

Einfluss der Proteinflexibilität auf die Identifikation neuer Antagonisten für die Glycin_B-Bindungsstelle des NMDA-Rezeptors

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

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

von Björn Krüger
aus Dillingen

Frankfurt 2009
(D30)

vom Fachbereich Biowissenschaften der Goethe-Universität als Dissertation
angenommen.

Dekan:	Prof. Dr. Volker Müller
Erster Gutachter:	Prof. Dr. Gisbert Schneider
Zweiter Gutachter:	Prof. Dr. Tanja Weil
Datum der Disputation:	18. September 2009

Inhaltsverzeichnis

Abkürzungsverzeichnis	VI	
Tabellenverzeichnis.....	VII	
Abbildungsverzeichnis.....	VIII	
1	Einleitung.....	10
1.1	Struktur und Pharmakologie des NMDA-Rezeptors	10
1.1.1	Die Familie der ionotropen Glutamatrezeptoren	10
1.1.2	Struktur und Funktion des NMDA-Rezeptors.....	10
1.2	Die Glycin _B -Bindungsstelle der NR1-Untereinheit	12
1.3	Agonistischer und antagonistischer Bindungsmodus – die Kristallstrukturen von Glycin _B	13
1.3.1	Agonistischer Bindungsmodus von Glycin _B	14
1.3.1.1	Affinitätsbestimmende Interaktionen der Agonisten.....	16
1.3.1.2	Selektivitätsbestimmende Interaktionen der Agonisten	16
1.3.2	Antagonistischer Bindungsmodus von Glycin _B	17
1.3.3	Zusammenhang zwischen Domänenbewegung und Rezeptoraktivierung.....	19
1.4	Virtuelle Wirkstoffidentifizierung (<i>Virtual Screening</i>).....	21
1.5	Das Problem der Proteinflexibilität.....	22
1.6	Ziele dieser Arbeit.....	24
2	Material und Methoden	26
2.1	Sammlung einer Bibliothek von Glycin _B -Referenzliganden	26
2.2	Etablierung einer automatisierten Plattform zur <i>in silico</i> Aufarbeitung von Screeningsubstanzen	26
2.2.1	Zusammenstellung der Substanzkataloge.....	26
2.2.2	Aufbereitung der zweidimensionalen Molekülstrukturen... ..	29
2.2.3	Aufbereitung der dreidimensionalen Molekülstrukturen	30
2.3	Distanzmetriken.....	32
2.3.1	Tanimoto-Ähnlichkeit	32
2.3.2	Abweichung vom quadratischen Mittel	33
2.3.3	Receiver-Operating-Characteristics.....	33
2.3.4	Matthews-Korrelationskoeffizient.....	34
2.4	Diversitätsanalyse der Substanzbibliotheken – der chemische Raum von Glycin _B	34
2.5	Ligandenbasiertes zweidimensionales Virtual Screening	36
2.5.1	Laplace-modifizierter naiver Bayes-Klassifikator	37
2.5.2	Ähnlichkeitssuche mit einem 2D-Pharmakophordescriptor	39
2.6	Virtual Screening mit dreidimensionalen Methoden.....	40

2.6.1	Dreidimensionaler Formvergleich	40
2.6.2	Pharmakophorsuche und QSAR-Analyse.....	41
2.6.3	Molecular Docking im eingeschränkt flexiblen Rezeptor...	42
2.7	Modellierung eines flexiblen Proteins am Beispiel von Glycin _B	45
2.7.1	Induced Fit Docking	45
2.7.2	Moleküldynamiksimulationen.....	48
2.7.2.1	Vorbereitung der Proteinstruktur für die MD-Simulation ...	49
2.7.2.2	Durchführung der MD-Simulation	50
2.7.3	Relation von Temperaturfaktor und Mobilität von Aminosäureresten.....	51
2.7.4	Berechnung entropischer und enthalpischer Energieterme durch Wasserverdrängung	52
2.8	Experimentelle Methoden.....	54
2.8.1	[³ H]-MDL-105,519 Bindungsassay.....	54
2.8.2	Funktionaler Bindungsassay.....	54
3	Ergebnisse und Diskussion	55
3.1	Analyse der verwendeten Strukturdatenbanken.....	55
3.1.1	Diversität im zweidimensionalen Strukturraum	55
3.1.2	Diversität im dreidimensionalen Konformationsraum.....	57
3.2	Ligandenbasiertes Virtual Screening ohne Rezeptor.....	59
3.2.1	Ligandenbasiertes zweidimensionales Virtual Screening .	59
3.2.1.1	Laplace-modifizierter naiver Bayes-Klassifikator	59
3.2.1.2	Ähnlichkeitssuche mit einem 2D-Pharmakophordescriptor	64
3.3	Ligandenbasiertes Virtual Screening im implizit flexiblen Rezeptor...	65
3.3.1	Dreidimensionaler Formvergleich	65
3.3.2	Pharmakophorsuche und QSAR-Analyse.....	66
3.4	Rezeptorbasiertes Virtual Screening im rigiden Rezeptor	69
3.4.1	Molecular Docking	69
3.5	Ergebnisse des prospektiven Virtual Screenings.....	73
3.5.1	Neuheitsanalyse	73
3.5.2	Leistungsvergleich der Methoden des Virtual Screenings	75
3.5.3	Aktivität im <i>in vitro</i> -Bindungsassay	76
3.6	Identifizierung unterschiedlicher Serien von Glycin _B -Antagonisten mit neuartigen Scaffolds.....	78
3.6.1	Struktur-Wirkungsbeziehungen der Azo-Verbindungen....	78
3.6.2	Funktionaler Bindungsassay.....	86
3.7	Etablierung eines flexiblen Rezeptormodells.....	87
3.7.1	Analyse der antagonistischen Kristallstrukturen	87
3.7.2	Simulation der Flexibilität von Glycin _B	89
3.8	Auswirkung des flexiblen Rezeptors auf das Wirkstoffdesign.....	93
3.8.1	Induced Fit Docking mit Standardprotokoll	93
3.8.2	Induced Fit Docking mit erweitertem Protokoll.....	95

3.8.3	Induced-Fit-Rezeptorkonformationen der Azo-Liganden ..	99
3.8.4	Thermodynamikänderungen durch Wasserverdrängung als Scoringfunktion	101
3.8.5	Hypothetischer Bindungsmodus der Azo-Liganden unter Berücksichtigung der Wasserverdrängung	104
3.8.6	Vom antagonistischen zum agonistischen Zustand – die funktionale Bedeutung von Rezeptorflexibilität	107
4	Zusammenfassung	110
5	Ausblick	112
6	Literaturverzeichnis	114
	Danksagung	134
	Lebenslauf	136
	Eidesstattliche Erklärung	138
Anhang	139
A.1	Literatur-Referenzliganden	139
A.2	Hann-Filter	166

Abkürzungsverzeichnis

2D	zweidimensional
3D	dreidimensional
ACBC	1-Aminocyclobutancarbonsäure
ACPC	1-Aminocyclopropancarbonsäure
ADME	Absorption, Distribution, Metabolismus, Exkretion
CATS	Chemically Advanced Template Search
DCKA	Dichlorkynurensäure
DRC	Dosierungs-Wirkungs-Kurve (<i>dose-response-curve</i>)
EC_{50}	mittlere effektive Konzentration
ECFP	<i>Extended Connectivity Fingerprint</i>
FCFP	<i>Functional Connectivity Fingerprint</i>
FEP	<i>Free Energy Perturbation</i>
Glycin _B	strychnin-insensitive Glycin-Bindungsstelle in der NR1-Untereinheit des NMDA-Rezeptors
GPCR	G-Protein gekoppelter Rezeptor (<i>G-protein coupled receptor</i>)
IC_{50}	mittlere inhibitorische Konzentration
IFD	Induced-Fit-Docking
iGluR	ionotroper Glutamatrezeptor
MD	Moleküldynamik
NMDA	N-Methyl-D-Aspartat
NMR	Kernresonanzspektroskopie (<i>nuclear magnetic resonance spectroscopy</i>)
OPLS	<i>Optimized Potentials for Liquid Simulations</i>
PCA	Hauptkomponentenanalyse (<i>principal component analysis</i>)
PLS	<i>Partial Least Squares</i>
QSAR	Quantitative Struktur-Wirkungsbeziehung (<i>quantitative structure-activity relationship</i>)
REOS	rapid elimination of swill
RMSD	mittlere quadrierte Abweichung (<i>root mean square deviation</i>)
RMSF	mittlere quadrierte Fluktuation (<i>root mean square fluctuation</i>)
ROC	<i>Receiver-Operating-Characteristics</i>
TM	Transmembrandomäne
uHTS	Ultrahochdurchsatzscreening (<i>ultra high-throughput screening</i>)
VCR	verfügbarer chemischer Raum
vdW	van-der-Waals
ZNS	zentrales Nervensystem

Tabellenverzeichnis

Tabelle 1: Die Familie der metabotropen und ionotropen Glutamatrezeptoren	10
Tabelle 2: Öffentliche Röntgenkristallstrukturen von Glycin _B	14
Tabelle 3: Anbieter von Screeningsubstanzen	28
Tabelle 4: Statistiken der 3D-Konformerendatenbanken	58
Tabelle 5: Retrospektive Validierung verschiedener Bayes-Klassifikatoren	62
Tabelle 6: Anzahl an berechneten Dockingposen	72
Tabelle 7: Leistungsvergleich der Methoden des Virtual Screenings	75
Tabelle 8: Struktur-Wirkungsbeziehungen des Azo-Scaffolds 1.....	78
Tabelle 9: Struktur-Wirkungsbeziehungen des Azo-Scaffolds 2.....	81
Tabelle 10: Struktur-Wirkungsbeziehungen des Azo-Scaffolds 3.....	82
Tabelle 11: Struktur-Wirkungsbeziehungen des Azo-Scaffolds 4.....	83
Tabelle 12: Struktur-Wirkungsbeziehungen des Azo-Scaffolds 5.....	84
Tabelle 13: Messdaten weiterer identifizierter Substanzen	85
Tabelle 14: Korrelation funktionaler Assay und Bindungsassay	86
Tabelle 15: Referenzstrukturen für das Induced Fit Docking mit erweitertem Protokoll	96

Abbildungsverzeichnis

Abbildung 1: Aufbau des NMDA-Rezeptors	11
Abbildung 2: Modularer Aufbau einer iGluR-Untereinheit.....	13
Abbildung 3: Liganden der Glycin _B -Röntgenkristallstrukturen	14
Abbildung 4: Agonistischer Bindungsmodus von D-Serin	15
Abbildung 5: Antagonistischer Bindungsmodus von DCKA.....	18
Abbildung 6: Überlagerung der agonistischen Kristallstrukturen von NR1	20
Abbildung 7: Überlagerung aller Kristallstrukturen von NR1	21
Abbildung 8: Theorien für die Protein-Ligand-Interaktion	23
Abbildung 9: Zyklische Generierung der Screeningdatenbank.....	27
Abbildung 10: Standardisierte Formalladungen.....	30
Abbildung 11: Hypothetische Receiver-Operating-Characteristics	34
Abbildung 12: Clusterzentren des externen generellen Testdatensatzes.....	38
Abbildung 13: Clusterzentren des externen nanomolaren Testdatensatzes	39
Abbildung 14: Shape-Referenzstruktur	41
Abbildung 15: Glycin _B -Bindungsstelle für das Molecular Docking	44
Abbildung 16: Histogramm der paarweisen Ähnlichkeiten Aktive/Inaktive	45
Abbildung 17: Verfügbarer chemischer, biologischer und „Glycin _B -Raum“.....	55
Abbildung 18: Am besten bewertete Features (Bayes-Modell 1).....	59
Abbildung 19: Am schlechtesten bewertete Features (Bayes-Modell 1).	60
Abbildung 20: Am besten bewertete Features (Bayes-Modell 3).....	60
Abbildung 21: Am schlechtesten bewertete Features (Bayes-Modell 3).	61
Abbildung 22: Receiver-Operating-Characteristics der Bayes-Klassifikatoren .	61
Abbildung 23: Korrelation Bayes-Score – experimentelle Messungen	63
Abbildung 24: Receiver-Operating-Characteristics - PH4FP2-Deskriptor	64
Abbildung 25: Receiver-Operating-Characteristics - Shape-Matching	66
Abbildung 26: Pharmakophorhypothese	67
Abbildung 27: Retrospektive Validierung des 3D-QSAR-Modells.....	67
Abbildung 28: Receiver-Operating-Characteristics - Pharmakophorsuche	68
Abbildung 29: Receiver-Operating-Characteristics - Docking.....	70
Abbildung 30: Receiver-Operating-Characteristics - Docking (300 beste Hits)	70
Abbildung 31: Die zehn häufigsten Glycin _B -Scaffolds	74
Abbildung 32: Histogramm der paarweisen Ähnlichkeiten	74
Abbildung 33: Ausgewählte Hits aus dem Virtual Screening.....	77
Abbildung 34: Strukturüberlagerung der antagonistischen Kristallstrukturen ...	87
Abbildung 35: Temperaturfaktoren der Kristallstruktur 1PBQ.....	89
Abbildung 36: Energieverlauf der Moleküldynamik-Simulation von NR1	90
Abbildung 37: RMSD und Temperaturfaktor der Aminosäuren	91
Abbildung 38: Über Zeit gemittelte RMSD-Werte	92
Abbildung 39: RMSD-Werte als Temperaturfaktoren	93

Abbildung 40: Fluktuation Kristallstruktur - mittlere IFD-Struktur	94
Abbildung 41: Histogramm der Molekülmassen der IFD-Ergebnisse	95
Abbildung 42: Alignment der Referenzliganden für das erweiterte Induced-Fit-Docking	97
Abbildung 43: Position der temporär mutierten Resten der Helix F	98
Abbildung 44: IFD-Rezeptorkonformationen der nanomolaren Literaturliganden	99
Abbildung 45: Induced-Fit-Rezeptorkonformationen der Azo-Liganden	100
Abbildung 46: Watermap der Glycin _B -Bindungsstelle	101
Abbildung 47: Watermap-Scoringfunktion: Korrelation Entropiegewinn – pIC_{50}	102
Abbildung 48: Korrelation Entropieterm (Watermap-Score) und pIC_{50}	103
Abbildung 49: Hypothetischer Bindungsmodus der Hits 1 und 24	104
Abbildung 50: Watermap des nanomolaren Virtual Screening Hits 1	106
Abbildung 51: Watermap des mikromolaren Virtual Screening Hits 24	106
Abbildung 52: Energieverlauf – Moleküldynamik-Simulation von D-Cycloserin in NR1	108
Abbildung 53: Rezeptorbewegung von NR1 im Übergang vom antagonistischen in den agonistischen Zustand	109

1 Einleitung

1.1 Struktur und Pharmakologie des NMDA-Rezeptors

1.1.1 Die Familie der ionotropen Glutamatrezeptoren

Der NMDA-Rezeptor, benannt nach seinem spezifischen nicht-natürlichen Agonisten *N*-Methyl-*D*-Aspartat (NMDA) bildet zusammen mit dem α -Amino-3-hydroxy-5-methyl-4-isoxazol-Propionsäure-Rezeptor (AMPA-Rezeptor) und dem Kainat-Rezeptor die Familie der ionotropen Glutamatrezeptoren (iGluR)¹ (Tabelle 1) und steht seit über 20 Jahren im Fokus wissenschaftlichen Interesses¹⁻⁵. Die Rolle des NMDA bei der nozizeptiven Übertragung in Neuronen des Rückenmarks wurde bereits in den 80-er Jahren aufgeklärt², allerdings wurde erst 10 Jahre später auch die Beteiligung von peripheren NMDA-Rezeptoren postuliert^{3,4}. Fehlfunktionen des NMDA werden als eine Ursache vieler weiterer neurodegenerativer Störungen wie der Parkinsonschen⁵ und Alzheimerschen Krankheit^{6,7} oder Schizophrenie⁸ betrachtet. Im Gegensatz zur benachbarten Familie der metabotropen Glutamatrezeptoren (mGluR), die G-Protein gekoppelte Rezeptoren (GPCRs) darstellen, bilden die iGluR Ionenkanäle⁹. Beide Familien zusammen gehören zur Gruppe der Glutamatrezeptoren¹⁰, deren gemeinsamer Neurotransmitter L-Glutamat ist. Sowohl metabotrope als auch ionotrope Glutamatrezeptoren sind Membranproteine, wodurch u.a. ihre Strukturaufklärung erschwert wird.

Tabelle 1: Die Familie der metabotropen und ionotropen Glutamatrezeptoren

Unterfamilie	Rezeptor-Untereinheiten
mGluR	mGluR1 – mGluR8
iGluR – AMPA	GluR1, GluR2a, GluR3, GluR4
iGluR – Kainat	GluR5a, GluR6a, GluR7, KA-1, KA-2
iGluR – NMDA	NR1, NR2A, NR2B, NR2C, NR2D, NR3A, NR3B

1.1.2 Struktur und Funktion des NMDA-Rezeptors

Der NMDA-Rezeptor besteht aus einem Zusammenschluss von verschiedenen Untereinheiten NR1, NR2 und NR3 zu einem hetero-tetrameren Komplex¹¹, welcher den Ionenkanal bildet¹². Er ist postsynaptisch lokalisiert und wird durch die Anwesenheit der beiden Neurotransmitter Glutamat und Glycin aktiviert¹³,

d.h. für die monovalenten Kationen Natrium und Kalium sowie bivalentes Calcium permeabel. Dabei ist Glutamat der hauptsächliche Agonist und interagiert vornehmlich mit der NR2-Untereinheit¹⁴, während sein Coagonist Glycin in der Strychnin-insensitiven Glycin-Bindungsstelle der NR1-Untereinheit bindet (Abbildung 1). Diese Bindungsstelle wird auch Glycin_B genannt, um sie von der Strychnin-sensitiven inhibitorischen Glycin-Bindungsstelle (Glycin_A) zu unterscheiden¹⁵.

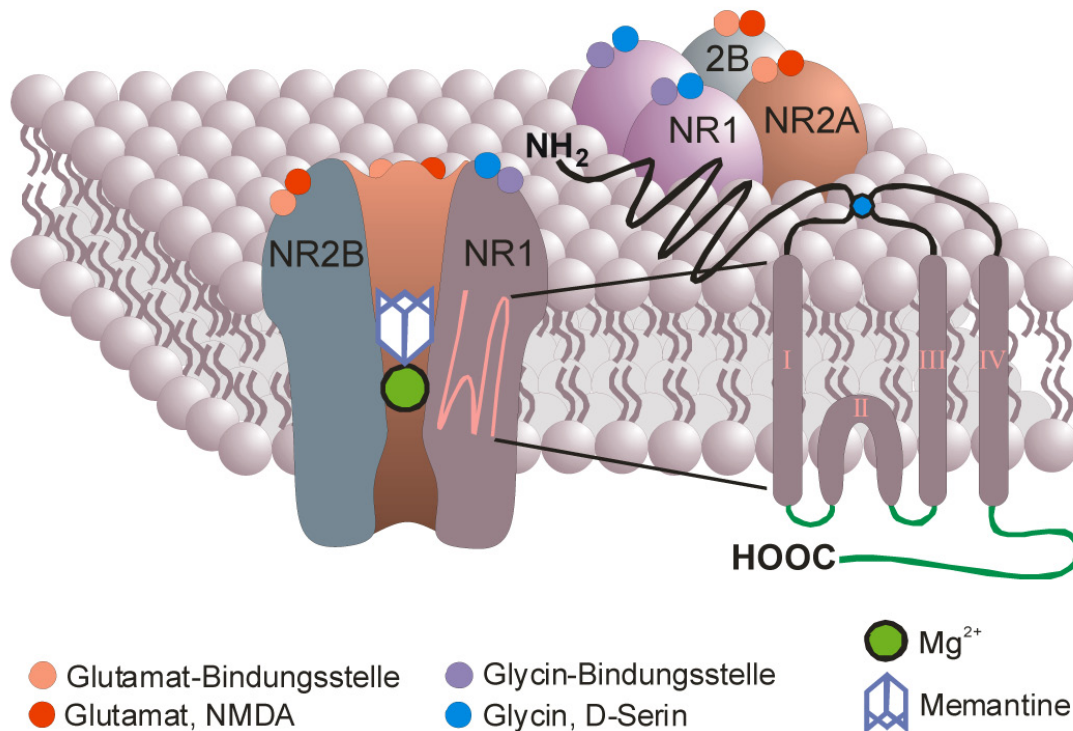


Abbildung 1: Aufbau des NMDA-Rezeptors
mit Genehmigung modifiziert nach Danysz & Parsons¹⁶

Die funktionale Diversität des NMDA-Rezeptors beruht auf seiner obligatorischen Zusammensetzung aus Heterotetrameren. Während die anderen ionotropen Glutamatrezeptoren funktionale Homotetramere bilden können, die nur durch Glutamat aktiviert werden, benötigt der NMDA-Rezeptor jeweils zwei NR1-Untereinheiten sowie NR2 und/oder NR3^{17,18}. Die Affinität von NR1 zu Glycin ist abhängig vom Subtyp der NR2-Untereinheit, was darauf schließen lässt, dass es eine allosterische Verbindung zwischen NR1 und NR2 im funktionalen NMDA-Rezeptor gibt¹⁹.

Im Ruhezustand wird der Kanal des NMDA durch extrazelluläres Mg²⁺ blockiert²⁰. Diese spannungsabhängige Blockade wird durch gleichzeitige Membrandepolarisierung und Agonistenbindung aufgelöst²¹. Dadurch unterscheidet sich der NMDA von allen anderen ligandengesteuerten Ionenkanälen. Die Rate, mit der die funktionale Antwort des Rezeptors auf die Gabe von Agonisten steigt bzw. nach dem Entfernen des Agonisten sinkt, ist deutlich langsamer als

bei nicht-NMDA-Rezeptoren^{22,23}. Insbesondere die langsame Deaktivierungsrate hat einen großen Einfluss auf das exzitatorische postsynaptische Potential²⁴ und damit auf die Dauer und Intensität von Nervenimpulsen. Eine Überaktivierung des NMDA-Rezeptors verursacht Exzitotoxizität, also den Tod der Nervenzelle durch Übererregung²⁵. Die Deaktivierung des NMDA durch Kanalblocker wie Memantine (3,5-Dimethylamin-adamantan-1-ylamin) ist ein erfolgreicher Ansatz zur Behandlung der moderaten bis schweren Demenz im Rahmen des Morbus Alzheimer^{7,26}. Memantine ist zurzeit der einzige NMDA-Rezeptorantagonist unter den zugelassenen Arzneimitteln zur Behandlung von Alzheimer-Symptomen²⁷. Ein weiterer Ansatz zur Modulation des NMDA-Rezeptors besteht in der Blockade der Glycin_B-Bindungsstelle durch kompetitive Antagonisten und erscheint somit als ein vielversprechender Ansatz für einen möglichen neuen NMDA-adressierenden Wirkstoff^{28,29}.

1.2 Die Glycin_B-Bindungsstelle der NR1-Untereinheit

Die NR1-Untereinheit des NMDA-Rezeptors folgt in ihrer Zusammensetzung dem prinzipiellen Aufbau einer iGluR-Untereinheit³⁰ (Abbildung 2). Das Protein durchspannt die Zellmembran, wobei der Amino-Terminus extrazellulär liegt und der Carboxy-Terminus in das Zellplasma reicht. Die Glycin-Bindungsstelle wird durch zwei Domänen S1 und S2 gebildet, wobei die S2-Domäne durch drei Transmembrandomänen (TM) unterbrochen wird (TM I, III und IV), während die hydrophobe TM II eine Haarnadelwende (*hair pin bend*) innerhalb der Membran vollzieht (Abbildung 1). Die Bindungsdomäne der iGluRs ist homolog zum bakteriellen Glutamin-Bindungsprotein QBP³¹ (25% Sequenzidentität zwischen QBP und der Bindungsstelle des Kainat-Rezeptors, 46% Identität bei Nichtberücksichtigung konservativer Mutationen).

Glycin_B spielt eine wichtige Rolle bei der Desensitivierung des NMDA-Rezeptors. Nach der Bindung eines Agonisten in der Glutamat-Bindungsstelle (NR2-Untereinheit) verringert sich die Affinität von Glycin_B zu seinem Agonisten Glycin oder D-Serin, welcher daraufhin dissoziiert. Glycin_B-Antagonisten erzielen einen ähnlichen Effekt^{28,32} und verfügen so über ein vielversprechendes therapeutisches Potential für die Behandlung verschiedener Erkrankungen des zentralen Nervensystems (ZNS). Die langjährige Entwicklung von selektiven Glycin_B-Antagonisten zeigte jedoch, dass die überwiegende Mehrzahl der Verbindungen eine begrenzte Bioverfügbarkeit im ZNS aufwies. Danysz und Parsons erklären dies mit einer schlechten Blut-Hirnschrankengängigkeit, aktivem Transport aus dem ZNS und hoher Plasma-eiweißbindung²⁸. Eine Betrachtung dieser und anderer pharmakokinetischer Parameter (ADME – Absorption, Distribution, Metabolismus, Exkretion)³³⁻³⁵ der Glycin_B-Antagonisten ist jedoch nicht Bestandteil der vorliegenden Arbeit. Alle

bekanntesten klinischen Entwicklungen an Glycin_B-Antagonisten sind eingestellt (Quelle: Thomson Pharma, Thomson-Reuters AG, <http://www.thomson-pharma.com>).

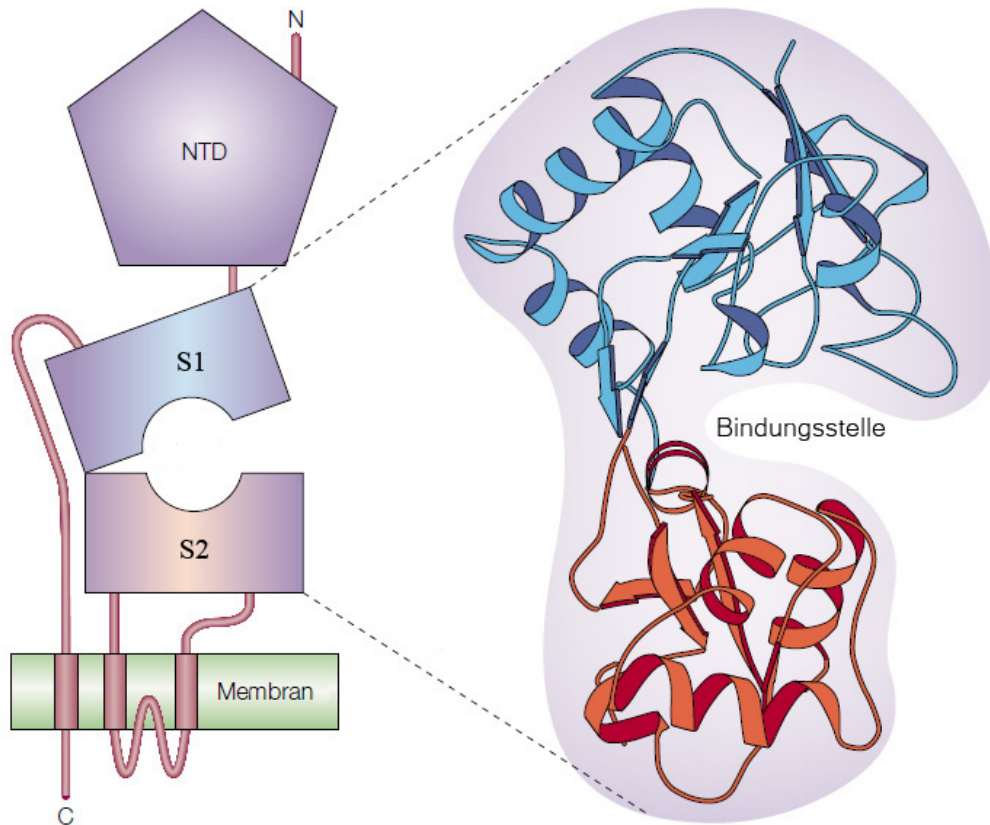


Abbildung 2: Modularer Aufbau einer iGluR-Untereinheit nach Madden *et al.*³⁰

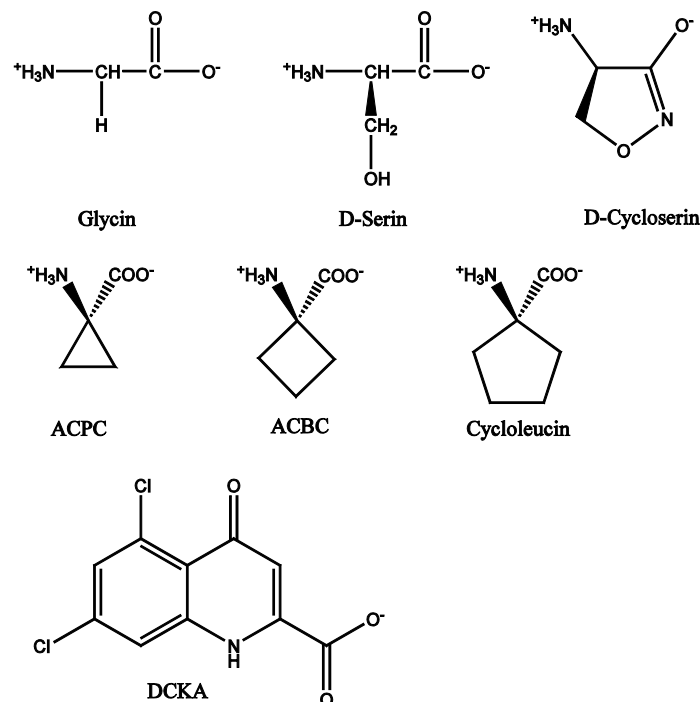
1.3 Agonistischer und antagonistischer Bindungsmodus – die Kristallstrukturen von Glycin_B

Röntgenkristallstrukturen von Glycin_B mit sieben verschiedenen Liganden sind in der Protein Data Base³⁶ (PDB) öffentlich verfügbar. Hierbei handelt es sich um rekombinante Konstrukte der Glycin-Bindungsstelle von NR1, die mit den Agonisten Glycin³⁷ (PDB-Code 1PB7 - 1,35 Å), D-Serin³⁷ (1PB8 - 1,45 Å), den partiellen Agonisten D-Cycloserin³⁷ (1PB9 - 1,6 Å), 1-Aminocyclopropan-säure ACPC³⁸ (1Y20 - 1,4 Å) und 1-Aminocyclobutansäure ACBC³⁸ (1Y1Z - 1,5 Å) und den Antagonisten Dichlorkynurensäure³⁷ (DCKA - 1PBQ - 1,9 Å) und Cycloleucin³⁸ (1Y1M - 1,8 Å) kristallisiert wurden (Tabelle 2). Die hohe Auflösung von 1,9 Å und mehr ermöglichte es, diese Röntgenkristallstrukturen für Methoden des molekularen Modellierens (*molecular modeling*) zu verwenden.

Tabelle 2: Öffentliche Röntgenkristallstrukturen von Glycin_B

PDB ID	Ligand	Pharmakologie
1PB7	Glycin	Agonist ($K_i = 26,4 \mu\text{M}$)
1PB8	D-Serin	Agonist ($K_i = 7,0 \mu\text{M}$)
1PB9	D-Cycloserin	partieller Agonist ($IC_{50} = 241 \mu\text{M}$)
1Y20	1-Aminocyclopropan-1-carbonsäure (ACPC)	partieller Agonist ($IC_{50} = 21,2 \mu\text{M}$)
1Y1Z	1-Aminocyclobutan-1-carbonsäure (ACBC)	partieller Agonist ($IC_{50} = 3,7 \text{ mM}$)
1Y1M	1-Aminocyclopentan-1-carbonsäure (Cycloleucin)	Antagonist ($IC_{50} = 67,4 \text{ mM}$)
1PBQ	Dichlorkynurensäure (DCKA)	Antagonist ($IC_{50} = 0,54 \mu\text{M}$)

Bereits in der Reihe der Strukturen ACPC, ACBC und Cycloleucin zeigt sich der Übergang von Agonisten zu Antagonisten (Abbildung 3). Mit jedem weiteren Methylenglied im Ring vergrößert sich das Volumen des Liganden um etwa 11 \AA^3 . Der Antagonist DCKA ist der größte Ligand der Glycin_B-Röntgenkristallstrukturen.

Abbildung 3: Liganden der Glycin_B-Röntgenkristallstrukturen

1.3.1 Agonistischer Bindungsmodus von Glycin_B

Die natürlichen Agonisten Glycin und D-Serin binden im Interface zwischen den Domänen S1 und S2 und sind umgeben vom N-Terminus der Helix D, den

Helices F und H sowie dem Beta-Faltblatt 14 (Abbildung 4). Glycin verfügt über acht direkte Wasserstoffbrückenbindungen und elektrostatische Interaktionen mit dem Rezeptor. Die Kristallwasser W1, W2, W4 und W5 vermitteln weitere Wasserstoffbrückenbindungen.

Die Alpha-Carboxygruppe von Glycin interagiert mit der Guanidiniumgruppe des Arg523. Dieser Rest ist über alle ionotropen Glutamatrezeptoren hinweg konserviert. Im GluR2 (AMPA-Rezeptor) bildet das entsprechende Arginin eine ionische Wechselwirkung mit der Alpha-Carboxygruppe des Glutamats³⁹. Die Alpha-Carboxygruppe bildet ebenso eine Wasserstoffbrückenbindung mit den Amidgruppen des Proteingrundgerüsts der Reste Thr518 und Ser688 sowie der Hydroxylgruppe von Ser688. Die positiv geladene Aminogruppe des Glycins interagiert mit dem Carbonylsauerstoff von Pro516, dem Hydroxyl von Thr518 und dem Carboxylat-Sauerstoff von Asp732. Auch dieser Rest ist über alle iGluRs hinweg als Aspartat oder Glutamat konserviert. Das sich in der Nähe des gebundenen Glycins befindliche Gln405 interagiert mit zwei Resten in Subdomäne S2: eine direkte Wasserstoffbrückenbindung mit dem Indolstickstoff von Trp731 sowie eine über das Wasser W3 vermittelte Wasserstoffbrückenbindung zum Asp732.

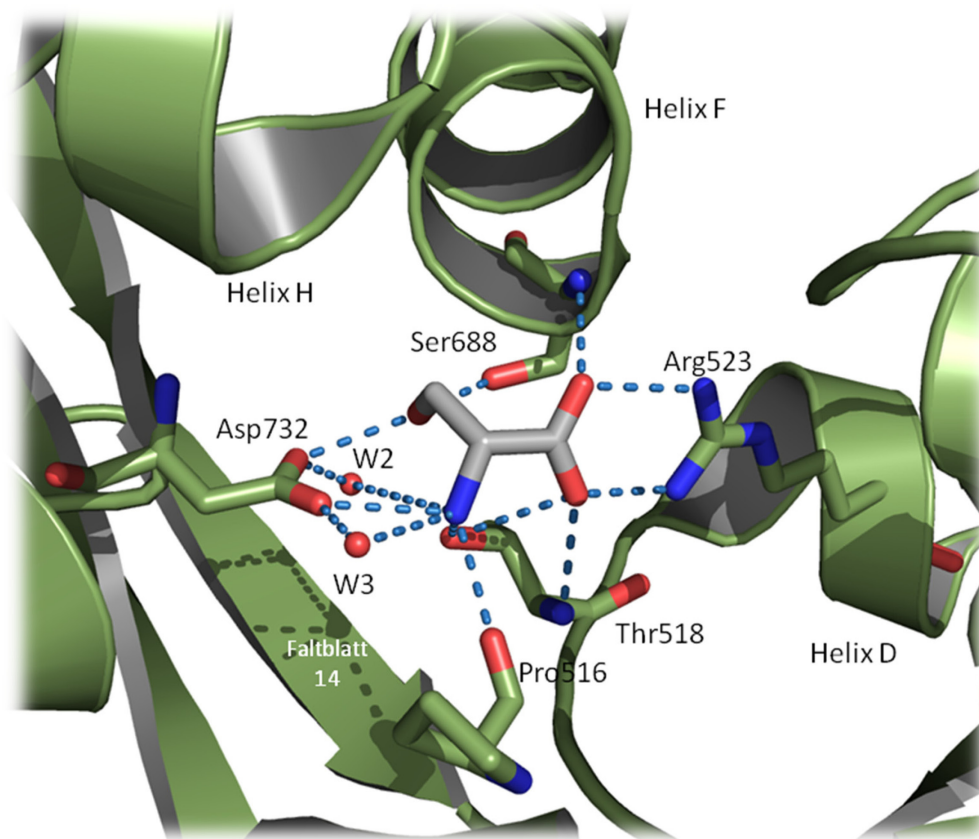


Abbildung 4: Agonistischer Bindungsmodus von D-Serin
Grau: D-Serin; blau: ionische Interaktionen und Wasserstoffbrückenbindungen

Die Proteinfaltung der Bindungsstellen von NR1 und GluR2 ist sich sehr ähnlich⁴⁰; die meisten der mit dem Agonist in Verbindung stehenden Reste sind identisch oder konservative Mutationen (NR1/GluR2: Pro516/478, Thr518/480, Arg523/485, Ser688/654, Asp732/Glu705) und zeigen verwandte Interaktionsmuster mit ihrem jeweiligen Liganden. Ein Unterschied besteht am Rest 655: dieser ist ein Threonin beim GluR2 und bildet eine Wasserstoffbrücke zum Sauerstoff des Gamma-Carboxylats des Glutamats aus. Bei NR1 ist dieser Rest ein Valin, das keine solche Interaktion ausbilden kann. Trp753 im NR1 entspricht Leu704 in GluR2. Der Indolring des Tryptophan zeigt im NR1 Richtung Bindungsstelle und würde mit dem Gammacarboxylat des Glutamats im GluR2 kollidieren. Das Leucin im GluR2 beansprucht nicht nur weniger Volumen, sondern ist auch unterschiedlich im Raum orientiert, so dass das Gamma-Carboxylat des Glutamats eine Interaktion mit dem Rezeptor eingehen kann. Bereits diese beiden Unterschiede definieren die Selektivitätskriterien zwischen den beiden ionotropen Glutamatrezeptoren NR1 und GluR2³⁷: während GluR2 Glutamat hochaffin bindet, ist die Bindungsaffinität von Glutamat zu NR1 sehr niedrig bis nicht mehr messbar⁴¹.

1.3.1.1 Affinitätsbestimmende Interaktionen der Agonisten

Die Affinität der Agonisten in Glycin_B wird hauptsächlich durch polare Interaktionen zwischen Ligand und Rezeptor definiert (Abbildung 4): Zwei elektrostatische Interaktionen mit Resten aus der Domäne S1 (Alpha-Carboxylat des Agonisten mit der Guanidiumgruppe des Arg523 sowie Alpha-Aminogruppe des Liganden mit der Carboxygruppe von Asp732) und eine Reihe von Wasserstoffbrückenbindungen zwischen dem Alpha-Carboxylat und der Domäne S2 bestimmen die Affinität des Liganden³⁷. Eine Deletion von Arg523 unterbricht die agonistische Antwort von NR1⁴². Die Mutation von Asp732 zu Asparagin oder Glutamat verringert den EC_{50} von Glycin um den Faktor 14.500 bzw. 4.200⁴³. Die korrekte Positionierung von positiven und negativen Ladungen in der Bindungsstelle ist somit essentiell für das Bindungsverhalten („ionischer Mechanismus“)⁴⁴.

1.3.1.2 Selektivitätsbestimmende Interaktionen der Agonisten

Die Selektivität von NR1 zu ihren korrespondierenden exzitatorischen Aminosäuren Glycin und D-Serin ist essentiell für die Erregungsleitung am NMDA-Rezeptor: die Glycin_B-Bindungsstelle muss in der Lage sein, hochselektiv zwischen den Neurotransmittern Glycin und Glutamat unterscheiden zu können. Zudem ist die Bindungsstelle von Glycin_B stereoselektiv: grundsätzlich binden nur D-Aminosäuren mit hoher Affinität (L-Serin bindet z.B. um den Faktor 300 schwächer als D-Serin)⁴¹.

Die Glycin-Selektivität von NR1 wird durch zwei Resten in der Bindungsstelle hergestellt: Val689 und Trp731³⁷. Beide verhindern die Bindung der Gamma-Carboxylgruppe von L-Glutamat, da sie keine Wasserstoffbrücken ausbilden können und zudem eine sterische Barriere bilden. Furukawa und Gouaux konnten weiterhin in ihren Kristallisationsstudien zeigen, dass die Spezifität für D-Aminosäuren eine Konsequenz der sterischen Hinderung durch den Rest Phe484 ist³⁷. Zusammen mit Trp731 bildet Phe484 eine hydrophobe Umgebung, die die Interaktion von L-Serin mit dem Rezeptor ungünstig beeinflusst, indem die polare Hydroxylgruppe des L-Serins durch das apolare Phe484 eine abstoßende Kraft erfährt. Im D-Serin ist diese Hydroxylgruppe hingegen so positioniert, dass sie Wasserstoffbrückenbindungen mit gleich drei Resten eingehen kann: Thr518, Asp732 und Ser688.

Furukawa und Gouaux zeigten zudem, dass der pK_a -Wert der Arg523-nahen Gruppe des Liganden einen essentiellen Beitrag zur Bindungsaffinität leistet³⁷: Die Affinität von D-Cycloserin ist um den Faktor 34 geringer als die Affinität von D-Serin. Die pK_a -Werte des exozyklischen Sauerstoffs und der Aminogruppe von D-Cycloserin betragen 4,5 bzw. 7,5. Die entsprechenden pK_a -Werte der Alpha-Carboxy- und Aminogruppe des D-Serins betragen hingegen 2,2 und 9,1⁴⁵. Die Konzentration der doppelt ionisierten Form von D-Cycloserin ist somit im Vergleich zur entsprechenden Form des D-Serins geringer. Zudem enthält D-Cycloserin keine Hydroxylgruppe, die eine Wasserstoffbrückenbindung mit dem Carboxylat des Asp732 ermöglichen kann. Die stärkere Bindung von D-Serin im Vergleich zu Glycin erklären Furukawa und Gouaux mit der Ausbildung von drei zusätzlichen Wasserstoffbrückenbindungen und der Verdrängung des Wassers W2³⁷.

