteht aus einer Eingabeschicht aus Eingabedatenpunkten und der Kohonenschicht. Die Kohonenschicht stellt ein zweidimensionales Gitter aus $k \times l$ Kohonen-Neuronen bzw. Gewichtsvektoren dar. Jedem Neuron ist ein Gewichtsvektor zugeordnet, der aus einzelnen Zahlenwerten besteht. Jeder Vektor der Eingabeschicht ist mit jedem Neuron verbunden, damit trifft jeder Datenpunkt beim Training auf jedes einzelne Neuron. Diese Ver-

bindung impliziert jedoch nicht, daß grundsätzlich eine Aktivierung und damit verbundene Veränderung des Neurons über sie erfolgt.

Ein mehrdimensionaler Raum wird topologieerhaltend auf die zweidimensionale Karte abgebildet. Unter Topologieerhalt versteht man, daß zwei im mehrdimensionalen Raum benachbarte Eingabedatenpunkte auf das gleiche oder zwei ebenfalls benachbarte Neuronen in der Neuronenschicht abgebildet werden. Die einzelnen Neuronen stehen somit über ein Nachbarschaftsgitter miteinander in Beziehung. Jedes Neuron erhält dabei eine gewisse Anzahl von benachbarten Neuronen ersten, zweiten oder dritten Grades usw. Diese Beziehung soll die Topologie des Netzes festlegen.^[38]

Das Nachbarschaftsgitter (Abbildung 4-8) ist üblicherweise so angeordnet, daß ein Neuron in der Kohonenschicht immer in der Mitte seiner nächsten Nachbarn zu liegen kommt. Innerhalb eines quadratisch planaren Gitters kann diese Nachbarschaftsbeziehung verwirklicht werden.

a) quadratisches Nachbarschaftsgitter

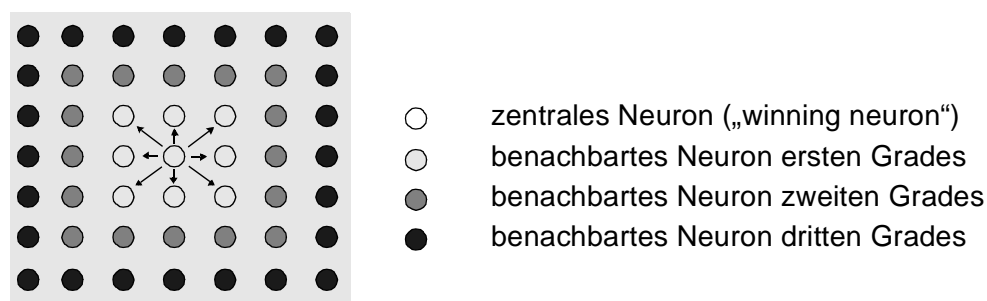


Abb. 4-8: Quadratisches Nachbarschaftsgitter.

Im quadratisch planaren Gitter hat jedes Neuron 8 direkte Nachbarn (Nachbarn ersten Grades), 16 Nachbarn zweiten Grades, 24 Nachbarn dritten Grades usw.

Hat die Karte die Gestalt eines Rechtecks, so reduziert sich an den Rändern der Karte die Anzahl der Nachbarn. Wird dagegen eine andere topologische Grundstruktur, z.B. die eines dreidimensionalen Torus gewählt, so kann jedem Neuron die gleiche Anzahl an Nachbarn zugeordnet werden. Durch zweifaches Aufschneiden des Torus wird die zweidimensionale Kohonen-Karte erhalten (Abbildung 4-9).

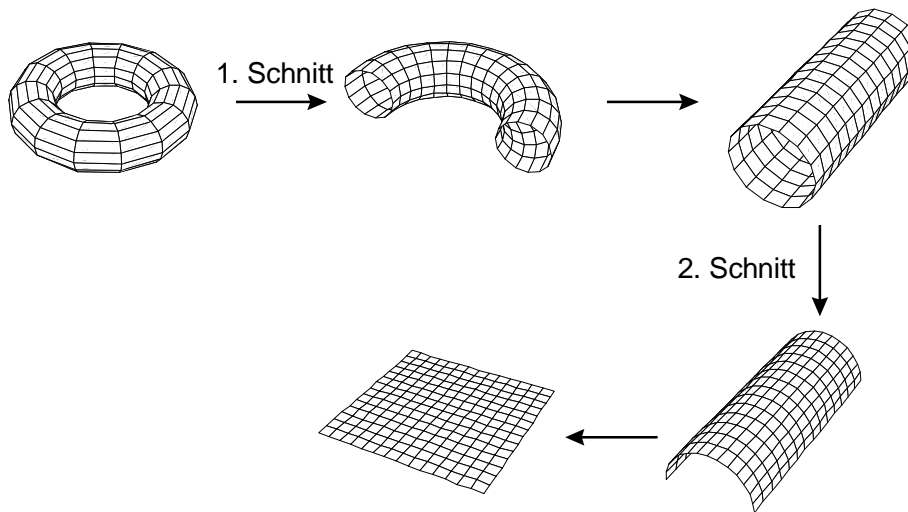


Abb. 4-9: Zweifaches Aufschneiden des toroidalen Kohonen Netzwerks ergibt eine zweidimensionale Karte.

Die Ähnlichkeit der Neuronen ist nicht abhängig von ihrer Nachbarschaftsbeziehung. Die Nachbarschaftsbeziehung legt die räumliche Position des Neurons im Netz fest, wohingegen die Ähnlichkeiten rechnerisch, z.B. durch die euklidische Distanz, d_{sj} , (Gl. 4-2) zwischen den Neuronen, ermittelt wird.^[91]

$$d_{sj} = \sqrt{\sum_{i=1}^m |w_{si} - w_{ji}|^2} \quad (\text{Gl. 4-2})$$

mit: m = Dimension des Eingabevektors
 w_{si} und w_{ji} = Gewichte der zu vergleichenden Neuronen s und j

4.2.3.3 Das Lernverfahren in Kohonen neuronalen Netzen

Das von Kohonen entwickelte Verfahren ist ein nicht überwachtes Lernverfahren. Dem Netz wird somit nur Information über die Eingabedaten, nicht über die Ausgabedaten, zur Verfügung gestellt.^[89]

Vor dem Training werden die Gewichtsvektoren der Kohonen-Neuronen zufällig initialisiert. Das Netz wird mit der Menge an Eingabedatenpunkten mehrmals trainiert, d.h. der im folgenden beschriebene Trainingszyklus wird mehrmals durchlaufen.

In jedem Trainingszyklus wird jeder Eingabevektor X_s mit allen Gewichtsvektoren W_j der Neuronen in der Kohonenschicht verglichen und der Gewichtsvektor W_{sc} des ähnlichsten

Neuron ermittelt. Dieses Neuron nennt man auch *zentrales (central oder winning) Neuron cs*. Da nur ein Neuron das Ähnlichste sein kann, spricht man von *kompetitivem Lernen*. Der Vergleich der Ähnlichkeiten kann in jeder beliebigen Metrik erfolgen, z.B. der euklidische Distanz der Neuronen (Gl. 4-2).

$$out_{sc} \leftarrow \min \left[\sqrt{\left(\sum_{i=1}^m |x_{si} - w_{ji}|^2 \right)} \right] = \min[\|X_s - W_j\|] \quad (\text{Gl. 4-3})$$

mit: $X_s \in \{X\}$, $s =$ aus der gegebenen Menge von Eingabevektoren $\{X\}$
 $W_j =$ Gewichtsvektor des Neurons j
 $x_{si}, w_{ji} =$ Komponenten der Eingabe- und Gewichtsvektoren
 $m =$ Länge bzw. Dimension des Eingabevektors

Die Gewichtsvektoren des ähnlichsten (zentralen) Neurons, cs , und, in geringerem Maße, seiner benachbarten Neuronen, j , werden dann auf eine Weise verändert, daß sie dem Eingabevektor, X_s , ähnlicher werden. Dabei ist zu beachten, daß die Gewichtsvektoren der nächsten Nachbarn nicht unbedingt am „zweitähnlichsten“ sein müssen. Dieses Verhalten steuert die *Propagierungsfunktion*.^[38]

Die Anpassung der Neuronen findet entlang der topologischen Umgebung statt. Die Aktivierung nimmt mit zunehmender Entfernung vom zentralen Neuron ab, so daß die Aktivierung eines Neurons, des zentralen Neurons, cs , nur in der nächsten Nachbarschaft eine Anpassung der Neuronen, j , veranlaßt. Mathematisch ausgedrückt, wird dabei der Gewichtsvektor des zentralen Neurons, W_{cs} , verändert, um W_{cs} noch stärker an den aktuellen Eingabevektor, X_s , anzugleichen und um eine topologische Organisation zu garantieren (Gl. 4-4).^[38] Mit dem Schritt des Anpassens der Neuronen an den Eingabevektor endet der Trainingszyklus.

$$W_j(t+1) = W_j(t) + \eta(t)h(t, z)[X_s - W_j(t)] \quad (\text{Gl. 4-4})$$

mit: $W_j =$ Gewichtsvektor des anzugleichenden Neurons j
 $X_s =$ Eingabevektor
 $0 < \eta(t) < 1$ zeitlich veränderbare Lernrate, die das Maß der Angleichung von W_j an X_s vorgibt
 $0 < h(t, z) < 1$ Distanz- oder Nachbarschaftsfunktion, die für eine Modifizierung der Gewichte des zentralen Neurons c und der benachbarten Neuronen sorgt

Es gibt verschiedene Arten von linearen Nachbarschaftsfunktionen, eine davon ist in Abbildung 4-10 dargestellt.

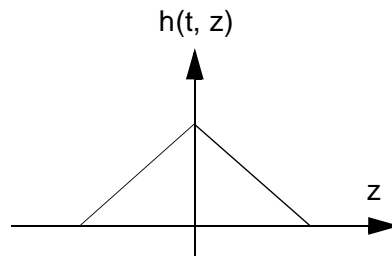


Abb. 4-10: Nachbarschaftsfunktionen: Die Gewichtsmodifizierung nimmt linear mit zunehmender Entfernung zum Gewinnerneuron ab.

Abbildung 4-11 zeigt schematisch den Ablauf des gesamten Lernvorgangs. Die Anpassung an den Eingabevektor X_s geschieht nicht nur für das zentrale Neuron sc (in Abbildung 4-11 schwarz) sondern auch für benachbarte Neuronen, die in Abbildung 4-11 grau markiert sind. Da ähnliche Datenpunkte in benachbarten Bereichen des Netzes eine sich gegenseitig verstärkende Wirkung zeigen, organisieren sich die Karten selbst, d.h. es bilden sich Gruppen ähnlicher Neuronen. Das unüberwachten Lernverfahren gründet sich daher darin, daß die Information über die Gruppenzugehörigkeit eines Datenpunktes beim Training nicht angegeben wird.

Mit Fortschreiten des Trainings muß ein „Übertrainieren“ verhindert werden. Die Anpassung der Neuronen erfolgt umso schwächer, je länger das Netz bereits trainiert wurde, d.h. je öfter ein Satz von Eingabevektoren zum Training verwendet wurde. Dies bestimmt die Lernrate η (Gl. 4-4). Die Anzahl an zu durchlaufenden Trainingszyklen wird vorgegeben.

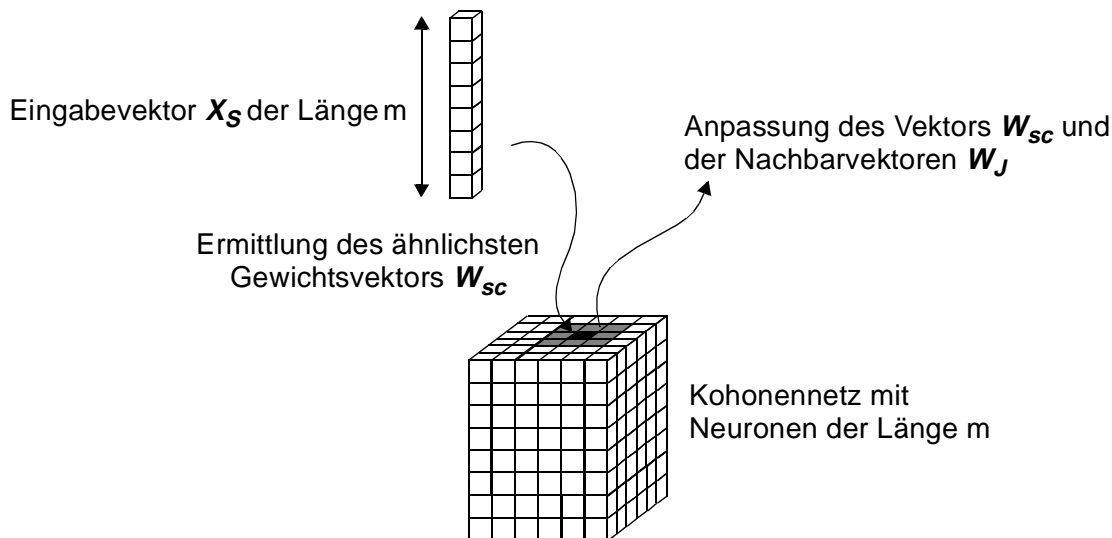


Abb. 4-11: Lernverfahren einer Kohonen neuronalen Netzes. Die grau markierten Neuronen werden zusätzlich zum zentralen Neuron sc an den Eingabevektor X_s angepaßt.

4.2.3.4 Vergleich von Moleküloberflächen

Eine bedeutende Anwendung von Kohonen neuronalen Netzen ist die Projektion der dreidimensionalen molekularen Oberfläche in zweidimensionale Karten, wie in Abbildung 4-12 dargestellt.^{[15]-[18],[92][93]} Dabei ist die Dimension des Eingabevektors, X_s , drei, entsprechend den x-, y- und z-Koordinaten der molekularen Oberfläche. Die Netzgröße für das Training betrug 50*50 Neuronen. Die entstehenden zweidimensionalen Karten sind sehr gut zum Vergleich molekularer Oberflächen geeignet.^{[17]-[18],[93]} Alle Kohonen Karten wurden mit Hilfe des Kohonen-Karten-Generators KMAP erstellt.^[94] Für das angeführte Beispiel in Abbildung 4-12 wurde das elektrostatische Oberflächenpotential (ESP) zur Projektion auf die Moleküloberfläche gewählt. Die physikochemische Eigenschaft wird auf die van der Waals-Oberflächen der Verbindungen projiziert und als Ausgabewert für die Anfärbung der Kohonen-Karten verwendet. Die Farbskalierung (Abbildung 4-12) läuft jeweils vom höchsten zum niedrigsten Wert der betrachteten physikochemischen Eigenschaft. Die Pfeile weisen explizit auf die entsprechenden Regionen von Moleküloberfläche und Karte hin.

Die van der Waals-Oberfläche, die auf den dreidimensionalen Koordinaten der Verbindungen basieren, wurden mit Hilfe des Programms SURFACE^[95] berechnet. Die Oberflächenpunktdichte war 50 Punkte/Å².

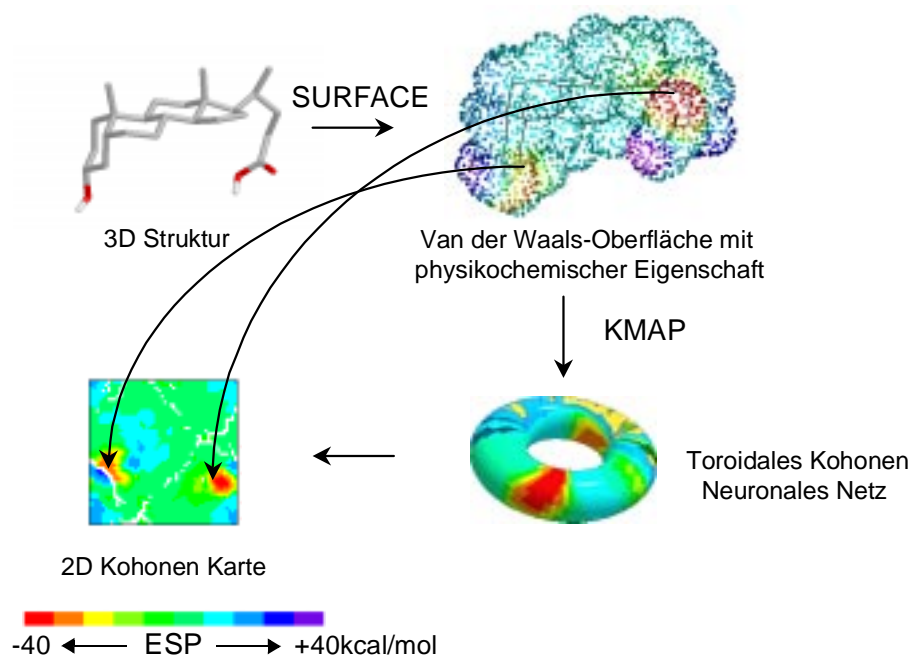


Abb. 4-12: Die Projektion einer dreidimensionalen van der Waals-Moleküloberfläche mit der Eigenschaft des ESP (elektrostatisches Potential) in kcal/mol auf eine zweidimensionale Karte mit Hilfe eines toroidalen neuronalen Netzes.

Es sind zwei generelle Methoden zum Vergleich von van der Waals-Oberflächen mit Kohonen Karten möglich.^{[16]-[92]} Erstens kann ein Netz für jede einzelne dreidimensionale Struktur trainiert werden und die Karten anschließend zum Vergleich nach strukturell ähnlichen Bereichen ausgerichtet werden. Zweitens kann die *Templat-Methode* angewandt werden (Abbildung 4-13).^[92] Die Templat-Methode verlangt eine vorherige Ausrichtung der dreidimensionalen Strukturen des Datensatzes (siehe Abbildung 4-13 oben). Das Netzwerk wird dafür ein einziges Mal mit den Daten eines Templatmoleküls trainiert, dies kann z.B. die aktivste Verbindung des Datensatzes sein, und die Daten jeder weiteren Struktur werden durch das erhaltene Netz geschickt. Strukturelle Regionen großer physikochemischer Äquivalenz, führen zu gefärbten Bereichen auf der Karte, wogegen sich unähnliche Regionen in weißen Bereichen ausdrücken. Weiße Bereiche in der Kohonen-Karte entsprechen leeren Neuronen, d.h. kein Gewinnerneuron wurde für den zugehörigen Eingabevektor gefunden.

Für das Beispiel aus Abbildung 4-13 wurde ein Kohonen-Netz mit den dreidimensionalen Oberflächenkoordinaten der Cholsäure, **3**, trainiert. Diese Verbindung bildet somit das Templat. Die Netzgröße für das Training betrug 50*50 (2500) Neuronen. Als Vergleichsverbindung wurde Desoxycholsäure, **24**, gewählt, da diese sich nur durch die fehlende Hydroxygruppe am

B-Ring von **3** unterscheidet (Abbildung 4-13 oben und Mitte). Zur Anfärbung der Kohonen-Karten wurde als Eigenschaft die Ordnungszahl der Atome verwendet (grün: Kohlenstoff, rot: Sauerstoff, hellblau: Wasserstoff). Durch das mit den Koordinaten von **3** trainierte Netz werden die Oberflächenkoordinaten von **24** geschickt. Es werden zwei Karten erhalten, die Templat-Karte und die Karte der Vergleichsverbindung (Abbildung 4-13 unten). Es ist deutlich zu erkennen, daß sich die fehlende Hydroxygruppe von **24** als weißer Bereich (leere Neuronen) auf der Karte zeigt. Die Anzahl der leeren Neuronen gibt ein quantitatives Maß für die Abweichungen in den beiden Strukturen (hier 375 von 2500 Neuronen). Damit ist ein anschaulicher Vergleich zweier dreidimensionaler Oberflächen möglich.