1.3.2 Antagonistischer Bindungsmodus von Glycin_B

Zum Zeitpunkt dieser Arbeit lagen zwei öffentliche Kristallstrukturen von Glycin_B mit kokristallisierten Antagonisten (DCKA und Cycloleucin) vor. Da beide Liganden im Vergleich zu den weiteren bekannten Glycin_B-Antagonisten nur über eine relativ geringe Molekülgröße verfügen, können die Kristallstrukturen nur Aussagen über den „Kernbindungsmodus“ der Antagonisten treffen.

Die weitaus höhere Anzahl an Wechselwirkungen gehen die Antagonisten DCKA und Cycloleucin mit den Resten der Domäne S1 ein. Wie die Agonisten, so bilden auch die Antagonisten eine ionische Wechselwirkung mit Arg523 über ihre Carboxylatfunktion aus (Abbildung 5). Die Aminogruppe von DCKA bildet zudem eine Wasserstoffbrücke mit dem Carboxylat von Pro516. Der Quinolinring des DCKA befindet sich in etwa 3,5 Å Entfernung von Phe484, hier findet, wie bereits in Modellierungsstudien vorgeschlagen, eine aromatische Wechselwirkung (Pi-Stacking) statt⁴⁶. Eine weitere hydrophobe Wechselwirkung geht

von den beiden Chloratomen aus, die in van-der-Waals-Kontakt mit den aromatischen Resten Phe408 und Trp731 stehen. Das Carbonylatom von DCKA ist Teil einer über Wasser vermittelten Wasserstoffbrückenbindung mit der Helix F.

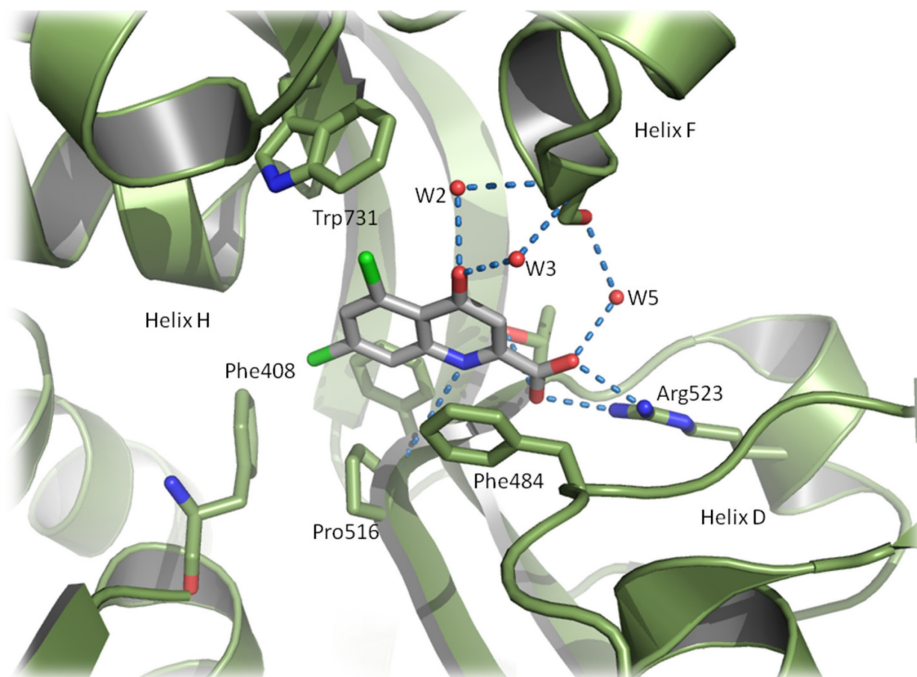


Abbildung 5: Antagonistischer Bindungsmodus von DCKA
Grau: Dichlorkynurensäure (DCKA); gestrichelte blaue Linien: ionische Interaktionen und Wasserstoffbrückenbindungen

In den agonistischen und partiell-agonistischen Rezeptorkonformationen existiert eine intramolekulare Wasserstoffbrückenbindung zwischen dem Amid-sauerstoff von Gln405 und dem Indolstickstoff von Trp731. Zusätzlich gibt es einen über das Wasser W3 vermittelten Kontakt zwischen Gln405 und Asp732. Diese beiden Interdomänkontakte werden durch die Bindung von DCKA unterbrochen. Eine Mutation von Gln405 zu Lysin erhöht durch verstärkte polare Interaktionen den EC_{50} von Glycin um den Faktor 14.230⁴⁷. Die funktionale Wirkung eines Glycin_B-Antagonisten besteht somit in der Unterbrechung essentieller Wechselwirkungen zwischen den Domänen S1 und S2. Dadurch kommt es zu einer Änderung der Konformation des Rezeptors (Kapitel 1.3.3).

Der antagonistische Bindungsmodus weist eine weitere Interaktion auf, welche die Agonisten nicht einnehmen: aromatische Pi-Pi-Interaktionen (*pi-stacking*) mit Phe484. Mutationen von Phe484 verringern die Bindungsaffinität von 7-Chlorkynurensäure erheblich⁴⁷. Im GluR2 (AMPA-Rezeptor) entspricht Tyr450 dem Phe484 des NR1. 5,6-Dinitroquinoxalindion (DNQX), ein GluR2-Antagonist geht dort ebenfalls ein Pi-Stacking mit Tyr450 ein³⁹. Laut Furukawa und Gouaux sind Pi-Stackings offenbar ein gemeinsames Merkmal der Antagonisten der ionotropen Glutamaterezeptoren³⁷.

Kynurensäure ist ein deutlich weniger potenter Antagonist als mono- oder dihalogensubstituierte Kynurensäure. In der Kristallstruktur mit DCKA befinden sich die beiden Chloride in einem Abstand von 3,3 bzw. 3,4 Å von den beiden aromatischen Resten Trp731 und Phe408 und gehen dort van-der-Waals-Interaktionen ein. Die Chlorsubstitution verringert zudem den pK_a der Substituenten am Quinolinring. Es wird vermutet, dass durch diesen Effekt die Bindungsaffinität von DCKA erhöht wird, da sich die Hydroxyl-/Carbonylgruppe nahe des N-Terminus der Helix F befindet, die dort Anionen bindet⁴⁴.

Eine Schlüsselrolle fällt dem Rest Trp731 zu: er „spürt“ das Volumen des Liganden. In den Kristallstrukturen von ACPC und ACBC (Agonisten) sowie Cycloleucin (Antagonist) bleiben die Abstände zwischen Ligand und dem Indol des Trp731 gleich (3,4, 3,5 und 3,5 Å). Als Reaktion auf die zunehmende Größe der Liganden vergrößert sich der Torsionswinkel des Indols von 4,8° (ACPC) über 6° (ACBC) zu 10° (Cycloleucin). Diese Bewegung von Trp731 „öffnet“ also die Bindungsstelle für Antagonisten mit höherem Molekularvolumen.

1.3.3 Zusammenhang zwischen Domänenbewegung und Rezeptoraktivierung

Die kristallographischen Analysen von Inanobe³⁸ und Furukawa^{12,37} ergaben, dass Agonisten NR1 in einer geschlossenen Rezeptorkonformation stabilisieren, Antagonisten hingegen in der offenen oder Apo-Konformation. Diese Eigenschaft hat NR1 mit GluR2 (AMPA-Rezeptor) gemein. Der GluR2-Antagonist DNQX stabilisiert dort die offene Apo-Konformation³⁹. Die durch die Bindung eines Agonisten induzierte Konformationsänderung zur geschlossenen Form aktiviert den Ionenkanal⁴⁸.

Der Unterschied im Winkel der S1 und S2 Domäne von NR1 zwischen agonistischem und antagonistischem Zustand beträgt 21°. In der Serie der kristallisierten Cycloliganden ändert sich der funktionelle Charakter der Liganden mit ihrem Volumen: ACPC und ACBC sind Agonisten, Cycloleucin ein Antagonist. Die Einführung eines zusätzlichen Methylens in den Ring verursachte eine Volumenerhöhung des Liganden um je etwa 11 Å³. Furukawa und Gouaux postulieren, dass durch die Vergrößerung des Rings die Fähigkeit von NR1 verringert wird, die komplett geschlossene Konformation der Glycin-gebundenen Form anzunehmen³⁷.

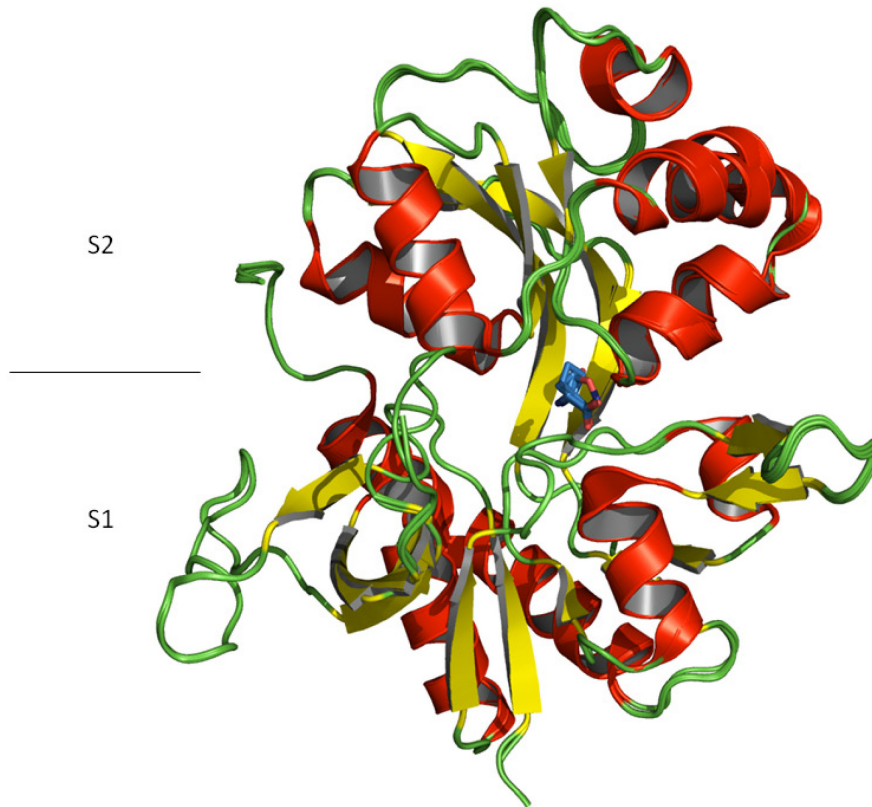


Abbildung 6: Überlagerung der agonistischen Kristallstrukturen von NR1
 Überlagerung der agonistischen Kristallstrukturen 1PB7 (Glycin), 1PB8 (D-Serin), 1PB9 (D-Cycloserin), 1Y1Z (ACPC) und 1Y20 (ACBC). Einfärbung des Proteins nach Sekundärstruktur (rot: Helices, gelb: Beta-Faltblätter, Grün: Loops). Die Liganden sind in blau dargestellt.

Die Kristallstrukturen von Glycin und D-Serin (Agonisten) sowie D-Cycloserin (partieller Agonist) zeigen den gleichen Grad an geschlossener Konformation, alle Kristalle sind isomorph. Eine Überlagerung der Kristallstrukturen ergab einen Unterschied (RMSD) von nur 0,1 bzw. 0,2 Å. In Bezug auf die Rezeptorkonformation unterscheiden sich somit volle und partielle Agonisten nicht (Abbildung 6). Die Überlagerung von agonistischer und antagonistischer Kristallstruktur hingegen zeigt große strukturelle Unterschiede (Abbildung 7). Trennt man jedoch die beiden Domänen S1 und S2 und überlagert jede individuell, so beträgt der RMSD weniger als 0,3 Å für S1 und 0,7 Å für S2³⁷. Es ist anzunehmen, dass die beiden Domänen eher rigide Körper sind, die nur über ein „Scharnier“ zueinander bewegt werden. Dieses Scharnier ist die Glycin_B-Bindungsstelle sowie die umgebenden Strukturelemente im Interface der Domänen S1 und S2.

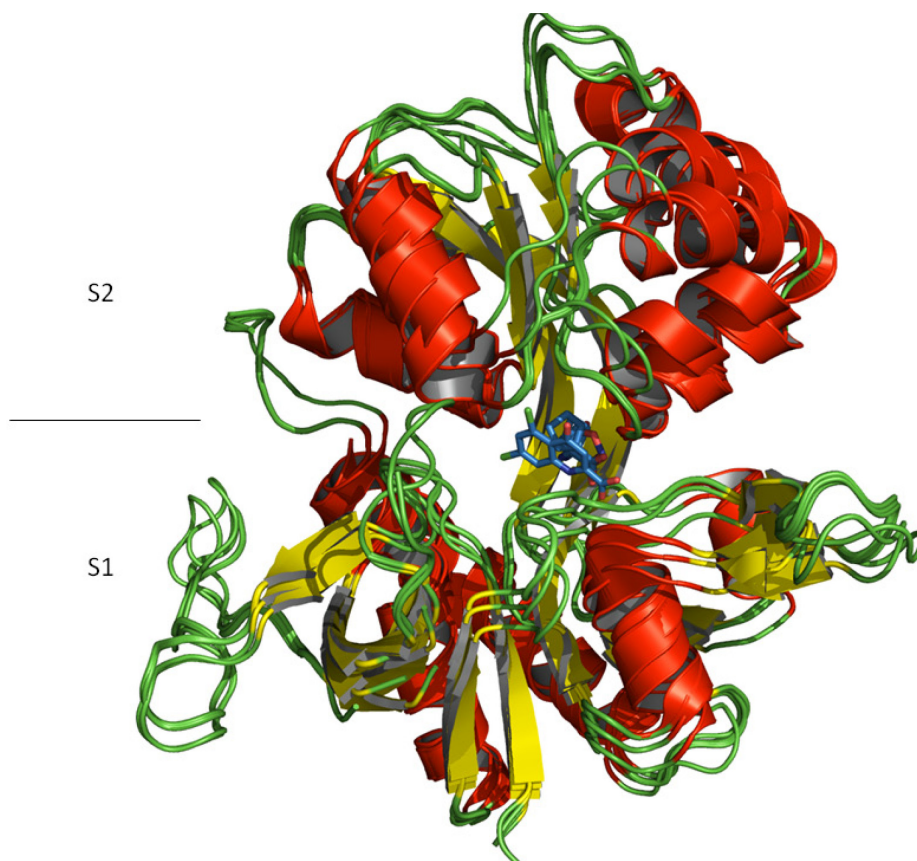


Abbildung 7: Überlagerung aller Kristallstrukturen von NR1
Überlagerung der agonistischen Kristallstrukturen 1PB7 (Glycin), 1PB8 (D-Serin), 1PB9 (D-Cycloserin), 1Y1Z (ACPC) und 1Y20 (ACBC) sowie der antagonistischen Kristallstrukturen 1Y1M (Cycloleucin) und 1PBQ (DCKA). Einfärbung des Proteins nach Sekundärstruktur (rot: Helices, gelb: Beta-Faltblätter, Grün: Loops). Die Liganden sind in blau dargestellt.

1.4 Virtuelle Wirkstoffidentifizierung (*Virtual Screening*)

Die zunehmend beschränkte Verfügbarkeit von freiem chemischen Raum und das dadurch bedingte Fehlen von Startstrukturen für die Synthese von potentiellen Wirkstoffmolekülen ist ein Problem, welches sich für viele Wirkstoffdesign-Projekte stellt⁴⁹. „Scaffold hopping“⁵⁰ oder „Rescaffolding“, also die Identifizierung isofunktionaler molekularer Kernstrukturen, die möglicherweise nicht (mehr) patentgeschützt sind, ist gewöhnlich eine der größten Herausforderungen in der frühen Entwicklungsphase von Wirkstoffen⁵¹.

Während der gesamte chemische Raum in seiner Dimension je nach Quelle auf 10^{60} bis 10^{100} Strukturen geschätzt wird⁵²⁻⁵⁵, beträgt die Anzahl an zurzeit kommerziell verfügbaren chemischen Substanzen nur etwa 6-8 Millionen unterschiedliche Moleküle^{56,57}. Diese deutlich geringere Anzahl an Verbindungen wird zunehmend für Methoden wie Ultrahochdurchsatz-Screening⁵⁸ (*ultra-High-Throughput Screening*, uHTS) zugänglich. Da jedoch der (u)HTS-Prozess sehr kostenintensiv ist⁵⁹ und zudem die Trefferraten gewöhnlich nur sehr gering sind⁶⁰, erfreut sich das Konzept des fokussierten Bibliothekendesigns zuneh-

menden Interesses⁶¹⁻⁶³. Hierbei werden Substanzen z.B. anhand chemieinformatischer Methoden ausgewählt, um eine mit potentiell aktiven Substanzen angereicherte Bibliothek von Screeningsubstanzen zu erstellen, die dann im biologischen Experiment getestet wird. Für pharmazeutische Unternehmen sind virtuelle Wirkstoffidentifizierung (*Virtual Screening*⁶⁴⁻⁶⁶) und Bibliotheksdesign zusätzlich aus finanziellen und logistischen Gründen attraktiv: anstelle riesige Sammlungen von Screeningsubstanzen im Hochdurchsatzverfahren zu testen, werden eigene Substanzbibliotheken oder Bibliotheken verschiedener externer Anbieter *in silico* durchsucht und nur die virtuellen Treffer (*hits*) erworben und *in vitro* auf biologische Aktivität gemessen^{67,68}.

Typische Methoden des Virtual Screenings sind Maschinenlernverfahren⁶⁹ (z.B. künstliche neuronale Netze⁷⁰, Support-Vector-Maschinen^{71,72} und evolutionäre Optimierungsverfahren⁷³), Ähnlichkeitssuchen^{74,75}, Docking⁷⁶⁻⁷⁸ oder Pharmakophorsuchen⁷⁹⁻⁸². In dieser Arbeit wurde eine Auswahl an Methoden des Virtual Screenings angewandt, um nach neuartigen Glycin_B-Antagonisten zu suchen.

1.5 Das Problem der Proteinflexibilität

Es gilt allgemein als anerkannt, dass Emil Fischers „Schlüssel-Schloss-Prinzip“⁸³ (Abbildung 8a) kein hinreichendes Modell für die Interaktion von Wirkstoffen mit ihren Rezeptoren darstellt^{84,85}. Modernere Theorien sprechen eher von einem „Hand-Handschuh-Prinzip“, bei dem sich der Rezeptor seinem Liganden während der Bindung anpasst, und umgekehrt. Beide Partner der Bindung sind flexibel und können ihre Form so verändern, dass sie komplementär zueinander sind⁸⁶. So werden günstige Interaktionen hergestellt, ungünstige Kontakte vermieden und schlussendlich die freie Bindungsenergie maximiert⁸⁷. Man spricht hierbei auch vom „induzierten Passform-Effekt“ (*induced fit*) (Abbildung 8b).

Andere Autoren gehen eher von einer Vielzahl an konformerer Zuständen aus, in denen die Rezeptormoleküle vorzufinden sind⁸⁸. Die Häufigkeit des Auftretens eines bestimmten Konformerer hängt sowohl von seiner Umgebung (z.B. pH, Lösungsmittel, etc.), als auch von der An- oder Abwesenheit von Liganden ab⁸⁹. Zudem sind nicht alle Rezeptorkonformationen gleichermaßen in der Lage, kleine Moleküle zu binden. Einige Konformere des Rezeptors können den Liganden aufnehmen, ohne größere Anpassungen in der Bindungsstelle z.B. durch Rotation von Seitenketten vorzunehmen⁹⁰, während andere Konformationen weniger bis gar nicht in der Lage sind, den Liganden aufzunehmen, da z.B. die Bindungsstelle sterisch blockiert ist⁹¹ (Abbildung 8c). Die Bindung eines Liganden beeinflusst das dynamische Gleichgewicht. Das Konformer, welches den Liganden am wahrscheinlichsten aufnimmt, bildet einen

Protein-Ligand-Komplex und wird somit aus der Population der freien Rezeptoren entfernt; andere Konformere passen sich an, um diesen nun leeren Konformationsraum abzudecken⁸⁸. So wird der Bindungsprozess vorwärtsgetrieben, bis schließlich alle Rezeptormoleküle mit ihrem Ligand gesättigt sind oder ein von der Bindungsstärke des Liganden abhängiges Gleichgewicht erreicht wird⁹¹. Diese Theorie wird gestützt von experimentellen Ergebnissen aus der Kernresonanzspektroskopie (*nuclear magnetic resonance spectroscopy* - NMR)⁹² und Kinetikstudien⁹³.

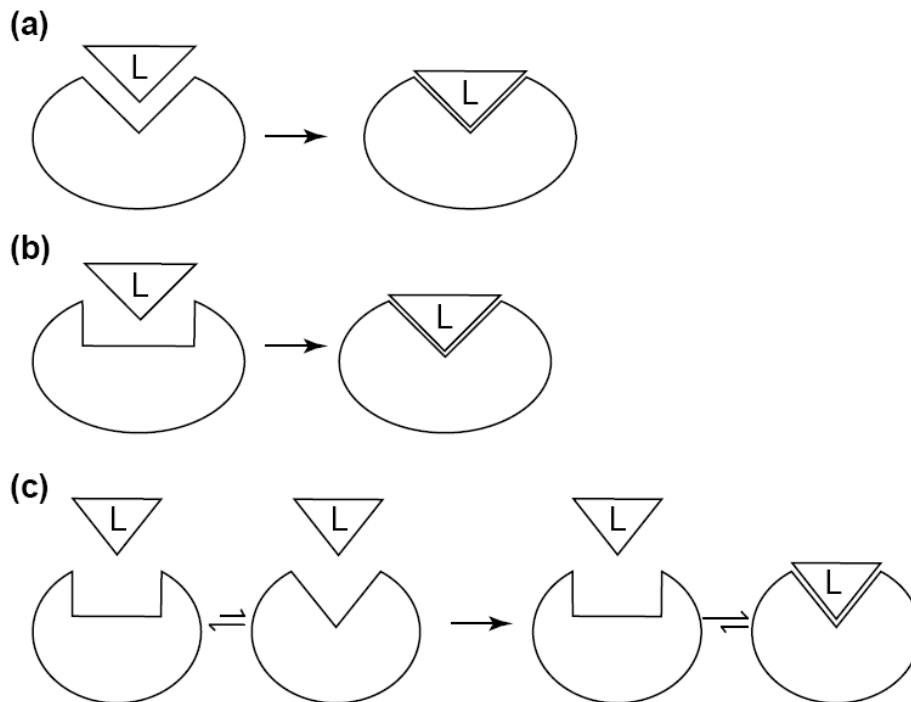


Abbildung 8: Theorien für die Protein-Ligand-Interaktion
 (a) Schlüssel-Schloss-Modell (b) Induced-fit Modell (c) Modell der Konformerengleichgewichte. Der Ligand ist hierbei als Dreieck mit Markierung "L" dargestellt. Zur Vereinfachung wurde die Flexibilität des Liganden nicht berücksichtigt. Abbildung nach Goh *et al.*⁹⁴

McCammon *et al.* führten basierend auf diesen Erkenntnissen das sog. „entspannter-Komplex-Schema“ (*relaxed-complex scheme*) ein⁹⁵, welches die Tatsache berücksichtigt, dass ein Ligand nur in einige energiearme Konformationen des Rezeptors binden kann. Hierzu wurde eine kurze Moleküldynamiksimulation der Apo-Struktur des Rezeptors durchgeführt, um den Konformerenraum des Proteins zu durchmustern. Daraufhin wurden kleine Ligandenbibliotheken in die verschiedenen Zwischenstadien der Simulation, die sog. *Snapshots* gedockt und die freien Energien der Komplexe berechnet. Korrekte Bindungsposen fanden sich nur in Komplexen mit der niedrigsten freien Energie. Das *relaxed-complex scheme* ist eines der ersten Verfahren, das versucht, Proteinflexibilität explizit berechnen und für das Virtual Screening nutzbar zu machen.

Das Prinzip multipler Rezeptorstrukturen macht sich auch das sog. Ensemble-Docking⁹⁶ zu eigen: Hierbei werden mehrere Proteinkonformationen parallel⁹⁷ oder in einer gemittelten Darstellung⁹⁸ als Rezeptorstruktur während des Dockings verwendet. Ein großes Problem hierbei ist jedoch eine starke Verringerung der Selektivität des Modells: mit steigender Anzahl an Rezeptorkonformationen erhöht sich die Anzahl der Ligandenmoleküle, die in eine der Konformationen gedockt werden können, obwohl diese nicht der bioaktiven Konformation entspricht. Somit kann auch für einen komplett inaktiven Liganden eine (falsche) passende Dockingpose erzeugt werden. Aktuelle Scoringfunktionen sind nur bedingt in der Lage, dieses Problem zu erkennen⁹⁹.

Luque & Freire konnten zeigen, dass die Bindungsstelle von Enzymen sowohl aus niedrig- als auch hochflexiblen Bereichen besteht¹⁰⁰. Flexibilität ist also keine intrinsische Eigenschaft eines Proteins, sondern vielmehr nur bestimmter Regionen oder sogar nur spezieller Aminosäuren, z.B. in sog. Hinge-Regionen, die als bewegliches „Gelenk“ zwischen eher rigiden Untereinheiten eines Proteins fungieren. Die Identifikation dieser mobilen Bereiche ist von hoher Wichtigkeit sowohl für die Modellierung von Proteinen und das Verständnis ihrer Funktion, als auch für die Suche nach neuen Wirkstoffen, z.B. im virtuellen Screening oder im rationalen Wirkstoffdesign¹⁰¹.

1.6 Ziele dieser Arbeit

Anhand des Beispielsystems der Glycin_B-Bindungsstelle des NMDA-Rezeptors wurden im Zuge dieser Arbeit die folgenden grundlegenden Fragestellungen beantwortet:

- Ist es möglich, mittels chemieinformatischer Methoden neuartigen chemischen Raum auf einem pharmakologischen Zielsystem zu identifizieren, welches seit über 20 Jahren Gegenstand intensiver medizinisch-chemischer und pharmakologischer Forschung ist? Welche Methoden des Virtual Screenings eignen sich besser für die Identifizierung neuartiger Antagonisten für die Glycin_B-Bindungsstelle?
- Können Aussagen über den Bindungsmodus zahlreicher Referenzliganden getroffen werden, die sich von den aus Röntgenkristallstrukturen bekannten Glycin_B-Antagonisten in Form und Größe stark unterscheiden und somit nicht in die Bindungsstelle in ihrer bekannten Form eingepasst werden können?
- Welchen pharmakologischen Sinn hat die hohe Rezeptorflexibilität der Glycin_B-Bindungsstelle? Gibt es Bereiche, die stärker oder weniger stark flexibel sind?

- Kann ein generelles Verfahren des Molecular Modelings erarbeitet werden, anhand dessen flexible Bereiche oder individuelle flexible Aminosäuren in Proteinen identifiziert werden können? Kann die so gewonnene Erkenntnis in die Vorhersage des Bindungsverhaltens neuartiger Liganden einfließen?
- Welche alternativen Möglichkeiten gibt es zu den aktuellen Scoringfunktionen des Molecular Dockings, um die Bindungsaffinität von Liganden zu quantifizieren? Kann eine solche neuartige Scoringfunktion die „Natur“ der Bindung zwischen Protein und Ligand in Bezug auf ihre Thermodynamik beschreiben und somit einen hilfreichen Beitrag für das Wirkstoffdesign leisten?

2 Material und Methoden

2.1 Sammlung einer Bibliothek von Glycin_B-Referenzliganden

Um Referenzliganden mit einer definierten biologischen Aktivität auf der Glycin_B-Bindungsstelle des NMDA-Rezeptor zu identifizieren wurde eine Literatur- und Patentrecherche durchgeführt; zusätzlich wurden die Wirkstoffdatenbank WOMBAT¹⁰² 2007.2 durchsucht. Obwohl biologische Aktivitäten in Publikationen und Datenbanken zwischen verschiedenen Experimentalsystemen (Assays) oder sogar zwischen unterschiedlichen ausführenden Forschungsgruppen stark variieren können, wurde eine Klassifizierung der gefundenen Substanzen in drei Gruppen durchgeführt: hochaffine Substanzen (annotierter IC_{50} unter 1 μ M), mittel- bis schwachaffine Substanzen (annotierter IC_{50} zwischen 1 und 50 μ M) und inaktive Substanzen (annotierter IC_{50} über 50 μ M oder keine Annotation für Glycin_B-Aktivität). Aufgrund der erheblichen Vielfalt an Literaturliganden (1123 Strukturen, davon 754 mit präziser Annotation der biologischen Aktivität) wurde das Risiko der gelegentlichen inkorrekten Klassifizierung einer hochaffinen Substanz in die Kategorie der Mittel- bis Schwachaffinen oder umgekehrt z.B. aufgrund experimenteller Abweichungen als nicht signifikant betrachtet. Insgesamt wurden 429 nanomolare sowie 288 weitere Substanzen mit einem IC_{50} unter 50 μ M identifizierten. Sämtliche Referenzsubstanzen mit ihrer entsprechenden Literaturannotation sind im Anhang aufgeführt.

2.2 Etablierung einer automatisierten Plattform zur *in silico* Aufarbeitung von Screeningsubstanzen

2.2.1 Zusammenstellung der Substanzkataloge

Die Größe des Raums an zurzeit kommerziell verfügbaren chemischen Substanzen beträgt etwa 6-8 Millionen verschiedene Moleküle^{56,57}. Dabei werden unterschiedliche Isoformen wie z.B. Salze nicht berücksichtigt. Die Verbindungen werden von diversen Anbietern in Katalogen verschiedener Größe angeboten. In einem grundlegenden Schritt wurden für diese Arbeit 46 Substanzkataloge zusammengetragen und eine Software entwickelt, die die Katalogdaten entsprechend automatisiert aufbereitet und in einer virtuellen Screeningbibliothek vereinigt.

Eine zusätzliche logistische Herausforderung war die Bevorzugung bestimmter Anbieter z.B. aus Gründen der generellen Zuverlässigkeit oder aufgrund von Rahmenverträgen. Hierfür wurde eine Priorisierungsfunktion implementiert, die

es erlaubt, Substanzen gemäß der Bevorzugung des Anbieters in den Screeningdatensatz aufzunehmen, d.h. war eine Substanz gleichzeitig bei einem Anbieter hoher und einem Anbieter niedriger Priorität verfügbar, so wurde automatisch die Substanz des hochprioritären Anbieters ausgewählt. Die Information, dass die jeweilige Substanz bei Anbietern niedriger Priorität ebenfalls verfügbar ist, wurde hierbei allerdings nicht verworfen, sondern ebenfalls in die Datenbank mit aufgenommen. Im Falle der Nichtverfügbarkeit beim Primäranbieter konnte so schnell auf einen Alternativanbieter zurückgegriffen werden.

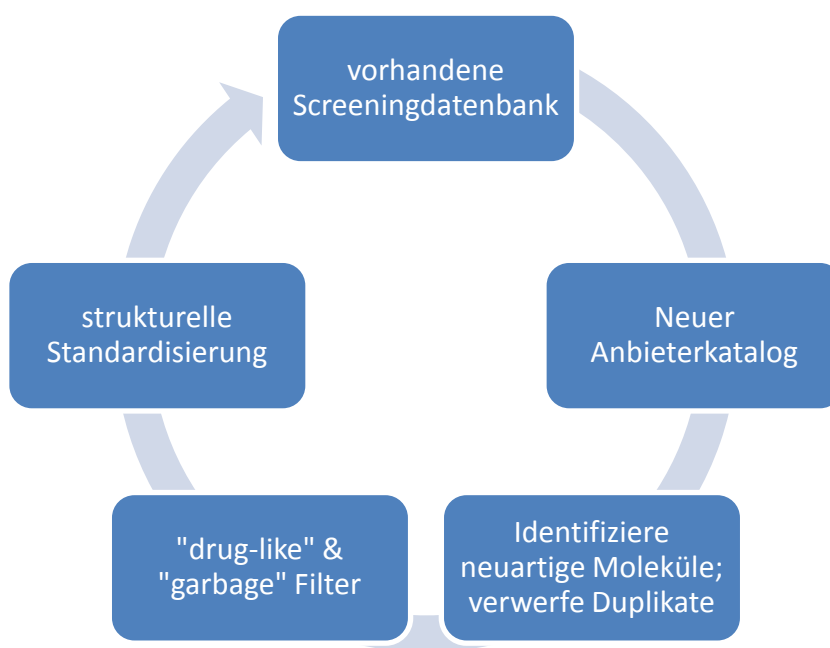


Abbildung 9: Zyklische Generierung der Screeningdatenbank

Insgesamt wurden etwa 8,9 Millionen Substanzeinträge aus 46 unterschiedlichen Katalogen von 33 verschiedenen Anbietern im SD-Datenformat zusammengetragen und anhand logistischer Kriterien (Zuverlässigkeit, Verfügbarkeit, Substanzkosten) priorisiert (Tabelle 3). Die Implementierung der automatisierten Katalogverwaltung erfolgte in der Software Pipeline Pilot 7.5 (Accelrys Software Inc., San Diego, 2008). In einem ersten Schritt wurde die Benennung der Strukturen standardisiert. Hierzu wurde das jeweilige Eigenschaftsfeld für den Substanznamen des jeweiligen Katalogs identifiziert und zusammen mit dem Namen des Anbieters in einem neuen Eigenschaftsfeld konkateniert, um bei mehreren Anbietern verfügbare Substanzen in der vereinigten Substanzdatenbank wiedererkennen zu können. Zusätzliche Informationen aus den Anbieterkatalogen wie Preise, Reinheit etc. wurden verworfen.

Jeder Anbieterkatalog wurde nun iterativ mit der bereits erstellten Screeningdatenbank verglichen: neue Strukturen wurden aufgenommen und Duplikate

verworfen. Hierfür wurde die Funktion „Find Novel Molecules“ von Pipeline Pilot verwendet, die im Wesentlichen auf einem Vergleich der Canonical SMILES^{103,104} der einzelnen Strukturen basiert (Abbildung 9).

Tabelle 3: Anbieter von Screeningsubstanzen
sortiert nach vorgenommener Priorisierung des Anbieters

Anbieter	Anbieter-Webseite	angebotene Substanzen	gefiltert / Duplikate *	verwendete Substanzen
Specs	www.specs.net	200.029	31.631	168.398
Asinex	www.asinex.com	354.629	74.223	280.407
Enamine	www.enamine.net	1.168.625	99.939	1.068.686
ChemDiv	www.chemdiv.com	761.189	177.072	584.117
AMRI	www.amriglobal.com	360.890	57.590	303.300
Chembridge	www.chembridge.com	454.761	192.046	262.715
Maybridge	www.maybridge.com	57.576	13.532	44.044
VitaSM	www.vitasmlab.com	361.461	241.858	119.603
Timtec	www.timtec.net	170.140	128.923	41.217
Keyorganics	www.keyorganics.ltd.uk	47.581	10.154	37.427
InterBioScreen	www.ibscreen.com	435.868	209.841	226.027
Analyticon	www.ac-discovery.com	25.270	3.021	22.249
Life Chemicals	www.lifechemicals.com	288.088	111.439	176.649
Pharmeks	www.pharmeks.com	185.327	104.441	80.886
Otava	www.otavachemicals.com	118.184	92.095	26.089
Nanosyn	www.nanosyn.com	65.325	53.209	12.116
Chemstar	www.chemstar.com	59.968	48.326	11.642
Acros	www.acros.com	15.706	6.657	9.049
TOSlab	www.toslab.com	16.419	9.763	6.656
NCH	www.nchlab.com	215.244	184.991	30.253
Akos	www.akosgmbh.de	3.402.761	2.364.722	1.038.039
MS Discovery	www.msdiscovery.com	2.000	1.045	955
Oakwood	www.oakwoodchemical.com	16.890	10.000	6.890
ACB Blocks	www.acbblocks.com	2.765	1.148	1.617
Prestwick	www.prestwickchemical.fr	1.120	757	363
Frinton	www.frinton.com	1.130	761	369
Ubichem	www.ubichem.com	470	146	324

* Duplikate wurden sukzessive gefiltert, d.h. für die Suche nach Duplikaten wurden die bereits gefilterten höher priorisierten Kataloge als Suchreferenz akkumuliert

Apollo	www.apolloscientific.co.uk	31.949	18.439	13.510
Bachem	www.bachem.com	2.956	1.814	1.142
Biofocus	www.biofocus.com	1.301	58	1.243
Combiblocks	www.combi-blocks.com	1.747	1.566	181
Alfa	www.alfa.com	27.556	24.079	3.477
Peptech	www.peptechcorp.com	1.503	35	1.468
Gesamt		8.856.428	4.275.321	4.581.108

Insgesamt wurden in verschiedenen Katalogen unterschiedlicher Anbieter etwa 8,8 Millionen Substanzen angeboten. Nach sukzessiver Entfernung von Duplikaten und der Anwendung von „drug-like“ und „garbage“ Filtern (Kapitel 2.2.2) verblieben noch etwa 4,6 Millionen unterschiedliche Substanzen (Tabelle 3).

2.2.2 Aufbereitung der zweidimensionalen Molekülstrukturen

Aufgrund der hohen molekularen Diversität der angebotenen Chemikalien (nicht alle Anbieter bieten ausschließlich Substanzen an, die für die Arzneimittelforschung in Frage kommen) wurden zwei relativ einfache Vorfilter implementiert, um die Wirkstoffähnlichkeit („*drug-likeness*“) der Screeningsubstanzen sicherzustellen.

Als erster Filterschritt wurden ein REOS-Filter¹⁰⁵ sowie eine Suche nach unerwünschten Substrukturen („*garbage*“) nach Hann *et al.*¹⁰⁶ angewandt. Diese Filter suchen nach Molekülen, die aufgrund ihrer strukturellen Eigenschaften eine hohe Wahrscheinlichkeit aufweisen, in der späteren Entwicklung aufgrund sicherheitspharmakologischer Kriterien durchzufallen oder falsch-positive Ergebnisse in biologischen Assays zu erzeugen¹⁰⁶. Hierzu gehören z.B. Toxikophore, also Strukturelemente, die bekanntermaßen eine toxische Wirkung im tierischen oder menschlichen Metabolismus entwickeln. Zur Filterung der Screeningdatenbank wurde eine Liste von unerwünschten Substrukturen im SMARTS-Format^{103,104} zusammengestellt und mit der Technik des Patternmatchings auf SMARTS-Ausdrücken von Pipeline Pilot (Komponente „Substructure Filter from File“) als Suchmuster angewandt. Die komplette Liste der entsprechenden SMARTS-Ausdrücke befindet sich im Anhang.

Im letzten Schritt wurde ein Filter entsprechend Lipinskis „Rule of five“¹⁰⁷ eingesetzt. Die „Rule of five“ stellt eine simple statistische Regel für die orale Bioverfügbarkeit einer Substanz dar:

- Molekülmasse ≤ 500 g/mol
- Anzahl Wasserstoffbrückenbindungs-Akzeptoren ≤ 10
- Anzahl Wasserstoffbrückenbindungs-Donoren ≤ 5
- Oktanol-Wasser-Verteilungskoeffizient ($\log P$) ≤ 5

Da eine blinde Anwendung der „Rule of five“ mittlerweile nicht mehr unumstritten ist¹⁰⁸, wurde eine „weichere“ Version implementiert, die die Verletzung von maximal einer der „Rule of five“ toleriert, um nicht zu viele potentiell interessante Substanzen bereits in diesem Schritt zu verwerfen.

Die verbliebenen Moleküle wurden anschließend standardisiert („gewaschen“): Ionen wurden entfernt und die verbleibende Struktur neutralisiert (Säuren protoniert, Basen deprotoniert). Die Implementierung dieses Arbeitsschritts erfolgte anhand der Funktion „Standardize Molecule“ in Pipeline Pilot. In Fällen, in denen keine explizite Annotation von Stereozentren in der Originalstruktur verfügbar war, wurden die Stereozentren mit einer Markierung für unbekannte Stereochemie versehen. Im umgekehrten Fall, wenn also ein Stereozentrum annotiert war, welches tatsächlich gar nicht zur Bildung von Enantiomeren fähig ist, wurde die falsche Annotation entfernt (Pipeline Pilot, „Standardize Stereo“). Formalladungen für bestimmte Gruppen wurden in ein Standardformat gebracht („Standardize Charges“, PLP) (Abbildung 10).

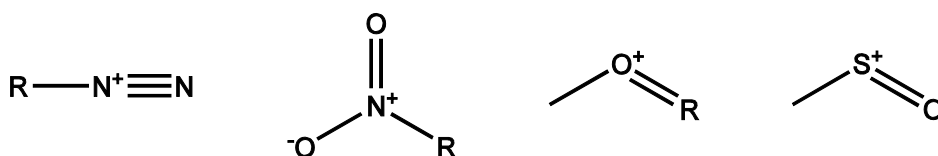


Abbildung 10: Standardisierte Formalladungen

Explizite Wasserstoffe wurden – so vorhanden – entfernt („Remove Hydrogens“, Pipeline Pilot), dreidimensionale Atomkoordinaten wurden gelöscht und die zweidimensionalen Atomkoordinaten aus der angegebenen Konnektivitätstabelle neu berechnet („Generate 2D Coordinates“, Pipeline Pilot). Die abschließend erzeugte 2D-Screeningdatenbank wurde im SD-Format abgespeichert.

2.2.3 Aufbereitung der dreidimensionalen Molekülstrukturen

Während die obige Prozedur zur Aufbereitung der zweidimensionalen Molekülstrukturen für alle in dieser Arbeit zur Anwendung gekommenen chemieinformatischen Verfahren ausreichend war, benötigte die Anwendung von dreidimensionalen (3D) Verfahren eine weitere Dimension nicht nur im kartesischen Raum, sondern auch in technischer Hinsicht aufgrund eines deutlich intensive-

ren Durchmustern des Suchraums. Für jede Substanz, die als zweidimensionale Struktur (Konnektivitätstabelle) vorlag, wurden alle wahrscheinlichen Protomerstadien sowie bis zu 8 Tautomere und bis zu 32 möglichen Stereoisomere im Bereich von pH 7 ± 3 berechnet. Für explizit annotierte Stereozentren wurde keinen zusätzlichen Stereoisomere erzeugt.

Zur Strukturberechnung wurden die Programme LigPrep 2.2 und Epik 1.6 (Schrödinger, LLC, New York, 2008) verwendet. Mithilfe von Epik, einem Programm zur Berechnung von pK_a -Werten für ionisierbare Gruppen in Liganden, wurden wahrscheinliche geladene und tautomerisierte Strukturen innerhalb des pH-Bereiches von pH 7 ± 3 generiert. Epik basiert dabei auf den empirischen Gleichungen von Hammett und Taft¹⁰⁹. Pro Struktur wurden bis zu 32 Stereoisomeren erzeugt, falls an Stereozentren keine explizite Stereoannotation vorhanden war. Explizit annotierte Stereozentren wurden respektiert, d.h. es wurden hier keine weiteren Stereoisomere generiert. Falls Moleküle Ringe enthielten, wurde eine Energieminimierung im OPLS2005¹¹⁰ Kraftfeld durchgeführt und die optimale (energieärmste) Ringkonformation behalten. Die 3D-Screeningdatenbank wurde anschließend im Maestro-Datenformat abgespeichert. Die Berechnung der Isomeren wurde „trivial parallelisiert“ über einen 26-CPU Cluster verteilt, d.h. die Ausgangsdateien wurden automatisiert in 26 gleich große Stücke zerteilt und dann auf einem Grid verteilt. Für diese und alle folgenden Parallelisierungen wurde die Sun Grid Engine¹¹¹ implementiert und verwendet.

Um den dreidimensionalen Strukturraum möglichst vollständig zu durchmustern, wurden für jede der zuvor berechneten Isomere bis zu 250 Konformere mittels ConfGen 2.0 (Schrödinger, LLC, New York, 2008) generiert. Der Algorithmus folgt dabei folgendem Muster:

Ein gegebenes Molekül wird in einen festen Kern und eine Rotamergruppe unterteilt. Die Rotamergruppe ist dabei definiert als eine endständige Gruppe, die selbst keine rotierbaren Einfachbindungen mehr verfügt, aber über eine solche an den Kern gebunden ist. Bindungen zu Wasserstoffen werden hierbei nicht berücksichtigt. Die Konformerensuche generiert zunächst iterativ alle möglichen Kern-Konfigurationen und variiert dann die peripheren Konfigurationen (Option „thorough sampling“, ConfGen). Redundante Konformere, also Konformere, deren Energie sich in einem definierten Energiefenster zu anderen Konformeren befand, wurden eliminiert. Ebenso wurden Konformere nicht weiter betrachtet, deren Energie über einem definierten Schwellenwert von 100 KJ/mol lag, da die Wahrscheinlichkeit hier sehr hoch war, dass es sich nicht um eine bioaktive Konformation handelte.

2.3 Distanzmetriken

2.3.1 Tanimoto-Ähnlichkeit

Ausgehend von der Repräsentation eines Moleküls durch seinen Deskriptor können in maschinellen Verfahren die Unterschiede zwischen zwei Molekülen als Funktion ihrer Deskriptoren berechnet werden. Hierbei kommen unterschiedliche Ähnlichkeitsmaße respektive Distanzmetriken zum Einsatz¹¹²:

Die Tanimoto-Ähnlichkeit^{113,114} (Gleichungen 1 und 2) vergleicht die eindimensionalen Fingerprints von zwei Molekülen A und B mit XOR-Arithmetik

$$T = \frac{c}{a + b - c}, \text{ wobei} \quad (\text{Gleichung 1})$$

a die auf „1“ gesetzten Bits im Molekül A bezeichnet, b die auf „1“ gesetzten Bits im Molekül B, und c die auf „1“ gesetzten Bits in beiden Molekülen. Die Tanimoto-Ähnlichkeit T befindet sich im Intervall $[0,1]$. 0 bezeichnet hierbei die maximale paarweise Unähnlichkeit und 1 die Identität zweier Strukturen.

Wird die Tanimoto-Ähnlichkeit für nichtbinäre Fingerprints berechnet, so gilt:

$$S_{A,B} = \frac{\sum_{j=1}^{j=n} X_{jA} X_{jB}}{\sum_{j=1}^{j=n} (X_{jA})^2 + \sum_{j=1}^{j=n} (X_{jB})^2 - \sum_{j=1}^{j=n} X_{jA} X_{jB}}, \text{ wobei} \quad (\text{Gleichung 2})$$

A, B die zu vergleichenden Moleküle sind und j die Position im jeweiligen Fingerprint. $S_{A,B}$ befindet sich im Intervall $[-1/3, 1]$, wobei auch hier 1 die Identität bezeichnet.

Die Minkowski-Distanz (Gleichung 3) wurde von Herrmann Minkowski im 19. Jahrhundert als eine Metrik erdacht, in der die Distanz zwischen zwei Punkten der Summe der absoluten Differenzen ihrer Koordinaten entspricht¹¹⁵. Der sich aus den jeweiligen Funktionen ergebende Wert befindet sich jeweils im Intervall $[0, \infty]$, wobei hierbei der Wert 0 bei Identität der verglichenen Deskriptoren erreicht wird und sich mit zunehmender Distanz immer weiter vergrößert.

$$D_{A,B} = \left[\sum_{j=1}^n \left(|X_{jA} - X_{jB}| \right)^t \right]^{\frac{1}{t}} \quad (\text{Gleichung 3})$$

Ein grundlegender Unterschied zwischen den Minkowski-basierten Distanzfunktionen auf der einen und der Tanimoto-Ähnlichkeit als fingerprintbasierter

Distanzmetrik auf der anderen Seite ist die nicht vorhandene Berücksichtigung fehlender Attribute in der Tanimoto-Ähnlichkeit. Hier werden somit nur vorhandene Attribute wie zum Beispiel molekulare Substrukturen betrachtet, wohingegen die Minkowski-basierenden Funktionen auch nicht vorhandene Fragmente in der Distanzberechnung berücksichtigen¹¹⁶. Im Zusammenhang chemischer Untersuchungen stellten James *et al.* die These auf, dass Distanzmetriken auf Deskriptoren nur als Maß „relativer“ (das heißt als Maß der Distanz zweier Moleküle zu einem Zielmolekül), aber nicht „absoluter“ Distanz (Unterschied zwischen zwei unabhängigen Molekülen) nützlich sind¹¹⁷.

2.3.2 Abweichung vom quadratischen Mittel

Die Abweichung des quadratischen Mittels (*root mean square deviation* - RMSD) erlaubt die Bestimmung der strukturellen Ähnlichkeit zweier dreidimensionaler Objekte. Der RMSD zweier Strukturen entspricht dabei der Quadratwurzel der Summe der quadrierten Distanzen zwischen äquivalenten Atomen *A* und *B* (Gleichung 4).

$$RMSD = \sqrt{\frac{\sum_{i=1}^N (x_{Ai} - x_{Bi})^2}{N}}, \text{ wobei} \quad (\text{Gleichung 4})$$

N die Anzahl der Atome, *A* und *B* unterschiedliche molekulare Strukturen beschreibt.

2.3.3 Receiver-Operating-Characteristics

Als Maß der Qualität der Ergebnisse verschiedener Methoden des Virtual Screenings können ROC Kurven^{118,119} (Receiver-Operating-Characteristics) erzeugt werden; hierbei wird die Sensitivität gegen die Spezifität (Gleichung 5) des Modells aufgetragen, Sensitivität und Spezifität werden hierbei durch die Anzahl an korrekt positiven (TP), korrekt negativen (TN), falsch positiven (FP) und falsch negativen (FN) Werten berechnet:

$$\text{Sensitivität} = \frac{TP}{TP + FN}; \text{ Spezifität} = \frac{TN}{TN + FP} \quad (\text{Gleichung 5})$$

Die Werte für Sensitivität und Spezifität liegen dabei im Intervall [0,1]. Ein auf einer Zufallsauswahl basierendes Modell wird demnach durch eine ROC-Kurve repräsentiert, die als Winkelhalbierende die Punkte (0,0) und (1,1) verbindet (Abbildung 11).

2.3.4 Matthews-Korrelationskoeffizient

Der Korrelationskoeffizient nach Matthews¹²⁰ (Matthews CC – Gleichung 6) ist eine weitere Möglichkeit, Sensitivität und Spezifität einer Klassifikationsmethode zu beurteilen, wobei der Matthews CC von -1 (antikorreliertes Modell) bis +1 (perfektes Vorhersagemodell) reicht; ein Modell, welches nur auf Zufall basierend klassifiziert, wird mit einem Matthews Score von 0 bewertet.

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} , \quad (\text{Gleichung 6})$$

wobei TP die Anzahl an korrekt positive Ergebnissen, TN Anzahl korrekt negativer Ergebnisse, FP falsch positive und FN falsch negative Ergebnisse bezeichnet.

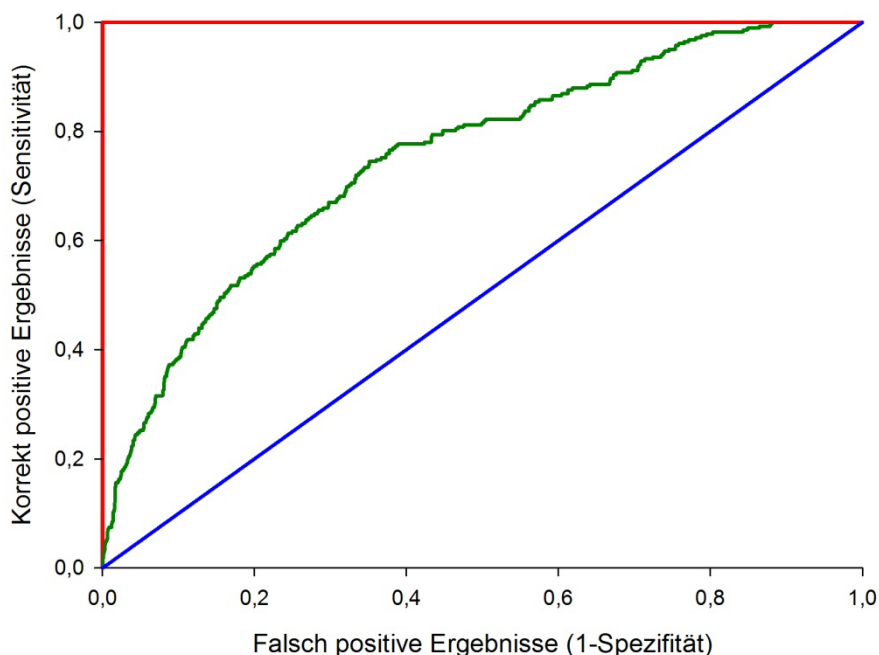


Abbildung 11: Hypothetische Receiver-Operating-Characteristics

Die rote Kurve entspricht einem Modell mit perfekter Vorhersagekraft; die blaue Kurve entspricht einem Zufallsmodell. Die ROC-Kurve eines Modells liegt gewöhnlich zwischen Zufallskorrelation und perfekter Vorhersage (grüne Kurve).

2.4 Diversitätsanalyse der Substanzbibliotheken – der chemische Raum von Glycin_B

Zur Ermittlung der Diversität des zusammengestellten Screeningdatensatzes des „verfügbaren chemischen Raumes“ (VCR) wurden für alle Moleküle der Functional Connectivity Fingerprint über die Länge von sechs Bindungen (FCFP-6)⁷⁵ anhand ihrer Implementierung in Pipeline Pilot 7.5 berechnet. Basis

der Connectivity Fingerprints als Fragmentdeskriptoren ist die Betrachtung ringförmiger Substrukturen. Dabei wird jedes Atom als String von „erweiterten Konnektivitätswerten“ (*extended connectivity*) repräsentiert, die anhand eines leicht modifizierten Morgan-Algorithmus¹²¹ ermittelt werden. Unterschieden werden Extended Connectivity Fingerprints (ECFP) und Functional Connectivity Fingerprints (FCFP), die jeweils von einer unterschiedlichen Repräsentation ihrer Atomtypen ausgehen: Während der Generierung des ECFP wird jedem Atom ein Code basierend auf der Anzahl seiner Bindungen, seinem Elementtyp, seiner Ladung und Masse zugewiesen. Beim FCFP hingegen wird jedem Atom einer von insgesamt sechs pharmakophoren Atomtypen zugewiesen (Wasserstoffbrückendonator, Wasserstoffbrückenakzeptor, positiv ionisierbar, negativ ionisierbar, aromatisch, Halogen). Für jedes Atom wird nun diese Codierung zusammen mit der Information über seine Bindungen an Nachbaratome und deren Atomtypen gespeichert. Dieser Schritt wird so lange iteriert, bis ein vorgegebener Abstand der Atome in Form von Bindungslängen erreicht ist.

Da die Größe des VCR-Datensatz mit mehreren Millionen Strukturen für eine anschauliche Analyse zu umfangreich ist, wurden die 100.000 maximal diversen Moleküle des Datensatzes identifiziert. Hierzu wurde der Maximum Dissimilarity-Algorithmus⁷⁴ verwandt, also ein Clusterverfahren, das darauf ausgelegt ist, die größtmögliche Distanz zwischen zwei Clusterzentren zu ermitteln, und somit den Datenraum möglichst optimal durch maximal voneinander entfernte Clusterzentren „aufzuspannen“:

1. Wähle zufällig ein Clusterzentrum
2. Berechne die Distanzen zu allen anderen Molekülen
3. Wähle das Molekül als nächstes Clusterzentrum, welches zu den vorherigen Clusterzentren den maximalen Abstand im Sinne der Ähnlichkeitsmetrik aufweist. Entferne das gewählte Molekül aus dem Datensatz.
4. Fahre mit Schritt 2 fort, bis die gewünschte Anzahl an Clusterzentren erreicht ist.
5. Ist die definierte Anzahl an Clusterzentren erreicht, so werden die Distanzen aller noch verbleibender Moleküle zu den Zentren berechnet und die Moleküle dem Clusterzentrum mit der geringsten jeweiligen Distanz zugeordnet.

Die Implementierung dieses Algorithmus erfolgte mittels der „Diverse Molecules“ Komponente in Pipeline Pilot, als Ähnlichkeitsmetrik wurde die Tanimoto-Ähnlichkeit auf den FCFP-6-Deskriptoren berechnet. Zur Beschleunigung der nachfolgenden statistischen Analyse wurde der Fingerprint anschließend auf ein 256 Byte langes Array mittels einer Implementierung in PilotScript reduziert. In diesem Array steht jedes Byte für das Vorhandensein einer gewissen molekularen Funktion oder Substruktur, zusätzlich wird noch die jeweilige Häufigkeit des Auftretens im Molekül gezählt.

Auf gleiche Weise wurden Deskriptoren für die 100.000 maximal diversen Moleküle der WOMBAT-Datenbank sowie für die Bibliothek aller Glycin_B-Referenzliganden berechnet und alle drei Datenquellen in einem Datensatz vereinigt. Dieser wurde dann mittels Hauptkomponentenanalyse¹²² (*Principal Component Analysis* – PCA) ausgewertet. Die PCA ist ein Verfahren der multivariaten Statistik, in dem ein hochdimensionaler Datenraum durch die Berechnung von Linearkombinationen, den sogenannten Hauptkomponenten, in einem niederdimensionalen Raum beschrieben wird und somit u.a. anschaulicher dargestellt werden kann. Dazu wird eine multidimensionale Matrix durch eine Projektionsmatrix mittels Linearkombination in eine niedrigdimensionale Scoringmatrix überführt. Die Spalten dieser Scoringmatrix entsprechen den Hauptkomponenten, die als orthonormale Vektoren in die Richtung der höchsten Datenvarianz zeigen.

Da in diesem Fall die PCA nicht als Methode des Maschinenslernens für die prospektive Wirkstofffindung angewandt wurde, sondern nur zur statistischen Analyse eines gegebenen Merkmalsraumes, findet das Prinzip der nicht-linearen Verteilung von Struktur-Wirkungs-Beziehungen¹²³ und die sich daraus ergebenden Probleme für lineare Verfahren¹²⁴ keine Anwendung.

Die zugrundeliegenden Daten, also die 256 Byte langen Strukturdeskriptoren wurden zur Verwendung in der PCA standardisiert, d.h. zentriert (Subtraktion des Mittelwertes eines Deskriptoren-Bytes von allen zugehörigen Werten, um stärkeres Gewicht auf die Varianz der Daten zu legen) sowie skaliert (Division der zentrierten Daten durch ihre Varianz). Für die so standardisierten Daten wurden die Hauptkomponenten berechnet. Die ersten drei Hauptkomponenten wurden dann in einer dreidimensionalen Darstellung visualisiert und geben einen ersten Überblick über Gemeinsamkeiten und Unterschiede des verfügbaren chemischen Raumes, des biologischen Raumes sowie des speziellen chemischen Raumes der Glycin_B-Antagonisten.

2.5 Ligandenbasiertes zweidimensionales Virtual Screening

Verschiedene Studien zeigen, dass sich das Virtual Screening komplementär zum klassischen empirischen Screening¹²⁵ verhält, indem z.B. unabhängig von physikalisch vorhandenen Screeningbibliotheken neuartige Substanzen entdeckt werden können. Auch die verschiedenen Methoden des Virtual Screening erzielen durchaus komplementäre Ergebnisse⁶⁵. Aus diesem Grund wurden in dieser Arbeit verschiedene zwei- und dreidimensionale Methoden der virtuellen Wirkstofffindung angewandt, um zu überprüfen, welcher Ansatz am besten geeignet ist, neuartige Glycin_B-Antagonisten zu identifizieren.

2.5.1 Laplace-modifizierter naiver Bayes-Klassifikator

Das Konzept der Bayes'schen Statistik wurde bereits erfolgreich auf verschiedenen Forschungsgebieten angewandt¹²⁶. In den vergangenen Jahren erfreut es sich auch zunehmender Beliebtheit als Werkzeug für die Wirkstofffindung aufgrund seiner konzeptionellen Einfachheit und der direkten Interpretierbarkeit der Ergebnisse durch den menschlichen Experten¹²⁷⁻¹³⁰. Im Kontext dieser Arbeit wurde ein Laplace-modifizierter naiver Bayes-Klassifikator in Pipeline Pilot implementiert. Der naive Bayes-Klassifikator ermöglicht es, die Zugehörigkeit eines Objektes zu einer Klasse zu ermitteln, indem er das Bayes'sche Theorem (Gleichung 7) anwendet, einen Satz über bedingte Wahrscheinlichkeiten. Dieser besagt, dass die Wahrscheinlichkeit für ein Ereignis A in Abhängigkeit von Ereignis B definiert ist durch:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (\text{Gleichung 7})$$

Hierbei ist $P(A|B)$ die Wahrscheinlichkeit für Ereignis A , unter der Annahme, dass Ereignis B eingetreten ist, und $P(A)$ resp. $P(B)$ die *a priori* Wahrscheinlichkeiten der Ereignisse A resp. B .

In dieser Arbeit wurden die FCFP-6 Deskriptoren der Glycin_B-Referenzverbindungen berechnet und dem Klassifikator als „gute“ Moleküle präsentiert. Moleküle des negativen Trainingsatz wurden als „schlecht“ markiert und ihre FCFP6-Deskriptoren ebenfalls zum Training des Modells verwendet. Der Bayes-Klassifikator „lernt“, indem er die Häufigkeit „guter“ molekularer Substrukturen, beschrieben durch den FCFP-6-Fingerprint im Vergleich zur Hintergrundhäufigkeit der Inaktiven analysiert. Der Score einer Struktur wird berechnet, indem die Laplace-gewichteten Teilscores der individuellen Substrukturen in einem Wahrscheinlichkeitswert aufsummiert werden. Dieser Wert stellt somit die relative Wahrscheinlichkeit dar, mit der ein gegebenes Molekül aus dem Satz der „guten“ Strukturen stammt.

Verschiedene Bayes-Klassifikatoren kamen hierbei zur Anwendung: Um die prinzipielle Fähigkeit eines Bayesschen Modells zu beurteilen, Aktive von Inaktiven zu trennen, wurde ein Klassifikator trainiert, dem als „gute“ Samples sämtlich Glycin_B-Referenzen aus der Literatur vorgelegt wurden. Die Rationale hierfür war, dass hochaffine Substanzen mit strukturellen Motiven angereichert sind, die für die biologische Aktivität wichtig sind. Diese erhalten vom Klassifikator einen hohen Bayes-Score zugeteilt. Substrukturen in mittel- oder kaum aktiven Substanzen hingegen sind eher divers verteilt, ihre individuelle Häufigkeit ist also eher gering und somit erhalten diese einen geringeren Score.

In einer retrospektiven Validierung dieses ersten Modelles wurden alle 1096 Referenzstrukturen aus Literatur und biologischen Datenbanken anhand des Maximum-Dissimilarity-Algorithmus auf dem FCFP-6 Deskriptor in 20 unter-

schiedliche Cluster eingeteilt. Vier dieser Cluster (73 Strukturen) mit der höchsten Unähnlichkeit ihrer Clusterzentren zu den übrigen Clustern (Abbildung 12) wurden als externer Testdatensatz gespeichert und während des Modelltrainings nicht benutzt. Die verbleibenden 1023 Referenzstrukturen wurden als „gute“ Samples während des Trainings verwendet.

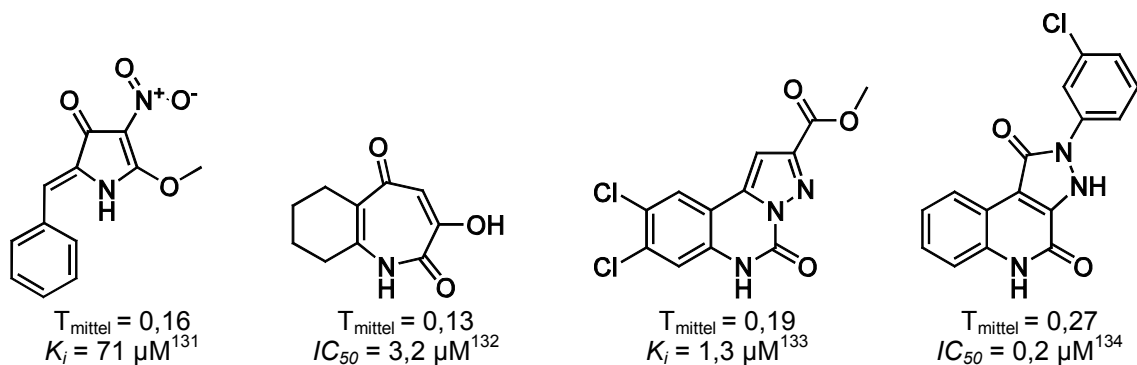


Abbildung 12: Clusterzentren des externen generellen Testdatensatzes. Zusätzlich annotiert sind ihre mittlere Entfernung zu allen anderen Strukturen im Testdatensatz (Tanimoto-Ähnlichkeit, FCFP-6 Fingerprint) und ihre jeweilige annotierte *in vitro* Affinität

Alle Moleküle der WOMBAT-Datenbank - exklusive der Verbindungen, die eine Annotation für das Target NMDA-NR1/Glycin_B aufwiesen - wurden als Negativsamples für das Training gesammelt. Liganden, die auf anderen Untereinheiten des NMDA-Rezeptors wie z.B. NR2A usw. Aktivität zeigten, wurden ebenfalls in den Trainingsdatensatz als Negativdaten aufgenommen, um die Subtyp-Selektivität des Modells zu erhöhen. Der Negativ-Trainingsatz wurde dann einer 50-prozentigen Zufallstrennung unterworfen, um Inaktive für den Trainings- und externen Testdatensatz zu erhalten. In der Summe enthielten beide Datensätze je etwa 67.000 Moleküle.

In einem zweiten Ansatz wurde ein weiterer Bayes-Klassifikator nur auf Referenzaktiven im nanomolaren Bereich trainiert (416 Strukturen). Verbindungen mit einer annotierten Affinität höher 30 μM wurden dem zuvor zusammengestellten Negativ-Trainingsatz hinzugefügt und ebenfalls während des Trainings vom Modell gelernt. Dieser zweite Ansatz sollte die Frage klären, ob ein Bayes-Klassifikator nicht nur prinzipielle Aktivität auf dem Zielrezeptor vorhersagen, sondern auch gezielt hochaffine Substanzen identifizieren kann. Die verbleibenden nanomolaren Aktiven wurden in sechs Cluster aufgeteilt. Zwei Cluster (54 Strukturen, Abbildung 13) wurden als externer Testdatensatz zurückbehalten, wohingegen die restlichen vier Cluster (363 Strukturen) für das Training verwendet wurden.

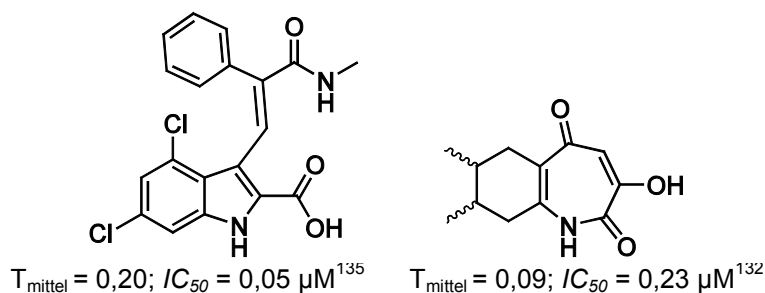


Abbildung 13: Clusterzentren des externen nanomolaren Testdatensatzes
Zusätzlich ihre mittlere Distanz zu allen anderen Molekülen im Testdatensatz (Tanimoto-Ähnlichkeit, FCFP-6 Fingerprint) und die jeweilige annotierte *in vitro* Affinität

Beide Modelle wurden nach dem o.a. Protokoll erneut trainiert, allerdings wurden nun zusätzlich zum Strukturdeskriptor weitere physikochemische Eigenschaften berechnet und als zusätzliche Deskriptoren während des Lernens verwendet. Diese Untersuchung sollte zeigen, ob Aktivität auf Glycin_B allein durch die Struktur der Liganden beschrieben werden kann, oder ob zusätzliche Parameter eine Rolle spielen und die Vorhersagekraft des Klassifikators erhöhen. Als physikochemische Deskriptoren wurden der AlogP, die Anzahl von Wasserstoffbrückendonoren und -akzeptoren sowie die Anzahl der drehbaren Bindungen aufgrund ihrer einfachen Berechenbarkeit ausgewählt.

Insgesamt wurden somit vier unterschiedliche Bayes-Klassifikatoren retrospektiv validiert und kamen danach zur prospektiven Anwendung.

2.5.2 Ähnlichkeitssuche mit einem 2D-Pharmakophordeskriptor

Als zusätzliches zweidimensionales Verfahren zur Identifizierung neuer potentieller Glycin_B-Antagonisten wurde eine Ähnlichkeitssuche auf einem zweidimensionalen Pharmakophorfingerprint durchgeführt. Auf Basis dieser Methode wurden bereits zahlreiche Studien erfolgreich durchgeführt¹³⁶⁻¹³⁸. Die erste Implementierung eines autokorrelierten zweidimensionalen Pharmakophordeskriptors wurde 1999 von Schneider *et al.* als CATS („*chemically advanced template search*“) beschrieben⁸¹. Die besondere Stärke dieses Deskriptors liegt in seiner impliziten Fähigkeit zum „Scaffold Hopping“, da er sehr stark von molekularer Struktur abstrahiert, und nur noch die Abstände von pharmakophoren Punkten zueinander, kodiert in einem Histogramm, betrachtet^{51,139}.

In dieser Arbeit wurde eine erweiterte Implementierung eines 2D-Pharmakophordeskriptors in Pipeline Pilot verwendet. Als Erweiterung zu den fünf pharmakophoren Atomtypen des CATS-Deskriptors wurden acht Pharmakophortypen definiert (Wasserstoffbrückendonator, Wasserstoffbrückenakzeptor, positive Ladung, negative Ladung, positive Ionisierbarkeit, negative

Ionisierbarkeit, Aromatizität und Hydrophobizität). Analog zum CATS-Deskriptor wurden alle möglichen Kombinationen von pharmakophoren Punkten in einem Molekül gebildet und mit der jeweiligen Distanz zueinander in Bindungslängen kodiert. Für die Distanzen der Punkte wurde keine obere Grenzen definiert; der ursprüngliche CATS-Deskriptor beschränkt sich auf maximal 15 Bindungen, was für die meisten „drug-like“ Moleküle ausreichend ist. Bei der Berechnung des Deskriptors konnten Paare, Triplets und Quartette von Pharmakophorpunkten gebildet werden. Für die retrospektive und prospektive Anwendung bei der Suche nach Glycin_B-Antagonisten wurden schließlich Paare von Pharmakophorpunkten (PHFP2-Deskriptor) verwendet, da sowohl die eigene retrospektive Validierung im Zuge dieser Arbeit als auch andere wissenschaftliche Veröffentlichungen hierbei gute Ergebnisse zeigten⁷⁵.

Als Suchstrukturen für die retrospektive Ähnlichkeitssuche mit dem PHFP2-Deskriptor wurden sämtliche nanomolaren Referenzmoleküle aus dem zuvor für das Training der Bayes-Klassifikator generierten Datensatz an Aktiven ausgewählt, und der ebenfalls zuvor generierte externe Testdatensatz durchsucht. Hier waren 54 bekannte nanomolare Glycin_B-Liganden in knapp 67.000 inaktiven Substanzen aus der WIOMBAT-Datenbank versteckt. Als Ähnlichkeitsmaß für alle Analysen wurde der jeweils beste Tanimoto-Score zwischen den PHFP2-Deskriptoren der Suchstrukturen und der zu analysierenden Screening-substanz verwendet.

2.6 Virtual Screening mit dreidimensionalen Methoden

Im Gegensatz zu den zweidimensionalen Methoden des Virtual Screenings, in denen Moleküle rein über die Konnektivität ihrer Atome beschrieben werden, benötigen dreidimensionale Methoden grundsätzlich eine intensivere Betrachtung. Moleküle im 3D-Raum sind keine starren Gebilde. Zusätzlich zur Kodierung der dreidimensionalen Anordnung der Atome zueinander kommen nun weitere Faktoren zum Tragen wie Konformationen (Rotationen um drehbare Einfachbindungen), Torsionen und Tautomerisierungen. Prinzipiell vorhanden sind weitere Phänomene wie Oszillationen von Bindungslängen oder Bindungswinkeln, die aber für das Virtual Screening nur selten beschrieben und eher die Domäne von Moleküldynamik-Simulationen (Kapitel 2.7.2) sind.

2.6.1 Dreidimensionaler Formvergleich

Die Analyse der dreidimensionalen Form (*shape*) eines Moleküls, bevorzugt in seiner biologisch aktiven Konformation, ist ein weiteres Konzept für das Virtual Screening, welches zunehmende Aufmerksamkeit in der Literatur erfährt^{105,140,141}. Die Struktur einer niedrig-nanomolaren Referenzverbindung¹⁴²

(Abbildung 14) wurde als Referenz für eine Shape-Suche in Schrödinger Phase ausgewählt. Das wahrscheinliche Isomer der Substanz bei pH 7,4 wurde mittels Schrödinger Epik 1.6 berechnet und die generierte 3D-Konformation durch MacroModel energieminiert (OPLS2005 Kraftfeld, implizite wässrige Lösung (GB/SA-Algorithmus), konstantes dielektrisches Feld, Polak-Ribiere konjugiertes Gradientenverfahren¹⁴³ (PRCG) mit 1000 Iterationen und einem Grenzwert von 0,01 Å).

Verschiedene Repräsentationen von molekularer Form mit wachsendem Informationsgehalt wurden als Suchparameter validiert:

- einfache Form (Volumen des Moleküls definiert als Summe der van-der-Waals-Radien der Ligandenatome)
- Form mit Elementkodierung (vdW-Radien der Ligandenatome, kodiert mit ihrem jeweiligen chemischen Element)
- Form mit Pharmakophorkodierung (vdW-Radien der Ligandenatome, kodiert mit ihrem Pharmakophortyp)
- Form mit elektrostatischer Kodierung (vdW-Radien der Atome, kodiert mit ihrem Kraftfeld-Atomtyp gemäß MacroModel 9.6 (Schrödinger, LLC, New York, 2008)).

Diese Untersuchung sollte zeigen, ob und welche Repräsentation von Molekülformen in der Lage ist, während einer Shape-Analyse zwischen Glycin_B-Liganden und inaktiven Molekülen zu unterscheiden.

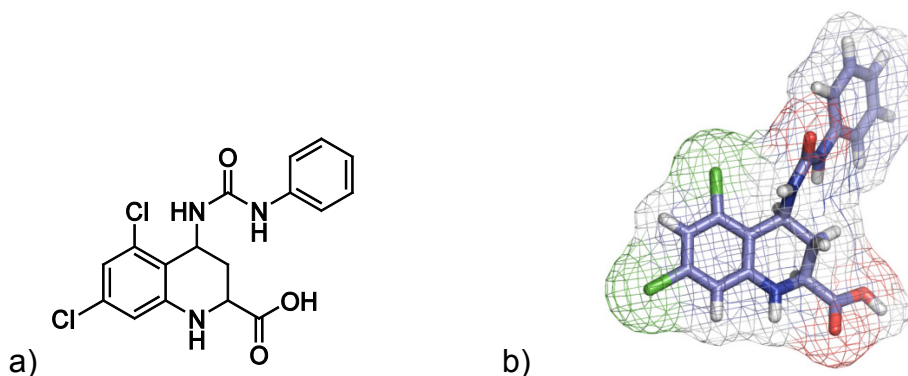


Abbildung 14: Shape-Referenzstruktur

a) Zweidimensionale und b) dreidimensionale Form der Shape-Referenz¹⁴², $IC_{50} = 0,01 \mu\text{M}$

2.6.2 Pharmakophorsuche und QSAR-Analyse

Auf der Grundlage von 194 nanomolaren Glycin_B-Referenzverbindungen mit der höchsten Affinität zum Rezeptor wurde eine dreidimensionale Pharmakophorhypothese erstellt. Für die Generierung der Hypothesen wurde Schrödinger Phase in der Version 3.0 verwendet. Wie in den vorangegangenen Experimenten etabliert, wurde molekulares Clustering durchgeführt, um einen externen Test-

datensatz zu erzeugen, der während der Erstellung der Hypothesen und des QSAR-Modells nicht betrachtet wurde. Um das Rauschen in den aus der Literatur entnommenen experimentell ermittelten Affinitäten zu minimieren, wurde der pIC_{50} (dekadischer Logarithmus des IC_{50} -Werts) für jeden Liganden berechnet. Ein Aktivitätsgrenzwert von $pIC_{50} = 8.5$ wurde gesetzt, um aktive Liganden zu definieren; Liganden mit einem $pIC_{50} \leq 6$ wurden als inaktiv definiert. Um Konformere für das Erstellen und Validieren der Hypothesen und QSAR-Modelle zu generieren, wurde ConfGen Version 2.0 (Schrödinger, LLC, New York, 2008) verwendet, um maximal 1000 Konformere pro Verbindung zu generieren. Dabei wurden höchstens 100 Schritte pro drehbare Bindung untersucht. Wie bei der Generierung der 3D-Datenbank der Screeningsubstanzen wurde der „thorough“ Suchalgorithmus verwendet, der ein gründlicheres Durchmusteren des Konformationsraum auf Kosten der Rechenzeit ermöglicht. Die Energien der einzelnen Konformere wurden während der Suche sowohl vorminimiert (maximal 100 Schritte in einem Steepest-Descent-Verfahren) als auch nachminimiert (50 Schritte, Steepest-Descent). Die Minimierung wurde im OPLS2005 Kraftfeld mit einem impliziten GB/SA („Generalized Born/Solvent Accessibility) Wassermode¹⁴⁴ vorgenommen. Redundante Konformere mit einem paarweisen RMSD von über 1 Å wurden eliminiert.

Verschiedene Pharmakophorhypothesen wurden durch Überlagerung der Referenzstrukturen automatisch generiert und retrospektiv bewertet, indem die Vorhersagekraft der Hypothese auf den Testdatensatz bewertet wurde. Zusätzlich wurde auf der am besten bewerteten Pharmakophorhypothese ein 3D-QSAR-Modell generiert und ebenfalls retrospektiv validiert. Hierzu wurde die Belegung eines dreidimensionalen Grids durch die einzelnen Atome der jeweiligen Liganden durch eine Partial Least Squares (PLS)-Analyse bewertet. Die Implementierung der PLS-Analyse erfolgt in Schrödinger Phase (Kapitel 3.3.2).

2.6.3 Molecular Docking im eingeschränkt flexiblen Rezeptor

Die Kristallstruktur von NR1 im Komplex mit DCKA (PDB-Code 1PBQ) wurde für die folgenden Dockingstudien ausgewählt, da DCKA der größte Ligand der Kokristallserie und die Bindungsstelle somit die weiträumigste der antagonistischen Zustände ist³⁷. Die Kristallisationsstudien von Inanobe, Furukawa und Gouaux zeigten, dass die NR1 Bindungsstelle in der Lage ist, Liganden unterschiedlicher Molekülgrößen aufzunehmen^{12,37,38}. Für NR1 wird zudem in der Literatur ein hoher Einfluss des induzierte-Passform-Phänomens (*induced fit*) diskutiert^{145,146}. Die im Vergleich zum Antagonisten DCKA höhere Molekülgröße der nanomolaren Referenzliganden zeigte ebenfalls, dass Glycin_B in der Lage ist, auch voluminösere Liganden sehr fest (nanomolar) zu binden: Die durchschnittliche molekulare Masse von 614 nanomolaren Literaturverbindungen betrug 361 Dalton, die molekulare Masse von DCKA hingegen nur 270 Dalton.

Das Phänomen des *induced fit* für Glycin_B ist ein größtenteils ungelöstes wissenschaftliches Problem. In diesem Virtual Screening-Ansatz wurde Rezeptorflexibilität in eingeschränktem Maße durch van-der-Waals-Skalierung implementiert. Während des Dockings der Liganden wurden sowohl die vdW-Radien der Ligandenatome als auch die der Atome der Bindungsstelle im Rezeptor auf bis zu 70% ihrer ursprünglichen Größe skaliert, um potentielle Kollisionen (*clashes*) zwischen Ligand und Atomen der Aminosäurereste der Bindetasche bei der Platzierung des Liganden in der Bindungsstelle zu vermeiden. Dies sollte den Fall simulieren, in dem sich bei einem Bindungsereignis die Seitenketten der Bindetasche leicht drehen, um den Liganden aufzunehmen. Eine weitergehende Implementierung des *induced fit*-Phänomens erfolgte in den späteren Teilen dieser Arbeit (Kapitel 2.7).

Die NR1-Bindungsstelle wurde definiert als der Raum 12 Å um das geometrische Zentrum des Liganden DCKA in der Kristallstruktur (Abbildung 15). Diese wurde anhand des Protein-Preparation-Wizard der Schrödinger Suite 2008 aufbereitet. Die in der PDB abgelegten Strukturen sind häufig aufgrund röntgenkristallographischer Gegebenheiten unvollständig. So fehlen in vielen Rezeptoren Atome in Seitenketten, die z.B. aufgrund hoher Temperaturfaktoren (Kapitel 2.7.3) nicht aufgelöst wurden. Auch annotieren Kristallographen in Strukturdateien keine Wasserstoffe (fehlende Auflösung der Dichte des Wasserstoffelektrons). So wurden an alle Proteinreste explizite Wasserstoffe hinzugefügt sowie Kristallwasser entfernt, da das verwendete Dockingprogramm mit impliziten Wassern arbeitet. Die wahrscheinlichen Tautomere der Aminosäurereste der Bindungsstelle wurde berechnet und alle Reste Asparagin, Glutamin und Histidin im Protein falls nötig um 180° gedreht, um die Ausbildung von Wasserstoffbrückenbindungen zu optimieren. Schließlich wurde das Protein einer kurzen Energieminimierung bis zu einem RMSD von 0,18 Å im OPLS2001¹¹⁰ Kraftfeld unterworfen, um die Struktur leicht zu relaxieren und nichtaufgelöste Kollisionen (*clashes*) zu entfernen.

Um zu überprüfen, ob diese strukturelle Relaxation einen Einfluss auf die Anreicherungsrate während des Virtual Screenings hat, wurden die Dockingexperimente sowohl mit der energieminierten Rezeptorstruktur als auch der nicht minimierten Struktur, also den unveränderten Originalkoordinaten der Kristallstruktur, durchgeführt. Die Dockingexperimente wurden mit Glide SP^{147,148} und Glide XP¹⁴⁹ (Glide, Version 5.0, Schrödinger, LLC, New York, 2008) durchgeführt.

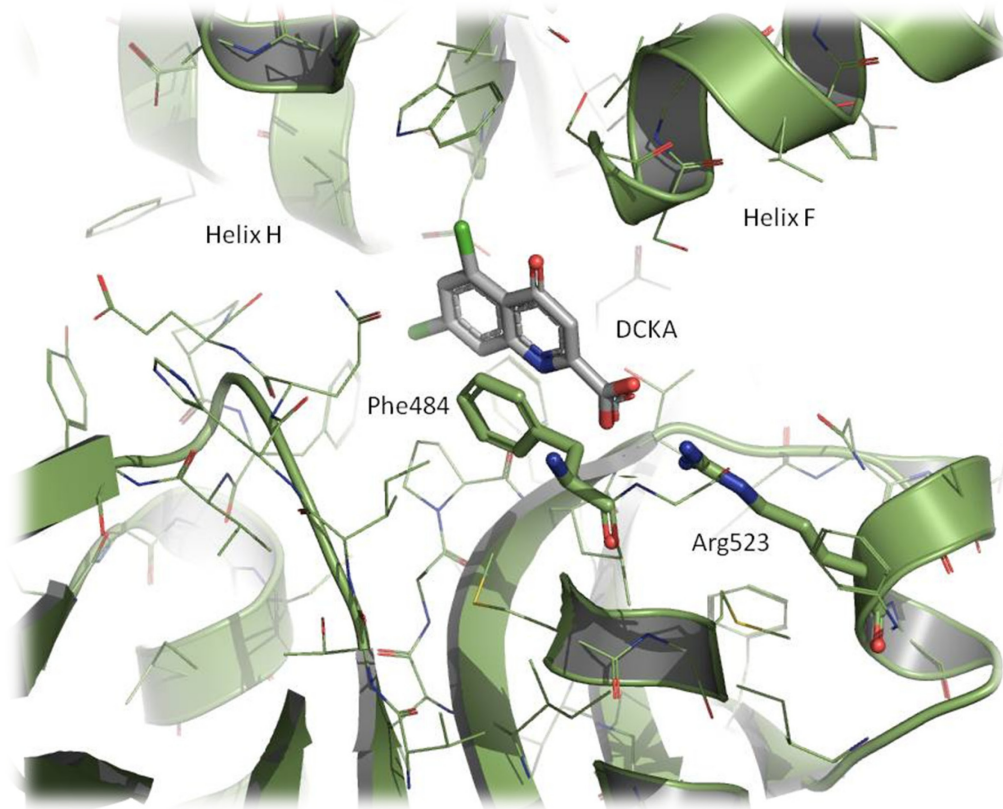


Abbildung 15: Glycin_B-Bindungsstelle für das Molecular Docking
 Die Bindungsstelle wurde definiert als die Aminosäuren im Umkreis von 12 Å (Linien-Darstellung) um den Liganden DCKA (grau). Die für die Bindung von Antagonisten essentiellen Aminosäuren Phe484 und Arg523 („Constraints“) sind in Stick-Darstellung gezeigt.