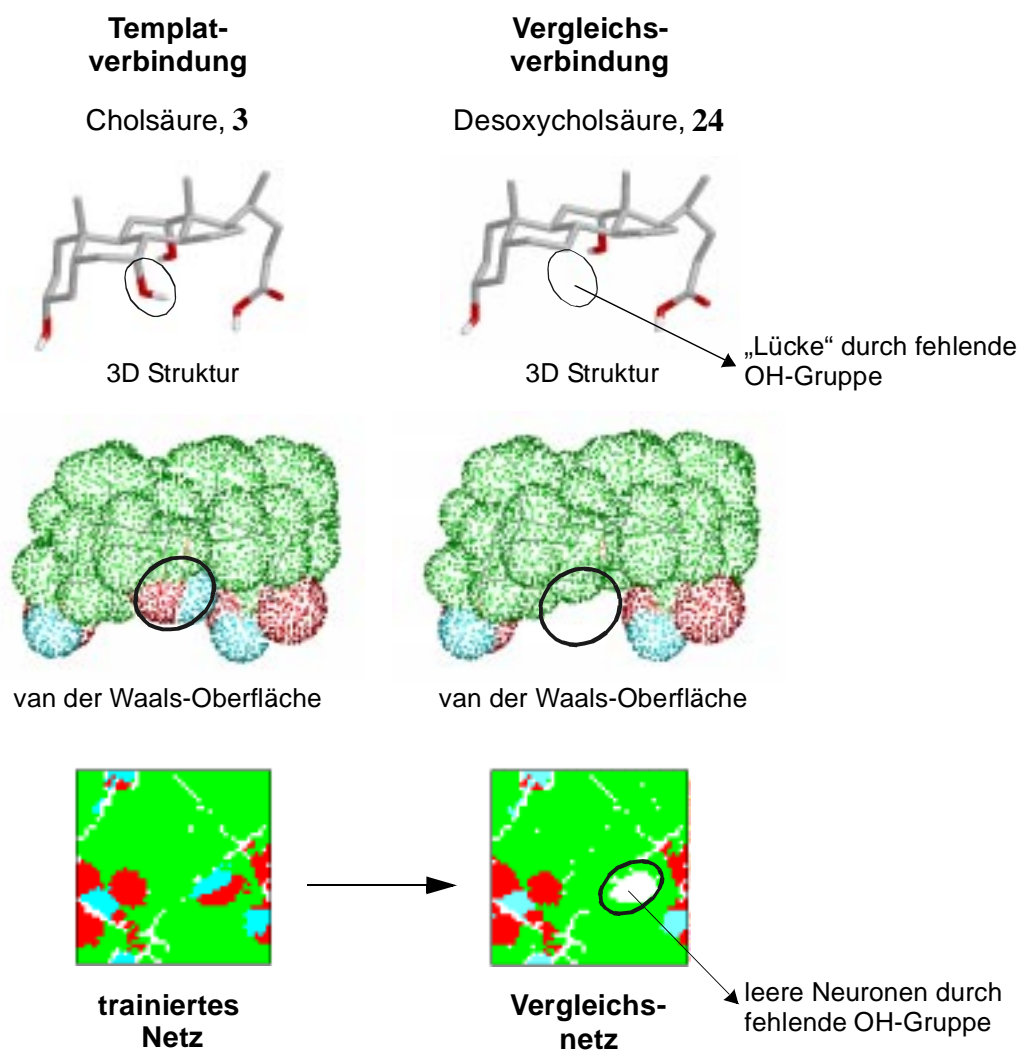


Abb. 4-13: Die Templat-Methode. Links ist die Kohonen Karte nach dem Training mit den Koordinaten der Templatverbindung, **3**, zu sehen, rechts die Karte, die entsteht, nachdem die Koordinaten der Vergleichsverbindung, **24**, durch das trainierte Netz geschickt wurden.

4.2.3.5 Berechnung des elektrostatischen Oberflächenpotentials

Das elektrostatische Oberflächenpotential wird auf Basis der partiellen Atomladung, q_{tot} , der Strukturen berechnet, die mit Hilfe der PEOE-Methode (Partial Equalization of Orbital Electronegativities)^{[96][97]} und deren Erweiterung auf konjugierte Systeme, erhalten werden.^[98] Bei der Berechnung des Potentials wird eine Probeladung von +1e über die gesamte molekulare Oberfläche abgerollt. Die Berechnung des molekularen elektrostatischen Oberflächenpotentials (ESP in kcal/mol) an diesem Punkt erfolgt durch das Coloumb-Gesetz. Es wird über alle Atompartialladungen im Molekül $q_{tot}(i)$ aufsummiert (Gl. 4-5).

$$V = \frac{1}{4\pi\epsilon_0} \sum_i \frac{q_{tot}(i)}{r_i} \quad (\text{Gl. 4-5})$$

mit: $q_{tot}(i)$ = Atompartialladung des Atoms i
 r_i = Abstand zwischen dem betrachteten Oberflächenpunkt und dem Atom i

4.2.3.6 Berechnung der Oberflächenhydrophobizität

Es gibt keine exakte physikalische Methode zur Berechnung der Hydrophobizität einer molekularen Oberfläche. Die bisherigen Methoden, relative Hydrophobizitätswerte zu definieren, gründen auf empirischen Untersuchungen. Aus dem Partitionierungskoeffizient ($\log P$) einer Probe in gemischten polaren und unpolaren Referenzsystemen können empirische Hydrophobizitätswerte erhalten werden. Da Fujita et al. den additiven Charakter des Partitionierungskoeffizient bewies, kann die Hydrophobie bzw. Lipophilie einer Verbindung als Summe der partiellen Lipophilizitäten, f_i , von i molekularen Fragmenten ausgedrückt werde (Gl. 4-6).^{[99][100]}

$$\log P = \sum_i f_i \quad (\text{Gl. 4-6})$$

mit: f_i = partielle Hydrophobie/Lipophilie des molekularen Fragments i

Die Vernachlässigung des Einflusses auf $\log P$ eines jeden Fragments, das sich in der Nachbarschaft des betrachteten Fragments i befindet, führt allerdings zu Fehlern in den Berechnungen der molekularen Hydrophobizität bezüglich der experimentell ermittelten Werte.

Heiden et al. führten die Berechnung des molekularen Hydrophobie- oder Lipophilie-Potential ein (engl.: molecular lipophilicity potential, MLP), das auf der Unterteilung der molekularen Lipophilie in fragmentale Beiträge und der Projektion dieser Lipophilie auf die

molekulare Oberfläche basiert.^[101] Eine distanzabhängige Beschreibung für den Einfluß verschiedener Fragmente auf die Lipophilie oder Hydrophobie eines Oberflächenpunktes war somit gefunden.

$$MLP = \sum_i \frac{f_i}{1+d_i} \quad (\text{Gl. 4-7})$$

mit: f_i = partielle Hydrophobie/Lipophilie des molekularen Fragments i
 d_i = Distanz zwischen Oberflächenpunkt und Zentrum des betrachteten molekularen Fragments i

4.3 Ergebnisse

4.3.1 Die Bindungskonfigurationen von Digoxin-verwandten Steroiden im Komplex mit dem Immunoglobulin-Rezeptor 26-10

In einer Dockingsimulation wurde der Bindungsmodus von Digoxin, **2**, innerhalb des 26-10-Murin monoklonalen Antikörpers mit Hilfe des automatischen Dockingsimulationsprogramms AutoDock berechnet. Bei der Simulation der Bindungsmodi werden die Liganden als flexibel, der Rezeptor dagegen als starr behandelt. Der Ligand wurde vorher manuell innerhalb des Programms InsightII/Discover^[76] in die Bindungstasche des Antikörpers eingelagert. Ein Vergleich mit der Kristallstruktur des 26-10-Fab-Digoxin Komplexes (Brookhaven Protein Database: 1IGJ),^[20] ergibt einen *rms*-Wert von 0.9Å, bezüglich der berechneten und experimentell ermittelten Konformation von Digoxin, **2**, im Komplex mit dem 26-10-Fab. Abbildung 4-14 zeigt die Lage des Digoxins innerhalb der 26-10-Fab Rezeptorbindungstasche. Einige der direkt an der Bindungstasche beteiligten Aminosäuren Trp^{H100}, Tyr^{H33} und Tyr^{H50} sind hervorgehoben. Die schwere Kette, V_H des Antikörpers, ist blau und die leichte Kette, V_L, grün eingefärbt.

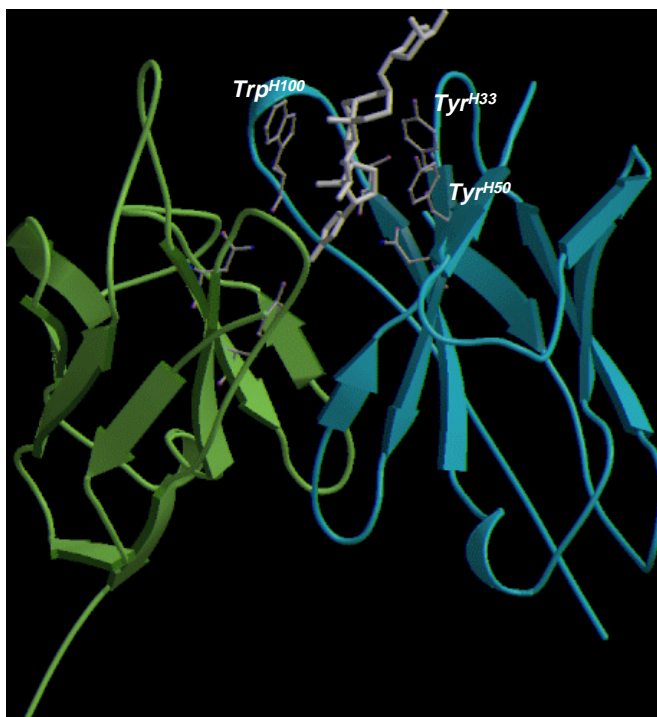


Abb. 4-14: Die berechnete Konformation von **2** innerhalb der Bindungstasche des 26-10-Fab. V_H des Antikörpers ist blau und V_L grün eingefärbt. Hervorgehoben sind einige Aminosäuren Trp^{H100}, Tyr^{H33} und Tyr^{H50}, die direkt an der Bildung der Rezeptorbindungstasche beteiligt sind.

In weiteren Schritten wurden die Konformationen der Steroide Digoxin, **21**, Digoxigenin, **22**, und Ouabain, **23**, innerhalb der 26-10 Bindungstasche bestimmt.

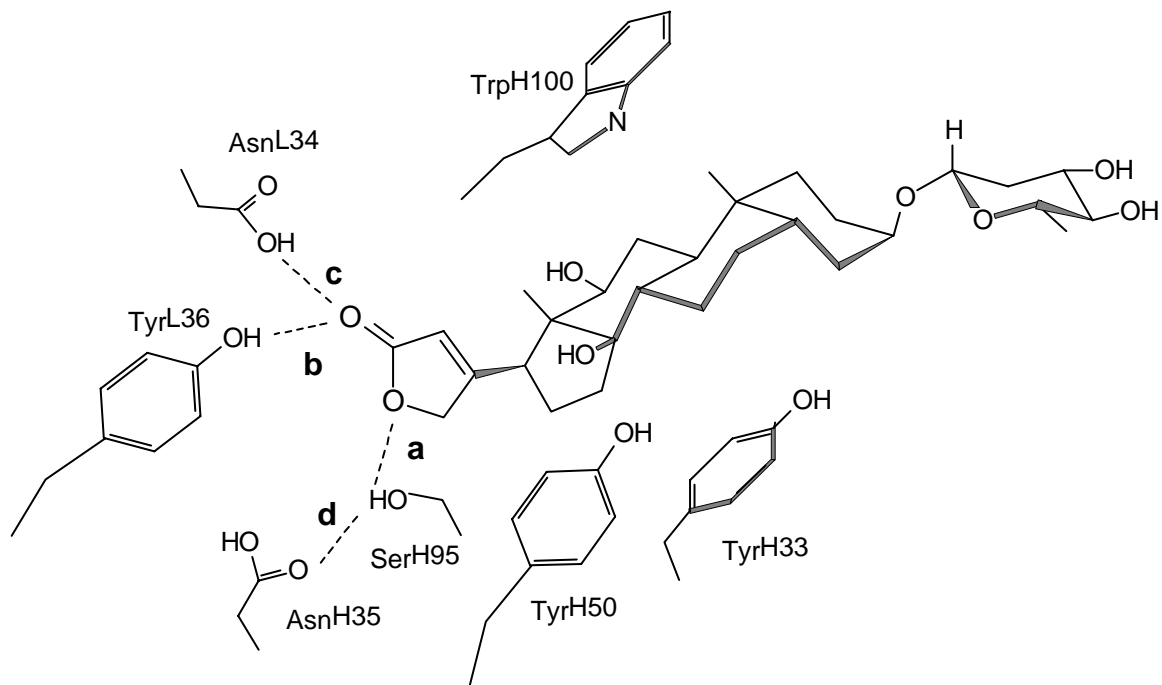
Bei der Simulation der Bindungsmodi werden die Liganden als flexibel behandelt, der Rezeptor dagegen als starr. Zusätzlich zur automatischen Erkennung aller rotierbaren Bindungen, ist in AutoDock die Möglichkeit zur expliziten Angabe von starren oder flexiblen Bindungen möglich. Die Steroidstrukturen **2**, **21**, **22** und **23** sind relativ starr mit einer maximalen Anzahl an 11 exocyclischen rotierbaren Bindungen im Falle von **23** (Abbildung 4-4). Zu diesen rotierbaren Bindungen zählen alle exocyclischen Bindungen, auch wenn sie nur weitere einatomige Substituenten aufweisen, wie z.B. die Bindungen zu Methylgruppen (–CH₃) oder Hydroxylgruppen (–OH). Für die Vorhersage der Bindungsmodi mit AutoDock, wurden nur jene Bindungen als rotierbar angegeben, die direkt am Steroidgrundkörper oder am Lactonring der Verbindungen lokalisiert sind (Tabelle 4-1). Die in Tabelle 4-1 angegebenen Atomindizes beziehen sich auf die Strukturformeln aus Abbildung 4-4.

Tab. 4-1: Listen von Bindungen der jeweiligen Liganden, die während der AutoDock Simulation als rotierbar angegeben wurden.

	Rotierbare Bindungen (Atomindizes, vgl. Abbildung 4-4)
Digoxin, 2	13 - 14 16 - 17 20 - 23
Digitoxin, 21	16 - 17 20 - 23
Digoxigenin, 22	13 - 14 16 - 17 20 - 23
Ouabain, 23	1 - 37 6 - 38 11 - 21 12 - 40 16 - 17 20 - 23 21 - 39

Jede AutoDock Simulation führt zu vielen unterschiedlichen Bindungskonfigurationen

der Steroide innerhalb der 26-10 Bindungstasche. Hier wird grundsätzlich der energetisch niedrigste ermittelte Bindungsmodus zum Vergleich herangezogen. Die Orientierung der drei Steroide Digitoxin, **21**, Digoxigenin, **22**, und Ouabain, **23**, ist der Orientierung von **2** innerhalb der Rezeptorbindungstasche sehr ähnlich. Die Lactonsubstituenten der D-Ringe der Steroide reichen am weitesten in das Innere der Bindungstasche hinein, hingegen ragen die Zuckersubstituenten der A-Ringe aus der Bindungstasche in Richtung Solvens heraus. Die Grundgerüste der Steroide sind zwischen den aromatischen Ringen der Aminosäurereste Tyr^{H33}, Tyr^{H50} und Trp^{H100} eingebettet (Abbildung 4-14 und Abbildung 4-15). Es wurden keine Wasserstoffbrücken zwischen den Liganden und den potentiellen Wasserstoffbrückendonatoren Asn^{L34}, Tyr^{L36} und Ser^{H95} gefunden. Die einzigen polaren Gruppen, die nicht durch Wechselwirkung mit dem Solvens stabilisiert werden, sind die Lactonring-O-21 und -O-23 Sauerstoffatome (Abbildung 4-15). Obwohl potentielle Wasserstoffbrückendonatoren, wie z.B. Ser^{H95}-O^γ im Bereich von O-21 liegen, ist in keiner Verbindung, mit Ausnahme von **22** ($d(\mathbf{a}) = 2.93\text{\AA}$, Abbildung 4-15), dieses Sauerstoffatom nah genug, um mit dem entsprechenden O-21 eine Wasserstoffbrücke zu bilden (Distanzen **a** in Abbildung 4-15). Weiterhin bildet Ser^{H95} O^γ bereits eine Wasserstoffbrücke ($d(\mathbf{d}) = 2.9\text{\AA}$, Abbildung 4-15) mit dem Carboxylsauerstoff der Aminosäuregruppe von Asn^{H35}. Analoges gilt für das Carbonylsauerstoffatom O-23. Die beiden in der Nähe befindlichen Hydroxylgruppen von Asn^{L34} und Tyr^{L36} als potentielle Wasserstoffbrückendonatoren sind in den Komplexen aller vier Verbindungen **2**, **21**, **22** und **23** zu weit entfernt um eine Wasserstoffbrücke zu O-23 auszubilden (Distanzen **b** und **c** in Abbildung 4-15).



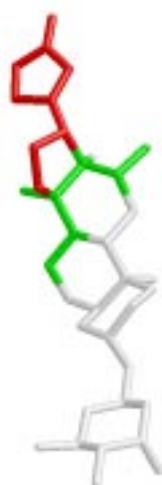
Distanzen: **a** ligand O-21 - Ser^{H95} O γ
b ligand O-23 - Tyr^{L36} OH
c ligand O-23 - Asn^{L34} OH
d Ser^{H95} O γ - Asn^{H35} O

	Distanzen [Å]			
	a	b	c	d
Kristallstruktur von 2	3.55	3.97	4.13	2.84
Digoxin 2	3.26	3.60	3.82	2.84
Digitoxin 21	3.03	3.83	3.88	2.84
Digoxigenin 22	2.93	4.05	4.11	2.84
Ouabain 23	3.11	3.83	3.62	2.84

Abb. 4-15: Schematische Darstellung der Bindungsmodi von **2**, **21**, **22** und **23** innerhalb der 26-10-Fab Bindungstasche. Gegeben sind die jeweiligen Distanzen der Sauerstoffatome der Liganden zu den Hydroxylgruppen der Aminosäuren Asn^{L34}, Tyr^{L36} und Ser^{H95} der 26-10-Fab Bindungstasche und die Distanzen der internen Wasserstoffbrücke von Ser^{H95} zu Asn^{H35}.

Die mit AutoDock berechnete absolute Potentialenergiefunktion ergibt eine van der Waals-Wechselwirkung zwischen Digoxin, **2**, und dem Antikörper von -50 kcal/mol. Für 70% dieser Wechselwirkung sind nur 39% der Atome der Gesamtstruktur verantwortlich. Zur Gesamtstruktur zählen dabei alle Kohlenstoff- und Heteroatome sowie an Heteroatome gebundene Wasserstoffatome. Der Prozentanteil der wechselwirkenden Atome (hier 39%) berechnet

sich als Quotient der Anzahl der Atome, die mit einem Beitrag von größer als 1.5kcal/mol wechselwirken und der Gesamtanzahl der Atome. Die wechselwirkenden 39% der Gesamtstruktur werden hauptsächlich von den B/C/D-Ringatomen und den Lactonringatomen gebildet. Abbildung 4-16 zeigt das Affinitätspotential von **2** mit der Antikörperbindungstasche des Fab 26-10 anhand der unterschiedlichen Einfärbung der verschieden stark wechselwirkenden Strukturbereiche. Die größten Interaktionen zeigen das Sauerstoffatom des Lactonrings und die Atome des Lacton- und des D-Rings mit einem absoluten Beitrag von mehr 2.0 kcal/mol (Abbildung 4-16: rot eingefärbt), weiterhin Teile des B- und C-Rings mit einem Beitrag von mehr 1.5 kcal/mol (Abbildung 4-16: grün eingefärbt). Die kleinsten Wechselwirkungen mit Beiträgen von weniger 0.5 kcal/mol, weisen der Zuckerrest und die Etherbrücke auf (Abbildung 4-16: grau eingefärbt).



Affinitätspotential von **2**

rot: Wechselwirkung mit > 2.0 kcal/mol

grün: Wechselwirkung mit \geq 1.5 kcal/mol

grau: Wechselwirkung mit < 1.5 kcal/mol

Abb. 4-16: Die dreidimensionale Struktur von Digoxin, **2**, mit dem Affinitätspotential bezüglich des 26-10-Fab. Strukturelle Bereiche hoher Wechselwirkung (>2.0 kcal/mol) zu der 26-10-Fab Bindungstasche sind rot, mittlerer Wechselwirkung (>1.5 kcal/mol) grün und keiner Wechselwirkung grau eingefärbt.

Die berechnete Zahl von insgesamt 16 Atomen als verantwortlich für die Interaktion mit der 26-10-Fab Rezeptorbindungstasche, entspricht weitgehend den bei Jeffrey et al.^[20] ermittelten 18 Atomen der gesamten Digoxinstruktur, die hauptsächlich für die Wechselwirkung mit der Antikörperbindungstasche verantwortlich sind.

Zum Vergleich des berechneten Bindungsmodus und des experimentell ermittelten Bindungsmodus^[20] von **2** innerhalb der 26-10 Bindungstasche ist in Abbildung 4-17 a) die Überlagerung der berechneten Konformation (blau) und der experimentell ermittelten Konformation (rot) dargestellt.

Die Überlagerung der Kristallstrukturdaten von Digoxin, **2**, (rot) mit der berechneten

Konformation von **2** (blau), Digitoxin, **21** (grün), Digoxigenin, **22** (grau) und Ouabain, **23** (magenta), in Abbildung 4-17 b) zeigt die ähnliche Orientierung der vier Steroide innerhalb der Ligandenbindungstasche des 26-10 Antikörpers.

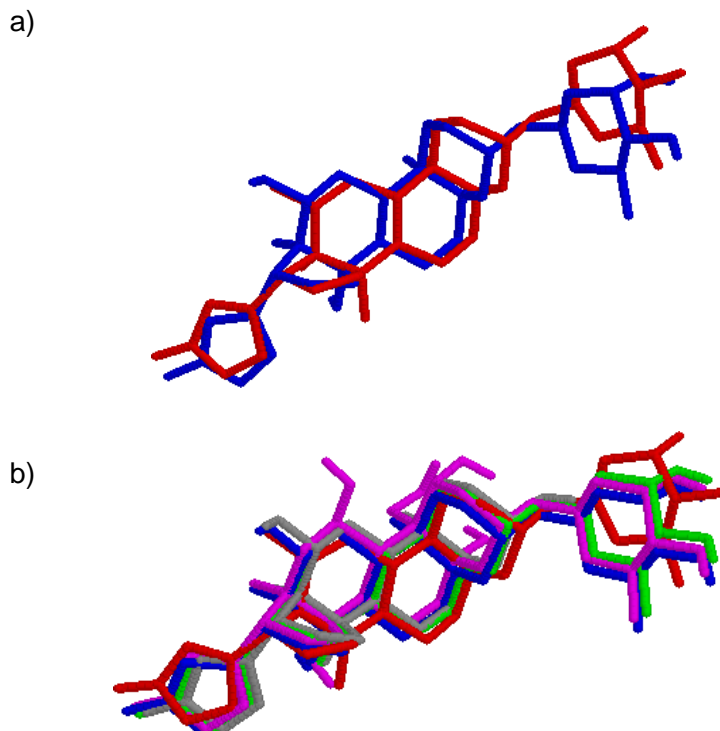


Abb. 4-17: a) Die Überlagerung der Kristallstrukturdaten von **2** (rot) mit der durch die AutoDock Rechnung ermittelten Konformationen von **2** (blau), innerhalb der 26-10 Bindungstasche. b) Die Überlagerung der Kristallstrukturdaten von **2** (rot) mit der berechneten Konformation von **2** (blau), **21** (grün), **22** (grau) und **23** (magenta).

Diese Dockingstudien zeigten vor allem, daß die Affinität der hier behandelten Steroide zu der Ligandenbindungstasche des betrachteten Immunoglobulins 26-10 hauptsächlich auf apolaren Wechselwirkungen und Oberflächenkomplementarität basieren, da nur in einem Fall eine Wasserstoffbrücke der polaren Lactongruppe der Steroide zu einer Aminosäurehydroxylgruppe im Inneren des Rezeptors gefunden werden konnte.

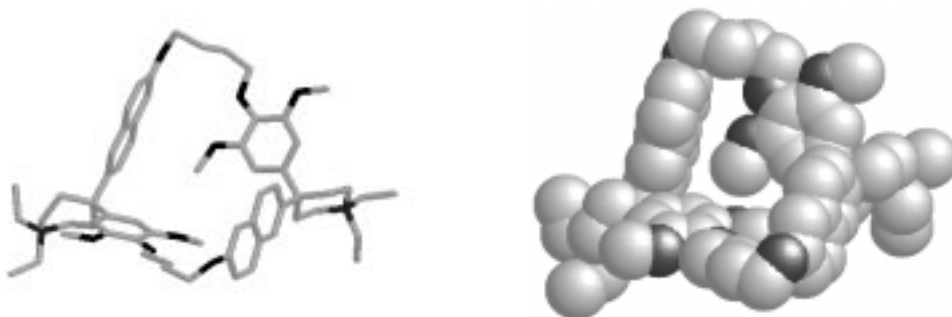
4.3.2 Die Bindungskonfigurationen verschiedener Steroide im Komplex mit einem synthetischen Paracyclophanrezeptor

Um die wahrscheinlichen Anlagerungsmodi der Steroide zu einem synthetischen Cyclophanrezeptor zu analysieren, wurden AutoDock Simulationen durchgeführt, wobei der Cyclophanring, **20**, als starrer Rezeptor, analog zum Immunoglobulin 26-10 in Kapitel 4.3.1

behandelt wird.

Die stabilste Konformation des Cyclophans, **20**, wurde durch Absuche des Konformationsraums mit Molecular Dynamics Methoden basierend auf einem MM3 Kraftfeld erhalten (siehe A.1.1 und A.1.2). In dieser Konformation ragen die Methoxygruppen der Benzolringe in das Innere des Rings (Abbildung 4-18 a) und verhindern daher die Möglichkeit zur Komplexierung eines potentiellen Liganden.

a) geschlossene Konformation des Cyclophans



b) geöffnete Konformation des Cyclophans

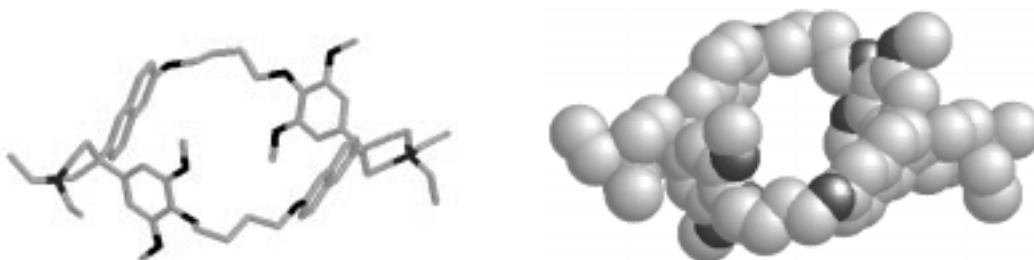


Abb. 4-18: a) Geschlossene Konformationen des Paracyclophanrings, **20**, b) geöffnete Konformationen des Paracyclophanrings, **20**. Links ist jeweils die normale Ansicht und rechts das Kalottenmodell der Struktur dargestellt. Die geöffnete Konformation erlaubt die Komplexierung eines Liganden im Inneren des Rings.

Eine alternative und besser für die Ligandenkomplexierung geeignete Konformation von **20** wurde durch weitere Molecular Dynamics Rechnungen erhalten, die, auf einem Cyclophan-Steroid-Komplex basierend, den Konformationsraum des Cyclophans, **20**, absuchten. Die so erhaltene Konformation (Abbildung 4-18 b) ist um 21 kcal/mol energiereicher als die Konformation, die ohne komplexiertes Steroid erhalten wurde (Abbildung 4-18 a). Aus den Kalottenmodellen in Abbildung 4-18 ist es jedoch ersichtlich, daß es die energieärmere, aber geschlossene Struktur nicht erlaubt, einen Liganden einzulagern. Infolgedessen wurde die weniger stabile, aber geöffnete Konformation von **20** als Ligandenbindungstasche für alle fol-

genden Simulationen verwendet.

In die Bindungstasche des Cyclophans werden vorerst die Cholsäure, **3**, und verwandte Verbindungen, wie Desoxycholsäure, **24**, Chenodesoxycholsäure, **25**, Ursodesoxycholsäure, **26**, und Lithocholsäure, **27**, eingelagert (Abbildung 4-5), weiterhin die Steroide Hydrocortison, **28**, Cortison, **29**, und Testosteron, **30** (Abbildung 4-6). Als letztes wurden Digoxin, **2**, Ouabain, **23** und Progesteron, **31** (Abbildung 4-4), die hohe Bindungsaffinitäten zu Immunoglobulinrezeptoren aufweisen^{[25][26]} in die Ligandenbindungstasche von **20** eingelagert. Die stabilste Konformation eines jeden Komplexes und die zugehörigen Strukturen mit Einfärbungen entsprechend der Stärke ihrer Wechselwirkung mit der Bindungstasche des Cyclophans sind in Abbildung 4-19 bis 4-22 dargestellt.

Um diejenigen strukturellen Regionen zu ermitteln, die für die Affinität zu dem Paracyclophan hauptsächlich verantwortlich sind, wurde der Energiebeitrag eines jeden Atoms zur gesamten van der Waals-Wechselwirkungsenergie bestimmt. Alle Beiträge mit einer absoluten Energie höher als 1.5 kcal/mol wurden für die folgenden Untersuchungen in Betracht gezogen. Die strukturellen Bereiche der Steroide, die eine absolute Wechselwirkungsenergie von größer als 2.0 kcal/mol aufweisen wurden in Abbildung 4-19 bis 4-22 rot eingefärbt, diejenigen die einen Beitrag zwischen 1.5 kcal/mol und 2.0 kcal/mol liefern, wurden grün eingefärbt. Blau eingefärbt wurden Atome, die einen anders gerichteten Beitrag zur gesamten Wechselwirkungsenergie liefern, d.h. zu einer Abstoßung führen.

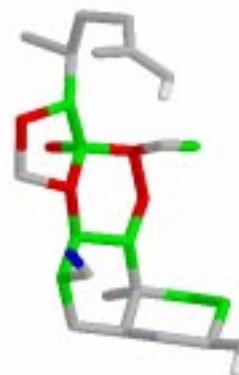
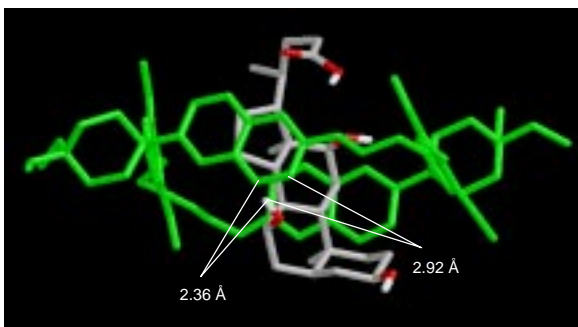
Cholsäure, **3**, eine Verbindung mit zwei axialen Hydroxylgruppen am zentralen Steroidgrundkörper bindet hauptsächlich mit zwei der A-Ringatome und neun der 13 B/C/D-Ringatome (Abbildung 4-19) an den künstlichen Rezeptor. Diese Substruktur repräsentiert 42% der Gesamtstruktur, ist aber für 68% der van der Waals-Interaktionen verantwortlich. Die α -Hydroxylgruppe am B-Ring führt zu einer van der Waals-Repulsion (in Abbildung 4-19 blau gekennzeichnet) aufgrund der kurzen Distanz von 2.36 Å und 2.92 Å des Hydroxylwasserstoffatoms zur Naphthalineinheit des Cyclophanrezeptors, **20**.

Desoxycholsäure, **24**, bindet in einem sehr ähnlichen Modus an den Cyclophanrezeptor, **20**, wie Verbindung **3**. Nahezu die gleichen Atome tragen zur größten Wechselwirkung mit dem apolaren Inneren des Cyclophanrezeptors bei (Abbildung 4-19). Die gesamte Wechselwirkungsenergie von **24** mit dem Cyclophanring, **20**, ist jedoch um 2.6 kcal/mol höher als die des Steroids **3**, aufgrund der fehlenden repulsiv wirkenden Hydroxygruppe am B-Ring.

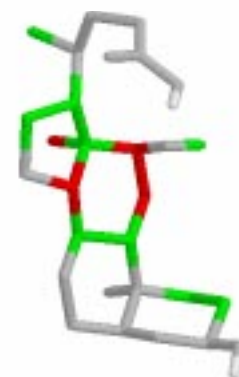
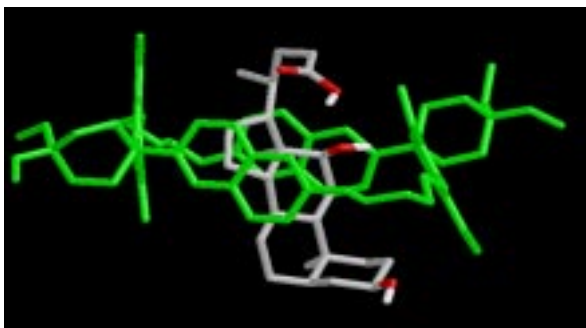
Chenodesoxycholsäure, **25**, bindet mit acht der neun Atome des C- bzw. D-Rings, einem Atom des A-Rings, der Methylgruppe an der C/D-Ringverbindung, der freien Methylgruppe des Pentylsäurerests, und dem Carboxylatoms des Pentylsäurerests (Abbildung 4-19). Diese Substruktur repräsentiert 39% der Gesamtstruktur und ist verantwortlich für 63% der gesamten

van der Waals-Wechselwirkungsenergie. Die α -Hydroxylgruppe, die im Falle von **3**, zu einer repulsiven Wechselwirkung führte, zeigt hier keinerlei Beitrag zur Wechselwirkung zwischen dem Steroid und dem künstlichen Cyclophanrezeptor.

Cholsäure, **3**



Desoxycholsäure, **24**



Chenodesoxycholsäure, **25**

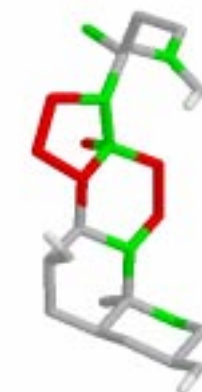
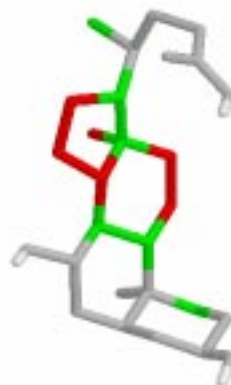
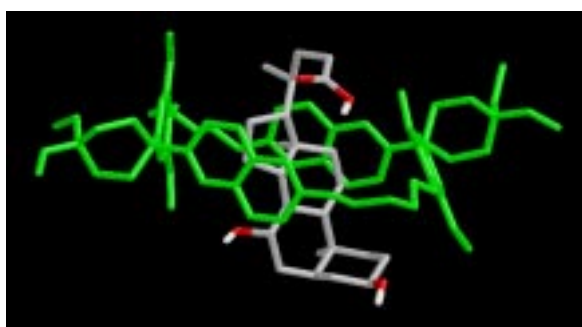


Abb. 4-19: Links: Simulierte Anordnung der Cyclophan-Steroid-Komplexe der Steroide **3**, **24** und **25** mit dem Cyclophan, **20**, rechts: das strukturelle Affinitätspotential. Rot eingefärbte Atome tragen zur van der Waals-Wechselwirkungsenergie mit >2.0 kcal/mol, grün eingefärbte mit >1.5 kcal/mol bei. Blau gefärbte Atome kennzeichnen einen repulsiven Beitrag zur Gesamtwechselwirkungsenergie.

Ursodesoxycholsäure, **26**, (Abbildung 4-20) bindet ebenfalls in einem ähnlichen Modus wie **3**, **24** und **25**. Die neun Atome des C- bzw. D-Rings und die Methylgruppe des Pentylsäurerests tragen am meisten zur Wechselwirkung mit dem Rezeptor bei, keinen Beitrag liefert allerdings das Carboxylatom des Pentylsäurerests. Die equatoriale Hydroxylgruppe am B-Ring der Ursodesoxycholsäure, **26**, führt zwar nicht zu einer Abstoßung, trägt aber zur gesamten van der Waals-Wechselwirkungsenergie nur mit 0.3 kcal/mol bei. Im Gegensatz zu einem axialen Substituenten am B-Ring der Steroide, führt ein Substituent in equatorialer Position offenbar nicht zu einer sterischen Hinderung der Wechselwirkung (vgl. Abbildung 4-19, Cholsäure, **3**).

Lithocholsäure, **27**, ein Steroid ohne polare Substituenten am Steroidgrundgerüst, bindet hauptsächlich wiederum über die C- bzw. D-Ringatome und die Methylgruppe des Pentylsäurerests (Abbildung 4-20). Diese Substruktur repräsentiert 38% der Gesamtstruktur und ist für 59% der van der Waals-Interaktion verantwortlich.

Ursodesoxycholsäure, **26**



Lithocholsäure, **27**

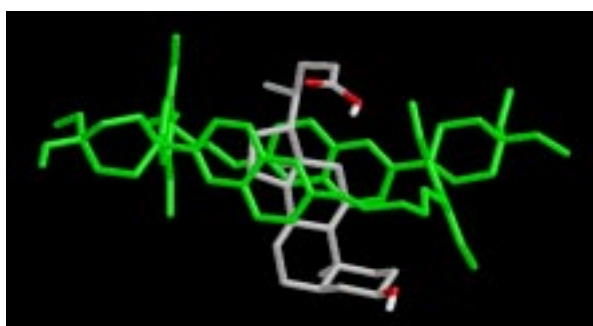


Abb. 4-20: Links: Simulierte Anordnung der Cyclophan-Steroid-Komplexe der Steroide **26** und **27** mit dem Cyclophan, **20**, rechts: das strukturelle Affinitätspotential. Rot eingefärbte Atome tragen zur van der Waals-Wechselwirkungsenergie mit > 2.0kcal/mol, grün eingefärbte mit >1.5 kcal/mol bei.

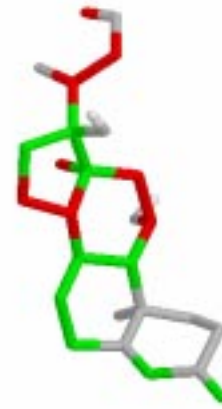
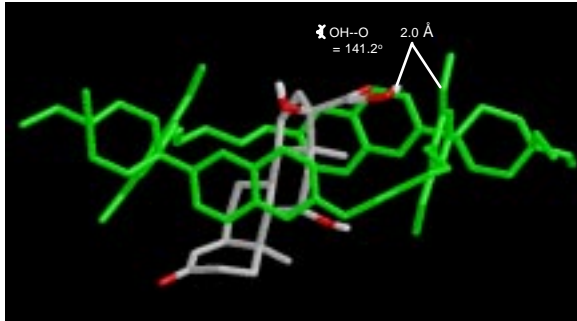
Zusätzlich zu den Cholsäureverbindungen **3** und **24** bis **27**, wurden Hydrocortison, **28**, Cortison, **29**, und Testosteron, **30**, in die Bindungstasche des Cyclophans, **20**, eingelagert (Abbildung 4-21).

Die niedrigste energetische Konfiguration des Hydrocortisonkomplexes zeigt die größten Wechselwirkungen im Bereich der B/C/D-Ringatome des Hydrocortisons, **28** (Abbildung 4-21). Die gesamte wechselwirkende Substruktur umfaßt 52% der Gesamtstruktur und ist für 72% der van der Waals-Interaktionen verantwortlich. Van-der-Waals Interaktionen von ungefähr 1.3 kcal/mol konnten für die Atome des Enonsystems des A-Rings ermittelt werden. Diese beiden Atome zeigen jedoch im Vergleich zu den elektrostatischen Wechselwirkungen von unter 0.1 kcal/mol, die in den Komplexen der Verbindungen **21** bis **27** auftreten, zusätzlich hohe elektrostatische Wechselwirkungen von 0.4 kcal/mol mit dem Cyclophanring. Der Hydrocortisonkomplex weist eine Wasserstoffbrücke der terminalen Hydroxygruppe des Pentylsäurerests am D-Ring zu dem Sauerstoffatom einer der Methoxygruppen der Naphthalinringe des Cyclophans, **20**, auf (d(H-O): 2.0Å, O-H--O Winkel: 141°, Abbildung 4-21).