Als erste Validierung dieser Methode wurde die Anreicherung der unterschiedlichen Dockingexperimente anhand von 416 nanomolaren Referenzsubstanzen überprüft, die in 4427 inaktiven Molekülen „versteckt“ wurden. Die inaktiven Strukturen wurden anhand einer 2D-Ähnlichkeitssuche (FCFP-6 Deskriptor, Tanimoto-Ähnlichkeit, Pipeline Pilot) mit allen nanomolaren Referenzliganden als Suchstruktur durchgeführt. Quelle von inaktiven Strukturen war die WOMBAT-Datenbank, die vorher auf Moleküle mit biologischer Aktivität auf dem Zielprotein Glycin_B gefiltert wurde. Die Anwendung von inaktiven Molekülen, die den aktiven Liganden strukturell ähnlich sind (Abbildung 16), wird in der Literatur als wertigere Validierung einer Methode angesehen als die Zusammenstellung eines Satzes an inaktiven Molekülen über Zufallsauswahl¹⁵⁰. Die Docking- und Scoringalgorithmen mussten nun zwischen Aktiven und ihnen sehr ähnlichen Inaktiven unterscheiden, und nicht nur zwischen aktiven Liganden und u.U. zu diesen sehr unähnlichen Inaktiven.

Eine zweite Validierung wurde mit der erweiterten Scoringfunktion von Glide XP durchgeführt, um potentiell falsch-positive Hits anhand eines strengeren Scorings zu identifizieren¹⁵¹. In einer dritten und letzten Validierung wurde der Einfluss von Interaktionseinschränkungen (*constraints*) festgestellt. In der Bin-

ungsstelle wurden vier Interaktionspunkte zwischen Rezeptor und Ligand definiert, die in der Literatur als notwendig für die Bindung beschrieben wurden³⁷: die Ausbildung von Wasserstoffbrückenbindungen zu den Resten Arg523, Pro516 und Thr516 sowie ein hydrophobes Gebiet an der Stelle des aromatischen Ringsystems von DCKA in der Kristallstruktur (Interaktion mit Phe484). Während die Bildung von Wasserstoffbrücken bestimmend für die eigentliche Bindungsaffinität ist, wird die Interaktion zwischen Ligand und Phe484 mittels Pi-Pi-Interaktion als essentiell für die antagonistische Aktivität der Liganden betrachtet³⁸. Während des Dockings mussten mindestens drei der vier Constraints erfüllt werden, andernfalls wurde eine generierte Ligandenpose verworfen.

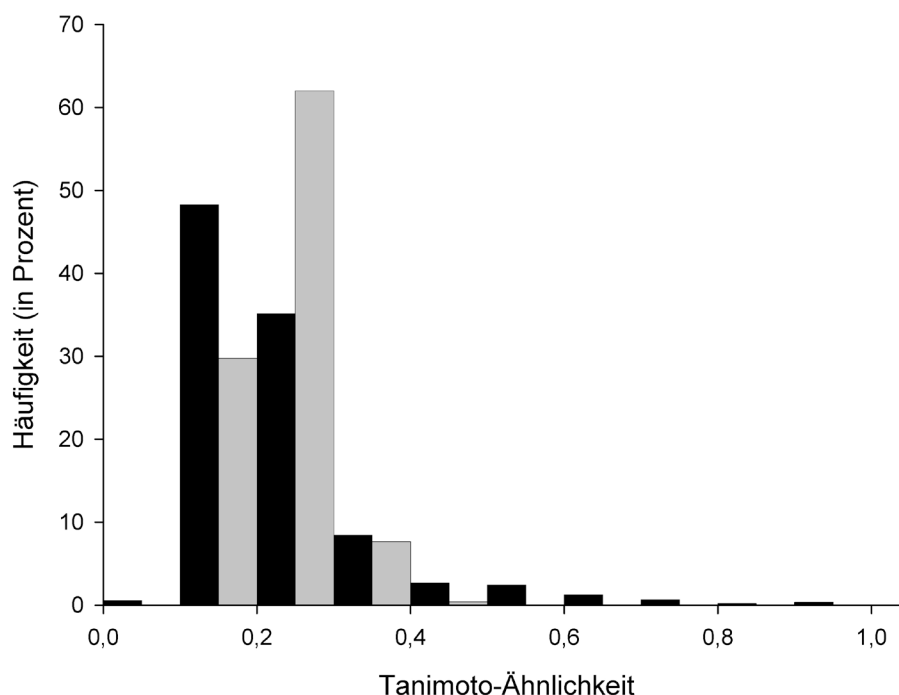


Abbildung 16: Histogramm der paarweisen Ähnlichkeiten Aktive/Inaktive Tanimoto-Ähnlichkeit, FCFP-6 Deskriptor; Ähnlichkeit der Aktiven zueinander (schwarz) sowie der Aktiven zu den Inaktiven (grau)

2.7 Modellierung eines flexiblen Proteins am Beispiel von Glycin_B

2.7.1 Induced Fit Docking

Obwohl es grundlegende Ansätze gibt, die intrinsische Flexibilität eines Proteins zu beschreiben (z.B. „relaxed complex scheme“⁹⁵, „SAR by NMR“¹⁵², „tethering“¹⁵³), ist die Verwendung eines rigiden Rezeptors während des Molecular Dockings derzeit die Standardmethode. Hierbei wird das Protein in einer Konformation eingefroren, und nur der Konformerer Raum des Liganden wird durchsucht. Die Annahme eines rigiden Rezeptors kann dabei zu falsch-negativen (Ligand mit bekannter Aktivität kann nicht gedockt werden, da seine

Form nicht zur Form der rigiden Bindetasche passt) Ergebnissen führen. Proteine können beim Binden von Liganden Seitenkettenbewegungen oder sogar Bewegungen des Proteinerückgrades (Backbone) ausführen, um sich der Form und Art des Liganden anzupassen (*Induced Fit*).

In dieser Arbeit wurde zur Modellierung von Bindungsposen von Liganden, die in ihrer Größe und Form nicht den Liganden der bekannten Kristallstrukturen entsprechen, ein *Induced-Fit-Docking*protokoll angewandt^{154,155}. Die Implementierung erfolgte in der Schrödinger Suite 2008 (Schrödinger Suite 2008 Induced Fit Docking Protokoll; Glide Version 5.0, Schrödinger, LLC, New York, NY, 2005; Prime Version 1.7, Schrödinger, LLC, New York, NY, 2005.). Hierbei wurden die Liganden iterativ in den implizit flexiblen Rezeptor (implizite Flexibilität durch van-der-Waals-Skalierung der Atomradien um bis zu 30%) gedockt und der Ligand-Rezeptorkomplex anschließend im Kraftfeld minimiert, um induzierte Konformationsänderungen zu berechnen. Somit können Bindungsposen von Liganden generiert werden, die unter anderen Umständen nicht unbedingt in die Bindungsstelle eingepasst werden konnten. Alle Berechnungen erfolgten auf Basis der Röntgenkristallstruktur von Glycin_B mit dem Antagonisten DCKA, welche entsprechend der Vorgehensweise für das Docking vorbereitet wurde (Kapitel 2.6.3).

Das Protokoll folgte iterativ den einzelnen Schritten:

- 1 Docking in den quasi-rigiden Rezeptor (implizite Flexibilität im lokalen Umfeld von Atomen des Proteins durch van-der-Waals-Skalierung um 30% – „softened potential“)
- 2 Sampling des Protein-Ligand-Komplexes im OPLS2005-Kraftfeld für jede generierte Bindungspose (maximal 30, bewertet nach GlideScore)
- 3 Erneutes Docking des Liganden in die energieminierte Rezeptorkonformation aus dem vorherigen Schritt
- 4 Bewertung der Pose anhand Dockingenergie (GlideScore) und Dehnungsenergie (*strain energy*) des Rezeptors sowie Solvatationstermen (Prime).
- 5 Der Algorithmus terminiert, wenn eine Anzahl an Wiederholungen überschritten ist, oder sich die Konformationen der generierten Protein-Ligand-Komplexe nicht mehr über einen definierten Wert (RMSD) verändern.

Die Bewertung einer Pose (IFDScore) folgte dabei dem Term:

$$IFDScore = GlideScore + 0.05 * PrimeEnergy,$$

d.h. bestimmendes Element der IFDScores war der Dockingscore, die Energie des Protein-Ligand-Komplexes ging nur mit 5% des ursprünglichen Wertes ein. Dies ist dadurch zu begründen, dass die intrinsische Energie eines Proteins im

Allgemeinen deutlich höher ist die freie Bindungsenergie eines Liganden und somit skaliert werden muss, um den deutlich geringeren Dockingterm nicht zu überlagern.

In einem erweiterten IFD-Protokoll konnten definierte Reste der Bindungsstelle zeitweise zu Alanin mutiert werden. Diese nicht-polare Aminosäure erlaubte es, aufgrund ihrer eigenen geringen Größe das Volumen der Bindungsstelle zu erhöhen und dabei keine zusätzlichen polaren Interaktionen mit dem Liganden einzugehen. Diesem stand nun mehr Raum zum Ausbilden seiner potentiellen bioaktiven Konformation zur Verfügung. Nach Generierung der entsprechenden Konformation des Liganden wurden die Reste der Bindungsstelle zurückmutiert, und der Protein-Ligand-Komplex im Kraftfeld energieminiert, um evtl. entstandene Clashes wieder zu neutralisieren. Auch mit dieser Modifikation terminierte der IFD-Algorithmus nach einer definierten Anzahl von Wiederholungen, oder bei Konvergieren des Dockings in eine bestimmte Pose (RMSD zur in der vorherigen Iteration generierten Pose von weniger als 0,18 Å).

Die Verwendung dieser Prozedur zur Betrachtung des Induced-Fit-Effekts brachte allerdings auch Nachteile mit sich: die Energieminimierung des Protein-Ligand-Komplex konnte u.U. falsch-positive Dockingposen begünstigen, indem einem inaktiven Liganden oder einer nicht-bioaktiven Konformation des Liganden künstlich Raum während der Mutation der Seitenketten der Bindungsstelle gegeben wurde. Dieses Problem wurde mit der Definition von essentiellen Rezeptorinteraktionen während des Dockings (*constraints*, Kapitel 2.6.3) sowie der visuellen Inspektion der erzeugten Dockingposen adressiert. Zusätzlich wurde im Verlauf dieser Arbeit eine Methodik entwickelt, um flexible Reste der Bindungsstelle und deren benachbarten Raum zu identifizieren und somit das fehleranfällige Mutieren aller Reste der Bindungsstelle unnötig zu machen, indem nur noch diese als flexibel identifizierten Reste zeitweise zu Alanin mutiert wurden.

Ein weiterer Nachteil dieser Methodik war das implizite Wassermode (*continuum solvation model*) der verwendeten Proteinmodellierungssoftware. Der PrimeEnergy-Score wurde entwickelt für die Energieminimierung von einzelnen Proteinen und nicht Protein-Ligand-Komplexen und kann somit die Änderungen in Entropie und Enthalpie des Systems durch Verdrängung von Wasser aus der Bindungsstelle des Proteins durch den Liganden nur annähern. Diese Terme sind aber häufig relevant für die freie Bindungsenergie. Aus diesem Grund wurde im weiteren Verlauf dieser Arbeit eine weitere Methode zur expliziten Berechnung von entropischen und enthalpischen Effekten durch Wasserverdrängung für die Betrachtung der Rezeptorflexibilität von Glycin_B angewandt (Kapitel 2.7.4).

2.7.2 Moleküldynamiksimulationen

Moleküldynamiksimulationen (MD-Simulationen) sind eine weit verbreitete Technik zur Beschreibung der Dynamik von Biomolekülen oder makromolekularer Komplexe¹⁵⁶⁻¹⁵⁸. Dabei werden die Prinzipien Newtonscher Mechanik angewandt, um die Energien bindender und nichtbindender Interaktionen innerhalb eines Moleküls und zwischen dem Molekül und seinem umgebenden System zu beschreiben:

$$E_{total} = E_{bindend} + E_{nichtbindend}, \text{ wobei}$$

$$E_{bindend} = E_{Bindung} + E_{Bindungswinkel} + E_{Torsionen}, \text{ und}$$

$$E_{nichtbindend} = E_{elektrostatisch} + E_{van-der-Waals}$$

Die Bindungs- und Winkelterme werden dabei gewöhnlich als harmonische Oszillatoren betrachtet, van-der-Waals Terme werden mit dem Lennard-Jones-Potential¹⁵⁹ (Gleichung 8) bezeichnet und elektrostatische Terme anhand des Coulombschen Gesetz (Gleichung 9) berechnet.

$$V(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right], \quad (\text{Gleichung 8})$$

wobei ϵ die Tiefe des Potentials, σ den endlichen Abstand, bei welchem das Potential 0 beträgt und r den Abstand der beiden Atome bezeichnet.

$$F = k_e \frac{q_1 q_2}{r^2}, \quad (\text{Gleichung 9})$$

wobei F die elektrostatische Kraft, q_1 und q_2 Ladungen, r die Distanz zwischen den beiden Ladungen und k_e eine Proportionalitätskonstante bezeichnet.

Bereits die erste MD-Simulation eines kompletten Proteins 1976 (Bovine Pancreatic Trypsin Inhibitor BPTI, 58 Reste, 9,2 Pikosekunden Simulationsdauer *in vacuo*) ermöglichte Einblicke in kürzeste Bewegungen von Biomolekülen¹⁶⁰⁻¹⁶². Aktuell erreichen MD-Simulationen erstmals die Mikrosekundengrenze¹⁶³. Generell sind MD-Simulationen aber auf wenige hundert Nanosekunden Simulationsdauer beschränkt, darüber hinaus muss mit starken Approximationen gerechnet werden¹⁶⁴. Das Problem der Zeitdauer der Proteinfaltung wird an Levinthals Paradoxon¹⁶⁵ deutlich: Ein Protein mit 100 Aminosäureresten würde mit nur 3 existierenden Konformationen je Rest und einer Dauer von 1 Pikosekunde pro Konformation eine unmöglich lange Zeit zum Erreichen seiner biologischen Konformation benötigen, nämlich $3^{100} * 10^{-13}$ s. Dies entspricht mehr Zeit als dem Alter unseres Universums.

Auch die Größe der betrachteten Systeme ist in den letzten Jahren kontinuierlich angestiegen, so können aktuell 10^4 bis 10^6 Atome simuliert werden^{166,167}. Die Betrachtung des Lösungsmittel von impliziten hin auf explizites Wasser un-

ter Addition von Gegenionen zur elektrischen Neutralisierung hat die Qualität der Ergebnisse von MD-Simulationen ebenfalls erheblich gesteigert¹⁶⁸.

In dieser Arbeit wurden MD-Simulationen verwendet, um die Stabilität und Flexibilität von Proteinstrukturen genauer zu analysieren. Zudem wurden die durch die schnelle, aber unpräzise Methode des Molecular Dockings erzeugten Bindungsposes neuer Liganden durch MD-Simulation präzisiert, da diese Methode explizit die Flexibilität sowohl des Liganden als auch des gesamten Rezeptors berücksichtigt. Die im Vergleich zum Docking genauere Beurteilung von Bindungsaffinitäten erlaubt zudem eine präzisere Vorhersage der freien Bindungsenergien^{169,170}, indem z.B. die Rolle des Wassers für die Stabilität der Bindung explizit betrachtet wird. Aufgrund des hohen Rechenbedarfs einer MD-Simulation kann diese Technik allerdings noch nicht als Standardmethode des Virtual Screenings eingesetzt werden. Selbst Vereinfachungen wie z.B. der „Multi-copy simultaneous Search“-Ansatz (MCSS)¹⁷¹ oder das sog. „locally enhanced sampling“¹⁷² sind immer noch um ein vielfaches langsamer als konventionelles Docking^{173,174}.

2.7.2.1 Vorbereitung der Proteinstruktur für die MD-Simulation

Alle MD-Simulationen der NR1-Bindungsdomäne des NMDAs wurden auf der Kette A der Röntgenkristallstruktur von Glycin_B im Komplex mit dem Antagonisten DCKA (PDB-Code 1PBQ) berechnet. In dieser Kristallstruktur sind einige Aminosäurereste nicht aufgelöst (Loop gebildet aus den Resten 47-58), zudem fehlt die Auflösung für eine Vielzahl von Atomen der verbleibenden 280 Reste. Diese wurden über die Funktion „Predict Side-chains“ von Schrödinger Prime modelliert. Die Reste des fehlenden Loop 1 wurden mithilfe der Build-Funktion von Schrödinger Maestro in die Struktur hinzugefügt und anschließend über die Funktion „Refinement“ von Prime energieminiert, um eine möglichst bioaktive Konformation zu erzeugen. Dabei wurden die vorhandenen Koordinaten des restlichen Proteins eingefroren, und nur die Reste des neu erzeugten Loops im Kraftfeld minimiert. Dazu wurden die Einstellungen des „Ultra extended loop samplings“ verwendet. Zu Beginn wurden die Rotamere der Seitenketten der zu betrachtenden Reste zufällig initialisiert. Der erste Rest wurde nun berechnet, indem sein Rotamer mit der niedrigsten Energie aus einer Rotamerbibliothek ausgewählt wurde, während alle verbleibenden Reste des Loops eingefroren wurden. Diese Prozedur wurde mit den verbleibenden Resten wiederholt, bis für alle Reste ein optimales Rotamer ausgewählt wurde. Darauf folgte eine iterative Energieminimierung des gesamten Loops: in jeder Iteration wurden die 4 besten Strukturen (im Sinne ihrer Energie) ausgewählt. Diese wurden dann mit einem immer kleiner werdenden Grenzwert (RMSD) energieminiert, bis ein definierter RMSD-Grenzwert unterschritten wurde. In einer Studie konnten mit dieser Methode Loops der gleichen Kettenlänge von 11

Resten in 79% der betrachteten Beispielsysteme mit einem RMSD von unter 1 Å korrekt modelliert werden¹⁷⁵.

Die nun vollständig modellierte Proteinstruktur von Glycin_B wurde daraufhin mittels des Desmond¹⁷⁶ System Builders (Desmond Molecular Dynamics System, Version 2.0, D. E. Shaw Research, New York, NY, 2008. Maestro-Desmond Interoperability Tools, Version 2.0, Schrödinger, New York, NY, 2008) für die MD-Simulation vorbereitet. Hierzu wurde eine orthorombische Simulationsbox der Dimensionen 68,7 * 78,1 * 72,9 Å³ um das Protein herum generiert. Da im Protein 4 positive Ladungen bestanden, wurden zur elektrischen Neutralisierung des Systems 4 Anionen (Chloride) zugegeben. Schließlich wurde die Simulationsbox mit 10913 expliziten Wassermolekülen vom Typ TIP4P geflutet, eines der zur Zeit komplexesten Wassermoleküle¹⁶⁸.

2.7.2.2 Durchführung der MD-Simulation

Alle in der vorliegenden Arbeit durchgeführten MD-Simulationen wurden nach dem gleichen Protokoll aufgesetzt:

- Die Simulation des Systems erfolgte in einem NPT-Ensemble, in dem die Anzahl der Atome (N), der Druck (P) sowie die Temperatur (T) konstant gehalten wurden. Die Simulation erfolgte in Raumtemperatur bei 300K, zur Regulierung wurde ein Nose-Hoover-Thermostat¹⁷⁷ eingesetzt. Der Druck wurde bei 1,01325 bar durch einen Martyna-Tobias-Klein Barostat¹⁷⁸ konstant gehalten. Diese Simulationsbedingungen entsprechen am ehesten einem *in vitro*-Experiment bei Raumtemperatur.
- Zur Beschleunigung der Rechenzeit wurde das zur Simulation benutzte Kraftfeld in drei Kategorien aufgeteilt, die in jeweils unterschiedlichen Zeitabständen berechnet wurden: Bindungsinteraktionen, nichtbindende Nahinteraktionen (van der Waals und kurze elektrostatische Interaktionen (Coulomb-Interaktionen)) und nichtbindende Ferninteraktionen (lange elektrostatische Interaktionen). Das Auftrennen des Kraftfelds in unterschiedliche Bereiche, die unabhängig voneinander berechnet werden, nennt man *Integration*. In den vorliegenden Simulationen wurde ein RESPA-Integrator¹⁷⁹ (*reference system propagator algorithm*) mit folgenden Zeitschritten verwendet:
 - Bindungsinteraktionen: 2 Femtosekunden
 - nichtbindende Nahinteraktionen: 2 Femtosekunden
 - nichtbindende Ferninteraktionen: 6 Femtosekunden

Coulomb-Interaktionen wurden ab einer Distanz von 9 Å nicht mehr berechnet, sondern durch die „*smooth particle mesh Ewald*“ Methode¹⁸⁰ approximiert. Dieser Algorithmus basiert auf der Ewald-Summation¹⁸¹

von Coulomb-Interaktionen, die die Summe der Interaktionen im konventionellen dreidimensionalen Raum mit einer äquivalenten Summe im Fourier-Raum ersetzt, die dort erheblich schneller konvergiert.

Kovalente Bindungen zwischen schweren Atomen und Wasserstoffen wurden dabei mittels des SHAKE-Algorithmus¹⁸² in ihrer Vibration stark eingeschränkt und somit aus der Berechnung während der MD-Simulation ausgenommen.

Durch die Summe dieser Optimierungen verringert sich die Laufzeit der Auswertung der paarweisen Interaktionsterme von $O(n^2)$ auf etwa $O(n \cdot \log(n))$.

- Alle Systeme wurden vor der eigentlichen MD-Simulation über eine Zeit von 10 Nanosekunden simuliert, um das gesamte System zu entspannen (Relaxation).
- Die anschließende freie MD-Simulation des Systems erfolgte über einen Zeitraum von 50 Nanosekunden. Die Trajektorien wurden dabei in einem Zeitschritt von einer Pikosekunde zur nachfolgenden Analyse abgespeichert.

2.7.3 Relation von Temperaturfaktor und Mobilität von Aminosäureresten

Die aus der röntgenkristallographischen Analyse resultierende Konformation eines Proteins ist kein rigides Gebilde, sondern thermalen Fluktuationen (Vibrationen) unterworfen¹⁸³. Der Grad der Vibrationen wird vom Kristallographen als Temperaturfaktor eines jeden Atoms des Proteins angegeben. In beschränktem Maße können diese Vibrationen als Modell für dynamische Bewegung im Protein dienen¹⁸⁴. Der Temperaturfaktor B (Gleichung 10) ergibt sich dabei aus

$$B = \frac{8\pi^2}{3} \cdot RMSF^2 \quad (\text{Gleichung 10})$$

Die Fluktuation um das quadratische Mittel der Atomkoordinaten (*root mean square fluctuation* – RMSF, Gleichung 11) ist definiert durch

$$RMSF = \sqrt{(r_{gesamt} - r_{atom})^2}, \text{ wobei} \quad (\text{Gleichung 11})$$

r_{gesamt} der mittleren Fluktuation aller Atomkoordinaten über eine definierte Zeit entspricht und r_{atom} der Fluktuation der individuellen Koordinaten eines Atoms.

In Studien konnte gezeigt werden, dass eine Strukturbestimmung mittels Röntgenkristallographie über die Zeit aufgelöst ähnlich der Kernresonanzspektroskopie erste Anhaltspunkte über die Dynamik eines Proteins liefern kann^{185,186}. Eine Auflösung größerer struktureller Änderungen wie z.B. die Be-

wegung ganzer Domänen zum Zeitpunkt der Bindung eines Liganden kann allein über die Temperaturfaktoren jedoch nicht wiedergegeben werden. Hier bedarf es der Durchführung mehrerer röntgenkristallographischer Experimente in verschiedenen Zuständen (z.B. ligandengebundener und Apo-Zustand) oder anderer experimenteller Vorgehensweisen¹⁸⁷ bzw. der Simulation von Proteindynamik *in silico*, um ein besseres Verständnis der Proteinmobilität zu erhalten.

2.7.4 Berechnung entropischer und enthalpischer Energieterme durch Wasserverdrängung

Scoringfunktionen sind approximative Methoden, um die Bindungsstärke nicht-kovalenter Interaktionen zwischen Ligand und Rezeptor beim Docking zu beurteilen. Obwohl sich seit Jahren Verbesserungen durch neue Algorithmen bemerkbar machen, sind alle bekannten Scoringfunktionen immer noch nicht in der Lage, experimentelle Bindungsaffinitäten im großen Maßstab korrekt vorherzusagen⁹⁹. Ein Grund hierfür ist, dass die Bindung zwischen Ligand und Protein von vielen Faktoren beeinflusst wird. So spielen nicht nur direkte Interaktionen der beiden an der Bindung beteiligten Partner für die freie Bindungsenergie (Gleichung 12) eine Rolle, sondern auch Faktoren wie entropische und enthalpische Energiegewinne. Die Verdrängung von Wasser bei der Bindung eines Liganden gilt als grundsätzlicher, wenn nicht sogar dominanter Faktor der freien Bindungsenergie¹⁸⁸. Wassermoleküle, die als Lösungsmittel die Bindungsstelle eines Proteins füllen, sind häufig entropisch und enthalpisch ungünstigen Energien aufgrund der lipophilen Eigenschaft der Proteinoberfläche ausgesetzt („hydrophober Einschluss“)¹⁴⁹. Sie können somit nur in unzureichender Anzahl Wasserstoffbrückenbindungen mit dem Protein eingehen. Die Verdrängung dieses „energetisch ungünstigen“ Wassers durch einen Liganden, der besser in der Lage ist, hydrophobe Interaktionen mit dem Protein einzugehen, führt somit zu einem Gewinn an Enthalpie¹⁸⁹. Zusätzlich führt eine Verdrängung dieses Wassers zu einem vorteilhafteren Entropieterm, da das Wasser im hydrophoben Einschluss nur eine eingeschränkte Zahl an Konfigurationen bei der Interaktion mit dem Protein und anderen Wassermolekülen der Bindungsstelle eingehen kann¹⁸⁹.

$$\Delta G = \Delta H - T\Delta S, \quad \text{(Gleichung 12)}$$

wobei ΔG die Änderung der Gibbs-Energie, ΔH die Enthalpieänderung, ΔS die Entropieänderung und T die Temperatur bezeichnet.

Die Methode der Störung der freien Energie (*free energy perturbation* – FEP)¹⁹⁰ ist in der Lage, den Änderung der Gibbs-Energie bei Bindung eines Liganden explizit bis zur Genauigkeit des zugrundeliegenden Kraftfeldes zu berechnen (Zwanzig-Gleichung – Gleichung 13).

$$\Delta G(A \rightarrow B) = G_B - G_A = -k_B T \ln \left\langle \exp \left(-\frac{E_B - E_A}{k_B T} \right) \right\rangle_A$$

(Gleichung 13)

wobei A und B zwei Zustände (ohne und mit Ligand), T die Temperatur, k_B die Boltzmann-Konstante und die eckigen Klammern den Mittelwert einer Simulation für den Zustand A (ligandenlos) bezeichnet. Allerdings erfordert die FEP-Methode einen hohen rechnerischen Aufwand, so dass sie aktuell nicht quantitativ z.B. für das Virtual Screening einsetzbar ist.

Eine weitere Methode, die Thermodynamik des Wassers der Bindungsstelle zu berechnen, ist die Berechnung einer „Wasserkarte“ anhand der Theorie der inhomogenen Solvation^{188,191}. Hierzu wurde das zugrundeliegende System einer kurzen Moleküldynamiksimulation (Kapitel 2.7.2) unterworfen, um die räumliche Verteilung des expliziten wässrigen Lösungsmittels der Bindungsstelle über Zeit zu analysieren. Ein Clusteringverfahren¹⁹² wurde angewandt, welches die Bindungsstelle in 1 Å große Kugeln unterteilt. Diese Kugeln entsprachen den „Hydratationsstellen“. Durch Anwendung der inhomogenen Solvationstheorie konnte dann diesen Hydratationsstellen Entropie- und Enthalpieterme zugewiesen werden. Die prinzipielle Funktionsweise dieses Verfahrens wurde von Abel *et al.* 2008 erfolgreich am Faktor Xa gezeigt¹⁹³.

In dieser Arbeit wurde die Implementierung durch Schrödinger Watermap verwendet (Desmond Molecular Dynamics System, Version 2.0, D. E. Shaw Research, New York, NY, 2008. Watermap, Version 1.0, Schrödinger, New York, NY, 2008). Hierzu wurde die Kristallstruktur von NR1 im Komplex mit DCKA (PDB-Code 1PBQ) verwendet und wie in Kapitel 2.7.2.1 beschrieben vorbereitet. Das geometrische Zentrum von DCKA wurde als das Zentrum der Bindungsstelle definiert, und alle Reste im Umkreis von 25 Å wurden zur Analyse der Wasserkarte betrachtet. Das Protein wurde in explizitem Wasser vom Typ TIP4P gelöst und für 10 Nanosekunden simuliert. Nach erfolgter Simulation wurden die Hydratationsstellen berechnet, indem die statistische Aufenthaltswahrscheinlichkeit von Wasser an diesen Stellen über den Zeitverlauf der Simulation ermittelt wurde. Für jede der Hydratationsstellen wurden dann die entropischen und enthalpischen Energieterme im OPLS2005 Kraftfeld berechnet. Dies erzeugte eine Karte von Wassermolekülen der Bindungsstelle, deren Verdrängung durch einen Liganden in besonders vorteilhaften Änderungen der freien Energie resultiert. Anhand dieser Wasserkarte wurde die Änderung der Gibbs-Energie bei Bindung verschiedener im Laufe der Arbeit identifizierter Glycin_B-Antagonisten abgeschätzt. Hierdurch konnten ebenfalls Aussagen über den Bindungsmodus der Substanzen getroffen werden.

2.8 Experimentelle Methoden

2.8.1 [³H]-MDL-105,519 Bindungsassay

MDL-105,519 ist ein selektiver hochaffiner Glycin_B-Antagonist¹⁹⁴. Er ist eine der am meisten genutzten Referenzsubstanzen für Glycin_B-Bindungsassays¹⁹⁵. Die Affinität einer Testverbindung für Glycin_B wurde anhand der Verdrängung von MDL-105,519 aus einer gesättigten Lösung von 2nM ermittelt. Da MDL-105,519 radioaktiv durch ³H markiert ist, kann die verdrängte Menge des Radioliganden anhand seiner Radioaktivität auf einem Szintillationszähler ermittelt werden. Zur Erzeugung von Dosis-Wirkungs-Kurven wurde die Konzentration von Protein und Radioligand konstant gehalten, während die Testsubstanz in steigenden Konzentrationen zugegeben wurde.

Das Protokoll des MDL-105,519-Bindungsassays wurde von Danysz *et al.* veröffentlicht¹⁹⁶. Die den Ergebnissen dieser Arbeit zugrundeliegenden *in vitro* Experimente wurden in Kooperation mit der Abteilung *In vitro* Screening bei Merz Pharmaceuticals durchgeführt.

2.8.2 Funktionaler Bindungsassay

Elektrophysiologische Patch-Clamp-Aufnahmen wurden von Neuronen des Hippocampus der Ratte nach 12 bis 15 Tagen *in vitro* durchgeführt. Die Messung erfolgte dabei an der vollständigen Zelle bei Raumtemperatur mit einer polierten Glaspipette (3 – 5 mU). Das vollständige Protokoll wurde von Parsons *et al.* veröffentlicht¹⁹⁷. Die in dieser Arbeit beschriebenen Patch-Clamp-Versuche wurden ebenfalls in Kooperation mit der Abteilung *In vitro* Screening bei Merz Pharmaceuticals durchgeführt.

3 Ergebnisse und Diskussion

3.1 Analyse der verwendeten Strukturdatenbanken

3.1.1 Diversität im zweidimensionalen Strukturraum

Die Deskriptoren der 100.000 maximal diversen Moleküle des Datensatzes verfügbarer kommerzieller Substanzen, der biologischen WOMBAT-Datenbank sowie der Sammlung von Glycin_B-Referenzliganden wurden berechnet und einer Hauptkomponentenanalyse unterzogen. Die ersten drei Hauptkomponenten wurden zur Veranschaulichung in einem dreidimensionalen Graphen aufgetragen (Abbildung 17).

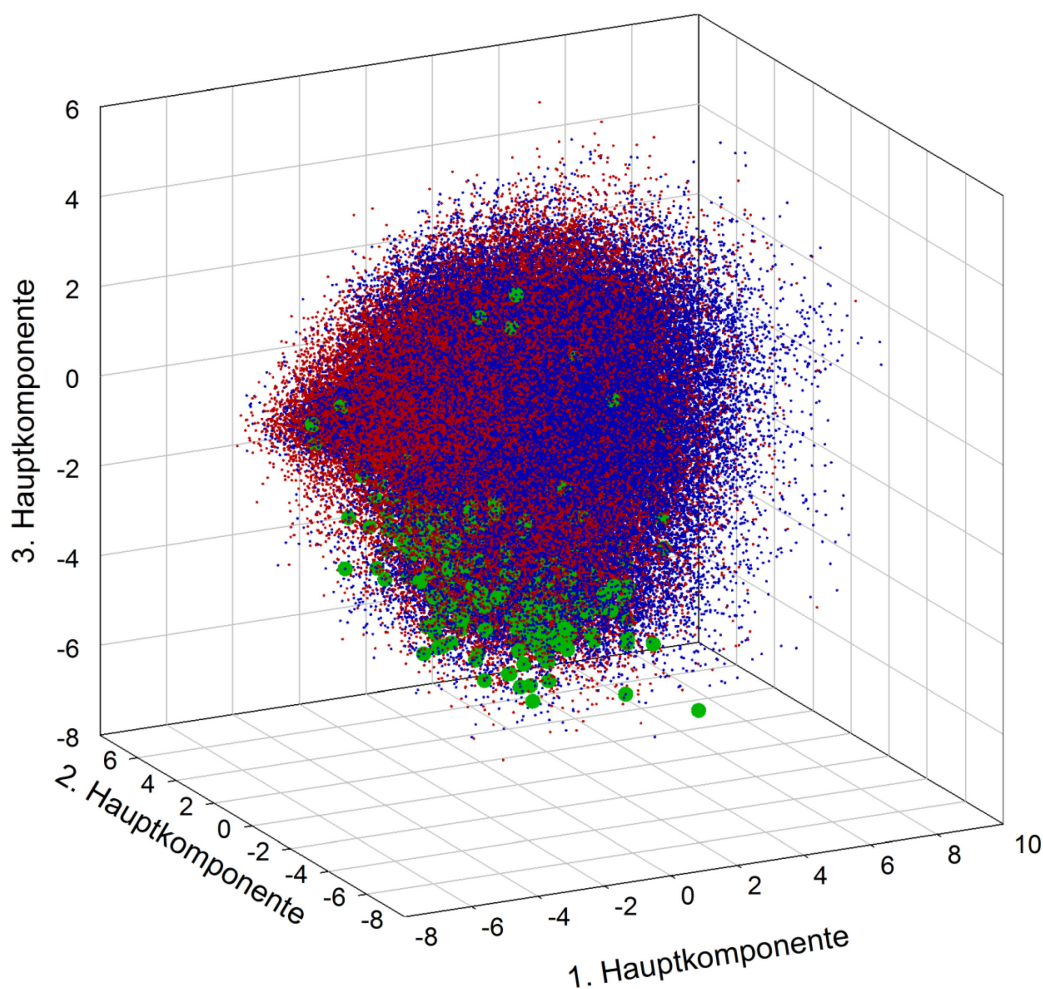


Abbildung 17: Verfügbarer chemischer, biologischer und „Glycin_B-Raum“
Rot: verfügbarer chemischer Raum; blau: biologischer Raum; grün: Glycin_B-Referenzantagonisten. Zur Verdeutlichung der Lage der Datenpunkte für die Glycin_B-Antagonisten wurden diese vergrößert dargestellt.

Aus diesem Ergebnis konnten mehrere Schlüsse, die einen direkten Einfluss auf das Wirkstoffdesign allgemein und insbesondere für das Design von Glycin_B-Antagonisten haben, gezogen werden:

Der Raum aktuell verfügbarer, bereits synthetisierter Substanzen (Abbildung 17, rote Punkte) deckte sich in den Dimensionen der ersten drei Hauptkomponenten seiner Moleküldeskriptoren zum überwiegenden Teil mit dem Raum biologisch aktiver Substanzen (blaue Punkte). Dennoch waren Bereiche zu erkennen, in denen der verfügbare chemische Raum nur dünn besiedelt ist, wohingegen der biologische Raum eher eine Gleichverteilung resp. eine Gauß'sche Verteilung in seinen ersten drei Hauptkomponenten erkennen ließ. Dies erklärt sich dadurch, dass viele Anbieter von chemischen Verbindungen zwar mittlerweile versuchen, ihre Kataloge an die Bedürfnisse der Arzneimittelforschung anzupassen, die Kataloge aber nach wie vor zu einem Teil von Serien dominiert sind, die eher nicht Wirkstoff-ähnlich („*drug-like*“) und damit nicht biologisch aktiv sind.

Im Umkehrschluss ließ sich jedoch folgern, dass sich viele Verbindungen aus Anbieterkatalogen durchaus für die virtuelle Wirkstofffindung eignen. Es besteht aber nach wie vor das Problem, dass viele für den biologischen Raum potentiell interessante chemische Strukturen kommerziell nicht verfügbar sind. Dies zeigt sich insbesondere dadurch, dass viele Bereiche des gleichmäßig besiedelten biologischen Raumes deutlich weniger dicht durch verfügbare chemische Strukturen abgedeckt sind. In den jeweiligen Extremwerten der drei Dimensionen, d.h. den „Randbezirken“ des chemischen und biologischen Raumes wurde die Dichte der bekannten Strukturen geringer. Dennoch waren hier Datenpunkte aus chemischem und biologischem Raum in etwa gleichverteilt. Hier besteht offenbar das größte Optimierungs- und Entdeckungspotential für neue Wirkstoffe.

Auch der durch die ersten drei Hauptkomponenten aufgespannte Raum der Glycin_B-Referenzsubstanzen überschneidet sich mit dem chemischen sowie dem biologischen Raum. Die größte Anzahl an Datenpunkte der Glycin_B-Referenzen war in einem bestimmten Unterraum konzentriert, dennoch gab es diverse Strukturen, die aus diesem Muster ausbrachen. Der „Glycin_B-Raum“ befand sich von seiner Datenverteilung her mehr in den dünn besiedelten Außenbezirken sowohl des chemischen als auch biologischen Raumes. Hieraus lässt sich schlussfolgern, dass die für Glycin_B-Aktivität benötigte Chemie durch weitere Synthesen neuartiger Verbindungen auch neue Aktivitätsinseln für Glycin_B erschließen kann.

Das Prinzip der Aktivitätsinseln war für Glycin_B generell gut erkennbar: Obwohl sich die meisten Strukturen in einem abgegrenzten Unterraum konzentrierten, gab es weiter entfernte, mehr isolierte Datenpunkte. Ob diese zum Gros der Referenzstrukturen sehr unterschiedlichen Glycin_B-Antagonisten hauptsäch-

lich „Ausreißer“ (*Outlier*) resp. Struktur-Singletons oder tatsächlich Repräsentanten neuer Aktivitätsinseln darstellten, ist Gegenstand aktueller Forschung und wird auch im Zuge des weiteren Verlaufs dieser Arbeit diskutiert.

3.1.2 Diversität im dreidimensionalen Konformationsraum

Insgesamt wurden aus den 4,6 Millionen unterschiedlichen 2D-Strukturen der Screeningdatenbank mehrere Milliarden 3D-Strukturen berechnet; nach Energieminimierung und Redundanzüberprüfung verblieben 750 Millionen 3D-Strukturen (Tabelle 4), die in einer Konformeren-Datenbank im Phase-Format gespeichert und im Folgenden durchsucht wurden. Die Berechnung der Konformere wurde wie zuvor auf einem 26-CPU Cluster parallelisiert und benötigte etwa zwei Wochen Rechenzeit. Die finale 3D-Datenbank belegte dabei eine Größe von 604 Gigabytes.

Hierbei zeigte sich eine durchaus unterschiedliche, anbieterspezifische Anzahl an Konformeren: während im Durchschnitt je Substanz etwa 150 Konformere vorlagen, wurde v.a. für Kataloge mit Naturstoffen und Naturstoffanaloge eine deutlich erhöhte Zahl an Konformeren beobachtet (z.B. Analyticon mit durchschnittlich 2053 Konformeren je Substanz). Dies kann als ein Hinweis gewertet werden, dass sich aktuelle Substanzen aus der konventionellen (kombinatorischen) chemischen Synthese in ihren Freiheitsgraden deutlich von den Naturstoffen und ihren Analoga unterscheiden. Aus Gründen der Substanzdiversität sollten demnach in jedem Virtual Screening auch eine höchstmögliche Anzahl an Naturstoffen durchsucht werden.

Tabelle 4: Statistiken der 3D-Konformerendatenbanken

Anbieter	Anzahl Substanzen	Isomere	Konformere	Konformere je Substanz
Specs	168.398	369.811	19.286.431	115
Asinex	280.407	719.561	45.322.723	162
Enamine	1.068.686	2.327.333	158.267.214	148
ChemDiv	584.117	1.384.589	85.193.355	146
AMRI	303.300	865.533	79.582.923	262
Chembridge	262.715	510.848	28.501.303	108
Maybridge	44.044	89.790	4.142.504	94
VitaSM	119.603	267.863	15.404.774	129
Timtec	41.217	41.002 [†]	1.899.660	46
Keyorganics	37.427	69.759	2.841.880	76
InterBioScreen	226.027	777.890	40.556.813	179
Analyticon	22.249	532.086	45.686.940	2.053
Life Chemicals	176.649	379.098	24.970.116	141
Pharmeks	80.886	421.157	25.753.591	318
Otava	26.089	64.438	3.397.236	130
Nanosyn	12.116	31.380	1.475.511	122
Chemstar	11.642	33.969	1.716.479	147
Acros	9.049	30.984	849.558	94
TOSlab	6.656	23.272	1.166.447	175
NCH	30.253	97.302	5.459.380	180
Akos	1.038.039	2.281.990	156.505.837	151
MS Discovery	955	10.277	507.849	532
Oakwood	6.890	14.776	457.805	66
ACB Blocks	1.617	11.324	287.082	178
Prestwick	363	5.801	292.980	807
Frinton	369	742	28.247	77
Apollo	13.510	24.765	754.446	56
Bachem	1.142	1.446	122.293	107
Biofocus	1.243	2.331	109.316	88
Alfa	3.477	9.740	397.924	114
Peptech	1.468	4.707	284.675	115
Gesamt	4.580.603	11.405.564	751.223.292	

[†] Einige Strukturen dieses Substanzkatalogs konnten nicht von Ligprep konvertiert werden, so dass insgesamt weniger Isomere generiert wurden als zweidimensionale Strukturen vorhanden waren.

3.2 Ligandenbasiertes Virtual Screening ohne Rezeptor

3.2.1 Ligandenbasiertes zweidimensionales Virtual Screening

3.2.1.1 Laplace-modifizierter naiver Bayes-Klassifikator

Verschiedene Bayes-Klassifikatoren wurden trainiert und retrospektive validiert. Im ersten Ansatz wurde überprüft, ob die Klassifizierung über die reine Molekülstruktur an sich ausreichend war, um aktive Glycin_B-Liganden und inaktive Substanzen voneinander zu trennen. Dazu wurde die Vorhersagekraft eines auf den FCFP-6 Strukturdeskriptoren der Moleküle trainierten Bayes-Modells mit einem Modell verglichen, welches auf FCFP-6 und zusätzlichen physikochemischen Deskriptoren basierte, namentlich dem AlogP, der Anzahl der Wasserstoffbrückendonoren und -akzeptoren sowie die Anzahl der drehbaren Bindungen. In diesem Modell wurden alle 1023 Literaturreferenzen verwendet (Abbildung 18 und Abbildung 19). Der gleiche Ansatz wurde für einen zweiten Vergleich gewählt; allerdings wurde hierbei der Bayes-Klassifikator nur auf nanomolaren Referenzsubstanzen (347 Strukturen) trainiert, während Literaturverbindungen mit einem IC_{50} über 30 μM dem Satz an Inaktiven hinzugefügt wurden (Abbildung 20 und Abbildung 21).

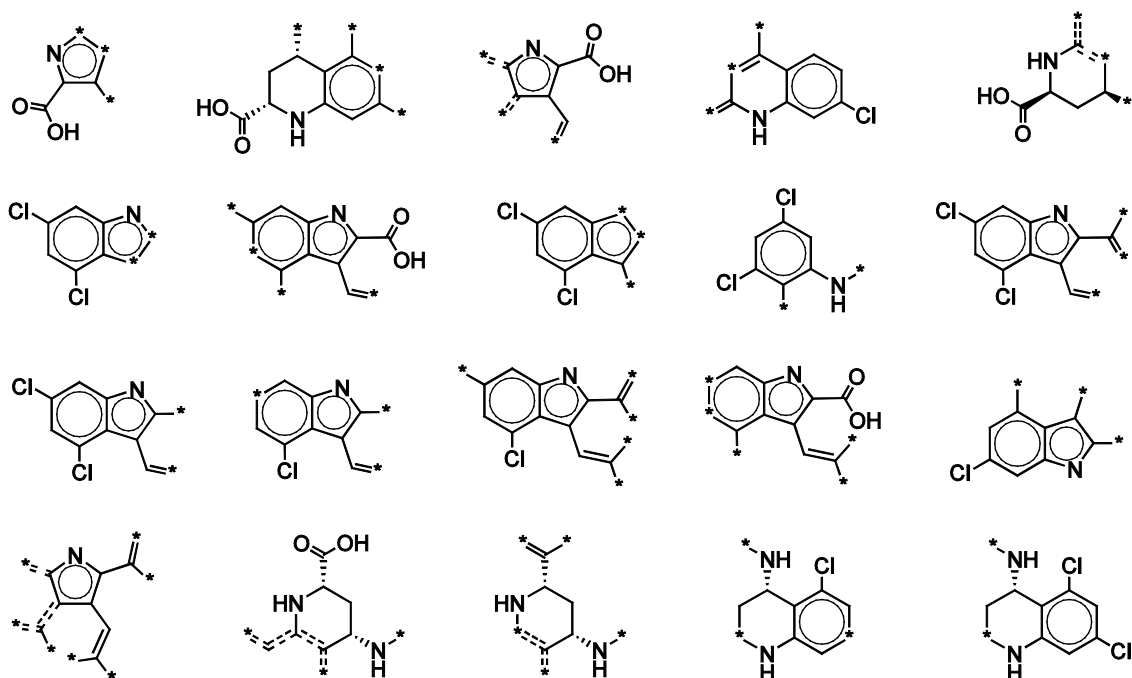


Abbildung 18: Am besten bewertete Features (Bayes-Modell 1)

Die statistische Auswertung wurde über 679 Referenzliganden mit einer Aktivität unter 30 μM sowie 61221 inaktivem Moleküle (*decoys*) aus der WOMBAT-Datenbank durchgeführt. Sterne (*) entsprechen Platzhalter für Bindungsstellen und Heteroatome. Doppelt gestrichelte Linien signalisieren Einfach- oder Doppelbindungen.

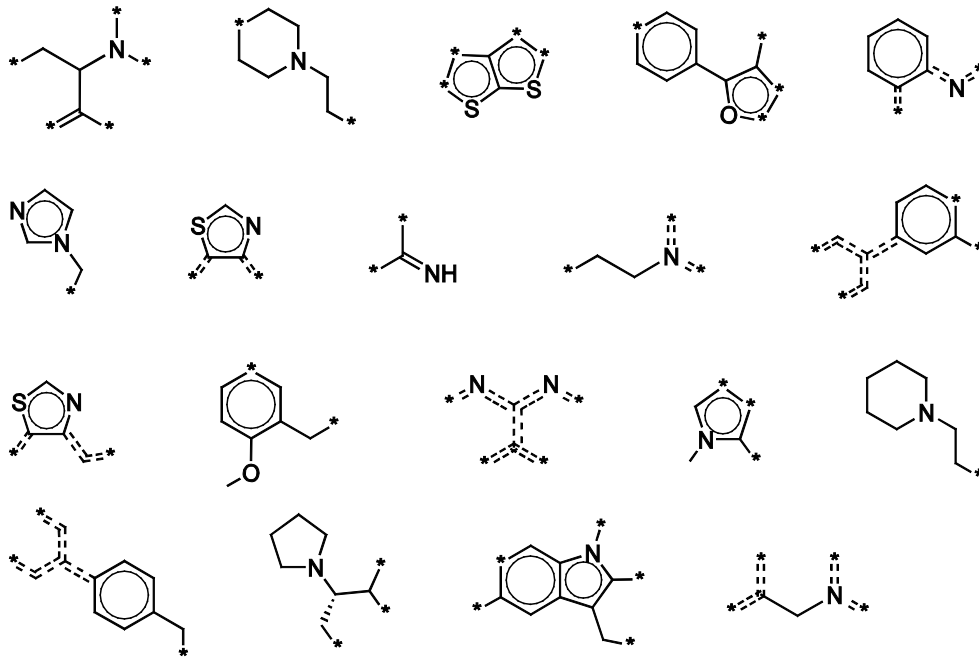


Abbildung 19: Am schlechtesten bewertete Features (Bayes-Modell 1).

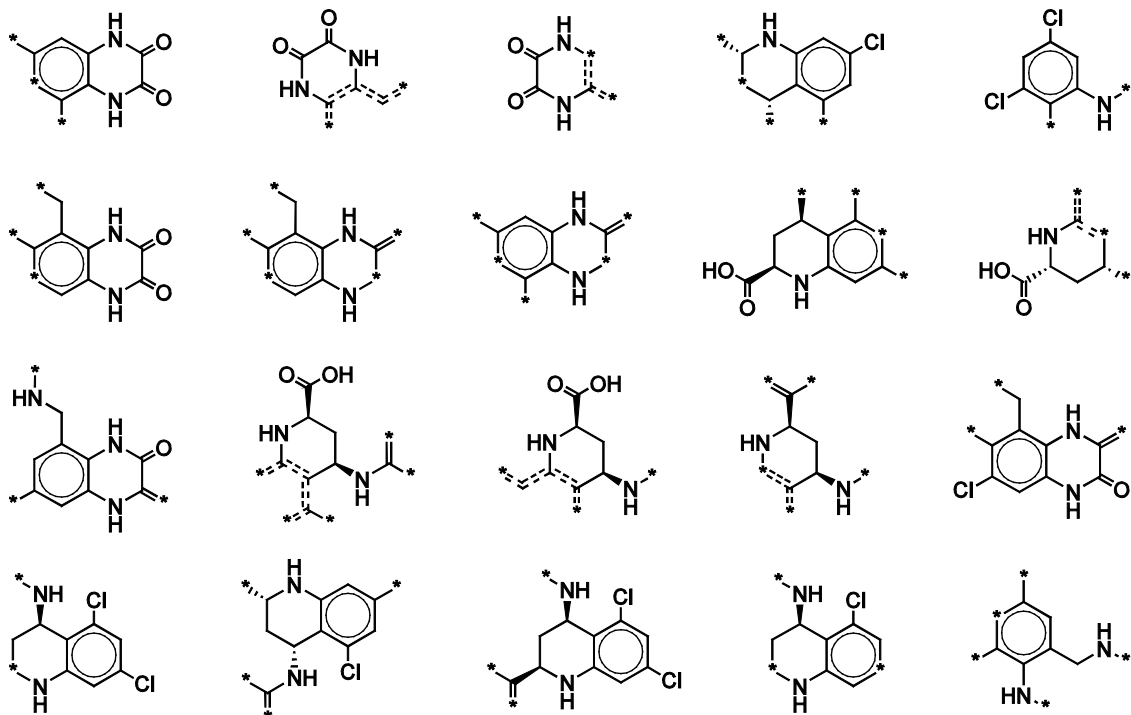


Abbildung 20: Am besten bewertete Features (Bayes-Modell 3)

Die statistische Auswertung wurde über 419 Referenzliganden mit einer Aktivität unter 1 μM sowie 61.221 inaktiven Moleküle (*decoys*) aus der WOMBAT-Datenbank durchgeführt.

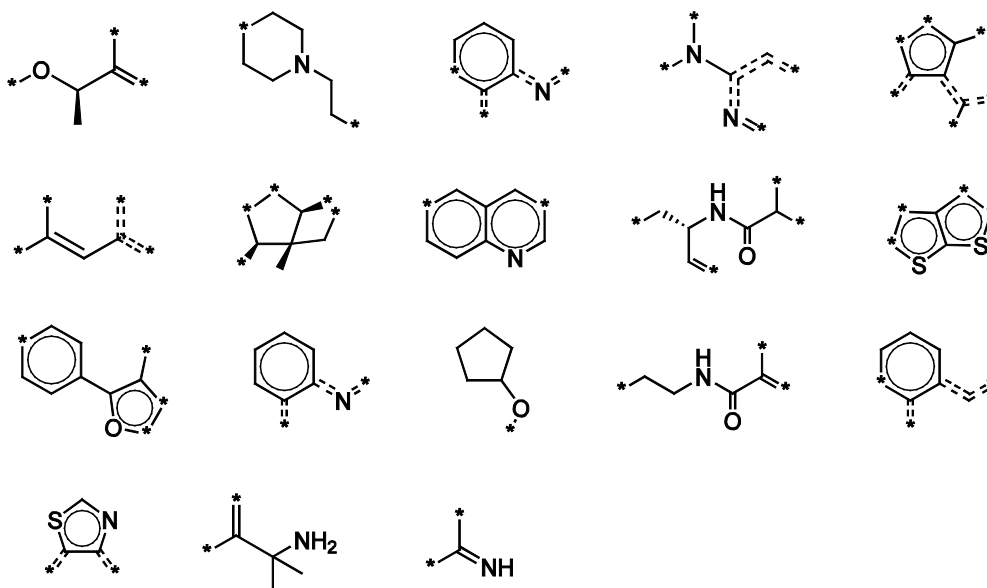


Abbildung 21: Am schlechtesten bewertete Features (Bayes-Modell 3).

Die Vorhersagekraft aller Modelle wurde mit einem zuvor erstellen externen Testdatensatz überprüft, welcher Moleküle enthielt, die der Bayes-Klassifikator nie zuvor „gesehen“ hatte. Die Korrektheit der Klassifikation wurde als ROC-Plot aufgetragen (Abbildung 22).

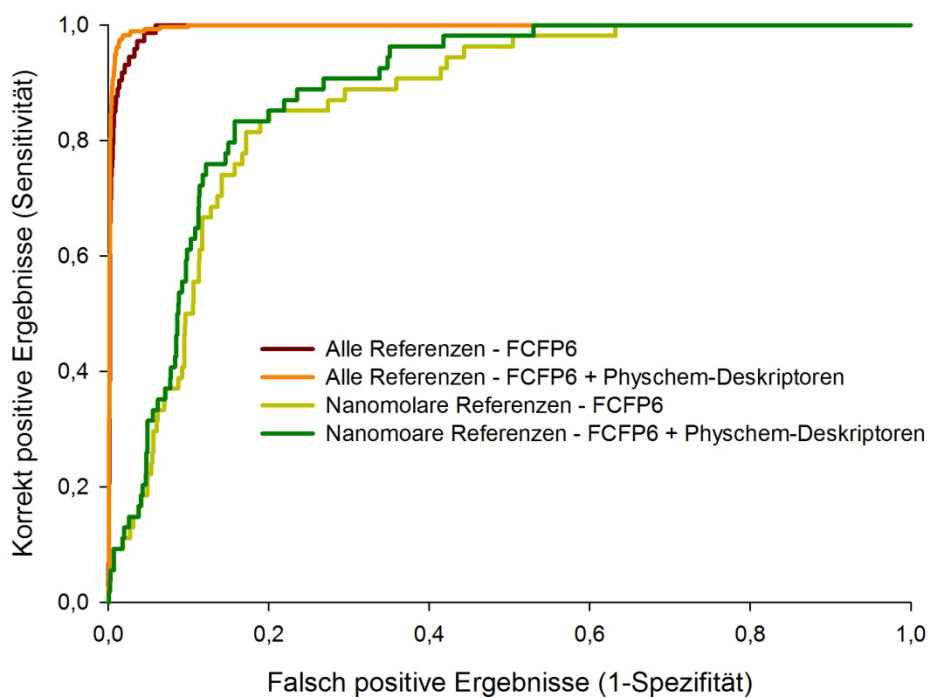


Abbildung 22: Receiver-Operating-Characteristics der Bayes-Klassifikatoren

Es zeigte sich, dass bereits die Verwendung von Referenzaktiven aller Aktivitätsklassen, also nicht nur hochaffiner Substanzen genügend Information beinhaltet, um eine Klassifikation zwischen Glycin_B-Liganden und inaktiven Substanzen vorzunehmen. Obwohl sich ein deutlicher Unterschied in der internen Modellvalidierung zeigte (Tabelle 5), führte das Hinzufügen weiterer physikochemischer Trainingsparameter im externen Testdatensatz nicht zu einer merklichen Verbesserung der Vorhersagequalität. Besonders deutlich wurde in dieser Untersuchung, dass ein deutlich „realistischeres“ Bild der Vorhersagegüte der Modelle durch das Verwenden eines externen Testdatensatzes gezeichnet werden konnte, als allein durch eine 10-fache interne Kreuzvalidierung durch *Leave-one-Out* auf den Trainingsdaten möglich war. Um diese Einschätzung des Modellbildungsprozesses zu relativieren, wurden zusätzlich für alle Bayes-Klassifikatoren die entsprechenden Korrelationskoeffizienten nach Matthews berechnet (Tabelle 5).

Tabelle 5: Retrospektive Validierung verschiedener Bayes-Klassifikatoren
 Modell 1 (alle Referenzstrukturen, FCFP-6 Deskriptor), Modell 2 (alle Referenzstrukturen, FCFP-6 plus Physchem-Deskriptoren), Modell 3 (nanomolare Referenzstrukturen, FCFP-6) und Modell 4 (nanomolare Referenzstrukturen, FCFP-6 plus Physchem-Deskriptoren)

Bayes-Klassifikator	interne Validierung		externe Validierung Matthews CC
	AUROC	Matthews CC	
1	0,99	0,67	0,47
2	0,99	0,80	0,56
3	0,99	0,60	0,34
4	0,99	0,61	0,37

Ein anderes Ergebnis zeigte sich beim nanomolaren Vorhersagemodell (Bayes-Klassifikatoren 3 und 4). Hier war der Klassifikator nur zum Teil in der Lage, ihm unbekannte nanomolare Glycin_B-Antagonisten aus dem externen Testdatensatz zu identifizieren. Eine Ursache hierfür dürfte die verstärkte Feinauflösung von Trainings- und Testdaten sein: die Anzahl an Strukturen, auf denen der Bayes-Klassifikator statistisch „lernen“ konnte, war im Vergleich zum ersten Modell stark reduziert, zudem waren die verschiedenen Chemotypen nun in weniger, aber dafür deutlich diverseren Clustern („Aktivitätsinseln“) aufgeteilt. Dem Bayes-Klassifikator standen somit insgesamt weniger Informationen über die Breite des chemischen Raumes der Glycin_B-Antagonisten während des Trainings zur Verfügung. Trotzdem konnte dieses Modell noch etwa 50% der nanomolaren Referenzverbindungen des externen Testdatensatzes, deren Chemotypen im Training des Modells nicht vorgelegt wurden, innerhalb der ersten 10% der Ergebnisse identifizieren. Auch hier führte das Verwenden

weiterer Parameter zum Lernen nicht zu einer deutlichen Verbesserung der prospektiven Vorhersagekraft des Modells (Tabelle 5).

Außerdem wurde überprüft, ob der Bayes-Score eines Moleküls mit seinem jeweiligen experimentell bestimmten Aktivitätswert korreliert. Der Bayes-Score entsprach dabei dem normalisierten Wahrscheinlichkeitswert, dass das Molekül aus der Menge der zuvor trainierten „guten“ Moleküle stammte. Dazu wurde von allen Literaturreferenzen und eigenen Substanzen der pIC_{50} (negativer dekadischer Logarithmus aus dem IC_{50} -Wert) berechnet, wenn für die jeweilige Substanz eine IC_{50} -Messung mit Dosierungs-Wirkungs-Kurve (Dose-Response-Curve, DRC) durchgeführt wurde. Geschätzte IC_{50} -Werte aus Einpunktmessungen wurden aufgrund ihres möglichen hohen Messfehlers nicht mit in die Analyse einbezogen. Hierbei war zu erkennen, dass sowohl retrospektiv als auch prospektiv keine lineare Korrelation zwischen Bayes-Score und experimentellem pIC_{50} gegeben war (Abbildung 23). Für das Bayes-Modell 1 (alle Referenzsubstanzen unter 30 μ M) wurde ein $R^2=0,18$ (retrospektiv) bzw. $R^2=0,02$ (prospektiv) ermittelt. Das Bayes-Modell 3 (nur nanomolare Referenzen) korrelierte nur unwesentlich besser mit $R^2=0,24$ (retrospektiv) bzw. $R^2=0,09$ (prospektiv).

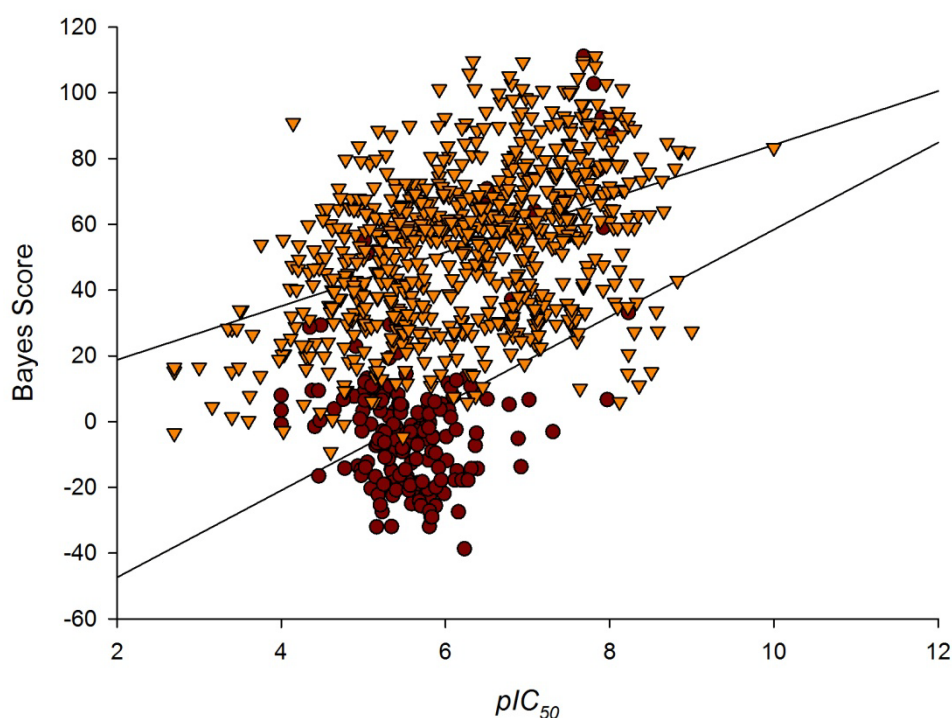


Abbildung 23: Korrelation Bayes-Score – experimentelle Messungen

Aufgetragen wurden die Score des nanomolaren Bayes-Klassifikators einer Substanz gegen ihre experimentelle pIC_{50} -Werten aus Literatur (retrospektive Studie - orange Dreiecke) und den Ergebnissen des Virtual Screening (prospektive Studie - rotbraune Kreise).

Das Modell mit der besten externen Vorhersagekraft (Modell 2) wurde prospektiv auf die Bibliothek der virtuellen Screeningsubstanzen angewandt. Aus den 4,6 Millionen untersuchten Substanzen wiesen 35.756 Verbindungen Sub-

strukturen auf, die allgemeines Merkmal der Glycin_B-Referenzliganden sind. Für diese identifizierten Verbindungen wurde mittels des Bayes-Klassifikators ein Wahrscheinlichkeitswert für die Zugehörigkeit der Moleküle zur Klasse der Glycin_B-Liganden berechnet. Die 500 Moleküle mit dem höchsten Bayes-Wahrscheinlichkeitswert wurden in einer vorläufigen Hitliste abgespeichert.

Sowohl die Modellgenerierung als auch die Anwendung des Bayes-Klassifikators im Virtual Screening der 4,6 Millionen käuflichen Substanzen zeigte sich als äußerst zeiteffizient: Das Training des Klassifikators mit 1023 aktiven und etwa 67.000 inaktiven Moleküle benötigte weniger als 3 Minuten auf einer Intel Xeon CPU mit 2,33 GHz, die prospektive Anwendung auf den 4,6 Millionen Screeningsubstanzen weniger als 20 Minuten. Der Bayes-Klassifikator war somit in der Lage, pro Sekunde mehr als 20.000 Moleküle zu analysieren und zu bewerten.

3.2.1.2 Ähnlichkeitssuche mit einem 2D-Pharmakophordeskriptor

Zur Validierung der Vorhersagekraft der Ähnlichkeitssuche mittels 2D-Pharmakophordeskriptors wurden die PHFP2-Fingerprints der Suchstrukturen sowie der Moleküle des externen Testdatensatzes berechnet. Als Maß der Ähnlichkeit wurde der Tanimoto-Score zwischen den Suchstrukturen und den Molekülen des Testdatensatzes ermittelt und eine nach Ähnlichkeit sortierte Liste erzeugt. Die Auswertung der Klassifikationsgüte als ROC-Kurve zeigte eine reduzierte Vorhersagekraft (Abbildung 24).

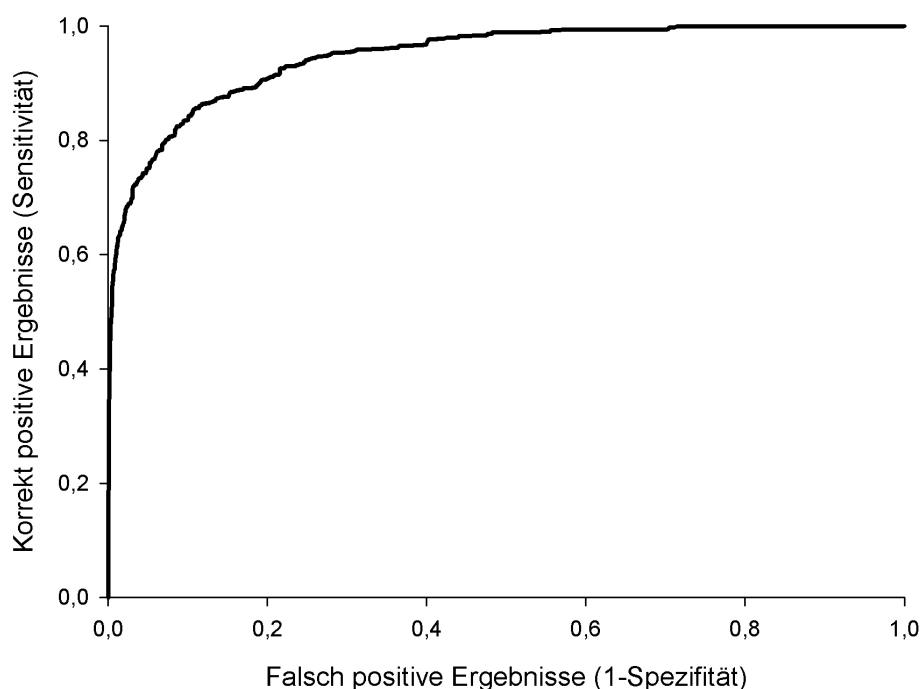


Abbildung 24: Receiver-Operating-Characteristics - PH4FP2-Deskriptor

3.3 Ligandenbasiertes Virtual Screening im implizit flexiblen Rezeptor

Im Gegensatz zu den rezeptorbasierten Verfahren des Virtual Screenings beinhalten ligandenbasierte dreidimensionale Methoden eine implizite Darstellung von Rezeptorflexibilität. Durch die Verwendung von möglichst nahe an der bioaktiven Konformation des Liganden liegenden dreidimensionalen Molekülstrukturen wird implizit der flexible Rezeptor dargestellt (Ligand-Rezeptor-Formkomplementarität). Diese Verfahren haben somit den Vorteil, erweiterte Konformationsräume der Liganden zu durchsuchen, ohne eine explizite Modellierung des Rezeptors mit allen damit verbundenen Schwierigkeiten zu erfordern.

3.3.1 Dreidimensionaler Formvergleich

In diesem Ansatz wurde die Technik des dreidimensionalen Formvergleichs von Liganden (*shape matching*) eingesetzt, um neuartige Glycin_B-Antagonisten zu identifizieren. Die Implementierung der Oberflächenanalyse erfolgte hierbei anhand Schrödinger Phase. Da hierbei verschiedene Arten der Atomtypisierung möglich sind (reines Volumen, Typisierung nach Element, Pharmakophortyp oder MacroModel-Atomkodierung (elektrostatistisch)), wurde die Anreicherung nach einer Shape-Suche mit den jeweiligen Atomtypen in zuvor generierten externen Testdatensatz überprüft. Hierbei erwies sich, dass die elektrostatische Atomtypisierung sowie das reine Volumen der Liganden eine bessere Anreicherung generierten als die beiden anderen Atomtypen (Abbildung 25). In der Summe zeigte die Anreicherung mit allen vier Typisierungen jedoch keine wesentlichen Unterschiede, und auch im Vergleich mit anderen getesteten Methoden des Virtual Screenings war die generelle Anreicherung hier eher schwach. Dies lässt sich dadurch begründen, dass die Form eines Moleküls, auch mit einer weitergehenden Beschreibung der Art der Form nicht ausreichend ist, um Glycin_B-Antagonisten von inaktiven Molekülen zu trennen. Bereits die Glycin_B-Referenzsubstanzen unterschieden sich zum Teil erheblich in ihren Formen (die 416 nanomolaren Literatur-Referenzliganden haben eine durchschnittliche molekulare Masse von 361,1 Dalton mit einer Standardabweichung von 76,2 Dalton). Auch das parallele Suchen mit mehreren Suchstrukturen in unterschiedlichen Formen („Consensus-Suche“) brachte keine Verbesserung des Ergebnisses.

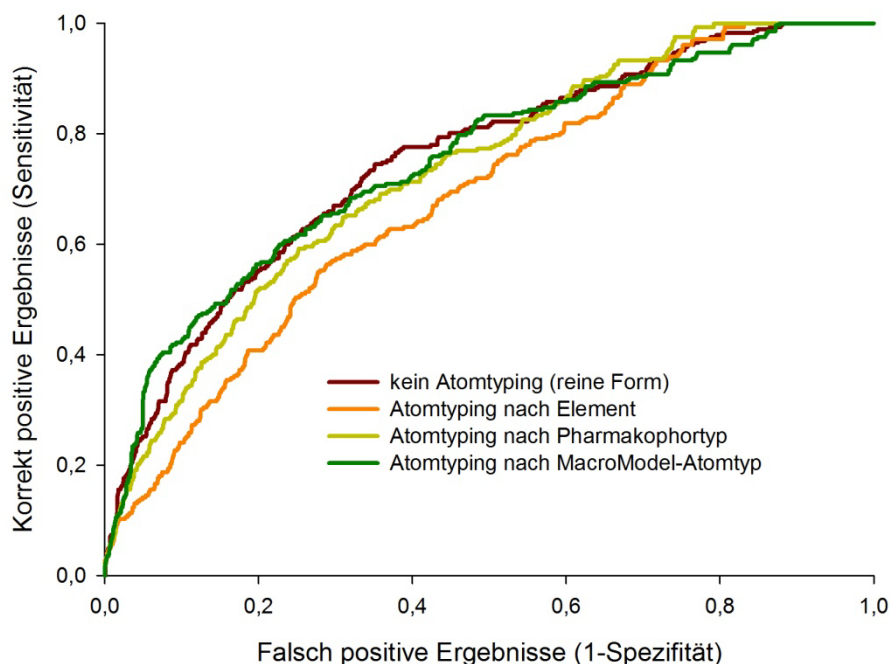


Abbildung 25: Receiver-Operating-Characteristics - Shape-Matching

3.3.2 Pharmakophorsuche und QSAR-Analyse

88 verschiedene Pharmakophorhypothesen wurden automatisch in Schrödinger Phase 3.0 durch Überlagerung von maximal sieben und minimal fünf Pharmakophorpunkten erzeugt. Das Modell mit der besten Vorhersagekraft (Abbildung 26) wurde als Pharmakophormuster für den externen Testdatensatz sowie die darauf folgende prospektive Suche angewandt. Das Modell erfüllt die Interaktionspunkte, die für Glycin_B-Antagonisten in der Glycin_B-Bindungsstelle benötigt werden (Kapitel 1.3.2).

Zusätzlich wurde auf der ausgesuchten Pharmakophorhypothese ein QSAR-Modell generiert. Hierzu wurde mittels einer Partial Least Squares-Analyse (PLS)^{198,199} die Belegung des dreidimensionalen Raumes durch die Pharmakophorpunkte des Liganden mit seiner jeweiligen Bindungsaffinität korreliert. Die Implementierung der PLS-Analyse erfolgte in Phase Schrödinger Phase 3.0. Eine Validierung auf einem internen Testdatensatz (50% Zufallsverteilung von Aktiven und Inaktiven um einen Trainings- und Testdatensatz zu erstellen) zeigte für das QSAR-Modell eine Vorhersagekraft von $R^2 = 0,75$ und $Q^2 = 0,59$. Die beste Vorhersage wurde unter Verwendung von drei PLS-Faktoren erreicht (Abbildung 27).

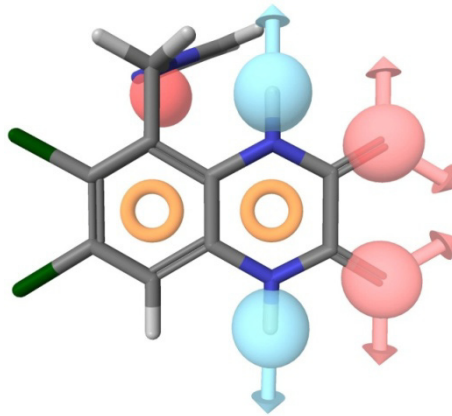


Abbildung 26: Pharmakophorhypothese

Hypothese mit der besten retrospektiven Vorhersagekraft. Braun: aromatische Systeme; blaue Kugeln mit Richtungsvektor: Wasserstoffbrückendonoren mit Richtung der Wasserstoffbrückenbindung; rote Kugeln mit Richtungsvektor: Wasserstoffbrückenakzeptoren mit Richtung der Wasserstoffbrückenbindung; dunkelrote Kugeln: negativ ionisierbar

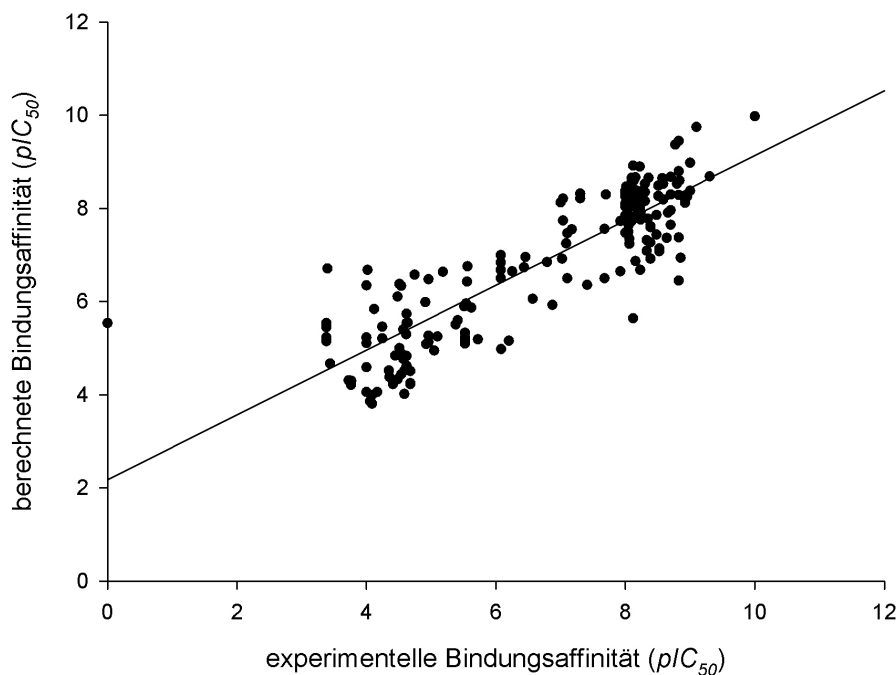


Abbildung 27: Retrospektive Validierung des 3D-QSAR-Modells

Berechnete (Partial Least Squares-Analyse, drei PLS-Faktoren) gegen experimentelle Bindungsaffinität, jeweils aufgetragen als pIC_{50} Werte. Die Vorhersage korreliert mit $R^2=0,75$ und $Q^2=0,59$. Phase Fitness-Score und Bindungsaffinität korrelieren mit nur $R^2=0,05$.

Nach Validierung im externen Testdatensatz zeige sich, dass die Nutzung des PLS-Modells als Scoringfunktion eine deutlich bessere Anreicherung erzeugte als der Phase Fitness-Score, der auf der reinen geometrischen Überlagerung der Pharmakophorpunkte basiert (Abbildung 28). Unter Verwendung des Fitness-Scores wurde nur die Hälfte der Glycin_B-Antagonisten des externen

Testdatensatzes erkannt, während die andere Hälfte zufällig bewertet wurde. Eine Untersuchung der falsch bewerteten resp. nicht erkannten Moleküle ergab, dass diese unterschiedlichen Serien angehörten, die Nichterkennung war somit nicht chemotypisch motiviert. Vielmehr lag eine hohe Übereinstimmung ihrer Pharmakophorpunkte mit den Pharmakophorpunkten des Modells vor. In der weiteren Analyse zeigte sich, dass sich diese fehlklassifizierten Moleküle erheblich im Phase Vektor-Score, einem Unterterm des Fitness-Scores, von den Vektor-Scores der korrekt identifizierten Glycin_B-Antagonisten unterschieden. Der Phase Vektor-Score addiert einen Term für die Richtionalität von Interaktionen, also z.B. Wasserstoffbrückenbindungen, zum globalen Fitness-Score. Es besteht somit die Möglichkeit, dass die Referenzliganden aufgrund von ungenügendem Konformerendurchmusterung falsch negativ klassifiziert wurden. Daher wurde das QSAR-Modell anstelle der Phase Fitnessfunktion als Scoringfunktion für das prospektive Virtual Screening eingesetzt.

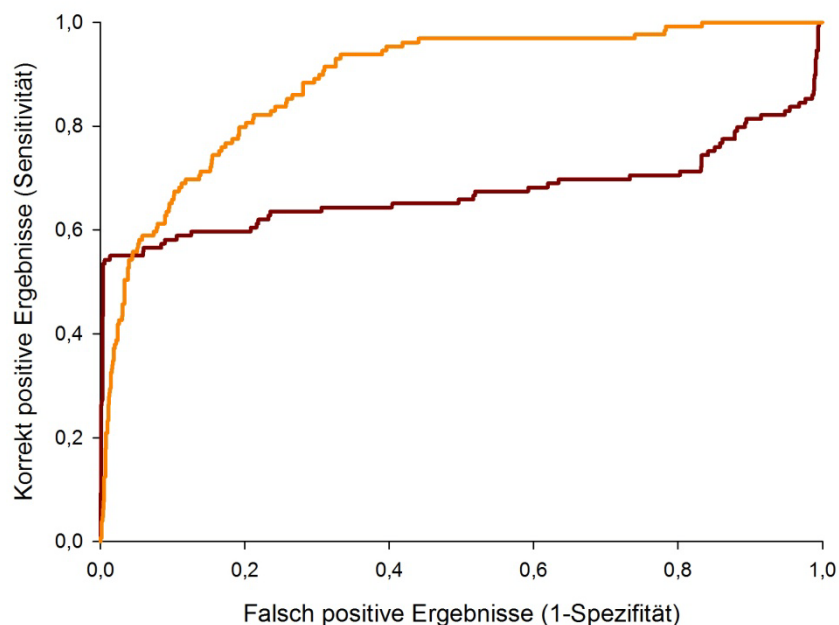


Abbildung 28: Receiver-Operating-Characteristics - Pharmakophorsuche
Braun: Phase Fitness-Score, orange: 3D-QSAR-Modell

Das Durchmustern der etwa 750 Millionen Konformere in der zuvor generierten 3D-Datenbank (Kapitel 2.2.3) benötigte mehrere Tage auf einem Computercluster mit 26 CPUs. Die besten 500 Substanzen, bewertet mit ihren jeweiligen QSAR-Scores, wurden in eine vorläufige Hitliste übernommen und visuell inspiziert.

3.4 Rezeptorbasiertes Virtual Screening im rigiden Rezeptor

3.4.1 Molecular Docking

Die Anreicherung an Glycin_B-Antagonisten in einem externen Testdatensatz bei der Verwendung verschiedener Scoringfunktionen und die Auswirkung der Definition von erforderlichen Interaktionen mit dem Rezeptor (*Constraints*) beim Molecular Docking war Gegenstand einer weiteren retrospektiven Studie. Zusätzlich wurde die Auswirkung der energieminierten (relaxierten) Rezeptorstruktur im Vergleich zur ursprünglichen Röntgenkristallstruktur untersucht. Hierbei zeigte sich, dass alle Methoden eine grundsätzlich hohe Anreicherungsrate aufwiesen (Abbildung 29), dabei aber durchaus Qualitätsunterschiede bestanden:

Die Verwendung der „weicheren“ Scoringfunktion von Glide SP ergab eine durchgehend bessere Anreicherung über den gesamten Testdatensatz als bei der „harten“ Scoringfunktion von Glide XP. Dies lässt sich dadurch erklären, dass Glide SP toleranter gegenüber kleineren Kollisionen (Clashes) zwischen Ligand und Rezeptor ist; dies schien besonders wichtig im Fall von Glycin_B zu sein, da für den Rezeptor eine hohe Auswirkung der Induced-Fit-Effektes postuliert wird. Glide SP erlaubte somit einen höheren Grad an impliziter Rezeptorflexibilität, wohingegen Glide XP Kollisionen explizit bestrafte, d.h. ein deutlich rigideres Rezeptormodell darstellt. Ebenso erzielte das Docking in die energieminierte Rezeptorstruktur ein besseres Ergebnis als die Verwendung der ursprünglichen Kristallstruktur. Das Relaxieren des Rezeptors im Kraftfeld vor dem eigentlichen Docken generierte dabei eine entspannte Bindetasche, in der offenbar mehr Glycin_B-Referenzsubstanzen optimal eingepasst werden konnten als in die originale Kristallstruktur. Das Setzen von Dockingconstraints wiederum hatte keine substantielle Auswirkung auf die Anreicherungen für Glide SP, Glide XP unter Verwendung von Constraints schnitt in der Auswertung des gesamten Testdatensatzes am schlechtesten ab (Abbildung 29).