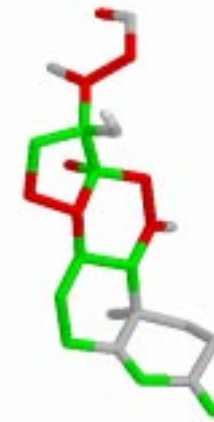
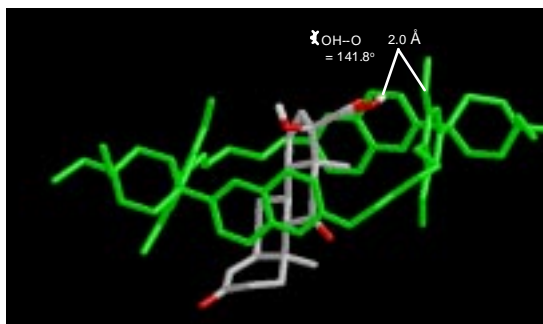
Ähnliche Ergebnisse wie für Hydrocortison, **28**, wurden für Cortison, **29**, erhalten (Abbildung 4-21). Die gleichen Atome wie in **28** tragen auch in **29** zu den stärksten Wechselwirkungen zwischen Steroid und Rezeptor bei. Eine Wasserstoffbrücke der terminalen Hydroxylgruppe des Pentylsäurerests am D-Ring zu einer der Methoxygruppen der Naphthalinringe des Cyclophans führt auch im Falle von **29** zu einer Stabilisierung des Komplexes (d(H-O): 2.0Å, O-H--O Winkel: 142°, Abbildung 4-21).

Testosteron, **30**, bindet hauptsächlich über die A-, B- und C-Ringatome zu dem Rezeptor (Abbildung 4-21). Die Ebenen des Steroids und der Rezeptornaphthalinringe sind nicht, wie in den bisherigen Fällen der Komplexe der Steroide **3** und **24** bis **29** parallel zueinander, sondern in einer ziemlich waagrechten Konfiguration in dem Cyclophanring eingebettet (Abbildung 4-21). 82% der Wechselwirkungen werden von 68% der Gesamtstruktur gestellt. Der elektrostatische Beitrag der Atome des Enonsystems zur insgesamt attraktiven Wechselwirkung zwischen Steroid und Cyclophan ist hier nur 0.1 kcal/mol. Die größte van der Waals-Wechselwirkung (> 2.0 kcal/mol) wird durch die gemeinsam dem A- und B-Ring angehörenden Atome und jenen, die direkt an diese gebunden sind, aufgebracht.

Hydrocortison, **28**



Cortison, **29**



Testosteron, **30**

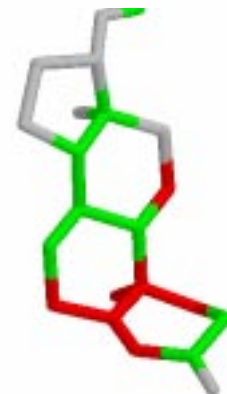
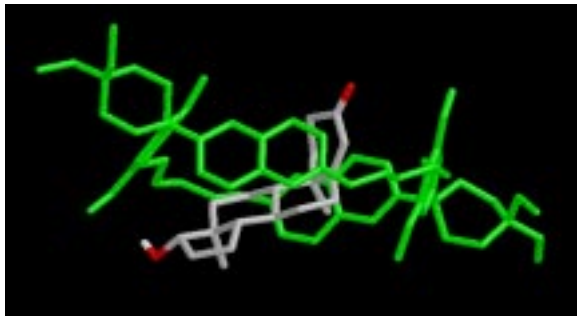


Abb. 4-21: Links: Simulierte Anordnung der Cyclophan-Steroid-Komplexe der Steroide **28**, **29** und **30** mit dem Cyclophan, **20**, rechts: das strukturelle Affinitätspotential. Rot eingefärbte Atome tragen zur van der Waals-Wechselwirkungsenergie mit >2.0 kcal/mol, grün eingefärbte mit >1.5 kcal/mol bei.

In weiteren Dockingsimulationen für die Steroide Digoxin, **2**, Ouabain, **23** und Progesteron, **31** (Abbildung 4-4 und 4-6) ließen sich deren Anlagerungsmodi im Cyclophanrezeptor, **20** vorhersagen (Abbildung 4-22). Von diesen Strukturen ist die hohe Affinität zu den monoklonalen Antikörperbindungsfragmenten des 26-10,^{[20][21]} 40-50,^{[25][26]} und DB3^[24] Immunglobulins bekannt. Für die Komplexe dieser Strukturen mit dem Paracyclophan, **20**, sind keine experimentellen Kristallstrukturdaten zum Vergleich verfügbar.

Digoxin, **2**, bindet über den größten Teil der Atome der A-, B- und C-Ringe, die Methylgruppe an der A/B-Ringverbindung und den Zuckerrest (Abbildung 4-22). 61.6% der gesamten van der Waals-Interaktionen resultieren dabei aus 34% der Atome der Gesamtstruktur. Die größten Interaktionen (>2.0kcal/mol) werden durch die A-Ringatome gestellt.

Ouabain, **23**, bindet in einem sehr ähnlichen Modus zu **2** an das Innere des Cyclophanrings. Der größte Teil der Atome der A-, B- und C-Ringe, die Hydroxylgruppen des A-Rings und der C/D-Ringverbindung sowie der Hydroxymethylsubstituent des B-Rings zeigen die höchsten Interaktionen mit dem Cyclophanring (Abbildung 4-22). Wechselwirkungen zwischen dem Inneren des Cyclophanrings und dem Zuckerrest des Steroids, **23**, konnten nicht ermittelt werden.

Progesteron, **31**, das ein Steroid mit kleinerer Atomanzahl, bindet fast über das gesamte Steroidgrundgerüst (Abbildung 4-22). 85% der gesamten van der Waals-Wechselwirkung wird durch 74% der Struktur gestellt. Die Orientierung des Steroids im Inneren des Rings ist analog zu Verbindung **30** (Abbildung 4-21) schräg bezüglich der Ebenen der Naphthalinringe (Abbildung 4-22). Die größten Interaktionen (>2.0kcal/mol) wurden für die A-Ringatome, eines der C-Ringatome und eines der B-Ringatome ermittelt.

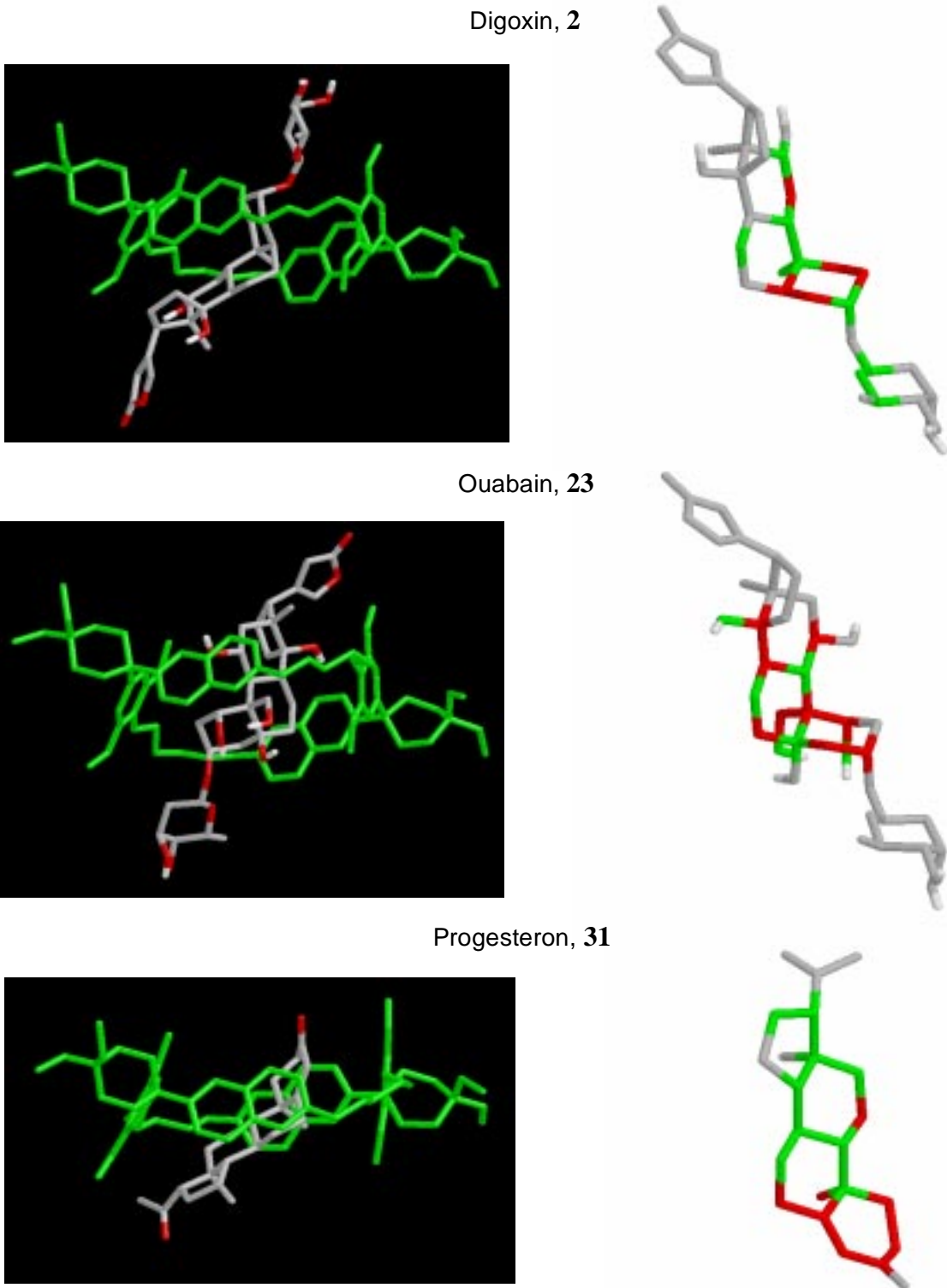


Abb. 4-22: Links: Simulierte Anordnung der Cyclophan-Steroid-Komplexe der Steroide **2**, **23** und **31** mit dem Cyclophan, **20**, rechts: das strukturelle Affinitätspotential. Rot eingefärbte Atome tragen zur van der Waals-Wechselwirkungsenergie mit $>2.0\text{kcal/mol}$, grün eingefärbte mit $>1.5\text{kcal/mol}$ bei.

Zusammenfassend kann festgestellt werden, daß Cholsäure, **3** und Desoxycholsäure, **24** eher über die Steroid-A- und B-Ringatome, als über die C- und D-Ringatome binden (Abbildung 4-19). Polare Substituenten am C-Ring von Verbindung **3** und **24** verhindern starke apolare Bindungen zu dem hydrophoben Inneren des Cyclophanrings, **20**, in diesen Bereichen. Im Gegensatz dazu scheinen die Steroide **25**, **26** und **27** die Wechselwirkung über ihre C- bzw. D-Ringe und Bereiche des Pentylsäurerests am D-Ring zu bevorzugen (Abbildung 4-20).

Die beiden Cortisonsteroide **28** und **29** und Testosteron, **30**, binden bevorzugt mit den A- und B-Ringen an den apolaren Rezeptor (Abbildung 4-21). Für die Stabilisierung dieser Komplexe sind zusätzliche elektrostatische Wechselwirkungen des Enonsystems des A-Rings mit dem elektronenreichen aromatischen System des Cyclophanrings verantwortlich.

Diese qualitativen Dockingstudien erlauben die Ermittlung von strukturellen Bereichen der Liganden, die eine besonders hohe oder niedrige attraktive Wechselwirkung mit einem entsprechenden Rezeptor aufweisen. Um quantitative Stabilitätsdaten der Komplexe der Verbindungen **2** und **23** bis **31** zu erhalten, wurden die im folgenden dargestellten Berechnungen der freien Bildungsenthalpien der Komplexe der Steroide mit dem Paracyclophan, **20**, durchgeführt.

4.3.3 Die freien Bildungsenthalpien der Cyclophan-Steroid-Komplexe

Die freien Bildungsenthalpien ließen sich für die Steroid-Cyclophan-Komplexe der Steroide **2** und **23** bis **31** im Komplex mit dem Paracyclophan, **20**, mit Hilfe der *Linear Interaction Annäherung* berechnen. Eine Aufstellung der berechneten und experimentell^[22] ermittelten freien Bildungsenthalpien ΔG_{calc} und ΔG_{exp} ist in Tabelle 4-2 gegeben. Die freien Bildungsenthalpien wurden nach Gl. 4-1 berechnet. Für die Steroide **3** und **24** bis **27** wurde für α ein Wert von 0.31 ermittelt, wogegen für die ungesättigten Steroidverbindungen **28**, **29** und **30** ein höherer Wert von 0.52 für α berechnet wurde.

Tab. 4-2: Berechnete und experimentelle freie Bildungsenthalpien ΔG für die Komplexe der Steroide mit dem Paracyclophan, **20**, in kcal/mol.

Verbindung	$\langle V_{\text{Wasser}}^{\text{elek}} \rangle$	$\langle V_{\text{Antik}}^{\text{elek}} \rangle$	$\langle V_{\text{Wasser}}^{\text{vdW}} \rangle$	$\langle V_{\text{Antik}}^{\text{vdW}} \rangle$	ΔG_{exp}	ΔG_{calc}
3	-73.9±0.3	-55.7±0.9	-26.0±0.2	-46.0±0.5	-2.9	2.8±1.1
24	-56.5±0.3	-55.4±0.2	-27.4±0.3	-47.0±0.0	-3.2	-5.6±0.9
25	-55.8±0.5	-52.0±0.1	-28.2±0.0	-46.4±0.2	-3.9	-3.8±0.9
26	-55.5±0.3	-54.1±0.8	-28.9±0.2	-43.7±1.1	-4.4	-3.9±0.7
27	-41.1±0.4	-39.0±0.4	-30.7±0.0	-46.3±0.0	-5.2	-3.9±0.9
28	-64.2±2.6	-52.0±1.2	-23.9±0.1	-45.5±0.0	-4.1	-5.0±3.9
29	-64.7±0.8	-52.3±2.5	-24.2±0.3	-40.7±0.5	-4.3	-2.2±2.4
30	-69.7±0.8	-66.1±1.1	-21.8±0.3	-36.4±0.2	-4.8	-5.7±2.1
2	-92.3±1.8	-72.3±0.0	-35.5±0.1	-57.3±0.0	-	2.2±0.8
23	-131.6±1.3	-102.9±2.8	-32.2±0.2	-54.9±0.7	-	6.9±2.9
31	-32.3±0.2	-32.3±0.0	-28.0±0.0	-42.3±0.2	-	-7.3±1.3

Die V_{vdw} und V_{elek} -Werte sind die durchschnittlich ermittelten Werte der van der Waals und elektrostatischen Wechselwirkung zwischen Ligand und Rezeptor während der Simulation. Diese durchschnittlichen Werte basieren auf *Molecular Dynamics* Simulationen, die einerseits für den Liganden umgeben von reinem Solvens und andererseits im Komplex mit dem Rezeptor durchgeführt wurden.

Die durchschnittlichen Werte für die freien Bildungsenthalpien ΔG_{calc} lassen folgende Schlüsse zu: Die Bindungsaffinität zum Paracyclophanrezeptor, **20**, steigt bezüglich der Cholsäure-ähnlichen Verbindungen von der Cholsäure, **3**, hin zur Lithocholsäure, **27**, leicht an, das gleiche gilt für Hydrocortison, **28**, bis Testosteron, **30**. Die mittlere absolute Abweichung zwischen den experimentellen und berechneten Werten ist 1.2 kcal/mol, mit Ausnahme von Verbindung **3**. Diese Abweichung entspricht den bei Åqvist et al. angeführten mittleren Fehler bei Berechnungen der freien Bildungsenthalpie solcher Komplexe. Da die experimentellen Werte in einem Solvensgemisch Wasser/Methanol 1:1 gemessen wurden,^[22] die Rechnungen dagegen auf einem reinen Wassersolvensmodell basierten, kommt es zu Abweichungen zwischen den experimentellen und berechneten Werten der freien Bildungsenthalpien. Der Grund dafür ist die Desolvatisierungsenergie, die in dem stärker polaren Solvens Wasser, im Vergleich zu dem weniger polaren Solvensgemisch Wasser/Methanol, zunimmt. Die Werte der vorhergesag-

ten freien Bildungsenthalpien der Komplexe der Steroide Digoxin, **2**, Ouabain, **23** und Progesteron, **31** mit **20** sind in Tabelle 4-2 dargestellt. Für diese Verbindungen konnten keine experimentellen Vergleichswerte erhalten werden.

Die größte Bedeutung für eine hohe Affinität zu dem Cyclophanrezeptor haben apolare Interaktionen der Oberflächen der Steroide mit dem Inneren des Cyclophanrings. Zieht man die Bindungskonfigurationen der Komplexe aus Abbildung 4-19 bis 4-22 und die jeweiligen Substrukturen der Steroide, die die größten Wechselwirkungen aufweisen in Betracht, so ist der strukturelle Bereich mit der höchsten Wechselwirkung grundsätzlich der Steroidgrundkörper selbst. Umso mehr Hydroxylgruppen am Grundgerüst gebunden sind, desto polarer wird die Verbindung und desto höher ist die Desolvatisierungsenergie in Wasser. Die Strukturen werden daher in drei Gruppen unterteilt, wobei sich die Zugehörigkeit nach der jeweiligen Anzahl an Hydroxylgruppen am Steroidgrundkörper ergibt (Tabelle 4-3). In jeder Gruppe ist die freie Bildungsenthalpie der jeweiligen Verbindungen im Komplex mit dem Paracyclophan, **20**, sehr ähnlich.

Die erste Gruppe umfaßt die Verbindungen mit den höchsten negativen freien Bildungsenthalpien, diese besitzen keine Hydroxylsubstituenten am Steroidgrundgerüst. Die Lithocholsäure, **27** weist eine experimentelle freie Bildungsenthalpie ΔG_{exp} von -5.2 kcal/mol ($\Delta G_{\text{calc}} = -3.6$ kcal/mol) auf und Testosteron, **30**, von -4.8 kcal/mol ($\Delta G_{\text{calc}} = -5.7$ kcal/mol), wobei die berechnete Bildungsenthalpie für **31** mit $\Delta G_{\text{calc}} = -7.3$ kcal/mol noch höher ist. Der Grund für die hohen berechneten freien Bildungsenthalpien der beiden Verbindungen **30** und **31**, liegt an ihrer Schrägstellung innerhalb des Cyclophanrings in der simulierten Dockingkonfiguration (Abbildung 4-21 und 4-22). Bei dieser Anordnung führt fast die gesamte Struktur (**30**: 68%, **31**: 74%) attraktive apolare Wechselwirkungen mit dem Rezeptor aus. Im Gegensatz dazu, ist die Stellung von **27** senkrecht innerhalb des Rings, so daß ein weit kleinerer Teilbereich der Struktur (38%) Kontakt zu der apolaren Rezeptorhülle hat, was ein Grund für die niedrigere berechnete Bildungsenthalpie ist.

Tab. 4-3: Einteilung der Verbindungen in drei Hauptgruppen, abhängig von der Anzahl an OH-Gruppen am Steroidgrundgerüst. Die freien Bildungsenthalpien ΔG sind für das Paracyclophan, **20**, in kcal/mol gegeben.

Gruppe 1	Verbindungen				OH-Gruppen
	27	30	31		keine
ΔG_{exp}	-5.2	-4.8	-		
ΔG_{calc}	-3.6	-5.7	-7.3		
Gruppe 2	Verbindungen				OH-Gruppen
	24	25	26	28	eine
ΔG_{exp}	-3.2	-3.9	-4.4	-4.1	
ΔG_{calc}	-5.6	-3.8	-3.9	-5.0	
Gruppe 3	Verbindungen				OH-Gruppen
	2	23	3		zwei oder mehr
ΔG_{exp}	-	-	-2.9		
ΔG_{calc}	> 0.0	> 0.0	> 0.0		

Die zweite Gruppe repräsentiert Verbindungen mit einer Hydroxylgruppen am steroidalen Grundkörper. Die Beträge der freien Bildungsenthalpien der Steroide **24**, **25**, **26** und **28** in Gruppe 2 sind niedriger als die in Gruppe 1 und bewegen sich zwischen -3.2 kcal/mol bis -4.4 kcal/mol in den experimentellen Werten und -3.8 kcal/mol bis -5.6 kcal/mol in den berechneten Werten. Bereiche von durchschnittlich ungefähr 40% der Strukturen in dieser Gruppe zeigen attraktive Wechselwirkungen mit der apolaren Bindungstasche (Abbildung 4-19, 4-20 und 4-21).