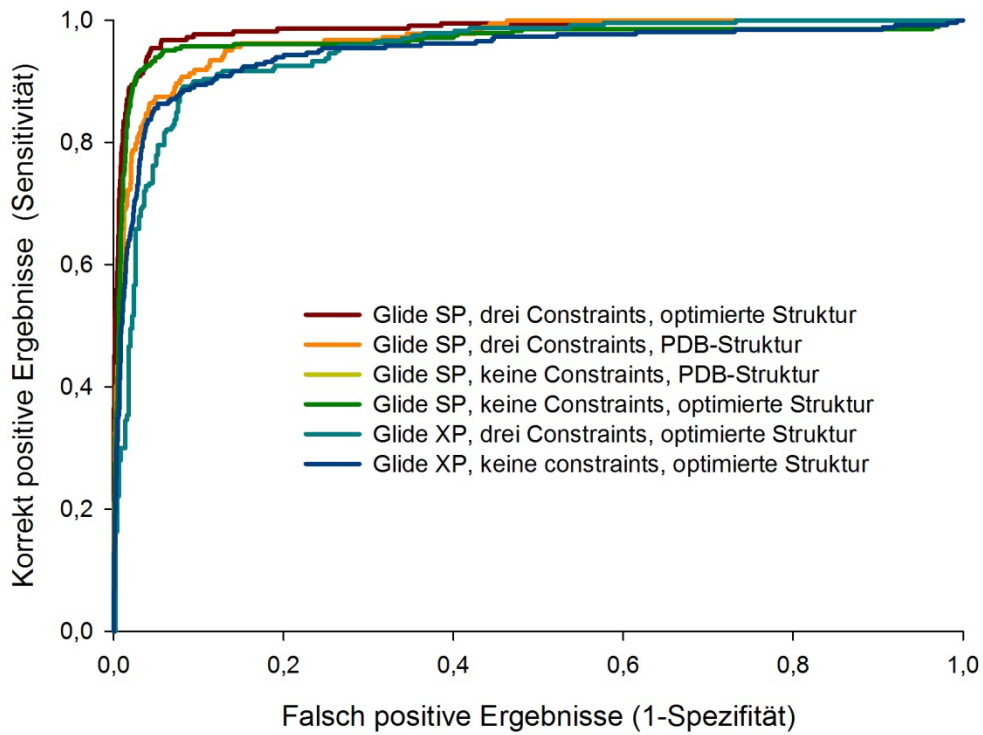


Abbildung 29: Receiver-Operating-Characteristics - Docking

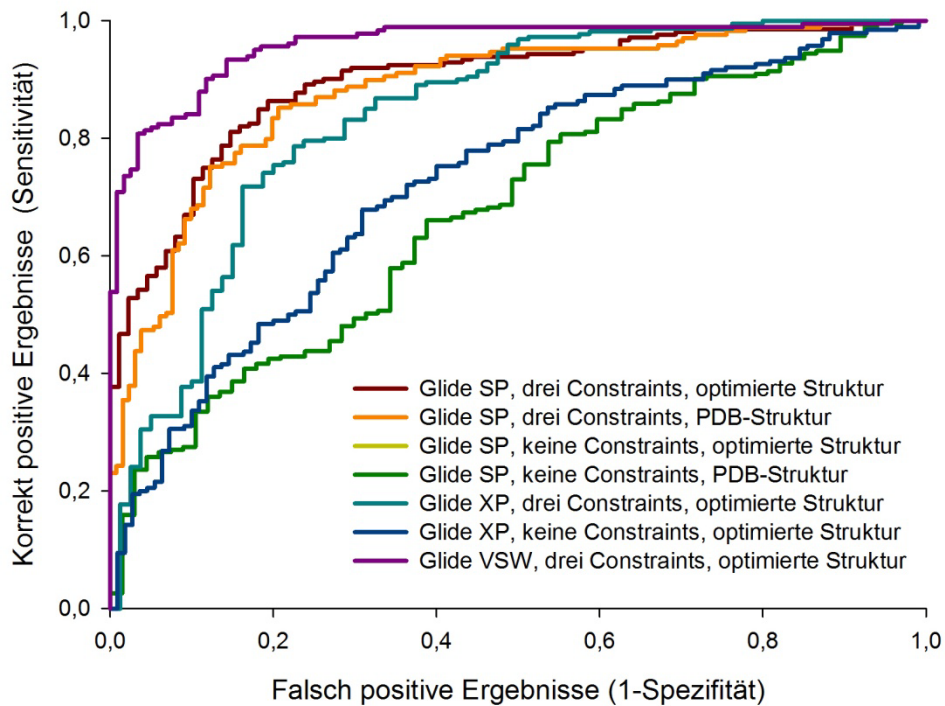


Abbildung 30: Receiver-Operating-Characteristics - Docking (300 beste Hits)

Zusätzlich wurde eine zweite Analyse mit den besten 300 Dockingergebnissen sortiert nach Dockingscore durchgeführt (Abbildung 30). Dabei zeigte sich, dass die Implementierung von Constraints essentiell für das Anreicherungsresultat ist. Sowohl Glide SP (im optimierten und nicht-optimierten Rezeptor) als auch Glide XP lieferten mit Constraints die besten Anreicherungsraten. Der positive Einfluss der Energieminimierung des Rezeptors auf die Anreicherungsrate war für Glide SP wie schon zuvor deutlich erkennbar.

Als Besonderheit in diesem Schritt der Validierung wurde der "Virtual Screening Workflow" (VSW) der Schrödinger Suite 2008 untersucht. Dieser Workflow führt verschiedene Stadien des Docking automatisiert sequentiell aus: Liganden, die ein initiales „Hochdurchsatz-Docking“ (*"high-throughput virtual screening"* (HTVS)) passierten wurden anschließend durch weitere Docking-schritte in Glide SP und schließlich Glide XP analysiert. Jeder dieser Schritte wendete dabei einen präziseren werdenden Scoringalgorithmus an. Das HTVS verfügte nur über eine grundlegende Beurteilung, ob ein Ligand prinzipiell in die Bindungsstelle passt (Formkomplementarität), Interaktionen mit dem Rezeptor wurden nur rudimentär beurteilt und ebenso wurde nur ein sehr beschränkter Satz von Konformeren für den Liganden berechnet. In den folgenden Dockings wurden mit fortschreitender Komplexität der verwendeten Scoringfunktionen auch die Details der Rezeptor-Liganden-Interaktion bewertet, von der Berechnung von Wasserstoffbrückeninteraktionen in Glide SP hin zur Beurteilung komplexer lipophiler Wechselwirkungen in Glide XP. Da in jedem Schritt die Anzahl der Liganden, die weitergereicht werden, stark reduziert wurde (nur die besten 10% aller Liganden erreichten jeweils die nächste Stufe), konnte im Folgeschritt mehr Zeit für eine detailliertere Berechnung von Konformeren und der Berechnung und Bewertung von Interaktionsmustern verwendet werden. Zusätzlich verringerte sich durch die schrittweise Einengung des Suchraumes an Liganden die Laufzeit im Vergleich zu einem vollen Durchlauf deutlich.

Bei der Analyse der besten 300 Dockingposen fiel zudem auf, dass unterschiedliche Docking- und Scoringfunktionen in der Lage waren, eine verschieden hohe Anzahl an Liganden erfolgreich zu docken. So erzeugte Glide XP in der Gesamtheit deutlich weniger Posen als Glide SP. Um das Dockingresultat von falsch-positiven Posen zu bereinigen, wurde der Matthews Korrelationskoeffizient für alle Dockingresultate berechnet. Jedem Dockingexperiment wurde der identische Testdatensatz von 232 Glycin_B-Antagonisten und 4.386 ihnen ähnlichen inaktiven Moleküle (*Decoys*) vorgelegt.

Die weitaus schlechteste Anreicherung wurde von Glide SP ohne Verwendung von Rezeptorconstraints erzielt: das Ergebnis war nur knapp besser als eine Zufallsauswahl (Matthews-CC von etwa 0). Der Algorithmus dockte zwar alle Glycin_B-Antagonisten, erzeugte aber auch für fast alle inaktiven Moleküle des Testdatensatzes Dockingposen. Die Einführung von Constraints verbesser-

te die Situation für Glide SP nur wenig. Dadurch reduzierte sich die Anzahl der falsch positiven Ergebnisse zwar auf 1/3, es wurden aber auch weniger korrekt positive Posen erzeugt. Auch Glide XP ohne Verwendung von Constraints generierte falsch positive Dockingposen für fast alle inaktiven Verbindungen (und damit einen Matthews-CC nahe 0). Mit Constraints reduzierte sich die Anzahl der Falsch-Positiven auf 1/8 und damit auf einen akzeptablen Matthews-CC von 0,47. Es lässt sich somit feststellen, dass die Verwendung von Constraints ein „Muss“ ist, wenn die Anreicherung in der Gesamtheit aller Dockingergebnisse ermittelt wird.

Tabelle 6: Anzahl an berechneten Dockingposen
Glycin_B-Referenzsubstanzen (korrekt positiv) und inaktive Moleküle (falsch positiv), daraus ergibt sich die Qualität der Anreicherung (Matthews Korrelationskoeffizient)

Algorithmus	alle Dockingposen			Top 300 Dockingposen		
	korrekt positiv	falsch positiv	Matthews	korrekt positiv	falsch positiv	Matthews
Glide SP, keine Constraints, PDB-Struktur	232	4050	0,06	190	70	0,76
Glide SP, keine Constraints, optimierte Struktur	232	4230	0,04	197	67	0,78
Glide SP, drei Constraints, PDB-Struktur	165	1278	0,20	151	131	0,57
Glide SP, keine Constraints, optimierte Struktur	195	1434	0,23	188	88	0,73
Glide XP, keine Constraints, optimierte Struktur	218	3987	0,02	161	110	0,62
Glide XP, drei Constraints, optimierte Struktur	206	501	0,47	189	80	0,74
VSW, drei Constraints, optimierte Struktur	-	-	-	182	118	0,67

Innerhalb der besten 300 Dockingposen erzielten alle Scoringfunktion ein sichtbar besseres Ergebnis als im Gesamtranking aller Posen. Die beste Anreicherung wurde für Glide SP mit und ohne Constraints (Matthews 0,73 bzw. 0,78) sowie für Glide XP mit Constraints (Matthews 0,73) erzielt. Beide Dockingalgorithmen erzielten mit der nicht-optimierten Rezeptorstruktur etwas schlechtere Ergebnisse. Bemerkenswert in dieser Untersuchung war die Anreicherung für den “Virtual Screening Workflow”, also das sukzessive Anwenden

immer feinerer Scoringfunktionen. Dieser produzierte eine ähnliche Anreicherungsrate in den 300 besten Dockingposen ($\text{Matthews}_{\text{CC}} = 0,67$) wie das durchgehende Verwenden der komplizierteren Scoringfunktionen, benötigte aber für den Durchlauf des kompletten Testdatensatzes nur etwa ein Zehntel der Rechenzeit.

Eine Begründung für die bessere Anreicherung der prinzipiell weniger präzisen Scoringfunktion von Glide SP könnte das Ignorieren leichter Clashes zwischen Ligand und Rezeptor darstellen: Glide XP bewertete diese Clashes deutlich strenger mit dem Ziel, ungünstige Posen und damit falsch positive Ergebnisse zu verringern. Im Falle von Glycin_B könnten allerdings die vermutete hohe Flexibilität des Rezeptors sowie die hohe Diversität der Glycin_B-Liganden in Bezug auf ihre Form den Ausschlag geben. Eine Scoringfunktion, die implizit Rezeptorflexibilität in beschränktem Maße ermöglicht, indem sie über geringe van-der-Waals-Kollisionen „hinwegsieht“, ist von Vorteil gegenüber einem Scoring, welches nur eine rigide Rezeptorkonformation betrachtet. Ein zusätzlicher Vorteil von Glide SP war die im Vergleich zu Glide XP deutlich reduzierte Laufzeit (Faktor 10).

3.5 Ergebnisse des prospektiven Virtual Screenings

3.5.1 Neuheitsanalyse

Um die Neuheit der im Virtual Screening identifizierten Substanzen zu untersuchen, wurde die Bibliothek von Glycin_B-Referenzliganden auf häufig auftretende Scaffolds untersucht. Dazu wurden mittels der Pipeline Pilot Funktion „Generate Fragments“ zusammenhängende Ringsysteme in den Referenzverbindungen identifiziert und die Häufigkeit ihres Auftretens gezählt. Scaffolds, die in den 754 Referenzliganden zehnmals und häufiger enthalten waren, wurden als „nicht neu“ definiert und in einer Liste der „bekannten Glycin_B Scaffolds“ (Abbildung 31) vereinigt. In der Summe wurden so 91 unterschiedliche Ringsysteme als bekannte Glycin_B-Scaffolds identifiziert.

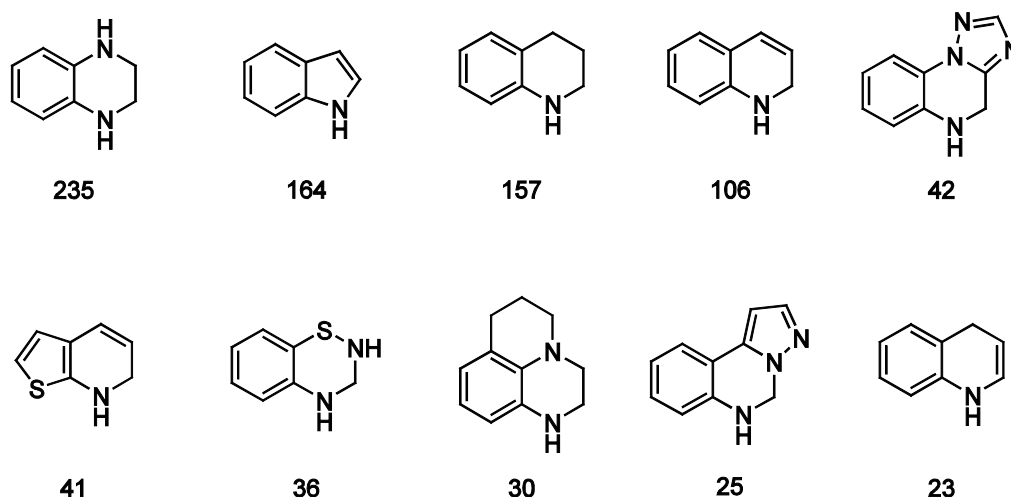


Abbildung 31: Die zehn häufigsten Glycin_B-Scaffolds ermittelt aus den Literatur-Referenzliganden, annotiert mit der jeweiligen absoluten Häufigkeit ihres Auftretens

Für jede der zuvor angewandten Virtual Screening Methoden wurde eine Substruktursuche nach den bekannten Glycin_B-Scaffolds durchgeführt. Alle virtuellen Hits, in denen ein solcher Scaffold enthalten war, wurden verworfen. Um zu ermitteln, welche Auswirkung diese Filterung auf die durchschnittliche Ähnlichkeit zwischen den Literaturreferenzen und den virtuellen Hits hatte, wurde die Verteilung der paarweisen Ähnlichkeiten vor und nach der Filterung berechnet (Abbildung 32). Die virtuellen Hits waren nach Filtern auf bekannte Scaffolds deutlich unähnlicher zu den vorhandenen Glycin_B-Referenzliganden.

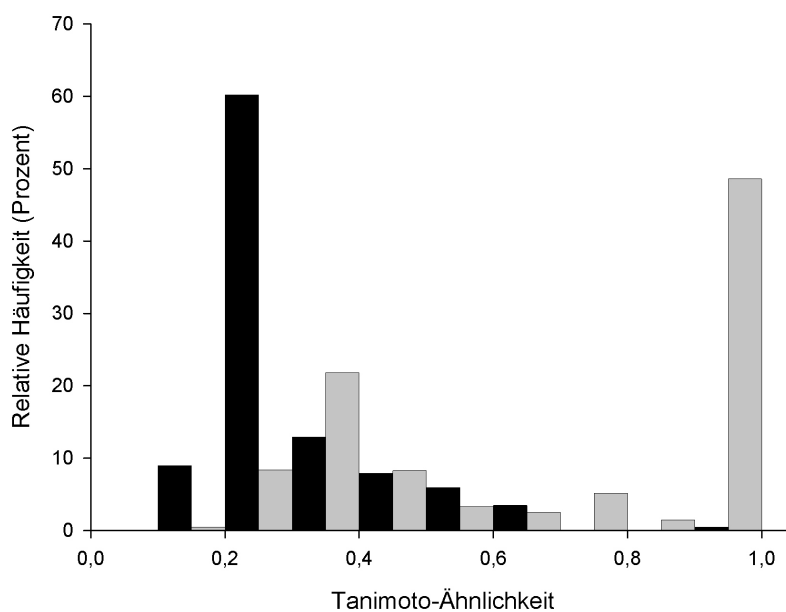


Abbildung 32: Histogramm der paarweisen Ähnlichkeiten zwischen nanomolaren Referenzstrukturen und den Virtual Screening Hits mit (schwarz) und ohne (grau) Filtern der „bekannten Scaffolds“ (FCFP6-Deskriptor, Tanimoto-Ähnlichkeit)

Diese Vorgehensweise des rigorosen Filterns auf bekannte Scaffolds ist in der Praxis ggf. nicht notwendig, um neuen patentierbaren chemischen Raum zu erschließen. Neue Strukturpatente lassen sich unter Umständen bereits durch eine leichte Modifikation einer Ringstruktur, welche durch das ursprüngliche Patent nicht geschützt ist, erzeugen. In dieser Arbeit wurde der Neuigkeitsfilter auf bekannte Glycin_B-Scaffolds vielmehr eingesetzt, um das „Scaffold-Hopping-Potential“ der einzelnen Virtual Screening-Methoden genauer zu analysieren.

3.5.2 Leistungsvergleich der Methoden des Virtual Screenings

Alle Methoden in dieser Virtual Screening-Studie waren in der Lage, neuartige Strukturen zu identifizieren. Der Prozentsatz an bekannten Scaffolds innerhalb der 500 jeweils am besten bewerteten Moleküle lag zwischen 10 und 66% (Tabelle 7). Während Methoden wie der Bayes-Klassifikator, Pharmakophorsuche sowie Shape-Matching eine starke Tendenz zum Scaffold Hopping aufwiesen (unter 20% der identifizierten Hits hatten einen bereits bekannten Scaffold), wiesen Docking und Ähnlichkeitssuche mittels Pharmakophorfingerprint eine höhere Anreicherung mit bekannten Scaffolds auf (35% bzw. 66%).

Tabelle 7: Leistungsvergleich der Methoden des Virtual Screenings bezüglich Rechenzeit, Neuheitserkennung als Prozentsatz von virtuellen Hits mit bekanntem Scaffold in den 500 jeweils am besten bewerteten Strukturen, Akzeptanz durch den menschlichen Experten (Prozentsatz an der Summe der real ausgewählten Substanzen), prospektive Aktivitätsvorhersage und Hitrate (Anteil tatsächlich aktiver Substanzen an der Menge der ausgewählten Hits der Methode)

Methode	Rechenzeit	Bekannte Scaffolds in den besten 500 Hits	Akzeptanz durch den menschlichen Experten	Substanzen mit <i>in vitro</i> Aktivität	Hitrate
Bayes-Klassifikator	30 Minuten (1 CPU)	16,8%	4,8%	6	50%
Docking	~2 Wochen (4 CPUs)	35,4%	16,7%	22	52%
Pharmakophor	~5 Tage (26 CPUs)	11,6%	42,2%	12	11,3%
Shape	~4 Tage (26 CPUs)	15,2%	16%	2	5%
Pharmakophorfingerprint	2 Stunden (1 CPU)	66,2%	20,3%	1	2%

Ein größerer Unterschied in den Methoden zeigte sich auch bei der Akzeptanz durch den menschlichen Experten. Hier wurden nur wenige vom Bayes-Klassifikator identifizierte Substanzen ausgesucht, während Pharmakophorsuche und Suche mit Pharmakophorfingerprints die als „interessanter“ erachteten Ergebnisse hervorbrachten. Ebenso wurde ein erheblicher Unterschied in der Vorhersagequalität deutlich: während durch Docking und Bayes-Klassifikator

eine Hitrate von 52% bzw. 52% bei unterschiedlicher absoluter Anzahl an Hits erzielt wurden, fielen die Methoden Pharmakophorsuche (11%), Shape-Matching (5%) und Ähnlichkeitssuche mittels Pharmakophorfingerprint (2%) deutlich zurück.

3.5.3 Aktivität im *in vitro*-Bindungsassay

Die im Virtual Screening identifizierten und ausgewählten Substanzen wurden in einem *in vitro* Bindungsassay auf Affinität zum Target überprüft. Eine erste Testung erfolgte mittels Einpunktbestimmung, und nur für die dort bestätigten Hitsubstanzen wurden vollständige Dosis-Wirkungs-Kurven (*Dose-Response-Curves*, DRC) mittels dreifach wiederholter Fünfpunktmessung erzeugt. Die untersuchten Substanzen wiesen Bindungsaffinitäten im niedrigen bis mittleren mikromolaren Bereich auf. Insgesamt hatten 44 Substanzen einen IC_{50} unter 100 μ M, davon 23 Substanzen unter 40 μ M. Obwohl Verbindungen mit mikromolarer Affinität meist nicht direkt als Leitstruktur in die weitere präklinische Entwicklung eingehen können, so können sie als Ausgangspunkte für weitere Startstrukturen dienen. Zudem waren viele der getesteten Substanzen fragmentartig, d.h. aufgrund ihres geringen molekularen Gewichtes eignen sie sich womöglich für eine direkte weitere chemische Optimierung.

Einige Ergebnisse des *in vitro* Screenings wurden ausgewählt und sind im Folgenden dargestellt (Abbildung 33). Viele der identifizierten Hits wiesen Glycin_B-untypische Scaffolds auf und sind bisher nicht als Glycin_B-Antagonisten beschrieben (Verbindungen **3** bis **23**). Die Verbindungen **1** und **2** stellten bisher nicht publizierte Analoge von Literaturreferenzen mit bekanntem Scaffold dar. Sie wurden in das Screening aufgenommen, um zu zeigen, dass auch niedrig nanomolare Substanzen von den Methoden identifiziert werden konnten.

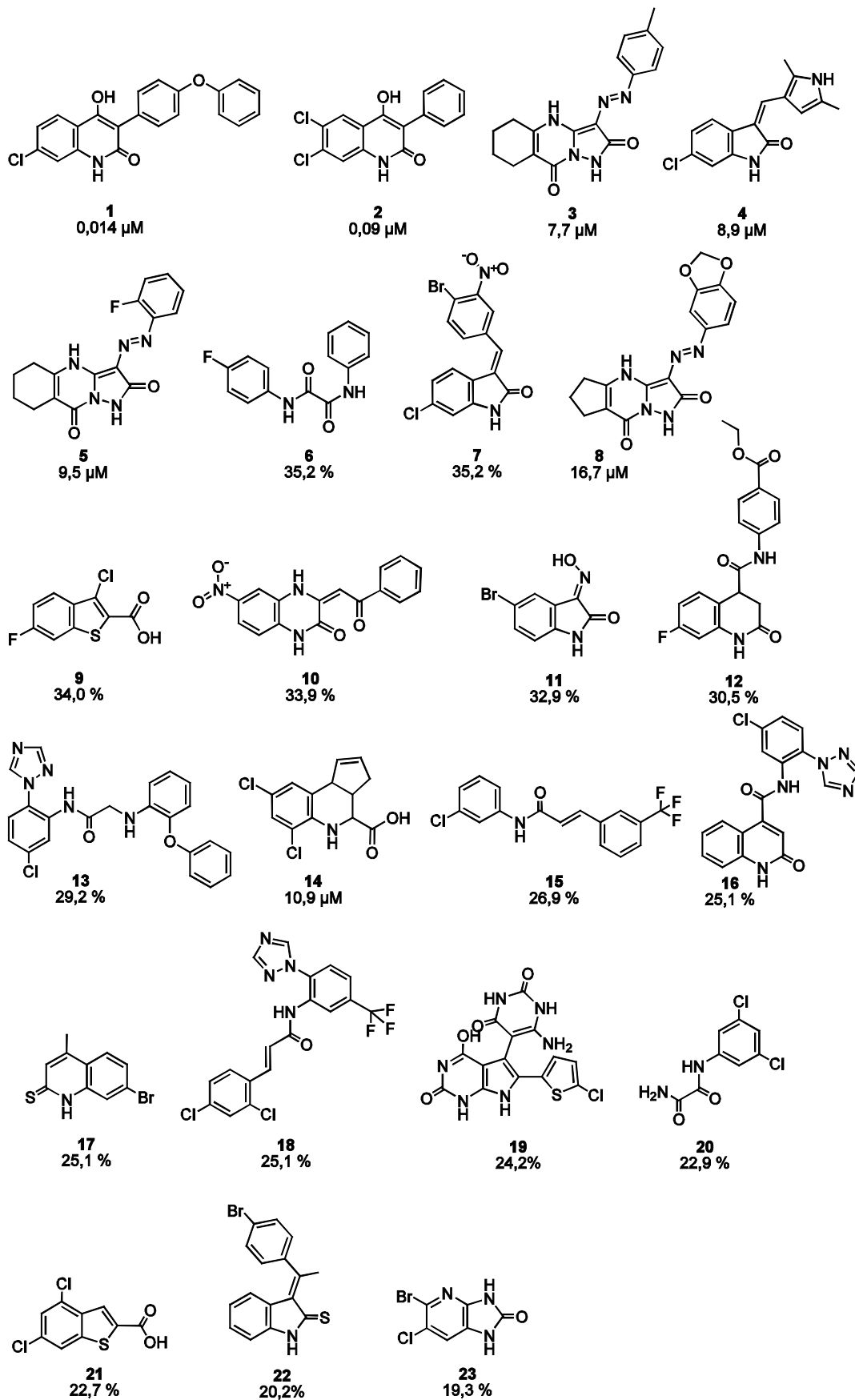


Abbildung 33: Ausgewählte Hits aus dem Virtual Screening mit ihren jeweiligen *in vitro* Bindungsaffinitäten. Messwerte aus dreifach wiederholten Einzelmessungen (geschätzter IC_{50} – Prozent Inhibition bei $10\mu\text{M}$) bzw. DRC-Experimenten

3.6 Identifizierung unterschiedlicher Serien von Glycin_B-Antagonisten mit neuartigen Scaffolds

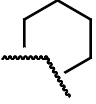
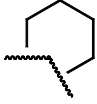
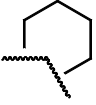
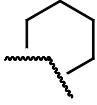
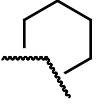
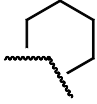
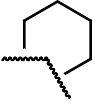
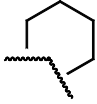
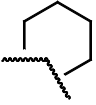
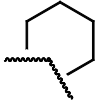


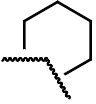
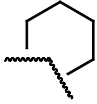
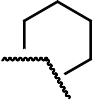
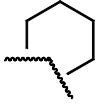

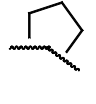
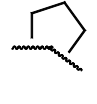
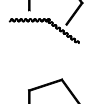
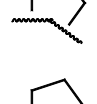
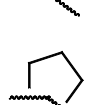
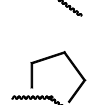
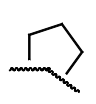
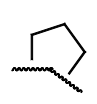
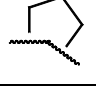
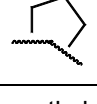



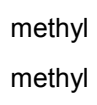
3.6.1 Struktur-Wirkungsbeziehungen der Azo-Verbindungen

Drei der während des virtuellen Screenings identifizierten Hits (Verbindungen **3**, **5** und **8**) wurden als besonders interessant erachtet, da sie einen für Glycin_B bisher unbeschriebenen Scaffold und eine ausreichend hohe Affinität aufwiesen. Aus diesem Grund sollte der chemische Raum um sie herum ebenfalls untersucht werden. Hierzu wurde eine weitere Ähnlichkeitssuche (FCFP-6 Fingerprint, Tanimoto-Ähnlichkeit) in den kommerziellen Substanzkatalogen mit diesen drei Verbindungen als Suchstrukturen durchgeführt. Dabei wurden 52 weitere Substanzen identifiziert, die erworben und pharmakologisch getestet wurden. Es zeigte sich, dass die initialen Hits aus dem Virtual Screening Teile einer größeren Serie von Azo-Farbstoffen waren, deren Struktur-Wirkungsbeziehungen im Folgenden beschrieben werden.

Der erste Scaffold der Azo-Verbindungen basiert auf dem 4H-Pyrazolo[1,5-a]pyrimidin-2,7-dion-Grundgerüst, das über einen Azo-Linker mit einem Phenylrest verbunden ist, welches optional in ortho- meta- und para-Stellung substituiert ist (Tabelle 8).

Tabelle 8: Struktur-Wirkungsbeziehungen des Azo-Scaffolds 1
Geschätzte IC_{50} -Werte (getestet bei 10 μ M) im Bindungsassay sind *kursiv*. Falls mehrere Wiederholungen der Messung vorgenommen wurden, wurde der Standardfehler (SEM) in Klammern angegeben. Sortierung nach IC_{50} .

Verbindung	R1	R2	R3	R4	R5	Glycin _B IC_{50} (μ M)
24		phenyl	fluor			1,6 (\pm 0,2)
25				chlor		5,3 (\pm 2,4)
26			ethoxy			5,6 (\pm 2,3)

27				methoxy	6,5
28				trifluormethyl	6,5 (± 0,5)
29					6,9 (± 3,9)
3				methyl	7,7
30					9,4
5			fluor		9,5 (± 0,2)
31					26 (± 17)
32				chlor	7,6
33				ethoxy	7,8 (± 2,8)
34				methyl	8,5
35					15,2 (± 0,3)
8					16,7
36				methoxy	21
37		methyl	fluor		25
38		methyl		ethoxy	28
39		methyl		methoxy	29
40		methyl			44
41		methyl		chlor	54 (± 45)
42		methyl		methyl	>100

Die aktivste Substanz dieser Reihe war die mit Fluor in R3-Stellung sowie mit einem Phenylrest an der Kernstruktur in R2-Position substituierte Verbindung **24**. Diese war zugleich die einzige Struktur der Serie mit einem nicht direkt annelierten Ringsystem bzw. einem Alkylrest am Kern. Die Synthese weiterer Verbindungen mit einem Phenylring an Position R2 sollte weitere aktive Verbindungen hervorbringen (Kapitel 4). Analog zur Dichlorsubstitution von DCKA wäre zu erwarten, dass eine Dichlorierung des Phenylrings die Affinität erhöht – vorausgesetzt, beide Substanzklassen nehmen den gleichen Bindungsmodus ein (Kapitel 3.8.3).

Weitere Substitutionen am Kerngerüst waren ein anneliertes Cyclohexanring bzw. Cyclopentanring gebildet aus R1 und R2 sowie Methyl an R2. Hierbei zeigte sich, dass ein Methylrest die Affinität negativ beeinflusst (Substanzen **37** etc.), während sich cyclische Substitutionen positiv auf die Aktivität auswirkten. Die Größe des annelierten Rings schien hierbei keinen Einfluss auf die Bindungsaffinität zu haben. Diese wurden eher durch die Substituenten an R3 bis R5 verursacht (z.B. **3/34**, **25/32** und **30/8**).

Affinitätssteigernd wirkten sich v.a. Substituenten mit negativem induktiven Effekt (Fluor **5**, Chlor **25/32**, Trifluormethyl **28**) resp. positivem mesomeren Effekt (Methoxy **27** und Ethoxy **26**) aus. Beide Effekte verringern bzw. delokalisieren die Elektronendichte am Benzolring und somit womöglich hydrophobe Wechselwirkungen mit dem Rezeptor.

Ein Methyl-Substituent an R5 (**3**) hat nur einen schwachen positiven induktiven Effekt und verursacht keine starke Affinitätsänderung. Ein Methylendioxy-Rest(**8**, **30**) sowie eine Sulfonamid-Gruppe in Position R5 (**31**) verringern die Affinität des Scaffolds zur Bindungsstelle.

Tabelle 9: Struktur-Wirkungsbeziehungen des Azo-Scaffolds 2
 Geschätzte IC_{50} -Werte (getestet bei 10 μ M) im Bindungsassay sind *kursiv*. Falls mehrere Wiederholungen der Messung vorgenommen wurden, wurde der Standardfehler (SEM) in Klammern angegeben.

Verbindung	R1	R2	R3	R4	R5	Glycin _B <i>IC</i> ₅₀ (μ M)
43					chlor	14 (\pm 5,5)
44			fluor			>100
45				chlor		>100
46					chlor	15
47				chlor		37
48			fluor			>100
49		methyl	fluor			>100

Scaffold 2 (Tabelle 9) unterscheidet sich durch Scaffold 1 nur durch eine Substitution der Carbonyl- durch eine Amino-Funktion im Kerngerüst. Dies reicht bereits aus, um ansonsten identische Strukturen in ihrer Affinität zum Rezeptor komplett zu deaktivieren (z.B. von $9,5 \pm 0,2 \mu$ M (**5**) auf $>100 \mu$ M (**44**)). Einzig ein Chloratom in meta- oder para-Stellung (R4 bzw. R5) behält eine moderate geschätzte Affinität bei (**46**, **47**). Verbindungen mit Chlor-Substituent in para-Position (R3) standen nicht zur Verfügung.

Tabelle 10: Struktur-Wirkungsbeziehungen des Azo-Scaffolds 3
 Geschätzte IC_{50} -Werte (getestet bei 10 μ M) im Bindungsassay sind *kursiv*.

Verbindung	R1	R2	R3	R4	R5	R6	Glycin _B <i>IC</i> ₅₀ (μ M)
50	amino				methoxy		19
51	amino				trifluormethyl		21
52	amino						30
53	amino						47
54	amino					methyl	86
55	amino			fluor			>100
56	methyl		methyl				>100
57	methyl		methyl	fluor			>100
58	methyl		methyl		trifluormethyl		>100
59	methyl		methyl		chlor		>100

Die Serie um Scaffold 3 (Tabelle 10) war in ihrer Größe im Vergleich zu den vorherigen beiden Serien deutlich reduziert; zudem wurden teilweise andere Substitutionsmuster eingesetzt, was die SAR-Analyse erschwerte. Die Carbonylfunktion in Position 2 wurde wieder eingeführt, dafür wurde die zweite Carbonylfunktion des Kerngerüsts durch eine zusätzliche R-Gruppe ersetzt, an der eine Amino- bzw. Methylgruppe eingeführt wurde. Die Amino-funktionalisierten Verbindungen haben zusätzlich einen Ethylester an R2; dieser fehlt den Methyl-Derivaten vollständig. Es konnte somit keine Aussage getroffen werden, ob sich Aktivitätsänderungen auf das Vorhandensein der Amino- oder

der Ethylestergruppe zurückführen lassen können. Die verbleibenden R-Gruppen waren äquivalent zu Scaffold 1.

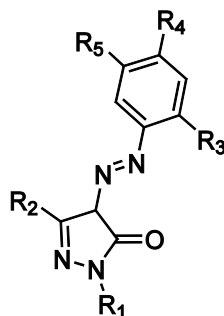
Während eine Methylgruppe an R1 die Aktivität komplett beseitigte, verblieb mit einer Aminofunktion an R1 sowie dem Ethylester an R2 Restaktivität. Grund hierfür kann der positive induktive Effekt des Methyls sein, welcher den pK_a des mit dem Rezeptor polar interagierenden Amids verringert. Positive mesomere (Methoxy, **50**) bzw. negative induktive Effekte (Trifluormethyl, **51**) und aliphatische Substitutionen steigerten die Aktivität im Vergleich zur Struktur ohne Substitution am Benzenring (**52**). Sulfonamid- und Methylsubstitution am Benzolring (**53** bzw. **54**) verringerten die Aktivität wie bereits bei Scaffold 1 sichtbar.

Tabelle 11: Struktur-Wirkungsbeziehungen des Azo-Scaffolds 4
Geschätzte IC_{50} -Werte (getestet bei 10 μ M) im Bindungsassay sind *kursiv*.

Verbindung	R1	R2	R3	R4	Glycin _B <i>IC₅₀</i> (μ M)
60	amino		chlor		<i>22</i>
61	amino		methoxy		<i>42</i>
62	amino	fluor			<i>53</i>
63	amino			methyl	<i>>100</i>

Mit Scaffold 4 (Tabelle 11) änderte sich die Kernstruktur der Azo-Verbindungen deutlich. Anstelle eines Zwei- oder Dreiringsystemes bildet hier ein Pyrazolon mit einem Phenylring am Stickstoff in Position 1 den Scaffold. An R1-Position befand sich eine Aminogruppe, so dass in diesem Scaffold nur Substitutionen am zweiten Benzolring in ortho-, meta und para-Stellung zum Azo-Linker erfolgten. Hierbei zeigte eine Chlorsubstitution in meta-Stellung einen moderat aktiven geschätzten IC_{50} , während eine Methoxygruppe in meta- und ein Fluoratom in ortho-Stellung nur noch schwach aktiv sind. Ein Methylsubstituent am Benzenring verringerte die Bindungsaffinität der Substanz wie bereits bei den vorherigen Scaffolds erkennbar um den Faktor 2 bzw. 2,5.

Tabelle 12: Struktur-Wirkungsbeziehungen des Azo-Scaffolds 5
 Geschätzte IC_{50} -Werte (getestet bei 10 μ M) im Bindungsassay sind *kursiv*.



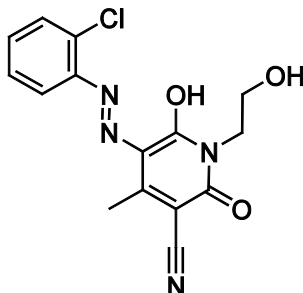
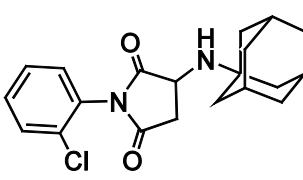
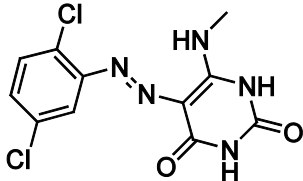
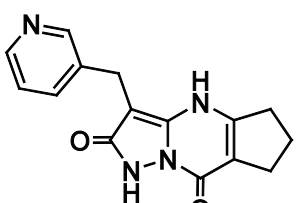
Verbindung	R1	R2	R3	R4	R5	Glycin _B IC_{50} (μ M)
64		methyl	chlor	chlor		<i>14</i>
65		methyl	chlor			<i>51</i>
66		methyl	chlor		chlor	<i>>100</i>
67	H	amino			chlor	<i>>100</i>
68	H	amino	fluor			<i>>100</i>

Kerngerüst des Scaffold 5 (Tabelle 12) war wiederum ein Pyrazolon, wobei der Phenylring an Position 1 entfiel. Substitutionen erfolgten am Stickstoff in Position 2 (R1, mit und ohne eine Thioamid-Gruppe) sowie an Position 5 (R2, Methyl- bzw. Aminogruppe) des Kerns. Das Substitutionsmuster des Benzolrings wurde ebenfalls geändert.

Es zeigte sich, dass ausschließlich Verbindungen mit Thioamid-Substituent an R1 in der Lage waren, Aktivität zu zeigen. Zwei Verbindungen ohne Thioamid waren inaktiv (**67**, **68**). Die Methylsubstitution an R2 hatte offenbar keinen Einfluss auf die Affinität der Substanzen. Während eine Dichlorsubstitution in ortho- und para-Stellung am Benzenring einen akzeptablen geschätzten IC_{50} von etwa 14 μ M erreichte, war sowohl die Monosubstitution in ortho-Stellung als auch eine Disubstitution in ortho- und meta-Stellung schwach aktiv bzw. inaktiv.

Durch die Ähnlichkeitssuche konnten auch vier weitere Substanzen identifiziert werden, die keiner Scaffoldklasse angehörten (Singletons) (Tabelle 13). Diese zeigten mittlere (**69**) bis schwache Affinitäten (**70**, **71**, und **72**) im Bindungsassay und wurden nicht weiter verfolgt.

Tabelle 13: Messdaten weiterer identifizierter Substanzen
Strukturen mit unterschiedlichen Scaffolds, geschätzte IC_{50} -Werte (getestet bei 10 μ M) im Bindungsassay sind *kursiv*.

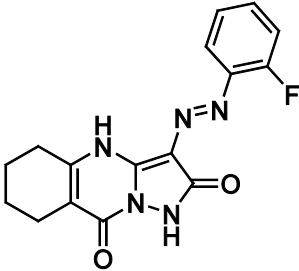
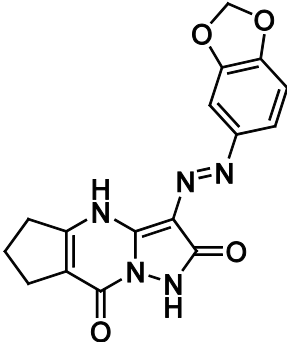
Verbindung		Glycin _B IC_{50} (μ M)
69		14
70		42
71		54
72		67

3.6.2 Funktionaler Bindungsassay

Zur Validierung der Korrelation zwischen Bindungsassay und funktionaler Antwort des zellulären NMDA-Rezeptors wurden für zwei Substanzen Patch-Clamp-Experimente durchgeführt (Tabelle 14). Hier zeigte sich, dass die Substanzen auch in einem Ganze-Zelle-Ansatz (*whole cell*) eine funktionale Rezeptorantwort mit vergleichbaren IC_{50} -Werten hervorriefen.

Dadurch wurde zusätzlich sichergestellt, dass es sich bei den Ergebnissen des Bindungsassays nicht um Artefakte handelte, die z.B. durch eine Farbigkeit der Azo-Verbindungen hervorgerufen wurde.

Tabelle 14: Korrelation funktionaler Assay und Bindungsassay

Verbindung	Glycin _B IC_{50} (μ M)	Patch-Clamp IC_{50} (μ M)
5 	9,53 (\pm 0,2)	6,1
8 	16,7	25,2

3.7 Etablierung eines flexiblen Rezeptormodells

3.7.1 Analyse der antagonistischen Kristallstrukturen

Die Röntgenkristallstrukturen von NR1 im Komplex mit den Antagonisten DCKA (PDB-Code 1PBQ) und Cycloleucin (1Y1M) wurden überlagert, um einen ersten Eindruck der konformellen Flexibilität der beiden Strukturen zu erhalten. Da beide Strukturen als Dimer kristallisiert wurden, standen für die Überlagerung insgesamt 4 Ketten zur Verfügung (Abbildung 34). Die Überlagerung der Kette 1 von 1PBQ und Kette 1 von 1Y1M ergab einen RMSD von 1,3 Å. Die einzelnen Monomere der jeweiligen Kristallstruktur unterschieden sich zueinander nur um 0,1 Å (1Y1M) bzw. 0,5 Å (1PBQ). Die kokristallisierten Liganden wurden nicht in die RMSD-Messung einbezogen.

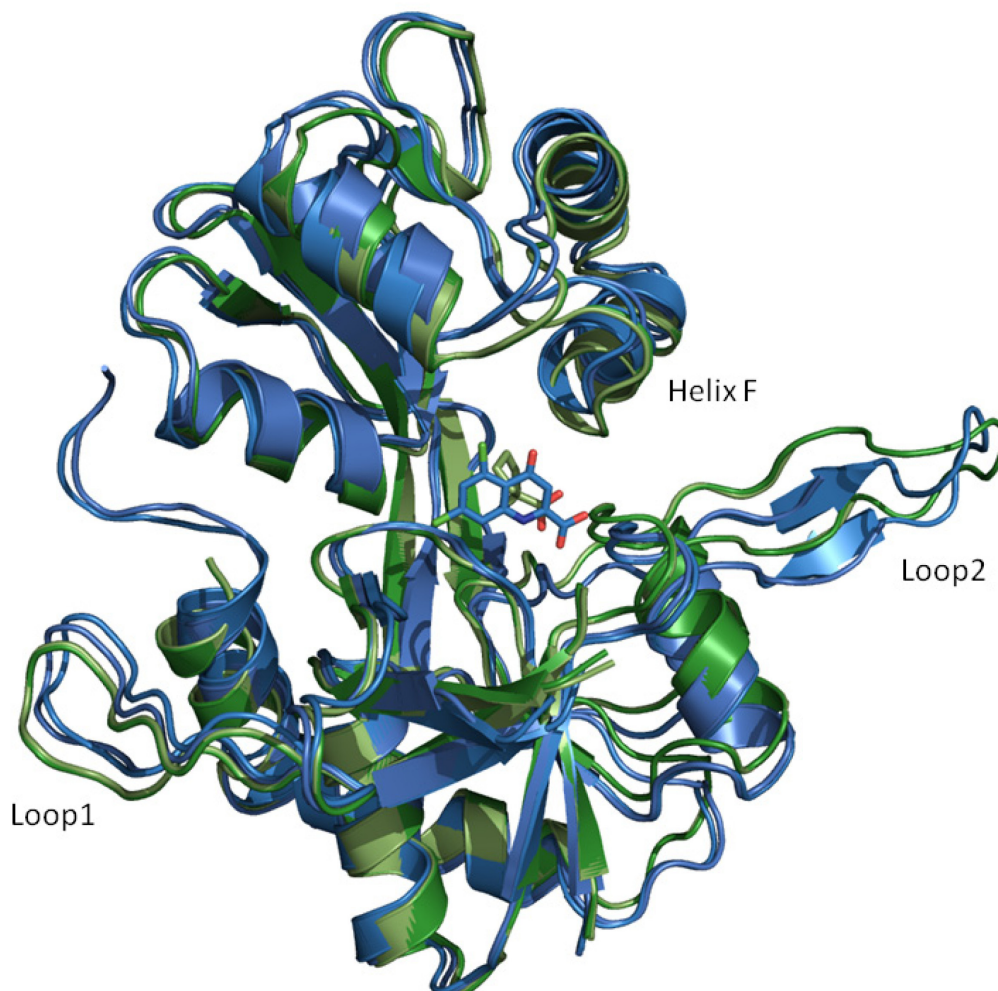


Abbildung 34: Strukturüberlagerung der antagonistischen Kristallstrukturen mit den kokristallisierten Antagonisten DCKA (blau) und Cycloleucin (grün). Überlagert wurden jeweils die beiden Ketten der Röntgenkristallstrukturen 1PBQ (Blautöne) und 1Y1M (Grüntöne).

Der mittlere Abstand der Strukturen wurde dabei hauptsächlich durch die beweglichen Loops 1 und 2 verursacht. In beiden Loops fehlte eine hohe Anzahl an atomaren Koordinaten der Reste, da diese während der Röntgenkristallisation nicht aufgelöst werden konnten. Beide Loops müssen demzufolge als hochflexibel angesehen werden, ohne dass diese Flexibilität durch ihre Distanz zur Bindungsstelle einen Einfluss auf das Bindungsverhalten hat.

Auch in der Helix F fanden sich hohe RMSD-Werte. Dies erklärt sich durch die Tatsache, dass sich hier Reste befinden, die Teil der Glycin_B-Bindungsstelle sind. Die Bindungsstelle passt sich somit nicht nur auf Ebene der einzelnen Aminosäurereste an, wie von Furukawa *et al.* beschrieben³⁷, sondern das Protein führt vielmehr in einigen Bereichen eine Umordnung seiner Backbone-Struktur durch, um auf die veränderte Größe des jeweiligen Antagonisten zu reagieren. Die genaue strukturelle Antwort der Glycin_B-Bindungsstelle auf die Bindung unterschiedlich großer Antagonisten ist bislang nicht aufgeklärt.

Eine Visualisierung der Temperaturfaktoren (B-Faktoren) der Atome der Kristallstruktur 1PBQ gemittelt über die einzelnen Reste ergab, dass sich die Flexibilität der beiden Loops 1 und 2 auch in den B-Faktoren widerspiegelt (Abbildung 35). Reste der Bindungsstelle zeigten hingegen nur einen sehr geringen mittleren Temperaturfaktor und können somit als rigide betrachtet werden. Eine Ausnahme hiervon bildeten wiederum die Reste der Helix F, deren B-Faktoren im Vergleich zu den Werten der übrigen Reste der Bindungsstelle erhöht waren.

Obwohl der B-Faktor durch seine direkte Korrelation mit dem gemittelten RMSF eines Aminosäurerestes einen ersten Hinweis auf flexible Bereiche im Protein geben kann (Kapitel 2.7.3), reicht er alleine nicht aus, um die Flexibilität des Rezeptors darzustellen. Die funktionale Antwort des Rezeptors auf die Bindung unterschiedlicher Antagonisten z.B. durch Konformationsänderungen des Protein-Backbones kann nur mit Temperaturfaktoren nicht hinreichend beschrieben werden. Aus diesem Grund wurde im Folgenden eine Methodik entwickelt, um die Konformationsänderung des Rezeptors über Zeit genauer zu quantifizieren und für das Virtual Screening und rationale Wirkstoffdesign nutzbar zu machen.

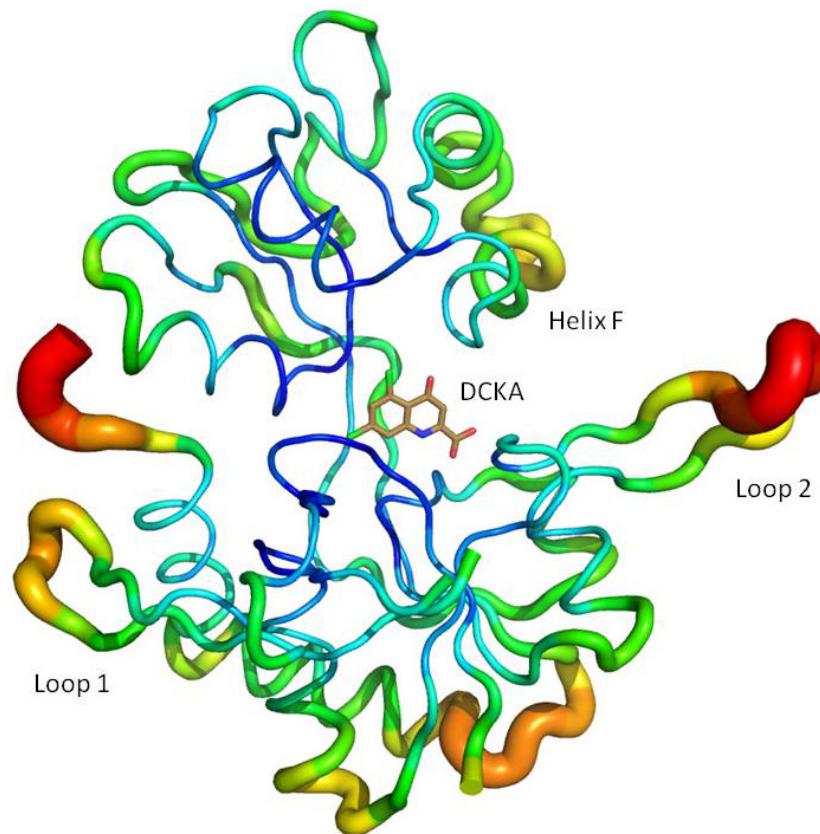


Abbildung 35: Temperaturfaktoren der Kristallstruktur 1PBQ
Visualisierung der Temperaturfaktoren von blau (niedrig) bis rot (hoch) sowie als RMSF (räumlich Ausdehnung) der jeweiligen Reste. Der Ligand DCKA ist in braun dargestellt.

3.7.2 Simulation der Flexibilität von Glycin_B

Zur genaueren Untersuchung der intrinsischen Flexibilität der Glycin_B-Bindungsstelle wurde die Kristallstruktur von NR1 zusammen mit dem kokristallisierten Antagonisten DCKA (PDB-Code 1PBQ) über einen Zeitraum von 50 Nanosekunden einer freien MD-Simulation unterworfen. Die Simulation fand dabei unter „Assay-Bedingungen“ statt, d.h. bei Raumtemperatur von 25°C, und 1 bar Atmosphärendruck. Die Rolle des Wassers wurde dabei durch das Hinzufügen von expliziten Wassermolekülen vom Typ TIP4P berücksichtigt. Die weiteren Parameter der Simulation wurden gemäß des MD-Protokolls (Kapitel 2.7.2.2) gesetzt.

Um zu überprüfen, ob die MD-Simulation in einem energetisch sinnvollen Zustand konvergierte, wurde die gesamte Energie des simulierten Systems über die Zeit in einem Graphen aufgetragen. Eine falsch aufgesetzte Simulation oder eine fehlerhafte Strukturaufbereitung des Proteins hätte einen Anstieg der gesamten Energie des Systems zur Folge gehabt. In der durchgeführten Simulation war dies allerdings nicht der Fall, die Energie des Systems verringerte sich während der 50 Nanosekunden Simulation und erreichte in den letzten fünf

Nanosekunden eine Plateauphase (Abbildung 36). Dies gilt allgemein als Indikator für eine erfolgreich durchgeführte MD-Simulation.

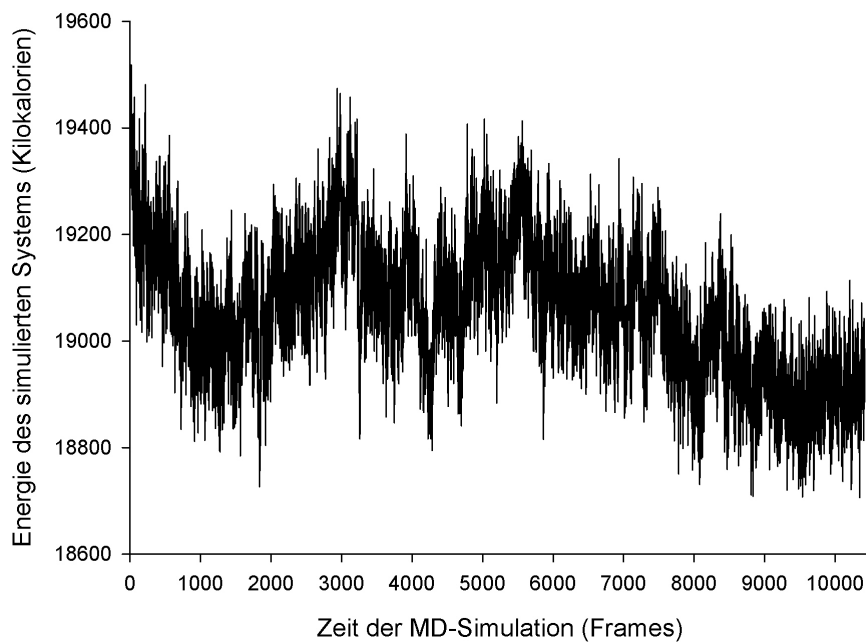


Abbildung 36: Energieverlauf der Moleküldynamik-Simulation von NR1
Freie MD-Simulation der Kristallstruktur von NR1 im Komplex mit DCKA über einen Zeitraum von 50 ns.

Die Trajektorien der einzelnen Atome des Proteins wurden über den Zeitraum der Simulation von 50 Nanosekunden alle 4,8 Pikosekunden als „Frames“ gespeichert. Somit wurden 10.419 Frames erzeugt. Für jeden dieser Frames wurde nun der RMSD aller Atome des Proteins in Referenz zum ersten Frame berechnet und über die zugehörigen Aminosäurereste gemittelt. Anschließend wurden die RMSD-Werte der Reste nochmals über alle Frames gemittelt, um so den endgültigen mittleren RMSD eines Restes über die Simulationszeit zu erhalten. Die RMSD-Werte wurden dann zusammen mit den aus der PDB-Annotation der Kristallstruktur entnommenen Temperaturfaktoren (Abbildung 37) bzw. einzeln (Abbildung 38) in einem Graphen aufgetragen.

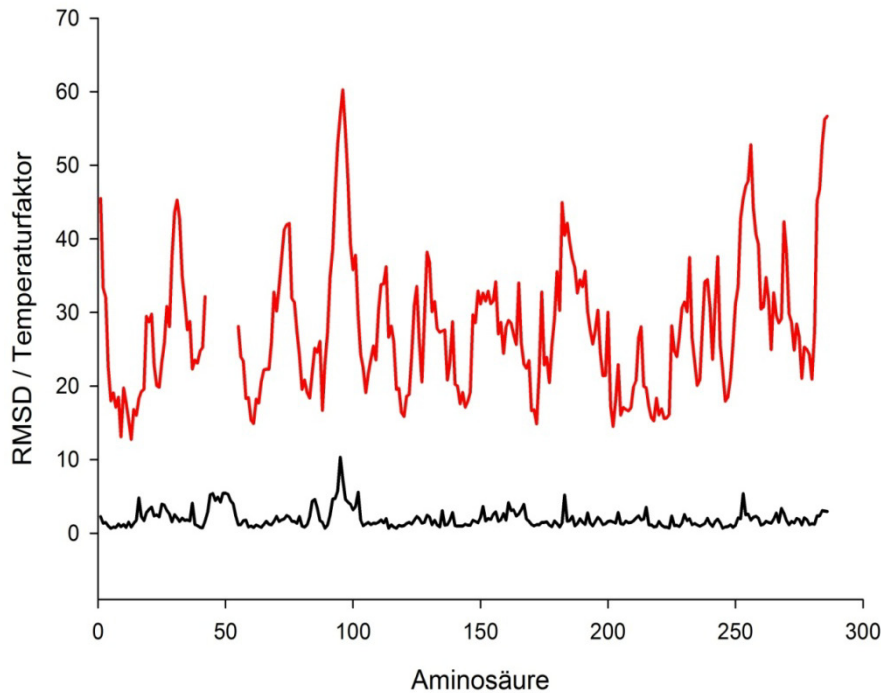


Abbildung 37: RMSD und Temperaturfaktor der Aminosäuren
Temperaturfaktoren (rot) der Aminosäuren der antagonistischen Kristallstruktur von Glycin_B sowie RMSD der Reste gemittelt über eine Dauer von 50ns während der MD-Simulation. Lücken in der Kurve der Temperaturfaktoren stammen von nicht aufgelösten Aminosäureresten der Kristallstruktur. Diese wurden für die MD-Simulation modelliert.

Es zeigte sich, dass MD-RMSD und Temperaturfaktor eines Restes in geringem direkten Zusammenhang stehen ($R^2=0,24$). Der Temperaturfaktor konnte somit die Beweglichkeit eines Restes über die Zeit und damit seine funktionale Flexibilität nicht vollständig wiedergeben. Hohe Übereinstimmung zwischen RMSD und B-Faktor ließen sich nur in zwei Bereichen erkennen: im Bereich des Loop 1 (Reste 45-50, in der Kristallstruktur nicht vollständig aufgelöst) und des Loop 2 (Reste 90-100). Da diese Loops aber mit der Glycin_B-Bindungsstelle nicht interagieren, war ihre Mobilität von eher geringer Relevanz für die funktionale Antwort des Rezeptors auf die Bindung eines Antagonisten.

Die gemittelten RMSD-Werte der Aminosäurereste wurden nun einzeln untersucht, um Reste zu identifizieren, die während der Simulation eine hohe Mobilität zeigten und sich in der Nähe der Glycin_B-Bindungsstelle befinden (Abbildung 38). Dabei zeigte sich, dass insbesondere die Reste Ser179 (RMSD 1,7 Å), Ser180 (RMSD 1,3 Å) und Ile183 (RMSD 5,2 Å) hohe RMSD-Werte erhielten. Diese befanden sich auf der durch die Domäne S2 gebildete Seite der Bindungsstelle. Reste auf der durch S1 gebildeten Seite zeigten hier geringere RMSD-Werte, was ein deutliches Indiz dafür war, dass sich diese bei Bindung eines Antagonisten nur wenig bewegen.

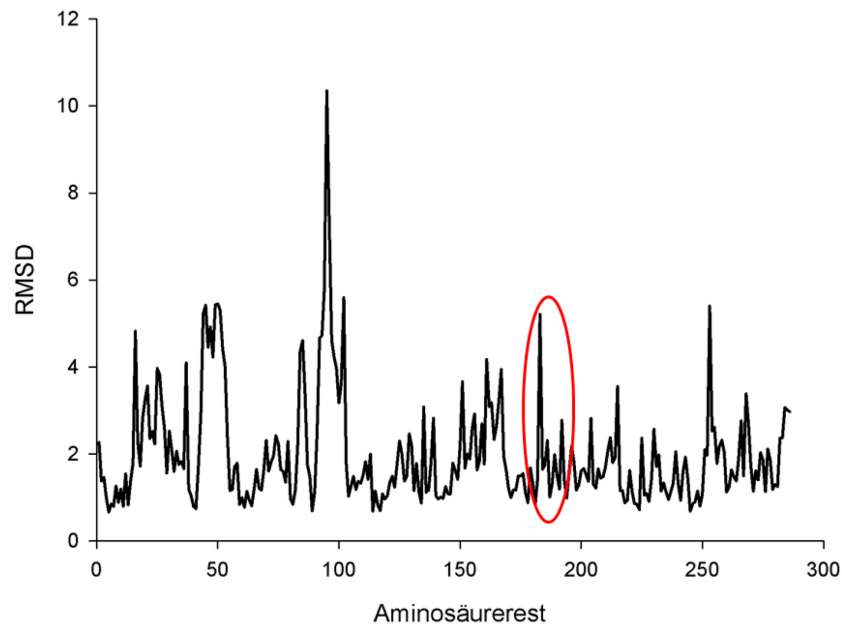


Abbildung 38: Über Zeit gemittelte RMSD-Werte

Das System der antagonistischen Kristallstruktur 1PBQ wurde für einen Zeitraum von 50 Nanosekunden frei simuliert und die über diesen zeitverlauf gemittelten RMSD-Werte der einzelnen Aminosäurereste aufgetragen. In rot markiert die als flexibel identifizierten Reste der Bindungsstelle.

Die berechneten RMSD-Werte der einzelnen Reste wurden nun als Temperaturfaktoren der Kristallstruktur visualisiert (Abbildung 39). Dabei wurde noch einmal deutlicher ersichtlich, dass die von der Domäne S2 gebildeten Reste der Helix F im Gegensatz zu allen anderen Resten der Bindungsstelle eine deutlich erhöhte Mobilität aufwiesen. Es konnten somit die am stärksten mobilitätsbestimmenden Reste der Bindungsstelle ermittelt werden. Die Reste der Helix F wurden in den folgenden erweiterten Induced-Fit-Dockings explizit als flexibel betrachtet.

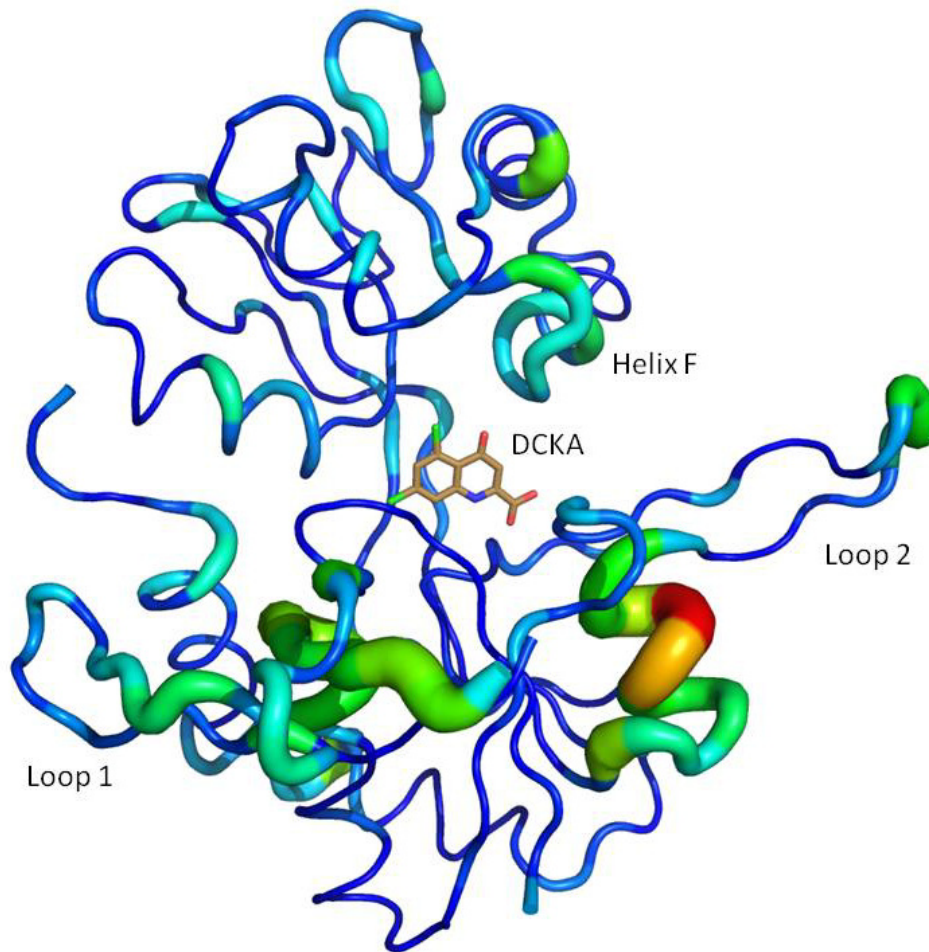


Abbildung 39: RMSD-Werte als Temperaturfaktoren
Visualisierung der aus der Moleküldynamik-Simulation gewonnenen RMSD-Werte der Reste von NR1 als Temperaturfaktoren der Proteinstruktur von blau (niedrig) bis rot (hoch).

3.8 Auswirkung des flexiblen Rezeptors auf das Wirkstoffdesign

3.8.1 Induced Fit Docking mit Standardprotokoll

Die 50 maximal diversen nanomolaren Glycin_B-Antagonisten wurden mittels des Maximum-Dissimilarity-Algorithmus (FCFP6-Deskriptor, Tanimoto-Ähnlichkeit, Pipeline Pilot) aus den 419 nanomolaren Referenzstrukturen identifiziert. Ihre isomeren Formen wurden bei pH 7±3 berechnet und mittels des Standard-IFD-Protokolls (Seite 45) in die Kristallstruktur von NR1 (PDB-Code 1PBQ) gedockt. Hierbei wurde ein Constraint auf die Interaktion (Wasserstoffbrückenbindung resp. ionische Interaktion) mit Arg523 gesetzt. Während des Induced-Fit-Dockings wurde zur impliziten Realisierung von Rezeptorflexibilität eine van-der-Waals-Skalierung von bis zu 50% erlaubt. Pro Ligand wurden maximal 20 Dockingposen berechnet und mittels Glide SP der jeweilige Dockingscore berechnet. Alle Reste im Radius von 12 Å um das geometrische Mittel von DCKA wurden auf Interaktionen mit dem zu dockenden Liganden untersucht.

Insgesamt wurden so 887 Dockingposen generiert. Die paarweise Distanzmatrix wurde anhand ihres jeweiligen RMSD über das gesamte Protein berechnet. Die Proteinstruktur mit dem geringsten RMSD zu allen anderen Strukturen wurde als durchschnittliche Proteinstruktur identifiziert. Die Spanne der RMSD-Werte der einzelnen Proteinstrukturen reichte dabei von 0,18 bis 0,38 Å. Die mittlere IFD-Proteinstruktur wurde mit der ursprünglichen Kristallstruktur von 1PBQ überlagert. Dabei ergab sich ein RMSD von 0,26 Å. Da der RMSD über das gesamte Protein errechnet wurde, allerdings nur Reste in direkter Nähe zum gebundenen Liganden während des IFD im Kraftfeld energieminiert wurden, erklärt sich der geringe RMSD, obwohl sich die Reste in der Glycin_B-Bindungsstelle deutlich in ihrer Anordnung unterscheiden (Abbildung 40). Relevante Fluktuationen waren v.a. an Phe484 („Antagonismus-Sensor“ der Bindungsstelle nach Furukawa³⁷ *et al.*) und Arg131 (affinitätsbestimmende Interaktion zwischen Ligand und Rezeptor über ionische Wechselwirkungen bzw. Wasserstoffbrückenbindungen) zu bemerken. Änderungen im Protein-Backbone waren nicht zu erkennen.

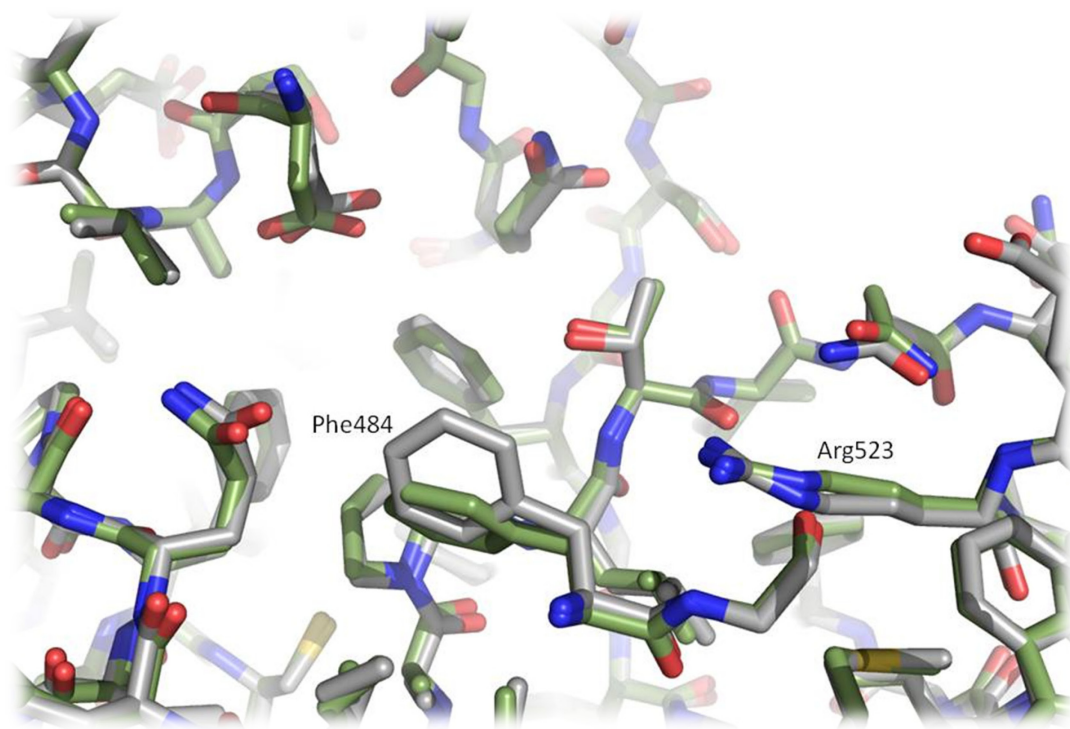


Abbildung 40: Fluktuation Kristallstruktur - mittlere IFD-Struktur

Die Induced-Fit-Struktur des „Kernbindungsgebietes“ von Glycin_B (grau) wurde über 887 Dockingposen gemittelt. Die Konformation der Kristallstruktur von Glycin_B ist in grün dargestellt. Während des Induced-Fit-Dockings wurden nur Reste in direkter Nachbarschaft zum Ligand energieminiert (Standard-IFD-Protokoll). Dadurch wurde nur eine vergleichsweise geringe Anpassung der Rezeptorstruktur an den Liganden ermöglicht.

Die automatische Erkennung zu mutierender Reste während des Induced-Fit-Dockings durch den Schrödinger Induced-Fit-Docking Workflow konnte keine geeigneten Reste identifizieren. Dieser Mechanismus analysierte Reste der Bindungsstelle im Abstand von bis zu 5 Å um den kokristallisierten Liganden. Die relevanten Reste für größere Liganden befanden sich jedoch in größerer Entfernung als fünf Å.

Dies hat einen Einfluss auf die Qualität der Dockingergebnisse: Nur für 19 der 50 maximal diversen nanomolaren Glycin_B-Referenzliganden konnten mit dem Standard-IFD-Protokoll Dockingposen erzeugt werden. Die verbleibenden 31 nanomolaren Antagonisten wurden verworfen. Eine Analyse der molekularen Masse der gedockten und nicht gedockten Liganden ergab eine klare Separierung: Das Standard-IFD-Protokoll war nur in der Lage, eher DCKA-ähnliche Liganden mit geringer molekularer Masse in die Bindungsstelle einzubringen (Abbildung 41). Größere Liganden fanden offenbar nicht genug Platz.

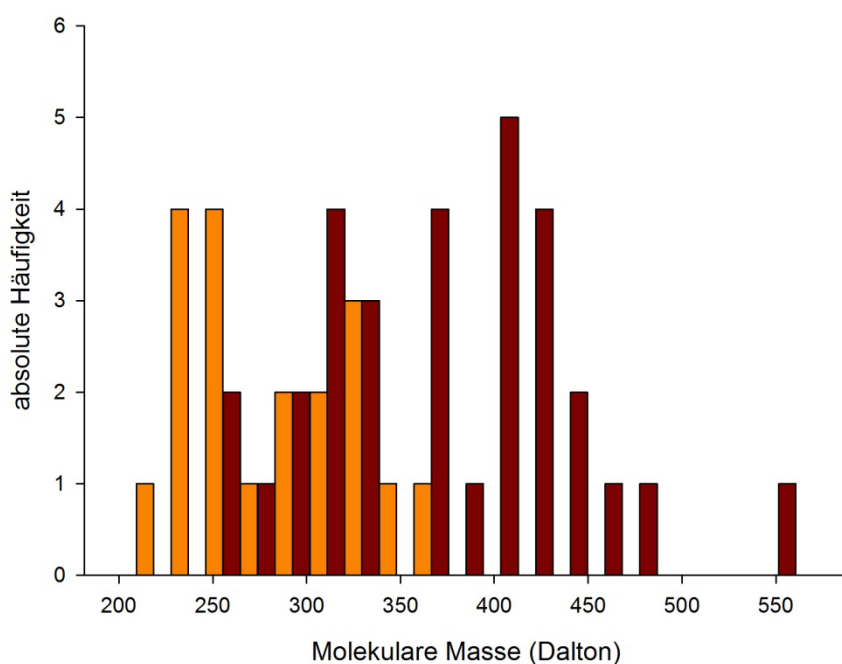


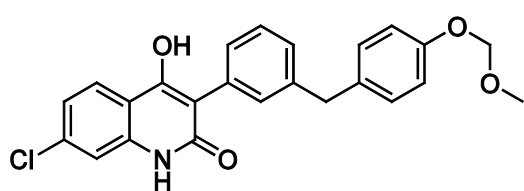
Abbildung 41: Histogramm der Molekülmassen der IFD-Ergebnisse. Gedockte (orange) und nicht gedockte Liganden (rotbraun) im Induced-Fit-Docking mittels IFD-Standardprotokoll

3.8.2 Induced Fit Docking mit erweitertem Protokoll

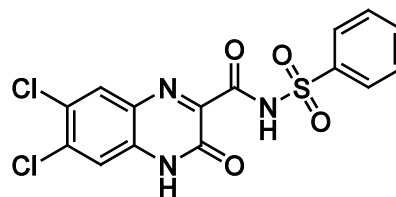
Zur Verringerung der benötigten Rechenzeit (für jede erzeugte Dockingpose musste eine vollständige Energieminimierung des Proteins durchgeführt werden) wurden vier mittels IFD-Standardprotokoll nicht gedockte Referenzliganden visuell selektiert (Tabelle 15). Allen diesen Referenzliganden war gemein, dass ihre Seitenkettensubstitution am Doppelringssystem nur reduzierte Frei-

heitsgrade aufwiesen, d.h. die Linker (Sulfonamid, Ester, etc.) waren eher rigide. Gleichzeitig verfügten sie über einen DCKA-ähnlichen Scaffold, so dass die Positionierung der Liganden in der Bindungsstelle einen DCKA-ähnlichen Bindungsmodus einnehmen sollte.

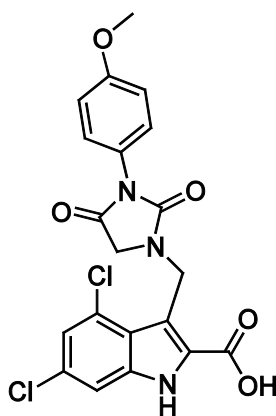
Tabelle 15: Referenzstrukturen für das Induced Fit Docking mit erweitertem Protokoll
Diese konnten zusammen mit 27 weiteren Substanzen nicht erfolgreich mit dem IFD-Standardprotokoll gedockt werden.



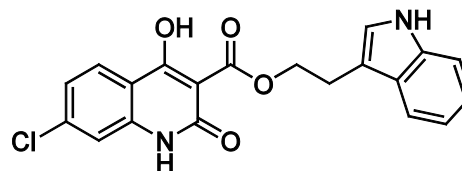
73
 $IC_{50} = 0,0022 \mu\text{M}^{200}$



74
 $IC_{50} = 0,12 \mu\text{M}^{201}$



75
 $IC_{50} = 0,1 \mu\text{M}^{202}$



76
 $IC_{50} = 0,43 \mu\text{M}^{203}$

Die Liganden-Rigidität stellte für das Induced-Fit-Docking eine größere Herausforderung dar, da der Ligand vom Dockingalgorithmus nicht mehr in eine falsch-positive Bindungspose positioniert werden kann, um dadurch Clashes mit dem Rezeptor zu vermeiden. Die Rigidität der Liganden erforderte somit eine korrekte Implementierung des Induced Fit-Effektes, da nur dadurch eine Bindungspose erzeugt werden konnte.

Während der Berechnung der isomeren Formen der Liganden im Bereich von $\text{pH } 7 \pm 3$ entstanden für Struktur **74** zwei isomere Formen durch Keto-Enol-Tautomerie am Ring. Insgesamt wurden so fünf Strukturen in den folgenden erweiterten Induced-Fit-Dockingexperimenten untersucht.

Zum besseren Verständnis des möglichen gemeinsamen Bindungsmodus wurden die fünf Liganden in MacroModel energieminiert (OPLS2005 Kraftfeld, implizites Wasser, Polak-Ribiere konjugierte Gradientenverfahren)¹⁴³

(PRCG) mit 1000 Iterationen und einem Konvergenzlimit von 0,01 Å) und rigide aligniert (Abbildung 42). Die energieminierten Konformationen wurden automatisiert überlagert, indem die jeweilige Kernstruktur (aromatisches Zweiringsystem) zur Deckung gebracht wurde. Die hohe Ähnlichkeit der Kernstrukturen zur antagonistischen Referenzsubstanz DCKA sprach dafür, dass alle Substanzen die Bindungsstelle von Glycin_B gleichermaßen adressieren. Die diversen, aber rigiden Seitenketten sollten dadurch mit hoher Wahrscheinlichkeit die gleichen Reste in der Umgebung der Kernbindungsstelle adressieren.

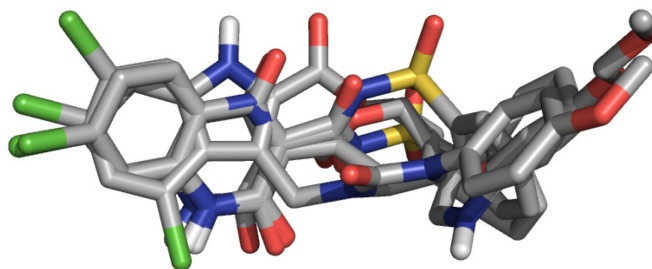


Abbildung 42: Alignment der Referenzliganden für das erweiterte Induced-Fit-Docking

Um diese Hypothese weiter zu verifizieren wurde ein Induced-Fit-Docking mit erweitertem Protokoll durchgeführt. Hierbei wurden nun die aus der MD-Simulation als mobil identifizierten Reste der Glycin_B-Bindungsstelle während des Dockings zu Alanin mutiert. Die Liganden hatten damit mehr Raum in der Bindungsstelle, um ihre bioaktive Konformation einzunehmen. Anschließend wurden die Reste der Bindungsstelle zu ihren ursprünglichen Aminosäuren zurückmutiert, und die neue Proteinkonformation energieminiert.

Mutiert wurden die Reste Ser179, Ser180, Val181, Asp182 und Ile183 der Helix F (Abbildung 43). Um die Erzeugung „sinnvoller“ Dockingposen zu forcieren, wurde ein Dockingconstraint auf die polare Interaktion mit Arg131 definiert. Im Gegensatz zum Standard-IFD-Protokoll ohne Mutation der Bindungsstelle konnten durch Anwendung des erweiterten IFD-Protokolls alle fünf Referenzstrukturen mit sinnvollen Bindungsposen (Interaktion mit Arg131, Einnahme des gleichen Kernbindungsbereiches wie DCKA durch die jeweiligen aromatischen Kernstrukturen) gedockt werden. Insgesamt wurden 80 gültige Posen berechnet. Von diesen wurde die paarweise Distanzmatrix anhand des durchschnittlichen RMSD über alle Reste gebildet. Als durchschnittliche Induced-Fit-Rezeptorstruktur wurde die Proteinstruktur mit dem geringsten gemittelten RMSD zu allen anderen Strukturen definiert. Die RMSD-Werte variierten dabei zwischen 0,16 Å und 0,40 Å. Die durchschnittliche Rezeptorstruktur wurde mit der antagonistischen Kristallstruktur von NR1 (1PBQ) überlagert, um die Veränderungen in der Bindungsstelle zu identifizieren. Der RMSD zwischen den beiden Strukturen betrug 0,26 Å.

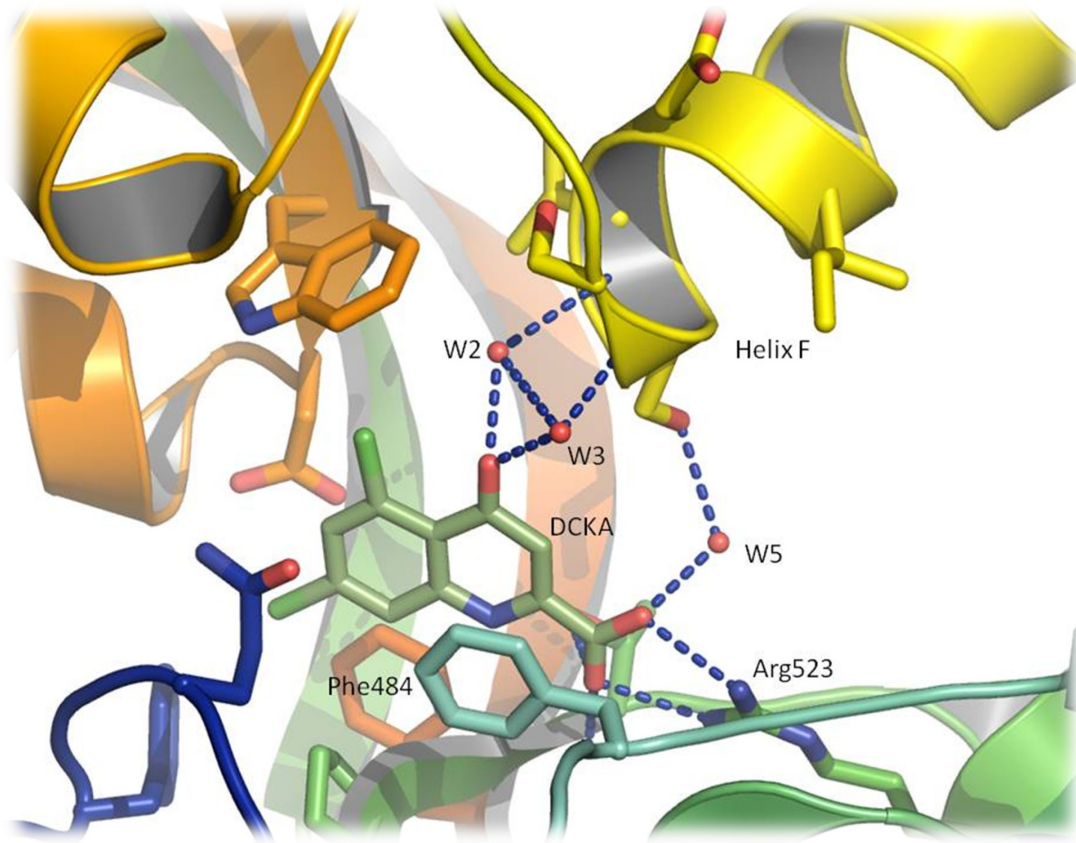


Abbildung 43: Position der temporär mutierten Resten der Helix F. Helix F dargestellt in Gelb. Die mit DCKA (grün) interagierenden Wasser der Kristallstruktur wurden ebenfalls eingezeichnet (rot). Diese vermitteln drei Wasserstoffbrückenbindungen (blaue Linien) zwischen DCKA und Helix F.