Die dritte Gruppe umfaßt alle Steroide mit zwei oder mehr Hydroxylgruppen am Steroidgrundkörper. Die Berechnungen der freien Bildungsenthalpien ergaben positive Werte; die Komplexierungen scheinen daher endergonische Prozesse zu sein. Im Falle von **3** ist die gemessene freie Bildungsenthalpie von -2.9 kcal/mol relativ klein und die Hydroxylgruppe in axialer Stellung am B-Ring führt sogar zu einer repulsiven Wechselwirkung (Abbildung 4-19). Wie bereits in Abbildung 4-22 gezeigt, bindet Ouabain, **23** mit dem A-, B- und C-Ring an den apolaren Rezeptor. Gerade dieser Strukturbereich weist eine sehr hohe Anzahl an polaren Substituenten auf. Wahrscheinlich verhindern die Hydroxylgruppen in diesem Bereich eine hohe

Rezeptor-Ligand-Affinität, die zu der berechneten positiven freien Bildungsenthalpie führt.

Ein Vergleich der gemessenen und berechneten freien Bildungsenthalpien zeigt, daß während der *Linear Response Approximation* anscheinend die polar-apolaren repulsiven Wechselwirkungen leicht überschätzt wurden. Zusätzlich ist die Desolvatisierungsenergie bei einer größeren Anzahl an polaren Substituenten höher. Die Bindungsaffinitäten der Steroide zu dem apolaren Cyclophanring, **20**, nehmen mit einer zunehmenden Anzahl an polaren Substituenten am Steroidgrundkörper ab. Eine starke Ligand-Rezeptor-Wechselwirkung basiert hier auf apolaren van der Waals-Interaktionen.

4.3.4 Die physikochemischen Oberflächeneigenschaften der Steroide

Um Ähnlichkeiten der physikochemischen Oberflächeneigenschaften der Liganden zu untersuchen, wurde ein Ansatz mit Hilfe von Kohonen neuronalen Netzen gewählt.

Die elektrostatischen Potentiale und Hydrophobizitäten auf der molekularen Oberfläche der Steroide **2**, **3** und **23** bis **31** wurden untersucht um eine Relation zu den Affinitäten dieser Steroide zum Cyclophanrezeptor, **20**, zu ermitteln.

Die Oberflächeneigenschaften der Cholsäure, **3**, wurden als Beispiel zur detaillierten Beschreibung der Kohonen-Karten gewählt. In Abbildung 4-23 sind die Kohonen-Karten des elektrostatischen Potentials (ESP) (oben) und des Hydrophobizitäts-Potentials (HYP) (unten) von **3** dargestellt. Die Oberflächen wurden mit einer Dichte von 50 Punkten pro Å² berechnet. Jede im folgende gezeigte Kohonen-Karte besteht aus 50 x 50 Neuronen. Die Netztopologie ist toroidal.

Rote Bereiche der ESP-Karte stehen für ein stark negatives elektrostatisches Potential, z.B. das eines Sauerstoffatoms einer Hydroxyl- oder Carbonylgruppe, wogegen violette Bereiche ein hohes positives elektrostatisches Potential repräsentieren, z.B. Kohlenstoffatome einer stark polarisierten Carbonylgruppe. Violette und blaue Bereiche der HYP-Karte stehen für hohe Hydrophobizitäten, z.B. eines reinen Kohlenstoffgerüsts ohne Heteroatome, wogegen rote Bereiche einer niedrigen Hydrophobizität entsprechen, z.B. einer Aminogruppe.

Die Skalierung des ESP und des HYP umfaßt die Bereiche zwischen dem maximalen und minimalen Wert der entsprechenden Oberflächeneigenschaften der Cholsäure, **3**. In diesem Fall liegen die Werte zwischen -46.89 kcal/mol und 36.62 kcal/mol für das ESP und zwischen -0.01 bis 0.1 für das HYP.

Die Cholsäure, **3**, weist vier Bereiche mit stark negativem ESP auf; dies entspricht den Sauerstoffatomen der drei OH-Gruppen und der COOH-Gruppe. Die Kohonen-Karte des ESP

in Abbildung 4-23 zeigt entsprechend dieser vier Bereiche, vier voneinander isolierte Regionen in rot-gelber Farbe. Das Carbonylsauerstoffatom der COOH-Gruppe hat das stärkste negative ESP und entspricht der roten Region auf der rechten Seite der Karte. Die COOH-Gruppe und die OH-Gruppen des B- und C-Rings sind durch größere apolare Regionen räumlich voneinander getrennt. Dieser apolare Bereich entspricht dem D-Ring, der keine polaren Substituenten aufweist und ist in der Mitte der Kohonen-Karte als grüne Region zu erkennen. Hohe positive ESP-Werte, die sich den Wasserstoffatomen der OH-Gruppen zuordnen lassen, erkennt man auf der Kohonen-Karte an violett eingefärbten Regionen.

Die polaren Hydroxylsubstituenten der A-, B- und C-Ringe und die COOH-Gruppe des Pentylsäurerests des D-Rings weisen sehr niedrige HYP-Werte auf, die auf der entsprechenden Kohonen-Karte rot eingefärbt sind. Der Steroidgrundkörper, der generell apolar und stark hydrophob ist, wird durch grüne, blaue und violette Bereiche der HYP-Karte repräsentiert.

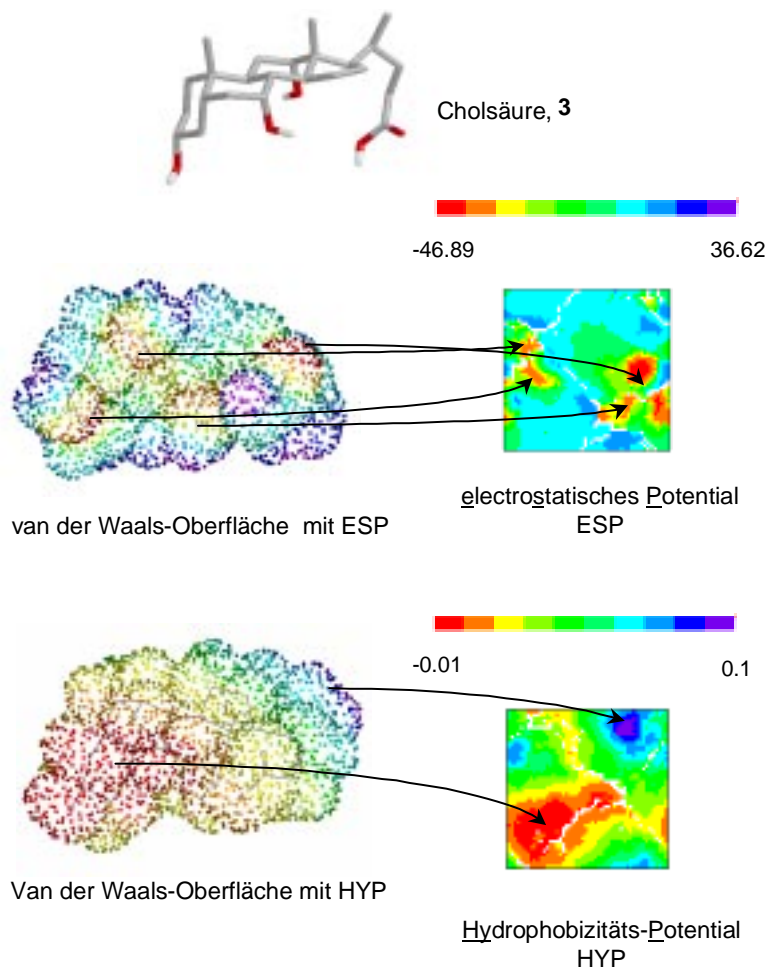


Abb. 4-23: Die Projektion des ESP und HYP auf der Oberfläche der Cholsäure, 3, in zweidimensionale Kohonenkarten.

Im folgenden sollen die elektrostatischen Oberflächenpotentiale der Liganden **2** und **23** bis **31** für den Paracyclophanrezeptor, **20**, näher untersucht werden.

Zum Training der Kohonen-Karten wurden die absoluten dreidimensionalen Koordinaten der behandelten Steroide den Ergebnissen der Docking Experimente mit AutoDock aus Kapitel 4.2.1 entnommen.

Abbildung 4-25 zeigt die einzelnen Kohonen-Karten der elektrostatischen Oberflächenpotentiale der Steroide **2** und **23** bis **31**. Für jede Struktur wurde ein einzelnes Netzwerk trainiert. Die einheitliche Skalierung des elektrostatischen Potentials auf den Kohonen-Karten läuft vom Minimal- (-46.89 kcal/mol) zum Maximalwert (36 kcal/mol) des ESP des gesamten Datensatzes.

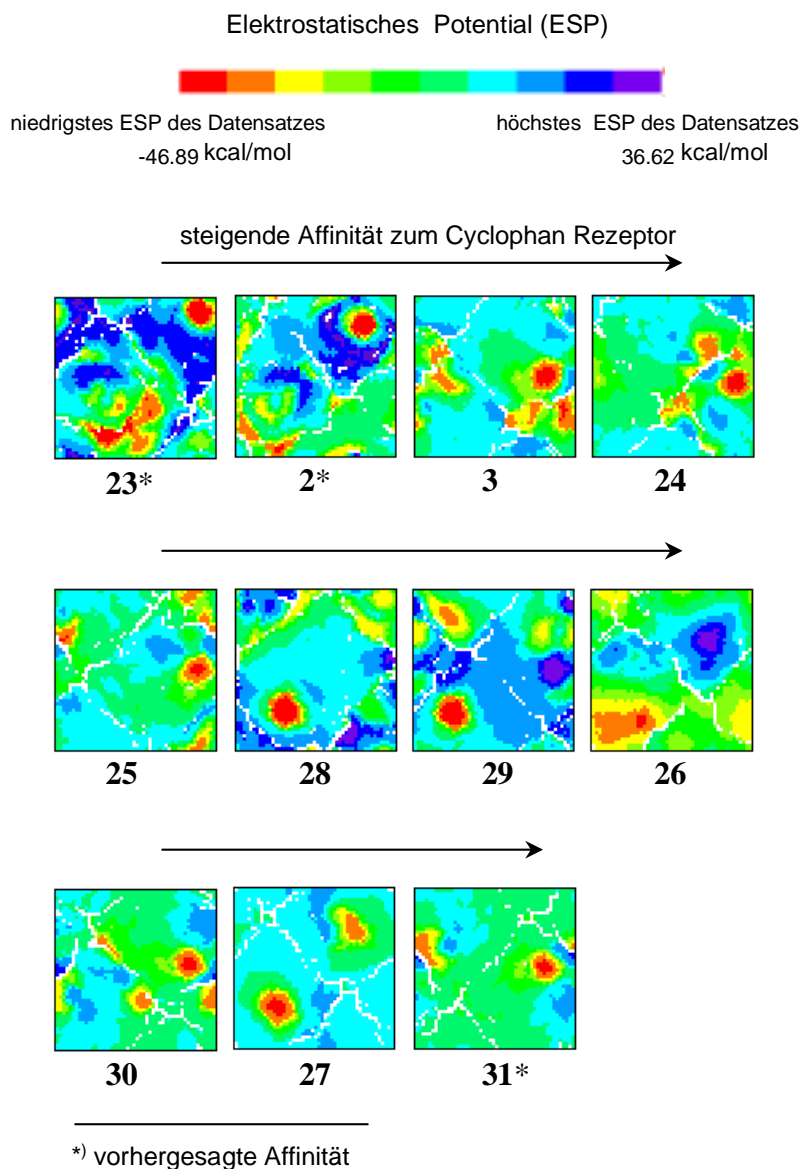


Abb. 4-24: Die Kohonen-Karten des elektrostatischen Oberflächenpotentials der Verbindungen **2** und **23** bis **31**. Die Reihenfolge der Karten der Liganden entspricht ihrer steigenden Affinität zum Cyclophanrezeptor, **20**. Rote und orange Bereiche entsprechen einem stark negativen ESP, violette und blaue Bereiche einem stark positiven ESP. Die mittlere Skala von hellgrün bis hellblau repräsentiert strukturelle Regionen mit geringem negativen oder positiven ESP.

Unter der Voraussetzung, daß die Rezeptorbindungstasche apolar ist, dürften Liganden mit hoher Affinität zu dem Rezeptor keine großen Bereiche von hohem negativem oder positivem elektrostatischem Potential aufweisen. In Tabelle 4-4 ist die freie Bildungsenthalpie ΔG der Verbindungen gegenüber der Anzahl der Neuronen, die auf den Karten aus Abbildung 4-24 geringe ESP-Werte präsentieren, aufgetragen. Diese sind hellgrün und hellblau eingefärbte Neuronen beziehungsweise Neuronen aus den Farbsegmenten fünf bis sieben (von links in der

Farbskala in Abbildung 4-24). Die ESP-Werte der betrachteten Neuronen laufen zwischen -14 kcal/mol bis +12 kcal/mol. Die Werte der freien Bildungsenthalpie entsprechen den experimentell gemessenen ΔG_{exp} , beziehungsweise für Verbindung **2**, **23** und **31** den berechneten ΔG_{calc} Werten (vgl. Tabelle 4-2).

Tab. 4-4: Anzahl der Neuronen, die nicht elektrostatische Bereiche der Kohonen-Karten aus Abbildung 4-25 markieren.

Verbindung	experimentelle ΔG_{exp} -Werte in kcal/mol	Anzahl der Neuronen wenig elektrostatischer Bereiche
23 *	6.9	1972
2 *	2.2	2107
3	-2.9	2078
24	-3.2	2184
25	-3.9	2180
28	-4.1	2002
29	-4.3	1957
26	-4.4	2176
30	-4.8	2191
27	-5.2	2230
31 *	-7.3	2274

*. berechnete freie Bildungsenthalpie ΔG

Es ist eindeutig der Trend zu erkennen, daß die Anzahl der Neuronen, die wenig elektrostatische Bereiche präsentieren, in Richtung Zunahme der Rezeptoraffinität ebenfalls ansteigt (1972 Neuronen für Verbindung **23** gegenüber 2274 Neuronen für Verbindung **31**). Bei der Betrachtung des ESP ist es jedoch schwierig, eindeutige Aussagen zu treffen und die Grenzen zwischen schwach und stark negativem oder positivem elektrostatischen Potential zu ziehen. Schließlich liegt einem hohen negativen Potential und einem starken positivem Potential eine grundverschiedene funktionelle Gruppe oder Atomanordnung zugrunde.

Daher werden zum Vergleich der physikochemischen Oberflächeneigenschaften der Liganden im folgenden das molekulare Hydrophobizitäts Potential herangezogen, das sich von stark hydrophilen bis stark hydrophoben Regionen erstrecken kann, das heißt nicht die mittle-

ren Bereiche der Hydrophobizitätsskala, sondern die äußeren Grenzbereiche sind für die Interpretation der Eigenschaft ausschlaggebend. Zusätzlich sollten sich wenig elektrostatische Bereiche mit hydrophoben Bereichen der Strukturen decken, da diese keine polaren Substituenten aufweisen dürfen.

Abbildung 4-25 zeigt die einzelnen Hydrophobizitätskarten der Steroide **2** und **23** bis **31**. Für jede Struktur wurde wiederum ein individuelles Netzwerk trainiert. Die einheitliche Skalierung der Hydrophobizität der Kohonen-Karten läuft vom Minimal- bis Maximalwert des HYP des gesamten Datensatzes, d.h. von $HYP = -0.09$ (rot) bis $HYP = 0.14$ (violett). Sofort ersichtlich wird die geringere Hydrophobizität von **2** und **23**, deren Karten sehr viele rote und grüne Bereiche aufweisen, im Gegensatz zu **27** und **31**, deren Karten vor allem große blaue und violette Bereiche zeigen (Abbildung 4-25).

Die Reihenfolge der Karten in Abbildung 4-25 entspricht einer steigenden Affinität der Liganden zu dem apolaren Cyclophanrezeptor, **20**, basierend auf den experimentellen Werten oder, falls diese nicht vorhanden waren (Verbindungen **2**, **23** und **31**), aus den berechneten Werten der freien Bildungsenthalpien ΔG aus Tabelle 4-2.

Ein quantitativer Wert der Hydrophobizität eines jeden Liganden wurde durch die Anzahl der Neuronen, die hydrophobe Bereiche in den Kohonen-Karten markieren, erhalten. Die Größe der hydrophoben Regionen ist proportional zur Anzahl an blau und violett eingefärbten Neuronen, das sind jene, die in der Skalierung des HYP-Wertes in den letzten drei Segmenten der Farbpalette liegen und damit Werte zwischen 0.05 und 0.14 aufweisen.

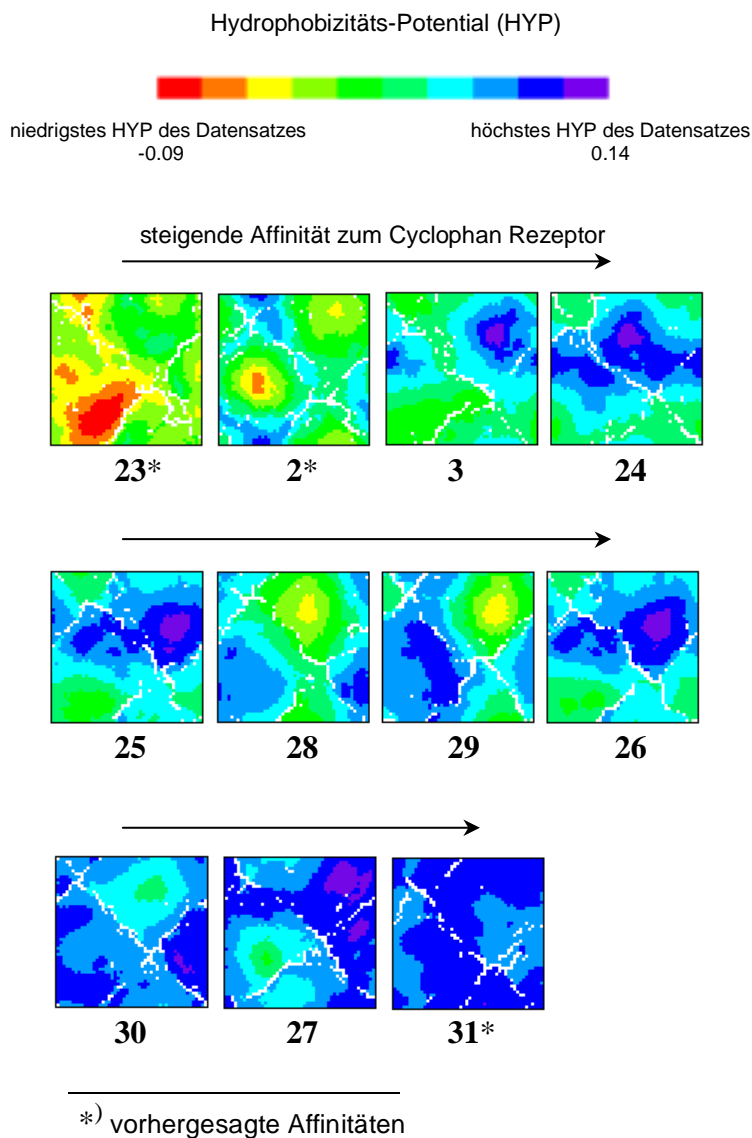


Abb. 4-25: Die Kohonen-Karten des Hydrophobizitätspotential der Verbindungen **2** und **23** bis **31**. Die Reihenfolge der Karten der Liganden entspricht ihrer steigenden Affinität zum Cyclophanrezeptor, **20**. Blaue Regionen markieren hohe Hydrophobie, wogegen rote Regionen hohe Hydrophilie darstellen.

In Tabelle 4-5 ist die jeweilige Anzahl der Neuronen jeder Karte aus Abbildung 4-25, die hoch hydrophobe strukturelle Regionen repräsentieren, aufgelistet.

Tab. 4-5: Anzahl der Neuronen, die hydrophobe Bereiche der Kohonen Karten aus Abbildung 4-25 markieren.

Verbindung	ΔG in kcal/mol	Anzahl hydrophober Neuronen
23 *	6.9	0
2 *	2.2	227
3	-2.9	536
24	-3.2	1105
25	-3.9	1121
28	-4.1	716
29	-4.3	1192
26	-4.4	1158
30	-4.8	1859
27	-5.2	1813
31 *	-7.3	2363

*. berechnete freie Bildungsenthalpie ΔG

Abbildung 4-26 zeigt die Auftragung von ΔG in kcal/mol gegen die Anzahl der hydrophoben Neuronen. Die Affinität der Steroide zu dem Cyclophanring, **20**, ausgedrückt in der freien Bildungsenthalpie ΔG der jeweiligen Verbindung, korreliert mit der Anzahl der hydrophoben Neuronen in den Kohonen-Karten. Die Berechnung des Korrelationskoeffizienten nach Pearson ergibt $r = -0.83$, das Quadrat des Korrelationskoeffizienten ergibt $r^2 = 0.69$. Der Standardfehler von ΔG läßt sich zu 2.3 kcal/mol unter Berücksichtigung aller Liganden aus Tabelle 4-5 berechnen und zu 0.5 kcal/mol ohne Einbezug der berechneten Bildungsenthalpien von **23**, **2** und **31**.

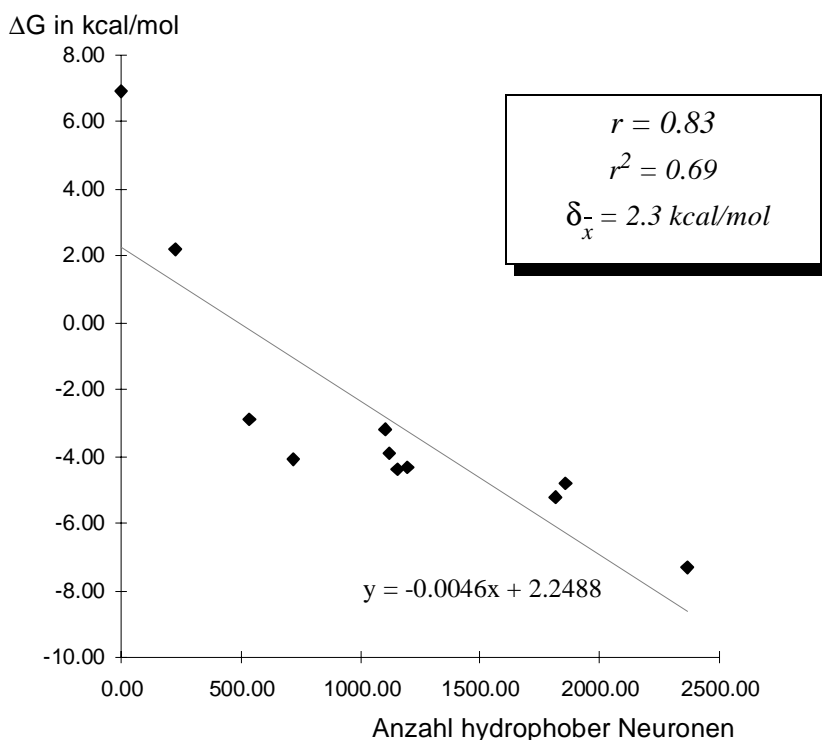


Abb. 4-26: Die freie Bildungsenthalpie ΔG [kcal/mol] angetragen gegen die Anzahl der hydrophoben Neuronen der Kohonen-Karten aus Abbildung 4-25. Der Korrelationskoeffizient r ergibt sich zu 0.83, r^2 zu 0.69 und der Standardfehler von ΔG zu 2.3 kcal/mol.

Die HYP Karten von Digoxin, **2** und Ouabain, **23** in Abbildung 4-25 zeigen, daß diese Verbindungen die hydrophilsten des Datensatzes sind. Der Minimalwert ihrer Hydrophobizität liegt bei -0.09. Die gelb-roten Bereiche in der Karte von **2** entsprechen dem Zuckerrest am A-Ring des Steroidgerüsts. Dieser Zuckerrest ist die hydrophilste Region der gesamten Verbindung. Dies entspricht auch den experimentellen und berechneten Bindungsmodi von **2** innerhalb der 26-10-Fab Bindungstasche. Die Bindungstasche des 26-10 Antikörpers ist ziemlich apolar, somit ragt der hydrophile und polare Zuckerrest in Richtung des polaren Lösungsmittels aus der Bindungstasche heraus und wird durch polare Wechselwirkungen mit dem Solvens stabilisiert (Abbildung 4-14 und 4-15). Die größeren roten Bereiche in der HYP-Karte von **23** entsprechen den drei Hydroxylgruppen am A-Ring des Steroidgrundkörpers und wiederum dem Zuckerrest.

Abbildung 4-27 zeigt die van der Waals-Oberfläche und zugehörige Kohonen-Karte des Paracyclophanrezeptors und des darauf projizierten Hydrophobizitäts Potentials. Die Skalierung läuft von HYP = 0.02 bis HYP = 0.14. Die Oberflächenhydrophobizität des Rezeptors bewegt sich daher im höher hydrophoben Bereich. Die van der Waals-Oberfläche wurde mit einer Dichte von 10 Punkten per \AA^2 berechnet. Die hydrophilsten Bereiche des Rezeptors (rot)

sind die beiden Ammoniumgruppen in den Cyclohexanresten. Strukturelle Bereiche hoher Hydrophobie (blau und violett) stellen die beiden aromatischen Naphthylmethaneinheiten dar.

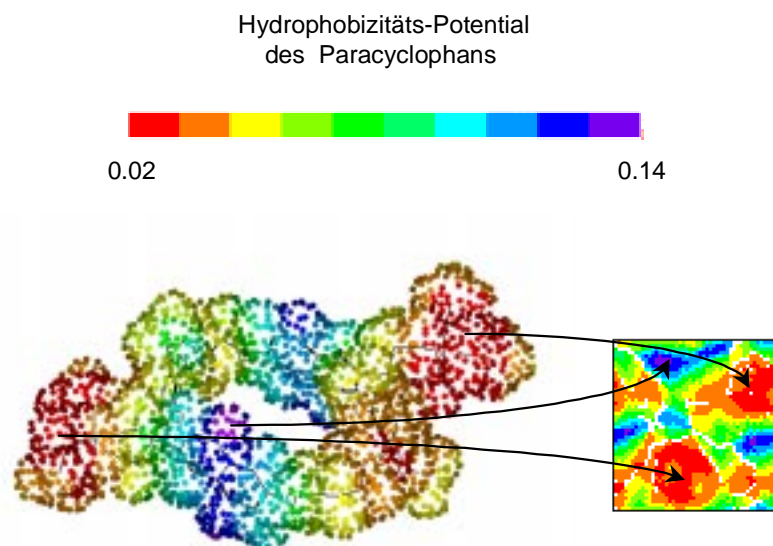


Abb. 4-27: Das Hydrophobizitätspotential des Cyclophanrezeptors, **20**. Links: auf der van der Waals-Oberfläche, rechts: in eine Kohonen-Karte projiziert.

Durch das mit den dreidimensionalen Koordinaten der van der Waals-Oberfläche des Rezeptors trainierte Netz (Abbildung 4-27), werden im folgenden sukzessive die Koordinaten der dreidimensionalen Oberflächen der Liganden, wie sie aus den Dockingsimulationen erhalten wurden, durchgeschickt. Die erstellte Kohonen-Karte des Rezeptors agiert somit als Templat. Die Anfärbung der erhaltenen Karten ergibt sich durch die Hydrophobizitäts Potentiale der Liganden (Abbildung 4-28). Die Skalierung des Hydrophobizitäts Potential läuft wiederum vom Minimal- bis zum Maximalwert der Hydrophobie des betrachteten Datensatzes unter Einbezug des Rezeptors.

Die erlaubte maximale euklidische Distanz zwischen den Oberflächenpunkten der Ligandenbindungstasche des Rezeptors und der eingelagerten Liganden wurde als 1.0\AA angegeben. Dies bedeutet, daß jeder Oberflächenpunkt des Rezeptors, der eine Distanz von größer 1.0\AA zu einem Oberflächenpunkt des jeweiligen Liganden aufweist zu leeren Neuronen führt. Leere Neuronen sind Neuronen für die kein ähnlichster Eingabevektor gefunden werden konnte; sie bilden daher weiße Flächen auf der Kohonen-Karte. Die gefärbten Bereiche der Karten aus Abbildung 4-28 repräsentieren die Kontaktfläche des jeweiligen Liganden mit dem Cyclophanrezeptor, **20**. Daher erlaubt die Templat-Methode die Visualisierung der Kontaktfläche und zusätzlich die physikochemischen Oberflächeneigenschaften der Liganden genau an

diesen Kontaktstellen.

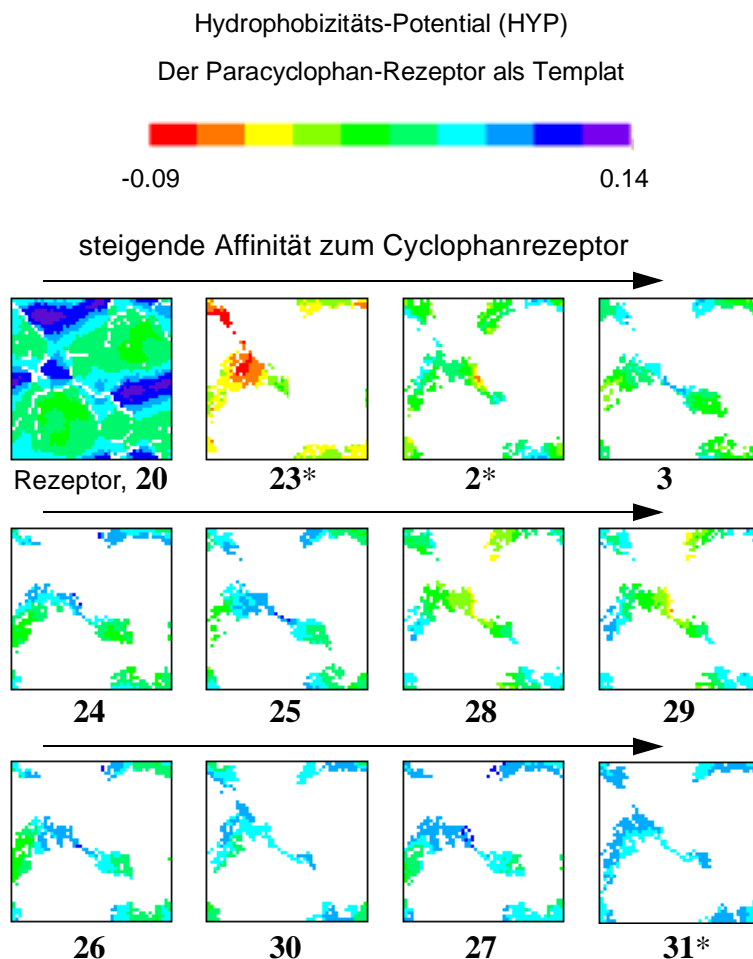


Abb. 4-28: Die Kohonen-Karten des Hydrophobizitäts Potentials der Liganden **2** und **23** bis **31**, erstellt mit Hilfe der Templat Methode. Als Templat wurde der Cyclophanrezeptor, **20**, gewählt.

In allen Fällen mit Ausnahme von Verbindung **23** liegt die physikochemische Eigenschaft an der Kontaktfläche im mittleren bis oberen Bereich der Hydrophobie. Dies zeigt wiederum, daß hier die Affinitäten zwischen Rezeptor und Ligand auf apolaren van der Waals-Wechselwirkungen basiert.

Die folgende Abbildung 4-29 zeigt eine Rückprojektion zweier Kohonen-Karten auf die dreidimensionale Oberfläche des Cyclophanrezeptors, **20**. Die Skalierung des Hydrophobizitäts Potentials läuft in diesem Fall vom niedrigsten bis zum höchsten HYP des Rezeptors und des betrachteten Liganden, **27**.

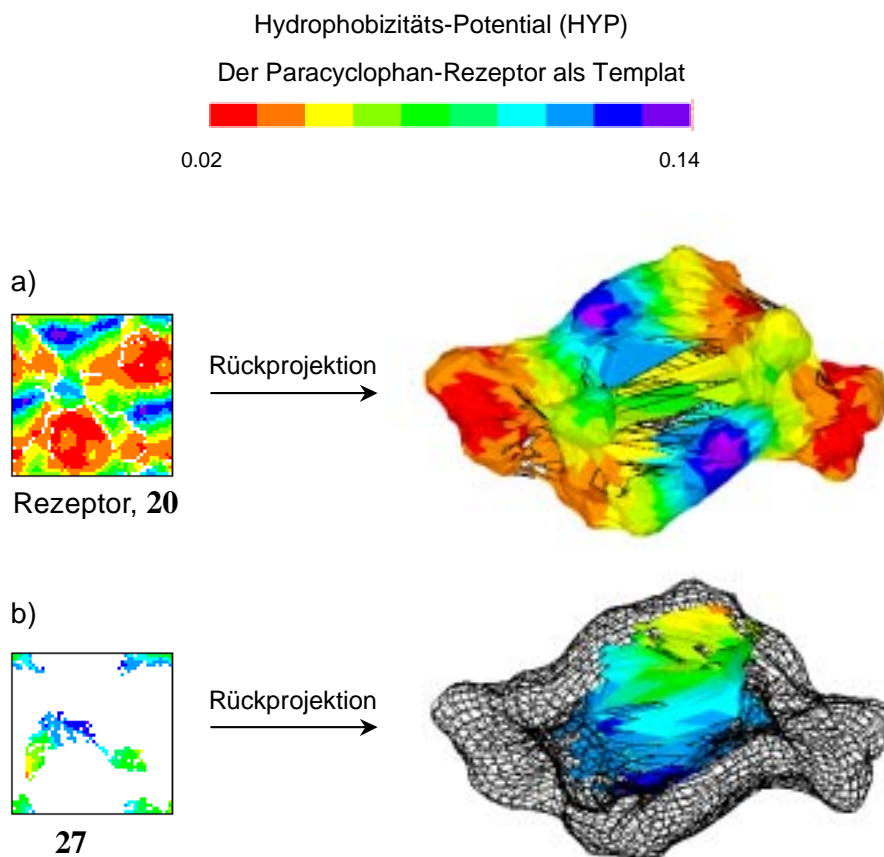


Abb. 4-29: a) Die Rückprojektion der HYP-Kohonen-Karte des Cyclophans, **20**, auf die dreidimensionale Oberfläche des Cyclophans. b) Die Rückprojektion der HYP-Kohonen-Karte von **27** auf die dreidimensionale Oberfläche des Cyclophans.

Im ersten Fall wird die Rückprojektion der Hydrophobizitäts Karte des Cyclophans auf die Oberfläche des Cyclophans gezeigt (Abbildung 4-29 a). Die genaue dreidimensionale Form des Cyclophans konnte zurück erhalten werden. Die entstandenen Abweichungen von der realen Form des Cyclophanrings, basieren auf der Verzerrung des Torus bei der ursprünglichen Projektion in die zweidimensionale Kohonen-Karte (vgl. Abbildung 4-9 und 4-12).^[18]

Im zweiten Fall wird die HYP-Karte des Steroids **27**, die durch die anhand Abbildung 4-28 beschriebene Templat-Methode erhalten wurde, auf die dreidimensionale Oberfläche des Cyclophanrezeptors zurückprojiziert (Abbildung 4-29 b). Diese Rückprojektion ist eine weitere anschauliche Methode, die Kontaktflächen zwischen Rezeptor und Ligand, hier zwischen dem Cyclophanrezeptor, **20**, und dem eingelagerten Steroid z.B. **27**, sichtbar zu machen.

In Bereichen geringer Distanzen oder Überlappung der Oberflächen von Ligand und Rezeptorring ist das Netz mit der entsprechenden Farbe des dort befindlichen Hydrophobizitätspotentials eingefärbt (Inneres des Cyclophanrings), wogegen kontaktlose Stellen durch ein

durchsichtiges Gitternetz dargestellt werden (Äußeres des Cyclophanrings). Die Durchsichtigkeit entspricht leeren Neuronen, das heißt weißen Flächen der ursprünglichen Karte. Das Hydrophobizitäts Potential des Liganden **27** innerhalb des Rezeptornetzes zeigt, daß der hydrophobe Körper des Liganden in blau und violett innerhalb des Rezeptorrings zum liegen kommt, wogegen die eher hydrophilen Bereiche in gelb und orange in Richtung der äußeren Rezeptorbereiche zeigen (Abbildung 4-29 b).

5 Zusammenfassung

5.1 Der Hybridalgorithmus zur Überlagerung konformativ flexibler Strukturen

In der Aufstellung von Struktur-Aktivitätsbeziehungen für die Suche nach neuen Leitstrukturen in der pharmazeutische Wirkstoffforschung ist die Analyse molekularer Ähnlichkeiten ein grundlegendes Werkzeug. Molekulare Ähnlichkeiten können durch eine strukturelle Überlagerung von chemischen Verbindungen analysiert werden, durch die sich Gemeinsamkeiten und Unterschiede in den Strukturen herausstellen lassen.

Das im Rahmen der vorliegenden Arbeit entwickelte Programm GAMMA (genetic algorithm for multiple molecule alignment) dient zur Überlagerung mehrerer dreidimensionaler Molekülmodelle in einem simultanen Optimierungsprozeß. Damit kann von einer größeren Anzahl an Molekülen die größte gemeinsame dreidimensionale Substruktur (engl.: maximum common substructure MCSS) ermittelt werden. Die Suche nach der dreidimensionalen MCSS startet mit einer Konformation für jede Struktur und untersucht den gemeinsamen konformativen Raum der Verbindungen während des Optimierungsprozesses. Die entstandenen Konformationen werden dabei auf eine Überlappung der van der Waals-Radien nicht gebundener Atome untersucht. Das automatische Auffinden struktureller Ähnlichkeit eines Datensatzes an Verbindungen basiert auf einer Hybridmethode, deren Effizienz in der Kombination eines genetischen Algorithmus mit einer numerischen Optimierungsmethode, dem *Directed-Tweak* liegt. Die Effektivität des Optimierungsprozesses mit dem genetischen Algorithmus wurde durch die Einführung zweier zusätzlicher Operatoren, *Creep* und *Crunch*, die direkt auf das Problem der MCSS Suche zugeschnitten sind, erhöht. Die hier beschriebene *Restricted Tournament*-Selektion verhindert den Verlust der genetischen Vielfalt während des Optimierungsprozesses. Da die Suche nach der MCSS ein multidimensionales Problem mit gegenläufigen Optimierungskriterien (Größe der Substruktur und geometrische Anpassung) darstellt, wurde die sogenannte *Pareto-Optimierung* integriert. Diese Art der Optimierung führt nicht zu einer Lösung, sondern zu einer Menge von Lösungen. Damit läßt sich ein Satz von Überlagerungen mit steigender Substrukturgröße und jeweils optimierter geometrischer Anpassung erhalten.

Es wurden Möglichkeiten geschaffen, die es erlauben die pharmakophoren Bereiche der chemischen Strukturen zu analysieren. So können als Überlagerungskriterium statt der Ord-

nungszahl der Atome, physikochemische Atomeigenschaften angegeben werden. Weiterhin ist es möglich Atome im voraus auszuwählen, die auf jeden Fall Teil der Substruktur sein sollen oder auf die sich die Suche der Substruktur beschränken muß. Rotierbare Bindungen der Strukturen werden innerhalb des Programms automatisch erkannt oder können alternativ durch den Benutzer angewählt werden. Zudem kann benutzerdefiniert ein Molekül während des Prozesses starr gehalten werden. Die gewählte Verbindung spielt damit die Rolle eines Templates, an das sich die Konformationen der anderen Verbindungen des Datensatzes anpassen.

Die Hybridmethode wurde hier anhand zweier Anwendungsbeispiele, einerseits der Überlagerung von Angiotensin-II-Antagonisten am AT₁-Rezeptor und andererseits der Überlagerung von Cytochrom P450c17-Inhibitoren evaluiert.

Die Angiotensin-II-Antagonisten wurden sowohl starr als auch durch Rotation der Einfachbindungen überlagert. Ein Vergleich der nach diesen beiden Überlagerungen erhaltenen Substrukturen ergab starke Verbesserungen der geometrischen Anpassungen der Strukturen bei flexibler Überlagerung. Dabei werden auch größere Substrukturen erhalten, die es erlauben strukturelle Ähnlichkeiten in scheinbar sehr unterschiedlichen Strukturen zu finden. Weiterhin wurden für binäre und multiple Überlagerungen verschiedene Substrukturen analysiert, die sich durch Angabe unterschiedlicher Überlagerungskriterien erhalten ließen. Die ermittelten Pharmakophorgruppen, die aus der Analyse der Substrukturen der flexiblen Überlagerungen der Angiotensin-II-Antagonisten erhalten wurden, sind die beiden lipophilen Phenylringe, der basische Imidazolring und die wiederum lipophilen Alkylseitenketten des Imidazolrings.

P450c17-Inhibitoren können in Naphthyl-, Biphenyl und Steroidsysteme unterteilt werden. Anhand der Cytochrom P450c17-Inhibitoren wurden vor allem Programmeigenschaften getestet, die es erlauben Wissen über die Bindungseigenschaften des Rezeptors zu integrieren. Für die Affinität der Verbindungen zum Cytochrom P450c17 ist eine basische Gruppe zur Koordination an das zentrale Eisenatom des Enzyms und am anderen Ende ein Wasserstoffbrückenakzeptor nötig. So wurden Überlagerungen unter Angabe fester Atomtupel, die unbedingt an der Substruktur beteiligt sein sollen untersucht. Die Analyse der hier berechneten Überlagerungen zeigt, daß für die Gruppe der Biphenyl-Verbindungen von einer eher steroidähnlichen Anlagerung ausgegangen werden kann. Weiterhin konnte bewiesen werden, daß die Affinität der behandelten Verbindungen von der Position des Wasserstoffbrückenakzeptors am A-Ring des Steroidgrundgerüsts abhängt. Eine 3'-Position des Wasserstoffbrückenakzeptors ist gegenüber der 4'-Position bevorzugt.

5.2 Relevante Wechselwirkungen in Steroid-Rezeptorsystemen

Molekulare Ähnlichkeiten basieren nicht nur auf rein geometrischen Ähnlichkeiten, sondern auch auf Ähnlichkeiten der Oberflächenformen und der physikochemischen Oberflächeneigenschaften der Verbindungen. Um diese zu untersuchen, wurden Bindungsmodi einer Reihe von Steroiden zu einem Immunoglobulin-Rezeptor und einem künstlichen Cyclophanrezeptor berechnet. Weiterhin wurden die physikochemischen Oberflächeneigenschaften der Steroide miteinander verglichen.

Durch Dockingexperimente mit Hilfe des automatischen Dockingprogramms AutoDock, wurden eine Reihe von Bindungsmodi von Digoxinsteroiden an das antigen-bindende Fragment des 26-10-Antikörpers und von Gallensäuresteroiden, Cortisonsteroiden, Testosteron und Progesteron sowie der zwei der Digoxinsteroide an einen künstlichen Cyclophanrezeptor vorhergesagt.

Die Experimente zeigten, daß die Affinität der Immunoglobulin-Rezeptoren zu Steroide auf apolaren Wechselwirkungen basiert. Die an der Bindungstasche des 26-10-Antikörpers beteiligten Aminosäuren konnten dabei ermittelt werden. Die Spezifizierung zwischen verschiedenen Steroiden basiert auf der Oberflächenkomplementarität zwischen Ligand und Rezeptor, die hier am größten bei den Lactonringen und dem Steroid-D-Ring ist und in Richtung des A-Rings und des Zuckerrests des Steroidliganden stark abnimmt. Die größten Bindungskräfte entstehen aus hydrophober Desolvatisierung und van-der-Waals-Wechselwirkung. Wasserstoffbrückenbildungen spielen eine eher untergeordnete Rolle.

Für die Wechselwirkung der Liganden mit dem Immunoglobulin-Rezeptor, sowie mit dem Cyclophanrezeptor wurden Affinitätspotentiale erstellt. Diese setzen sich aus den genauen Beiträgen der einzelnen Atome zu der gesamten Interaktion mit der jeweiligen Rezeptorbindungstasche zusammen. Bei der Analyse der Bindungsmodi und Affinitätspotentiale der Steroide an dem künstlichen Cyclophanrezeptor ließen sich strukturelle Bereiche der Liganden herausstellen, die jeweils besonders relevant für die Affinität zu dem Rezeptor sind. Nur bei wenigen Verbindungen konnten höhere elektrostatische Wechselwirkungsenergien ermittelt werden.

Die berechneten freien Bildungsenthalpien ΔG der Steroide mit dem Cyclophanring entsprechen den experimentell bestimmten freien Bildungsenthalpien mit einer mittleren Abweichung von 1.2 kcal/mol. Es konnte eine indirekte Beziehung zwischen der Höhe der freiwerdenden Komplexbildungsenthalpie und der Anzahl an Hydroxylgruppen am Steroidgrundkörper ermittelt werden. Weiterhin ließ sich ein Zusammenhang zwischen der freien Bildungsenthalpie der Komplexe und der Oberflächenhydrophobizität der Liganden herstellen. Zur Berechnung eines quantitativen Maßes für die Oberflächenhydrophobizität wurden mit den dreidimensionalen Koordinaten der van-der-Waals-Oberflächen, die mit der Eigenschaft

der Hydrophobizität kodiert waren, Kohonen neuronale Netze trainiert. Ein Maß für die Hydrophobizität ergab sich aus der jeweiligen Anzahl der „hydrophoben Neuronen“, also der Neuronen, die Datenpunkte beinhalten, denen Oberflächenpunkte mit hoher Hydrophobizität entsprechen. Die jeweilige Anzahl der hydrophoben Neuronen korreliert mit einem Korrelationskoeffizienten r von 0.83 mit der Affinität der betrachteten Steroide zu dem Cyclophanrezeptor.

Durch die Anwendung der Templat-Methode, bei der der künstliche Rezeptor selbst als Templat agiert, wurde eine Möglichkeit gefunden, die Region, in der Rezeptor und Ligand wechselwirken, sowie die dort herrschenden physikochemischen Oberflächeneigenschaften sichtbar zu machen.

In beiden beschriebenen Fällen waren apolare Wechselwirkungen für den größeren Teil der Affinität zwischen Ligand und Rezeptor verantwortlich. Die van-der-Waals-Wechselwirkungsenergien sind bei fast allen Atomen um ein zehnfaches größer als die Beiträge der elektrostatischen Wechselwirkungsenergien. Polare Gruppen am Ligandgrundkörper führten teilweise sogar zu einer Destabilisierung der Komplexe. Diese Beobachtungen verdeutlichen, daß spezifische hydrophobe Wechselwirkungen zur Affinität und Selektivität bei natürlichen sowie künstlichen Rezeptorsystemen beitragen können. Relativ zu den meist hoch bewerteten elektrostatischen Wechselwirkungen oder Wasserstoffbrückenbindungen dürfen hydrophobe Wechselwirkungen somit nicht vernachlässigt werden.

6 Literaturverzeichnis

- [1] Sadowski, J.; Gasteiger, J. From Atoms and Bonds to Three-Dimensional Atomic Coordinates: Automatic Model Builders. *Chem. Reviews* **1993**, *93*, 2567 - 2581.
- [2] Sadowski, J.; Gasteiger, J.; Klebe, G. Comparison of Automatic Three-Dimensional Model Builders Using 639 X-Ray Structures. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1000 - 1008.
- [3] Sheridan, R.P.; Nilakantan, R.; Rusinko, III, A.; Baumann, N.; Haraki, K.S.; Venkataraghavan, R. 3DSEARCH: A System for Three-Dimensional Substructure Searching. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 255-260.
- [4] Greene, J.; Kahn, S.D.; Savoio, H.; Sprague, P.; Teig, S. Chemical Function Queries for 3D Database Search. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1297-1308.
- [5] Martin, Y.C.; Danaher, E.B.; May, C.S.; Weininger, D. MENTHOR, a Database System for the Storage and Retrieval of Three-Dimensional Molecular Structures and Associated Data Searchable by Substructural, Biological, Physical or Geometric Properties. *J. Comput.-Aided Mol. Des.* **1988**, *2*, 15-29.
- [6] Van Drie, J.H.; Weininger, D.; Martin, Y.C. ALADDIN: An Integrated Tool for Computer-Assisted Molecular Design and Pharmacophoric Pattern Recognition from Geometric, Steric and Substructure Searching of Three-Dimensional Molecular Structures. *J. Comput.-Aided Mol. Des.* **1989**, *3*, 225-251.
- [7] Pepperrell, C.A.; Willett, P. Techniques for the Calculation of Three-Dimensional Structural Similarity using Inter-Atomic Distances. *J. Comput.-Aided Mol. Des.* **1991**, *5*, 455-474.
- [8] Sheridan, R.P.; Miller, M.D.; Underwood, D.J.; Kearsley, S.K. Chemical Similarity Using Geometric Atom Pair Descriptors. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 128-136.
- [9] Bath, P.A.; Poirrette, A.R.; Willett, P.; Allen, F.H. Similarity Searching in Files of Three-Dimensional Chemical Structures: Comparison of Fragment-Based Measures of Shape Similarity. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 141-147.
- [10] Lauri, G.; Bartlett, P.A. CAVEAT: A Program to Facilitate the Design of Organic Molecules. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 51-66.
- [11] Fisanick, W.; Cross, K.P.; Rusinko, III, A. Similarity Searching on CAS Registry Substances. 1. Global Molecular Property and Generic Atom Triangle Geometric Searching. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 664-674.

- [12] Wagener, M.; Gasteiger, J. Die Bestimmung größter deckungsgleicher Teilstrukturen mit einem genetischen Algorithmus: Anwendung in der Syntheseplanung und zur strukturellen Analyse biologischer Aktivität. *Angew. Chem.* **1994**, *106*, 1245-1248, The Determination of Maximum Common Substructures by a Genetic Algorithm: Application in Synthesis Design and for the Structural Analysis of Biological Activity. *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1189-1192.
- [13] Handschuh, S.; Wagener, M.; Gasteiger, J. Superposition of Three-Dimensional Chemical Structures Allowing for Conformational Flexibility by a Hybrid Method. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 220-232.
- [14] Kohonen, T. Self-organized Formation of Topologically Correct Feature Maps. *Biol. Cybern.* **1982**, *43*, 59-69; *Self-Organization and Associative Memory*; third Edition; Springer Verlag: Berlin, Germany, 1989; *Self-Organizing Maps*; Huang, T. S., Kohonen, T., Schröder, M.R.; Eds., Springer Verlag: Berlin, Germany, 1995.
- [15] Gasteiger, J. Li, X. Mapping the Electrostatic Potential of Muscarinic and Nicotinic Agonists with Artificial Neural Networks. *Angew. Chem.* **1994**, *33*, 671-674, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 643-646.
- [16] Anzali, S.; Barnickel, G.; Krug, M.; Sadowski, J.; Wagener, M.; Gasteiger, J.; Polanski, J. The Comparison of Geometric and Electronic Properties of Molecular Surfaces by Neural Networks: Application to the Analysis of Corticosteroid Binding Globulin Activity of Steroids. *J. Comp.-Aided Mol. Design* **1996**, *10*, 521-534.
- [17] Polanski, J.; Gasteiger, J.; Wagener, M.; Sadowski, J. The comparison of molecular surfaces by neural networks and its application to quantitative structure activity studies. *Quant. Struct. Act. Relat.* **1998**, *17*, 27-36.
- [18] Gasteiger, J., Li, X., Rudolph, C., J. Sadowski, J. and Zupan, J. Representation of Molecular Electrostatic Potentials by Topological Feature Maps. *J. Am. Chem. Soc.* **1994**, *116*, 4608-4620.
- [19] Wallimann, P.; Marti, T.; Fürer, A.; Diederich, F. Steroids in Molecular Recognition. *Chem. Rev.* **1997**, *97*, 1567-1608.
- [20] Jeffrey, D.P.; Strong, R.K.; Sieker, L.C., Chang, C.Y.Y.; Campbell, R.L.; Petsko, G.A.; Haber, E.; Margolies, M.N.; Sheriff, S. 26-10 Fab-digoxin complex: Affinity and specificity due to surface complementarity. *Proc. Nat. Acad. Sci. USA* **1993**, *90*, 10310-10314.
- [21] Schildbach, J.F; Near, R.J.; Bruccoleri, R.E.; Haber, E.; Jeffrey, P.D.; Ng, S.-C.; Novotny, J.; Sheriff, S.; Margolies, M.N. Heavy Chain Position 50 Is a Determinant of Affinity and Specificity for the Anti-digoxin Antibody 26-10. *J. Biol. Chem.* **1993**, *268*, 21739-21747.
- [22] Peterson, B.R., Wallimann, D.R.; Carcanague, Diederich, F. Steroid Complexation by Cyclophane Receptors in Aqueous solution: Substrate Selectivity, Enthalpic Driving

- Forces for Cavity Inclusion, and Enthalpy-Entropy Compensation. *Tetrahedron* **1995**, *51*, 401-421.
- [23] Peterson, B.R.; Mordasini-Denti, T.; Diederich, F. Cavity depth and width effects on cyclophane-steroid recognition: molecular complexation of cholesterol and progesterone in aqueous solution. *Chemistry and Biology* **1995**, *2*, 139-146.
- [24] Arevalo, J.H.; Stura, E.A.; Taussig, M.J.; Wilson, I.A. Three-dimensional Structure of an Anti-steroid Fab' and Progesterone-Fab' Complex. *J. Mol. Biol.* **1993**, *231*, 103-118.
- [25] Burks, E.A.; Chen, G.; Georgiou, G.; Iverson, B.L. In vitro scanning saturation mutagenesis of an antibody binding pocket. *Proc. Nat. Acad. Sci. USA* **1997**, *94*, 412-417.
- [26] Jeffrey, D.P.; Schildbach, J.F.; Chang, C.Y.Y.; Kussie, P.H.; Margolies, M.N.; Sheriff, S. Structure and Affinity of the Anti-Digoxin Antibody 40-50. *J. Mol. Biol.* **1995**, *248*, 344-360.
- [27] Kawakami, H.; Yoshino, O.; Odashima, K.; Koga, K. Design, Synthesis, and Properties of Novel Water-Soluble Cyclophanes Having Naphthylphenylmethane Units as Hosts for Aliphatic and Aromatic Guests. *Chem. Pharm. Bull.* **1985**, *33*, 5610-5613.
- [28] Klaerner, F.-G.; Benkhoff, J.; Boese, R.; Burkert, U.; Kamieth, M.; Naatz, U. Molecular Tweezer as Synthetic Receptors in Host-Guest Chemistry: Inclusion of Cyclohexane and Self-Assembly of Aliphatic Side Chains. *Angew. Chem.* **1996**, *108*, 1195-1198, *Angew. Chem. Int. Engl.* **1996**, *35*, 1130-1133.
- [29] Wallimann, P.; Seiler, P.; Diederich, F., Dendrophanes: Novel Steroid-Recognizing Dendrite Receptors. *Helv. Chim. Acta* **1996**, *79*, 779-788.
- [30] Peterson, B.R.; Diederich, F. Dissolution of Cholesterol in Water by a Synthetic Receptor. *Angew. Chem.* **1994**, *106*, 1688-1690, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1625-1628.
- [31] Gouverneur, V.E.; Houk, K.N.; Pascual-Teresa, de B.; Beno, B.; Janda, K.D.; Lerner, R.A., Control of the *exo* and *endo* Pathways of the Diels-Alder Reaction by Antibody Catalysis. *Science* **1993**, *262*, 204-208.
- [32] Hurst, T. Flexible 3D Searching: The Directed Tweak Technique. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 190-196.
- [33] Sealy, J.E.; Laragh, J.H. The Renin-Angiotensin-Aldosterone System for Normal Regulation of Blood Pressure and Sodium and Potassium Homeostasis. In *Hypertension: Pathophysiology, Diagnosis and Management*; Laragh, J.H.; Brenner, B.M., Eds.; Raven Press: New York, 1990, 1287-1317.
- [34] Böhm, H.J.; Klebe, G.; Kubinyi, H. *Wirkstoffdesign*, Spektrum Akademischer Verlag: Heidelberg, 1996, S. 143.
- [35] Bosche v.d. H. Inhibitors of P450 dependent Steroidbiosynthesis: From research to medical treatment. *J. Steroid. Biochem. Molec. Biol.* **1992**, *43*, 1003-1021.

- [36] Wachall, B. G.; Hector, M.; Hartman, R. W. Imidazole Substituted Biphenyls - A New Class of Highly Potent and In Vivo Active Inhibitors of P450 17 as Potential Therapeutics for Treatment of Prostate Cancer. *Bioorg. Med. Chem.* **1999**, in press.
- [37] Wachall, B. Vom Substrat Mimetikum zum in vivo aktiven Hemmstoff der P450 17: Synthese, Struktur-Wirkungs-Beziehungen, theoretische und pharmakologische Untersuchungen. Dissertation, Universität Saarbrücken, 1999.
- [38] Zupan, J. and Gasteiger, J., *Neural Networks for Chemists: An Introduction*, VCH-Publishers: Weinheim, Germany, 1993.
- [39] Auterhoff, H.; Knabe, J.; Höltje, H.-D. *Lehrbuch der Pharmazeutische Chemie*, Wissenschaftliche Verlagsgesellschaft mbH: Stuttgart, 1994.
- [40] Goldberg, D.E.; Deb, K. An Investigation Of Niche and Species Formation in Genetic Function Optimization. In *Proceedings of the 3rd International Conference on Genetic Algorithms*; Schaffer, D., Ed.; Morgan Kaufmann Publishers, Inc.: San Mateo, CA, 1989, 42-50.
- [41] Fonseca, C.M.; Fleming, P. J. Genetic Algorithm for Multiobjective Optimization: Formulation, Discussion and Generalization. In *Proceedings of the 5th International Conference on Genetic Algorithms*; Forrest, S., Ed.; Morgan Kaufmann Publishers, Inc.: San Mateo, CA, 1993, 416-423.
- [42] Jones, G. Genetic and Evolutionary Algorithms. In *Encyclopedia of Computational Chemistry*; Schleyer, P.v.R.; Allinger, N.L.; Clark, T.; Gasteiger, J.; Kollman, P.A.; Schaefer, III, H.F.; Schreiner, P.R., Eds; John Wiley & Sons, Inc.: Chichester, UK, 1998, 1127-1136.
- [43] Venkatasubramanian, V.; Sundaram, A. Genetic Algorithms: Introduction and Applications. In *Encyclopedia of Computational Chemistry*; Schleyer, P.v.R.; Allinger, N.L.; Clark, T.; Gasteiger, J.; Kollman, P.A.; Schaefer, III, H.F.; Schreiner, P.R., Eds; John Wiley & Sons, Inc.: Chichester, UK, 1998, 1115-1127.
- [44] Payne, A.W.R.; Glen, R.C. Molecular Recognition Using a Binary Genetic Search Algorithm. *J. Mol Graphics* **1993**, *11*, 74-91.
- [45] Fontain, E. Application of Genetic Algorithms in the Field of Constitutional Similarity. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 748-752.
- [46] Choy, W.Y.; Sanctuary, B.C. Using Genetic Algorithm with a Priori Knowledge for Quantitative NMR Signal Analysis. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 685-690.
- [47] Chen, H.; Zhou, J.; Xie, G. PARM: A Genetic Algorithm to Predict Bioactivity. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 243-250.
- [48] Gillet, V. J.; Willett, P.; Bradshaw, J. Identification of Biological Activity Profiles Using Substructural Analysis and Genetic Algorithms. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 165-179.

- [49] Nair, N.; Goodman, J.M. Genetic Algorithm in Conformational Analysis. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 317-320.
- [50] Wehrens, R.; Pretsch, E.; Buydens, L.M.C. Qualitative Criteria of Genetic Algorithms for Structure Optimization. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 151-157.
- [51] Brown, R.D.; Jones, G.; Willett, P. Matching Two-Dimensional Chemical graphs Using Genetic Algorithm. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 63-70.
- [52] Jones, G.; Willett, P.; Glen, R.C. A Genetic Algorithm for Flexible Molecular Overlay and Pharmacophore Elucidation. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 532-549.
- [53] Goldberg, D.E. *Genetic Algorithms in Search Optimization and Machine Learning*, Addison-Wesley Publishing Company: New York, 1989.
- [54] Jones, G.; Willett, P.; Glen, R.C., Leach, A.R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* **1997**, *276*, 727-748.
- [55] Jones, G.; Willett, P.; Glen, R.C. Molecular Recognition of Receptor Sites Using a Genetic Algorithm with a Description of Desolvation. *J. Mol. Biol.* **1995**, *245*, 43-53.
- [56] Wild, D.J.; Willett, P. Similarity Searching in Files of Three-Dimensional Chemical Structures. Alignment of Molecular Electrostatic Potential Fields with a Genetic Algorithm. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 159-167.
- [57] Thorner, D.A.; Wild, D.J.; Willett, P.; Wright M. Similarity Searching of Three-Dimensional Chemical Structures: Flexible Field-Based Searching of Molecular Electrostatic Potentials. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 900-908.
- [58] Clark, D.E.; Jones, G.; Willett, P.; Kenny, P.W.; Glen, R.C. Pharmacophoric pattern Matching in Files of Three-Dimensional Chemical Structures: Comparison of Conformational -Searching Algorithms for Flexible Searching. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 197-206.
- [59] DeJong, K.A. An Analysis of the Behaviour of a Class of Genetic Adaptive Systems. Dissertation, University Michigan, 1975.
- [60] Harik, G.R. Finding Multimodal Solutions Using Restricted Tournament Selection. In *Proceedings of the 6th International Conference on Genetic Algorithms*; Eshelman, L.J., Ed.; Morgan Kaufmann Publishers, Inc.: San Francisco, CA, 1995, 24-31.
- [61] Oliver, I.M.; Smith, D.J.; Holland, J.R.C. A Study of Permutation Crossover Operators on the Travelling Salesman Problem. *Genetic Algorithms and their Applications*: In *Proceedings of the 2nd International Conference on Genetic Algorithms*; Grefenstette, J.J., Ed.; Lawrence Erlbaum Associates: Hillsdale, NJ, 1987, 224-230.
- [62] Davis, L. Adapting Operator Probabilities in Genetic Algorithms. In *Proceedings of the 3rd International Conference on Genetic Algorithms*; Schaffer, D., Ed.; Morgan Kaufmann Publishers, Inc.: San Mateo, CA, 1989, 61-69.

- [63] Press, H.W., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. *Numerical Recipes in C*, Cambridge University Press: Cambridge 1997.
- [64] Huheey, J.E., Keiter, E.A., Keiter, R.L. *Inorganic Chemistry: Principles of structures and reactivity*, 4. Auflage, Harper Collins: New York, 1993.
- [65] Atkins, P.W. *Physikalische Chemie*, VCH Verlagsgesellschaft: Weinheim, 1990.
- [66] Gasteiger, J. Empirical Methods for the Calculation of Physicochemical Data of Organic Compounds. In *Physical Property Prediction in Organic Chemistry*; Jochum, C.; Hicks, M.G.; Sunkel, J., Eds.; Springer Verlag: Heidelberg, 1988, 119-138.
- [67] Saller, H. Quantitative Empirische Modelle für Elektronische Effekte in π -Systemen und für die Chemische Reaktivität. Dissertation, Universität München, 1985.
- [68] *The Merck Index*, Budavari, S.; O'Neil, M.J.; Smith, A., Eds., Merck & Co., Inc.: Whitehouse Station, NJ, 1996.
- [69] Wexler, R.R.; Greenlee, W.J.; Irvin, J.D.; Goldberg, M.R.; Prendergast, K.; Smith R.D.; Timmermans, P.B.M.W.M. Nonpeptide Angiotensin II Receptor Antagonists: The Next Generation in Antihypertensive Therapy. *J. Med. Chem.* **1996**, *39*, 625-656.
- [70] Levens, N.R.; de Gasparo, M.; Wood, J.M.; Bottari, S.P. Could the Pharmacological Differences Observed between Angiotensin II Antagonists and Inhibitors of Angiotensin Converting Enzyme be Clinically Beneficial? *Pharmacol. Toxicol.* **1990**, *71*, 241-249.
- [71] Lin, H.-S.; Rampersaud, A.A.; Zimmerman, K.; Steinberg, M.I.; Boyd, D.B. Nonpeptide Angiotensin II Receptor Antagonists: Synthetic and Computational Chemistry of N-[[4-[2(2H-Tetrazol-5-yl)-1-cycloalken-1-yl]phenyl]methyl]imidazole Derivatives and Their in Vitro Activity. *J. Med. Chem.* **1992**, *35*, 2658-2667.
- [72] Branden, C.; Tooze, J. *Introduction to Protein Structure*, Garland Publishing: New York and London, 1991.
- [73] Mohamadi, F.; Richards, N.G.J.; Guida, W.C.; Liskamp, R.; Lipton, M.; Canfield, C.; Chang, G.; Hendrickson, T.; Still, W.C., Macromodel - An integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comp. Chem.* **1990**, *11*, 440-467.
- [74] Allinger, N.L.; Yuh, Y.H.; Lii, J.-H. Molecular Mechanics. The MM3 Force Field for Hydrocarbons. *J. Am. Chem. Soc.* **1989**, *111*, 8551-8566.
- [75] a) Martin, A.C.R.; Cheethan, J.C., Rees, A.R., Modeling Antibody Hypervariable Loops: A Combined Algorithm. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9268. b) Pederson, J.; Searle, S.; Henry, A.; Rees, A.R., Antibody Modeling: Beyond Homology. *Immunomethods* **1992**, *1*, 126.
- [76] InsightII/Discover, Biosym/Molecular Simulations Inc., 9685 Scranton Road, San Diego, CA 92121-2777, USA.

- [77] a) Goodsell, D.S.; Olson, A.J., Automated Docking of Substrates to Proteins by Simulated Annealing. *Protein: Structure, Function and Genetics* **1990**, *8*, 195-202. b) Morris, G.M.; Goodsell, D.S.; Huey, R.; Olson, A.J., Distributed Automated Docking of Flexible Ligands to Proteins: Parallel Application of AutoDock2.4.. *J. Comp.-Aided. Mol. Des.*, **1996**, *10*, 293-304.
- [78] Rauhut, G.; Alex, A.; Chandrasekhar, J.; Steinke, T.; Sauer, W.; Beck, B.; Hutter, M.; Gedeck, P.; Clark, T., VAMP Version 6.1, Erlangen 1996.
- [79] Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham III, T.E.; Ferguson, D.M.; Seibel, G.L.; Singh, U.C.; Weiner, P.K.; Kollman, P.A. Amber4.1, University of California, San Francisco, 1995.
- [80] Gilson, K.M.; Given, J.A.; Bush, B.L.; McCammon, J.A., The Statistical Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. *Biophysical Journal* **1997**, *72*, 1047.
- [81] Fox, T.; Scanlan, T.S., Kollman, P.A., Ligand Binding in the Catalytic Antibody 17E8. A Free Energy Perturbation Calculation Study. *J. Am. Chem. Soc.*, **1997**, *119*, 11571-11577.
- [82] Lamb, M.L.; Jorgensen, W.L. Computational Approaches to Molecular Recognition. *Chemical Biology* **1997**, *1*, 449-457.
- [83] Åqvist, J.; Medina, C.; Samuelsson, J.-E., A New Method for Predicting Binding Affinity in Computer-Aided Drug Design. *Protein Engineering* **1994**, *7*, 385-391.
- [84] Cornell, W.D.; Ciepak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
- [85] Gaussian 94, Gaussian Inc., Carnegie Office Park, Building 6, Pittsburgh, PA 15106 USA.
- [86] Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926-935.
- [87] Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J.R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684-3690.
- [88] Ryckaert, J.P.; Ciccotti, G. Introduction of Andersen's Demon in the Molecular Dynamics of Systems with Constraints. *J. Chem. Phys.* **1983**, *78*, 7368-7374.
- [89] Zell, A. *Simulation Neuronaler Netze*, Addison-Wesley Publishing: Bonn, 1994.
- [90] Zupan, J.; Gasteiger, J. Neural Networks: A New Method for Solving Chemical Problems or Just a Passing Phase? *Anal. Chim. Acta* **1991**, *248*, 1-30.

- [91] Gasteiger, J. and Zupan, J. Neural Networks in Chemistry. *Angew. Chem*, **1993**, *105*, 510-536, *Angew. Chem. Int. Ed. Engl* **1993**, *32*, 502-527.
- [92] Anzali, S.; Gasteiger, J.; Holzgrabe, U.; Polanski, J.; Sadowski, J.; Teckentrup, A.; Wagener, M. The Use of Self-Organizing Neural Networks in Drug Design. In *3D QSAR in Drug Design*, Volume 2; Kubinyi, H.; Folkers, G.; Martin, Y.C., Eds.; Kluwer/Escom: Dordrecht, 1998, 273-299.
- [93] Gasteiger, J., Li, X. and Uschold, A. The beauty of molecular surfaces as revealed by self-organizing neural networks. *J. Mol. Graphics* **1994**, *12*, 90-97.
- [94] KMAP, Version 2.1, Li, X., Wagener, M. and Gasteiger, J., 1996.
- [95] SURFACE program, Version 1.0, Sadowski, J. and Gasteiger, J., 1994.
- [96] Li, X., Gasteiger, J. and Zupan, J. On the Topology Distortion in Self-Organizing Feature Maps. *Biol. Cybern.* **1993**, *70*, 189-198.
- [97] Gasteiger, J. and Marsili, M. Iterative Partial Equalization of Orbital Electronegativity - A Rapid Access to Atomic Charges. *Tetrahedron*, **1980** *36*, 3219-3228.
- [98] Gasteiger, J. and Saller, H. Calculation of the Charge Distribution in Conjugated Systems by a Quantification of the Resonance Concept. *Angew. Chem.*, 1985, *97*, 699-701, *Angew. Chem. Int. Ed. Engl.* **1995**, *24*, 687-689.
- [99] Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.* **1964**, *86*, 5175.
- [100] Crippen, G.M.; Havel, T.F. Distance Geometry and Molecular Conformation. Research Studies Press; Letchworth, U.K., 1988.
- [101] Heiden, W.; Moeckel, G.; Brickmann, J. A new approach to analysis and display of local lipophilicity/hydrophilicity mapped on molecular surfaces. *J. Comp.-Aided Mol. Design* **1993**, *7*, 503-514.

A Anhang

A.1 Grundlagen rechnerischer Methoden

A.1.1 Molecular Mechanics (MM)

Berechnung von Eigenschaften und Struktur einer chemischen Verbindung basierend auf der Theorie der klassischen Newtonschen Mechanik und Elektrostatik.

Gleichungen zur Beschreibung der intramolekularen und nichtbindenden Energie eines Moleküls basierend auf dem Modell des harmonischen Oszillators unter Annahme der Born-Oppenheimer Näherung. Dies ergibt einen Satz von Gleichungen zur Beschreibung der potentiellen Energie.

MM-Rechnungen wurden innerhalb des Programmpakets MacroModel^[73] durchgeführt.

A.1.2 Molecular Dynamics (MD)

Simulation der Bewegungen in Vielteilchensystemen durch Lösung der zeitabhängigen Hamiltonschen Bewegungsgleichungen, basierend auf den klassischen Newtonschen Gesetzen der Mechanik ($F = ma$)

Sequenz der Simulation:

- a) Datensammlung zur Definition eines Anfangsstatus
- b) Lösung der Hamiltonschen Bewegungsgleichungen für einzelne Zeitschritte (Zeitfenster)

MD-Rechnungen wurden innerhalb des Programmpakets MacroModel^[73] durchgeführt.

A.1.3 Monte Carlo (MC)

Eine Methode zur Optimierung von Zuständen, basierend auf der anfänglichen Zufallsinitialisierung mehrerer Lösungen, Änderung dieser Lösungen und Akzeptanz der neuen Lösungen mit vorgegebener Wahrscheinlichkeit nach der Boltzmannverteilung.

MC-Rechnungen wurden innerhalb des Programms AutoDock^[77] durchgeführt.

A.1.4 Simulated Annealing

Eine Methode zur Optimierung von Lösungen in multidimensionalen Suchräumen, die auf der Grundlage der Boltzmann Verteilung von Zuständen basiert. Die Wahrscheinlichkeit (WK) zur Besetzung der Zustände ist proportional zur Energie E des Zustands nach:

$$WK(E) \approx e^{\left(-\frac{E}{kT}\right)}$$

Die Simulation wird mit Hilfe von MC oder MD unter Angabe einer entsprechenden Energiefunktion durchgeführt.

Sequenz der Simulation:

- a) Initiale Simulation bei eingestellter hoher Starttemperatur, Zeitdauer ist festzulegen
- b) langsame Abkühlung mit vorgegebenen Temperaturschritten
- c) Einfrieren des Endzustands

Ein wichtiges Anwendungsgebiet des Molecular Modellings ist die Absuche des Konformationsraums einer chemischen Verbindung.

In Kombination mit MC kam die Simlated Annealing Methode innerhalb des Programms AutoDock^[77] zur Anwendung.

A.1.5 Free Energy Perturbation

Berechnung der freien Energie aufgrund der Störungstheorie (engl.: free energy perturbation).

Eine Methode zur Berechnung der freien Energie bzw. Enthalpie eines Systems ΔG_A , basierend auf der Gibbsschen Gleichung nach der statistischen Thermodynamik:

$$\Delta G_A = -kT \ln K_A$$

mit: $K_A = \frac{[XY]}{[X][Y]}$

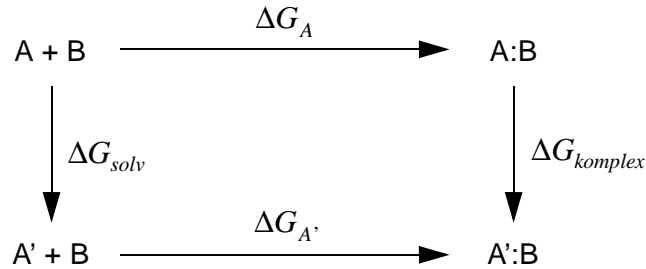
k	=	Boltzmann Konstante
T	=	Temperatur

Die Differenz der freien Enthalpien ist proportional zum Logarithmus der relativen Wahr-

scheinlichkeit, das System in einer von zwei Gleichgewichtszuständen anzufinden:

$$\Delta G_{A \rightarrow B} = kT \ln \langle \exp[-E_B - E_A] / kT \rangle$$

Die Berücksichtigung aller Zustände für die Berechnung der absoluten freien Enthalpie ist schwer möglich. Daher wird eine Störung am System angebracht, die den Übergang von einem Zustand zum nächsten Zustand hervorruft ($A \rightarrow A'$). Diese Zustandsänderung kann in folgendem thermodynamischen Kreislauf simuliert werden (Abbildung A-1).



$$\Delta \Delta G_{bind} = \Delta G_{A'} - \Delta G_A = \Delta G_{komplex} - \Delta G_{solv}$$

Abb. A-1: Der thermodynamische Kreislauf zur Berechnung der Differenz der freien Bildungsenergien $\Delta \Delta G_{bind}$ zwischen zwei Gastmolekülen A und A' im Komplex mit B. Für die vertikalen Schritte werden jeweils die freien Enthalpien berechnet. Molekül A wird zu A' einerseits innerhalb eines Solvens (linker vertikaler Pfeil) und andererseits gebunden im Komplex mit B modifiziert (rechter vertikaler Pfeil).

A.1.6 TIP3P Wassermodell

Das TIP3P Wassermodell ist ein von Jorgensen entwickeltes einfaches kraftfeldbasiertes Modell zur Simulation einer Lösungsmittelbox.^[86] Drei elektrostatische Wechselwirkungsbe-
reiche werden durch zwei partielle positive Ladungen an den Wasserstoffen und einer negativen Partialladung am Sauerstoff gebildet.

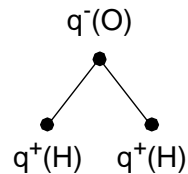


Abb. A-2: Das TIP3P Wassermodell.

Die van der Waals-Wechselwirkung zwischen den Wassermolekülen wird durch die Lennard-Jones-Funktion berechnet mit nur einem Wechselwirkungspunkt lokalisiert auf dem Sau-

erstoffatom. Die Wasserstoffe werden somit bei der Berechnung der van-der-Waals Wechselwirkung nicht berücksichtigt.

Das TIP3P Wassermolekülmodell ist innerhalb des Amber4.1 Programmpaket^[79] implementiert.

A.1.7 SHAKE-Algorithmus

Der Shake Algorithmus wird angewandt um die Bindungslängen in einem System konstant zu halten.^{[87][88]} Das geschieht durch Eliminierung der höchst frequenten Bewegung im System. Da die Größe der Zeitfenster der MD-Simulation durch die höchst frequente Bewegung limitiert ist, erlaubt diese Eliminierung die Definition größerer Zeitfenster. Die Länge der Bindung r wird auf die Länge des Grundzustands r_0 beschränkt, somit liefert die Bindungslängenänderung bei der Simulation der freien Enthalpie keinen Beitrag.

Der Shake-Algorithmus ist innerhalb des Amber4.1 Programmpaket^[79] implementiert.

A.2 Publikationsliste

- [1] S. Handschuh, C.H. Schwab, J. Sadowski, A. Teckentrup, M. Wagener, J. Gasteiger, P. Levi, T. Will, A. Zell, H. Siemens, G. Klebe, T. Mietzner, F. Weber, G. Barnickel, S. Anzali, M. Krug, Stepwise Database Screening as a Tool for Systematic Drug Development, *Software Entwicklung in der Chemie II*, G. Fels; V. Schubert, Eds.; GDCh, Paderborn, 1997, 183-192.
- [2] C.H. Schwab, S. Handschuh, A. Teckentrup, M. Wagener, J. Sadowski, J. Gasteiger, P. Levi, T. Will, A. Zell, H. Siemens, G. Klebe, T. Mietzner, F. Weber, G. Barnickel, S. Anzali, M. Krug, A Systematic Approach to Finding New Lead Structures Having Biological Activity, *Lecture Notes in Computer Science*, R. Hofestädt; T. Lengauer; M. Löffler; D. Schomburg, Eds.; Springer Verlag, Berlin, 1997, 166-177.
- [3] S. Handschuh, M. Wagener, J. Gasteiger, Superposition of Three-Dimensional Chemical Structures Allowing for Conformational Flexibility by a Hybrid Method, *J. Chem. Inf. Comput. Sci.*, **1998**, 38, 220-232.
- [4] J. Gasteiger, S. Handschuh, M.C. Hemmer, T. Kleinöder, C.H. Schwab, A. Teckentrup, J. Sadowski, M. Wagener, 3D Structure Descriptors for Biological Activity, *Molecular Modelling and Prediction of Bioactivity*. F.S. Jørgensen; K. Gundertofte, Eds.; Plenum Press, Kopenhagen: 1998, im Druck.
- [5] S. Handschuh, J. Gasteiger, Pharmacophores Derived from the 3D Substructure Perception, *Pharmacophore Perception, Development, and Use in Drug Design*. O. Güner, Ed.; MSI, San Diego: 1999, im Druck.
- [6] S. Handschuh, J. Chen, B. Goldfuss, K. N. Houk, J. Gasteiger, Steroid Binding by Antibodies and Artificial Receptors: Exploration of Theoretical Methods to Determine the Origins of Binding Affinities and Specificities, *J. Comp.-Aided Mol. Design*, **1999**, eingereicht.

A.3 Lebenslauf

Name Sandra Handschuh
Geburtsdatum und -ort 10.07.1972 in Selb
Staatsangehörigkeit deutsch
Familienstand verheiratet

Schulausbildung

1978 - 1982 Grundschohle Eckental-Eckenhaid
1982 - 1991 Staatliches Gymnasium Lauf a. d. Pegnitz
Abschluß: allgemeine Hochschulreife

Hochschulausbildung

November 1991 - April 1996 Studium der Chemie an der Friedrich-Alexander-Universität Erlangen-Nürnberg
November 1993 Abschluß der Diplomvorprüfungen
Oktober 1995 Abschluß der Diplomhauptprüfungen
November 1995 - April 1996 Diplomarbeit bei Prof. Dr. J. Gasteiger am Institut für Organische Chemie der Universität Erlangen-Nürnberg
Thema: Ermittlung molekularer Ähnlichkeiten von Östrogen-Mimetika mit Hilfe Genetischer Algorithmen und Neuronaler Netze.
seit Mai 1996 Promotionsarbeit bei Prof. Dr. J. Gasteiger am Institut für Organische Chemie der Universität Erlangen-Nürnberg

Auslandsaufenthalte

März - Juni 1998 Forschungsaufenthalt an der University of California, Los Angeles, Department of Chemistry & Biochemistry, Arbeitskreis Prof. K.N. Houk.
Thema: Theoretische Untersuchung der Bindungsaffinitäten von Steroiden zu Antikörpern und artifiziellen Rezeptoren.

Stipendien

Oktober 1996 - September 1998 Chemiefonds-Stipendium für Doktoranden des Verbandes der Chemischen Industrie e.V.