Zur Quantifizierung der strukturellen Änderungen des Rezeptors während des erweiterten IFD-Protokolls wurden alle 80 erzeugten Rezeptorkonformationen überlagert (Abbildung 44). Hierbei zeigte sich, dass die Mutation der Reste Ser179, Ser180, Val181 und Ile183 während des ersten Schritts des erweiterten IF-Dockings eine Relaxierung der der Bindungsstelle zugewandten Seite der Helix F ermöglicht. Die Veränderung des Protein-Backbones an dieser Stelle ist offenbar notwendig, um diese Referenzliganden binden zu lassen. Weiterhin konnte beobachtet werden, dass auch zwei Loops an der durch Domäne S1 gebildeten Seite der Bindungsstelle unterschiedliche Konformationen auswiesen. Hier fanden direkte Interaktionen mit dem aromatischen Kerngerüst der Liganden statt, der Rezeptor passte sich somit auch hier an die Form des gebundenen Liganden an.

Eine Ausweitung der zu mutierenden Reste auf den Bereich Gln178 bis Phe185 (acht der Bindungsstelle von DCKA zugewandte Reste der Helix F) zeigte keine Veränderung der erzeugten Rezeptorstrukturen. Die zusätzlichen Reste liegen offenbar zu weit von den mit den Liganden interagierenden Resten der Helix F entfernt, um konformelle Änderungen zu erfahren.

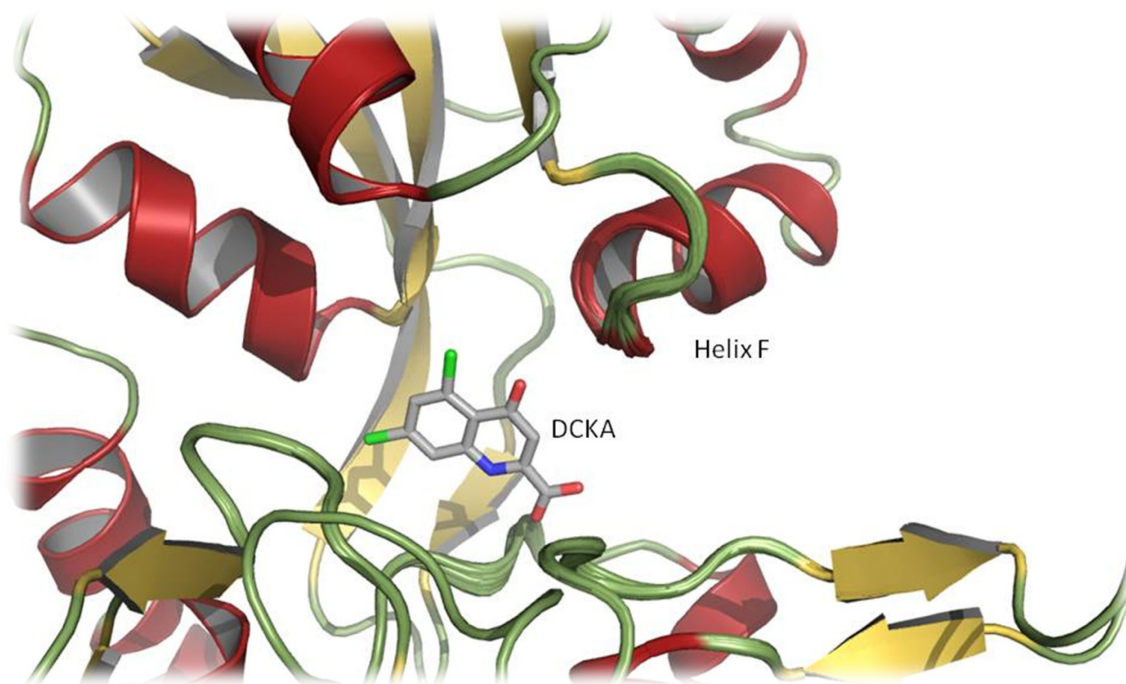


Abbildung 44: IFD-Rezeptorkonformationen der nanomolaren Literaturliganden
Überlagerung von 80 anhand des erweiterten IFD-Protokolls berechneten Rezeptorkonformationen. Die Mutation der Reste Ser179, Ser180, Val181, Asp182 und Ile183 ermöglicht die Generierung von Dockingposen, die essentielle Interaktionspunkte im Rezeptor adressieren. Ohne die Mutation dieser Reste (Standard-IFD-Protokoll) konnten für die betrachteten nanomolaren Referenzliganden keine Bindungsposen erzeugt werden. Der Antagonist DCKA (grau) wurde mit seinen ursprünglichen Koordinaten aus der Kristallstruktur eingefügt, um das Ausmaß der Erweiterung der Bindungsstelle darzustellen.

3.8.3 Induced-Fit-Rezeptorkonformationen der Azo-Liganden

Zur Analyse des potentiellen Bindungsmodus der im Virtual Screening identifizierten Azo-Verbindungen wurden alle Strukturen ausgewählt, die einen durch Dosis-Wirkungskurven bestätigten niedrig-mikromolaren IC_{50} aufwiesen. Hierbei handelte es sich um die Verbindungen **3**, **5**, **24 – 30**, **32** und **34**. Als zusätzliche Referenzen wurden für die niedrig-nanomolaren Virtual Screening Hits **1** ($IC_{50} = 0,014 \mu\text{M}$) und **2** ($IC_{50} = 0,9 \mu\text{M}$) ebenfalls Bindungsposen anhand des erweiterten IFD-Protokolls berechnet.

Wie bereits beim Docking der nanomolaren Referenzverbindungen mit erweitertem IFD-Protokoll zeigten sich in der Serie der Azo-Verbindungen deutliche Fluktuationen in der Konformation des Proteinbackbones der Helix F sowie in den beiden Loops der Domäne S2, die im direkten Kontakt mit dem aromatischen Kerngerüst standen (Abbildung 45). Aus den Induced-Fit-Rezeptorkonformationen wurde eine mittlere Rezeptorkonformation identifiziert und alle paarweisen Distanzen der Induced-Fit-Konformationen zu dieser berechnet. Die RMSD-Werte bewegten sich zwischen 0,14 und 0,53 Å. Der RMSD zwischen der mittleren IFD-Konformation und der Kristallstruktur von NR1 hingegen betrug nur 0,09 Å. Die Reste der Glycin_B-Bindungsstelle fluktuieren somit

während des Induced-Fit-Dockings in ihrer Position um die röntgenkristallographische Rezeptorkonformation von NR1 mit DCKA. Dies bedeutet auch, dass es im Falle von Glycin_B nicht ausreicht, den Durchschnitt der IFD-Strukturen zu ermitteln und dann als Rezeptorstruktur für rigides Docking zu verwenden (Prinzip des Ensemble-Dockings). Jeder Ligand erzeugte vielmehr seine eigene Induced-Fit-Proteinkonformation.

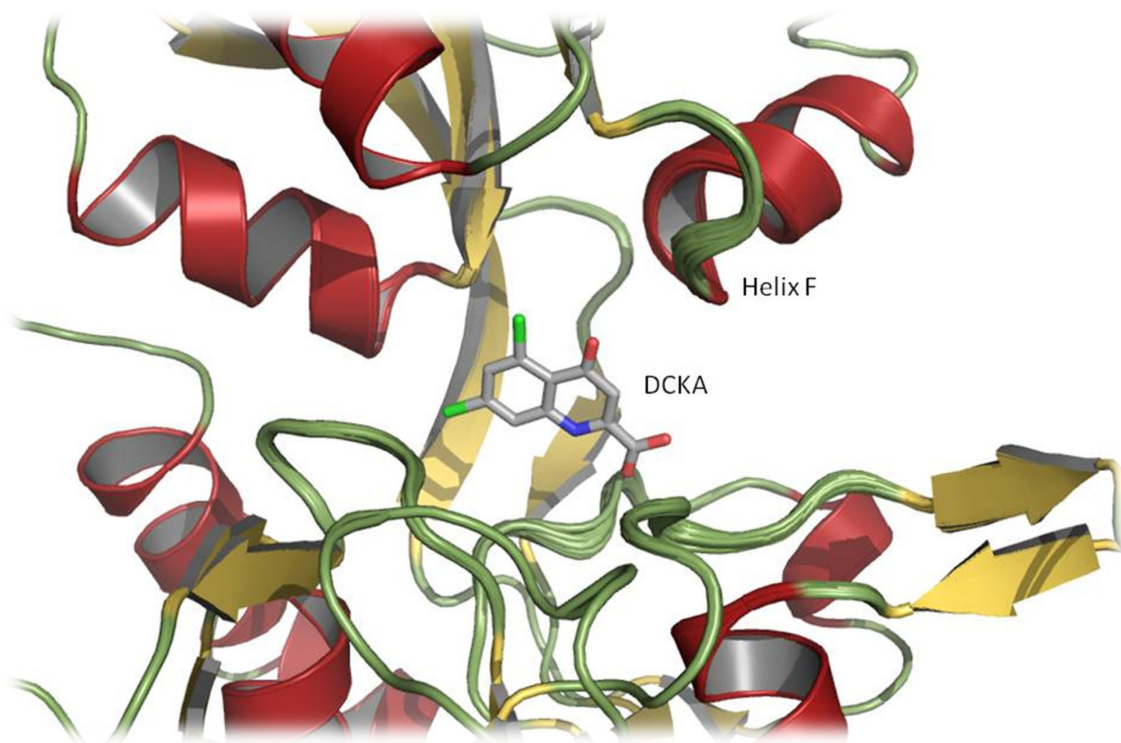


Abbildung 45: Induced-Fit-Rezeptorkonformationen der Azo-Liganden
Überlagerung von 250 anhand des erweiterten IFD-Protokolls berechneten Rezeptorkonformationen der niedrig-mikromolaren Azo-Verbindungen aus dem Virtual Screening
Während des IFD-Protokolls wurden die Reste Ser179, Ser180, Val181, Asp182 und Ile183 zu Alanin mutiert. Der Antagonist DCKA (grau) wurde wiederum mit seinen ursprünglichen Koordinaten aus der Kristallstruktur eingefügt, um das Ausmaß der Erweiterung der Bindungsstelle darzustellen

Es wurde zudem deutlich, dass der Dockingscore (hier: IFDScore, siehe Kapitel 2.7.1) kein Maß für die experimentell ermittelte Bindungsaffinität der jeweiligen Liganden ist. „Sinnvolle“ Dockingposen der niedrig-nanomolaren Referenzliganden, d.h. Posen, welche die gesetzten Dockingconstraints sowie den Kernbindungsmodus des Referenzantagonisten DCKA erfüllten, wurden anhand ihres Scores gleich oder teilweise schlechter eingeordnet als die nur niedrig-mikromolaren Liganden (drei Größenordnungen Unterschied in der Bindungsaffinität). Generell waren die Unterschiede in den Dockingscores nur gering, N=250 unterschiedliche Dockingposen der 13 Liganden wurden mit einem durchschnittlichen IFDScore von -572,17 mit einer Standardabweichung von nur 3,35 bewertet, d.h. das Scoring war nicht in der Lage, nanomolare von mikromolaren Liganden zu separieren.

In einer letzten Betrachtung der Induced-Fit-Bindungsmodi wurde überprüft, ob eine neuartige Scoringmethode besser geeignet ist, die Bindungsaffinität der Liganden genauer zu bewerten und damit mit hoher Wahrscheinlichkeit korrekte Bindungsposen von falsch-positiven zu unterscheiden.

3.8.4 Thermodynamikänderungen durch Wasserverdrängung als Scoringfunktion

Um zu analysieren, ob durch ein „thermodynamisches Scoring“ bessere Ergebnisse beim Bewerten der zuvor berechneten Induced-Fit-Dockingposen erzielt werden können, wurden die entropischen und enthalpischen Energieterme der Wasserverdrängung bei der Bindung neuer Liganden analysiert. Die Generierung der Wasserkarte folgte dabei dem zuvor beschriebenen Protokoll (Kapitel 2.7.4). Die Wasserkarte wurde in der Kristallstruktur von NR1 mit DCKA (PDB-Code 1PBQ, Auflösung 1,9 Å) erzeugt (Abbildung 46). Anhand dieser Karte wurden dann die Energieterme für alle auf dieser Proteinstruktur basierenden Dockingposen berechnet.

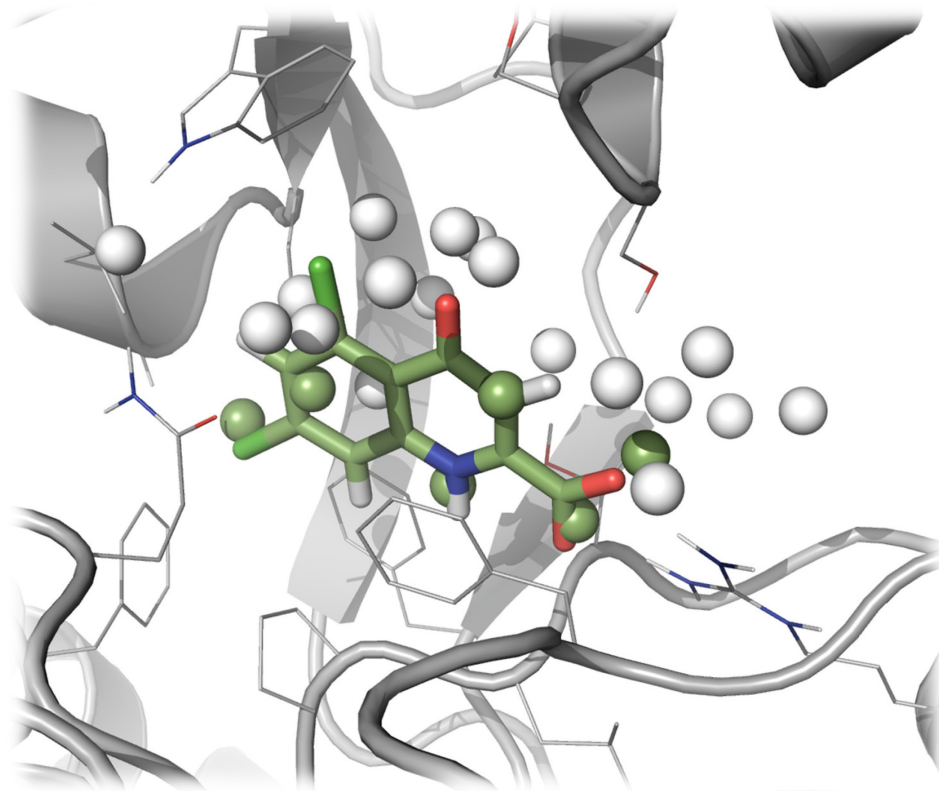


Abbildung 46: Watermap der Glycin_B-Bindungsstelle
Weiße Kugeln: Hydratationsstellen mit besonders hohen entropischen oder enthalpischen Energietermen. Grüne Kugeln: vom Antagonisten DCKA (grün) verdrängte Wasser auf energetisch ungünstigen Hydratationsstellen

Alle 250 aus dem erweiterten Induced-Fit-Docking stammenden Bindungsposen der 11 Azo- und zwei nanomolaren Analoga von Referenzsubstanzen wurden nun auf eine Korrelation zwischen dem experimentell ermittelten IC_{50} und der berechneten Änderung der Gibbs-Energie untersucht. Dazu wurde anhand der Wasserkarte von Glycin_B für jede Bindungspose die Terme der Entropieänderung ($-T\Delta S$), der Enthalpieänderung (ΔH) sowie der Änderung der Gibbs-Energie (ΔG) berechnet.

Die Korrelation zwischen den experimentell ermittelten IC_{50} -Werten und der berechneten Änderung des entropischen Terms der freien Bindungsenergie erwies sich als eher schwach ($R^2=0,53$) (Abbildung 47). Ein direktes Verhältnis aus berechnetem Entropiegewinn und tatsächlicher Bindungsaffinität konnte somit nicht angenommen werden. Zudem zeigten die berechneten Energiewerte eine zum Teil hohe Standardabweichung, so dass innerhalb der Serie der niedrig-mikromolaren Glycin_B-Antagonisten keine Unterscheidung möglich war. Auch die Änderungen der Enthalpie ($R^2=0,35$) und Gibbs-Energie ($R^2=0,03$) korrelierten schlecht bis gar nicht mit den experimentell ermittelten IC_{50} -Werten. Insgesamt betrachtet konnte der Watermap-Score für die betrachteten Hits aus dem Virtual Screening keine zuverlässige Vorhersage der Bindungsaffinität leisten. Der auf dem Entropieterm basierende Term der Scoringfunktion beschreibt die Affinität der betrachteten Glycin_B-Liganden mit Azo-Scaffold allerdings besser als die zuvor verwendete Scoringfunktion IFDScore.

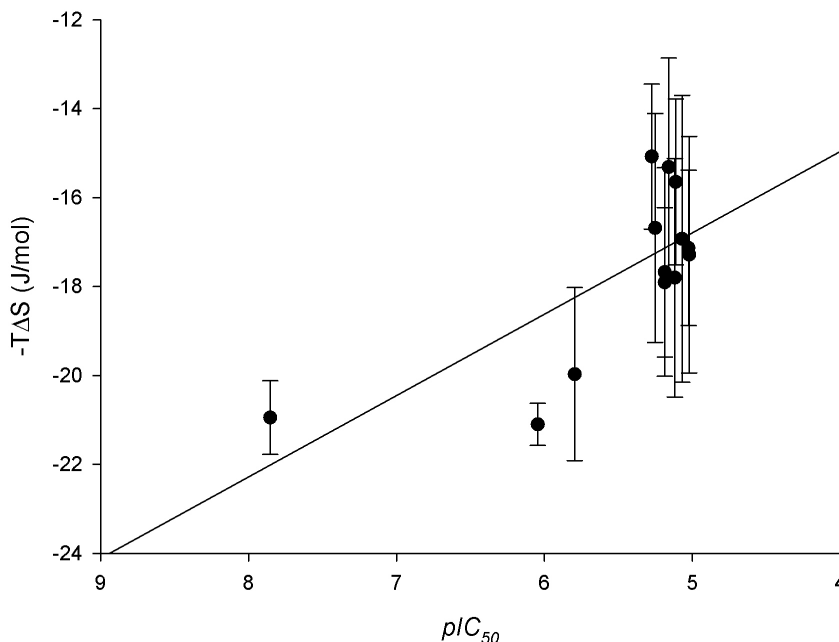


Abbildung 47: Watermap-Scoringfunktion: Korrelation Entropiegewinn – pIC_{50}
 Berechneter Entropiegewinn anhand Verdrängung von Wasser aus der Bindungsstelle durch den Liganden. Entropie und experimenteller IC_{50} korrelieren mit $R^2 = 0,52$. Die Entropiewerte wurden aus jeweils 20 mit erweitertem IFD-Protokoll berechneten Bindungsposen der ausgewählten Azoliganden sowie zwei nanomolarer Hits aus dem Virtual Screening gemittelt (Fehlerbalken: Standardabweichung).

Um zu überprüfen, ob diese Beobachtung auf die Summe aller Glycin_B-Antagonisten zutrif, wurden alle 754 Literaturverbindungen mit annotierten IC_{50} anhand des im Kapitel 2.6.3 beschriebenen Protokoll gedockt. Die so erzeugten Dockingposen wurden sowohl mittels der Glide XP-Scoringfunktion als auch anhand des Entropieterms des Watermap-Scores bewertet. Glide XP-Scoringfunktion und experimentell ermittelter pIC_{50} aller Literaturverbindungen korrelierten dabei nicht ($R^2=0,02$). Die Scoringfunktion des Dockingalgorithmus‘ stellt somit für Glycin_B kein hinreichendes Kriterium dar, um Aussagen über die tatsächliche Bindungsaffinität von Dockingposen unterschiedlich aktiver Liganden zu treffen. Als genereller diskriminierender Faktor zwischen aktiven und inaktiven Molekülen hingegen waren sie erfolgreich anwendbar (Kapitel 3.4.1).

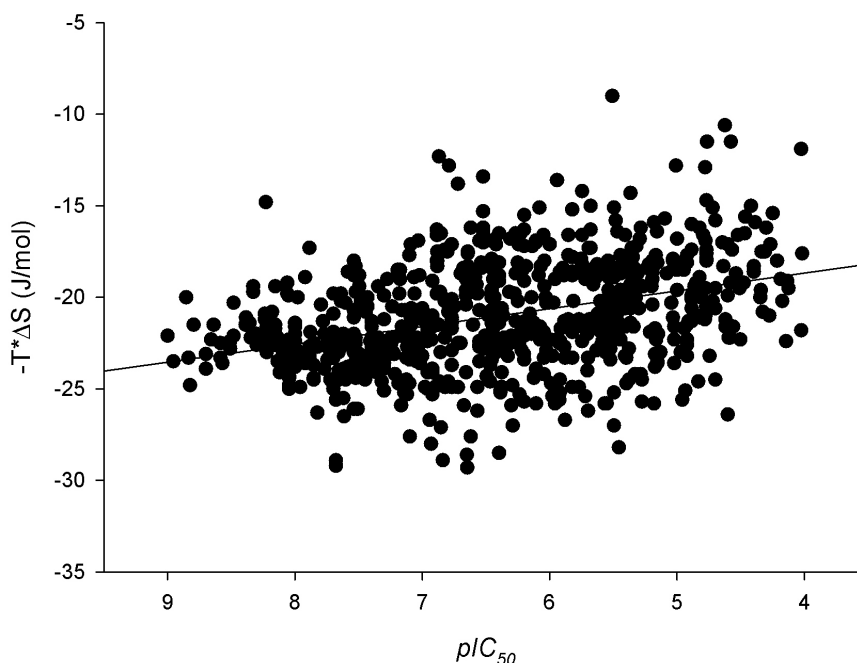


Abbildung 48: Korrelation Entropieterm (Watermap-Score) und pIC_{50}
Alle 754 Referenzliganden wurden rigide gedockt (GlideXP). Die Korrelation Watermap-Score (Entropieterm) – pIC_{50} ist nur schwach ($R^2=0,14$), allerdings stärker als die Korrelation zwischen Dockingscore und pIC_{50} ($R^2=0,02$)

Das Verhältnis zwischen Entropieterm des Watermap-Score und experimentell ermittelten IC_{50} zeigte eine um den Faktor 7 verbesserte, aber insgesamt immer noch unzureichende Korrelation ($R^2=0,14$) (Abbildung 48). Dies ist möglicherweise damit zu begründen, dass das Docking aufgrund der für alle 754 Glycin_B-Antagonisten benötigten extrem hohen Rechenkapazitäten für Induced-Fit-Dockings nur im quasi-flexiblen Rezeptor durchgeführt werden konnte. Für jede der im Induced-Fit-Docking erzeugten Rezeptorkonformation hätte dann eine eigene Watermap berechnet werden müssen, um vergleichbare

Ergebnisse zu erzeugen. Hier waren die zur Verfügung stehenden Rechenkapazitäten der limitierende Faktor.

Es konnte somit gezeigt werden, dass sowohl Induced-Fit-Docking (Kapitel 3.8.2) als auch Watermap-Scoring die Korrelation zwischen experimentell ermittelter Bindungsaffinität und *in silico* berechneten Affinitäten im Vergleich zu klassischen Dockingscores geringfügig verbessern. Für den Einsatz des Watermap-Scores als QSAR-Tool sind jedoch Rechenzeiten notwendig, die noch jenseits des vertretbaren Aufwandes für Virtual Screening liegen. Da aber erst dann wirklich überprüft werden kann, ob der Watermap-Score tatsächlich quantitativ angewandt werden kann, ist sein Einsatz ähnlich der Free-Energy-Perturbation aktuell auf kongenerische Serien von Liganden beschränkt, die einen identischen Bindungsmodus einnehmen.

3.8.5 Hypothetischer Bindungsmodus der Azo-Liganden unter Berücksichtigung der Wasserverdrängung

Um die Unterschiede im Bindungsverhalten der beiden Liganden **1** ($IC_{50} = 0,014 \mu\text{M}$) und **24** ($IC_{50} = 1,6 \mu\text{M}$) zu untersuchen, wurden ihre durch den Watermap-Score am besten bewerteten Induced-Fit-Dockingposen analysiert (Abbildung 49).

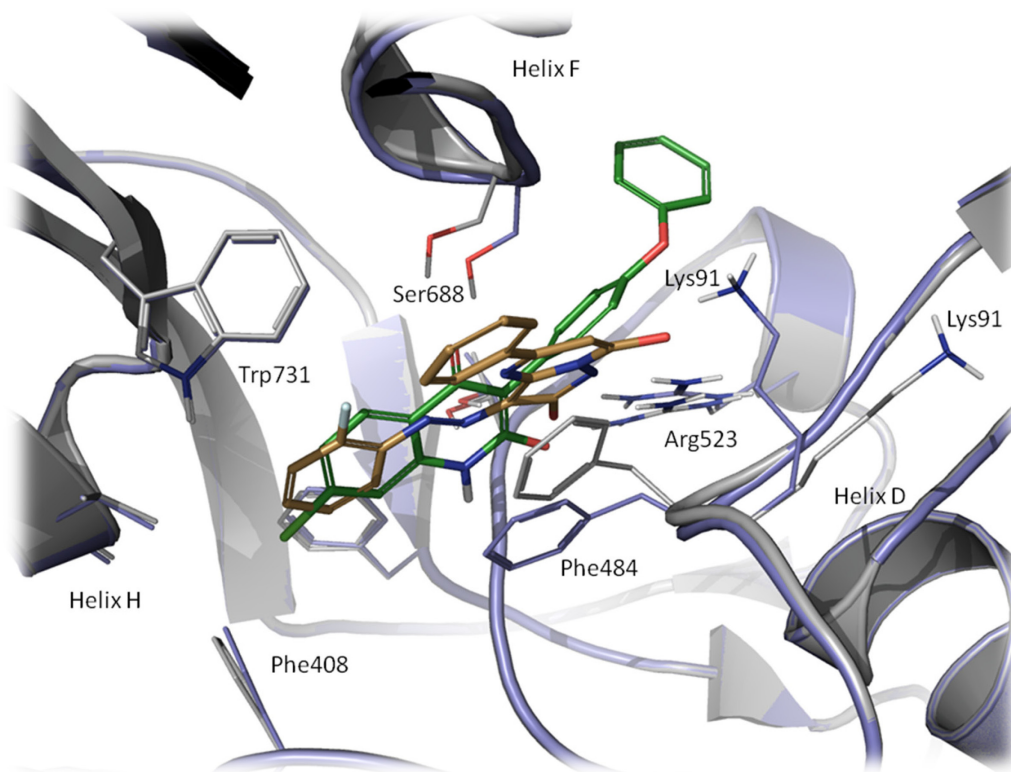


Abbildung 49: Hypothetischer Bindungsmodus der Hits **1** und **24**. Ligand **1** (nanomolar, grün, violetter Rezeptor) und **24** (mikromolar, hellbraun, grauer Rezeptor).

Auch hier konnten flexible und rigide Bereiche in der Glycin_B-Bindungsstelle identifiziert werden. Während die für die antagonistische Wirkung von Liganden essentiellen aromatischen Reste Trp731 und Phe408 keine nennenswerte Mobilität zeigten, wurde ein deutlicher Unterschied beim „Volumensensor“ Phe484 erkennbar. Wie bereits in den beiden antagonistischen Kristallstrukturen passte er sich auch hier der Position des aromatischen Kerngerüsts der Liganden an. Ebenso änderten die essentiellen polaren Reste Arg523 und Ser688 ihre Orientierung, um Wasserstoffbrückenbindungen resp. ionische Wechselwirkungen mit dem Liganden auszubilden. Während Ser688 in der Kristallstruktur 1PBQ noch eine über das Wasser W5 vermittelte Wasserstoffbrückenbindung mit DCKA eingeht, ist diese Interaktion aufgrund der erhöhten Molekülgröße der beiden Liganden **1** und **24** unvermittelt (beide verdrängen das Kristallwasser W5 - Abbildung 5). Ein ersichtlicher Unterschied tritt am Rest Lys91 (Nomenklatur nach Kristallstruktur) auf. Dieser Rest geht in der Kristallstruktur keine Interaktion mit DCKA ein, da er zu weit entfernt ist. Auch mit Ligand **24** findet keine Interaktion statt, da keine polaren Wechselwirkungen ausgebildet werden können. Ligand **1** hingegen geht eine ionische Wechselwirkung mit Lys91 ein. Bereits diese Interaktion könnte den Unterschied in der Bindungsaffinität beider Liganden erklären. Die übrigen Interaktionen der Liganden mit dem Rezeptor sind identisch mit dem zuvor beschriebenen antagonistischen Bindungsmodus (Kapitel 1.3.2)

Desweiteren wurde überprüft, ob beide Liganden eine unterschiedliche Anzahl von „virtuellen Kristallwassern“ (Watermap-Hydratationsstellen) verdrängen und somit eine zusätzliche Erklärung für die unterschiedlichen Bindungsaffinitäten der beiden Substanzen geben. Es stellte sich heraus, dass Ligand **1** zehn Watermap-Hydratationsstellen verdrängte (Abbildung 50), Substanz **24** jedoch nur neun (Abbildung 51). Zusammen mit den zuvor geschilderten Unterschieden im Bindungsmodus der beiden Substanzen lässt sich somit erklären, warum Ligand **1** um drei Größenordnungen affiner ist als Ligand **24**.

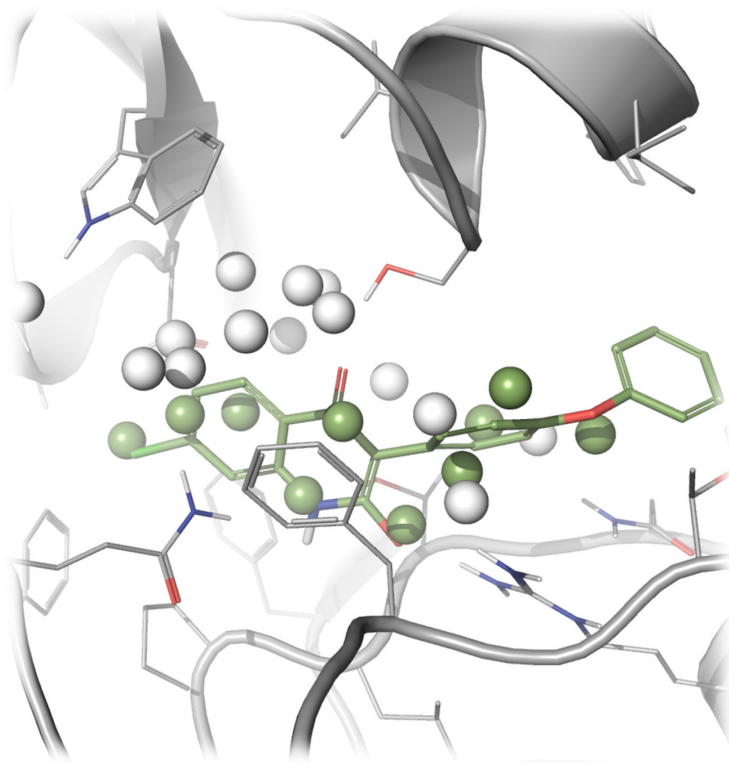


Abbildung 50: Watermap des nanomolaren Virtual Screening Hits **1**
Bindungsstelle von Ligand **1** (grün). Weiße Kugeln: nicht-verdrängte Watermap-Wasser; grüne Kugeln: zehn durch Ligand **1** verdrängte Watermap-Wasser

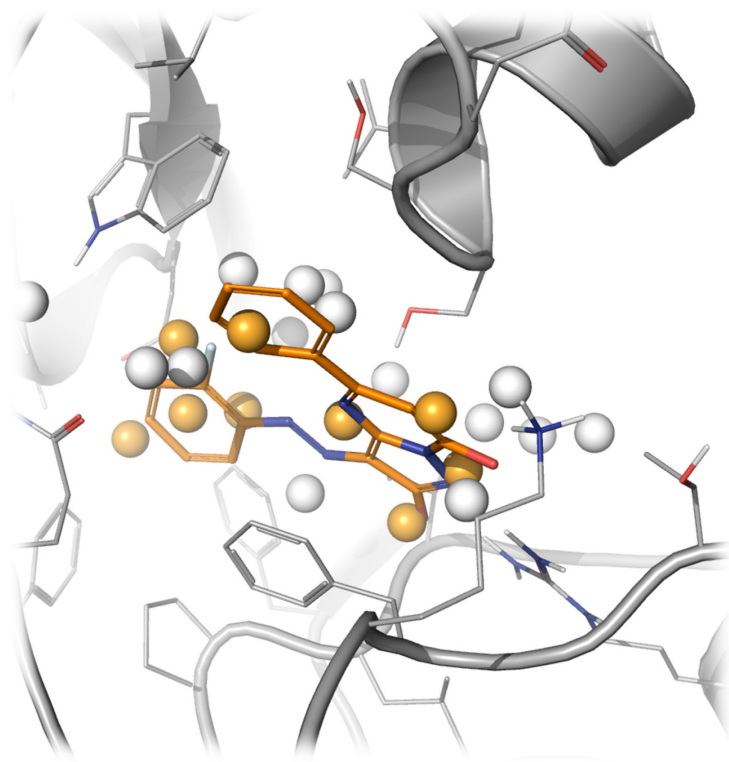


Abbildung 51: Watermap des mikromolaren Virtual Screening Hits **24**
Bindungsstelle von Ligand **24** (orange). Weiße Kugeln: nicht-verdrängte Watermap-Wasser; orange Kugeln: neun durch Ligand **24** verdrängte Watermap-Wasser

3.8.6 Vom antagonistischen zum agonistischen Zustand – die funktionale Bedeutung von Rezeptorflexibilität

In einem abschließenden Experiment wurde überprüft, ob anhand der entwickelten Methodik auch eine Aussage über die funktionale Antwort der NR1-Untereinheit des NMDA-Rezeptors auf die Bindung eines Liganden getroffen werden konnte. Hierzu wurde die offene Rezeptorstruktur im antagonistischen Zustand („*open cleft*“) als Ausgangsstruktur ausgewählt. Laut Furukawa *et al.*³⁷ entspricht diese dabei dem Apo-Zustand der NR1-Einheit, der Normalzustand ist somit der deaktivierte Rezeptor mit einer freien Bindungsstelle. Die Bindung eines vollen oder partiellen Agonisten verursacht eine Domänenbewegung, welche zu einer geschlossenen Konformation führt. Daraufhin wird der Ionenkanal des NMDA-Rezeptors geöffnet. Die Kristallstrukturen von Glycin, D-Serin (volle Agonisten) und D-Cycloserin (partieller Agonist) sind isomorph und unterscheiden sich nur um einen RMSD von 0,1 bzw. 0,2 Å.

Die Bindung eines Antagonisten bedeutet pharmakologisch gesehen die „Fixierung“ des Rezeptors in seinem deaktivierten offenen Zustand. Im Gegensatz zu den strukturell identischen geschlossenen Konformationen der Agonisten unterscheiden sich die antagonistischen Rezeptorstrukturen im Grad ihrer Öffnung. Unterschiedliche Antagonisten verursachen somit unterschiedliche offene Rezeptorkonformationen. Das Stören des Wasserstoffbrückennetzwerks in der Glycin_B-Bindungsstelle durch den Antagonisten kann daher als „Keilprinzip“ verstanden werden, mit dem der Antagonist die Schließbewegung der Domänen S1 und S2 blockiert.

Zur Simulation dieses Übergangs vom antagonistischen apo-Zustand (offen) zum agonistischen Zustand (geschlossen) wurde der partielle Agonist D-Cycloserin in die Kristallstruktur von NR1 (PDB-Code 1PBQ) gedockt und anschließend über einen Zeitraum von 50 Nanosekunden entsprechend des entwickelten MD-Protokolls (Kapitel 2.7.2.2) simuliert. Zur Überprüfung des korrekten Verlaufs der MD-Simulation wurde die Energie des Systems über den Verlauf der Simulation gemessen. Dabei verringerte sich die Energie geringfügig, das simulierte System blieb stabil und nahm einen insgesamt energetisch günstigeren Zustand ein (Abbildung 52).

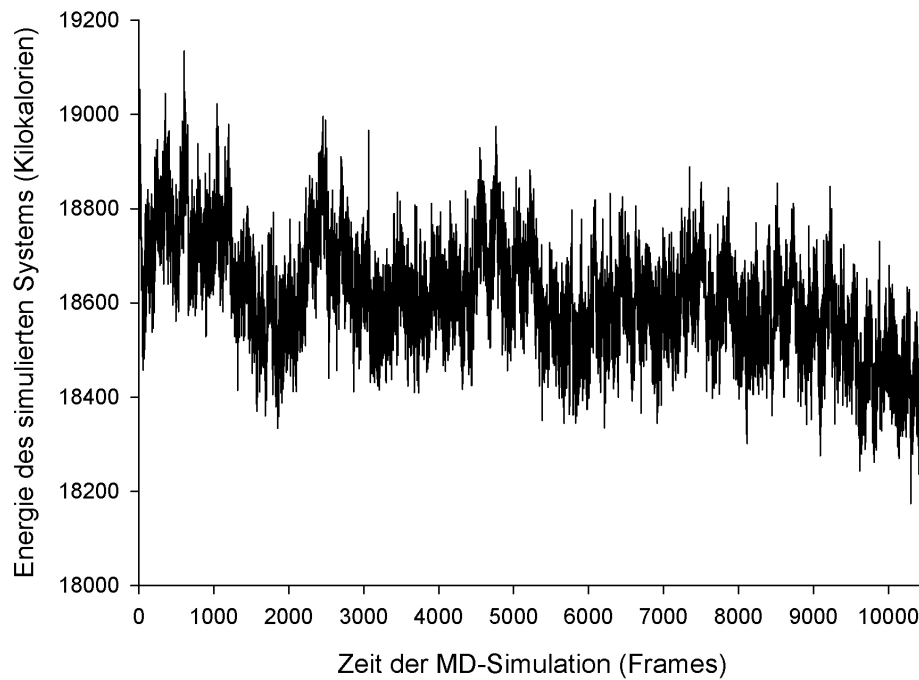


Abbildung 52: Energieverlauf – Moleküldynamik-Simulation von D-Cycloserin in NR1

Da auch nach 50 Nanosekunden Simulationszeit noch kein Konvergieren der Energie erkennbar war, sondern ein fortlaufender Trend hin zu niedrigeren Energiewerten, ist zu vermuten, dass das System noch nicht seinen endgültigen stabilen (agonistischen) Zustand erreicht hat. Die strukturellen Änderungen, die bei der Umwandlung der Apo-Konformation in den geschlossenen agonistischen Zustand notwendig sind, laufen offenbar jenseits der 50 Nanosekunden-grenze weiter ab. Für eine Übersicht, welche Reste in dieser Phase besonders mobil sind, war die Simulationsdauer jedoch ausreichend.

Der Verlauf der Beweglichkeit der einzelnen Reste von NR1, gemessen als RMSD zur Ausgangsstruktur der MD-Simulation, zeigte wie bei den bisherigen Simulationen eine hohe Mobilität der Loops 1 (Reste 45 bis 55) und 2 (Reste 90 bis 100). Auffällig war eine deutliche Bewegung im Bereich der zuvor als mobil identifizierten Reste der Glycin_B-Bindungsstelle (Reste 175 bis 185) (Abbildung 53).

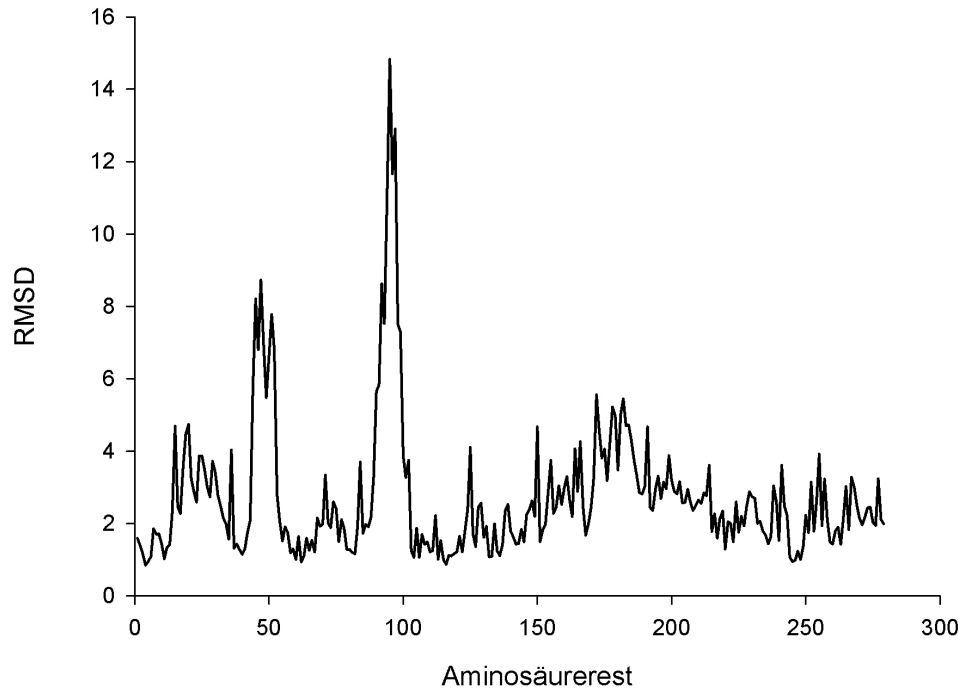


Abbildung 53: Rezeptorbewegung von NR1 im Übergang vom antagonistischen in den agonistischen Zustand

Freie MD-Simulation von D-Cycloserin (partieller Agonist) in der antagonistischen Kristallstruktur von NR1. Der RMSD der einzelnen Reste wurde über die Dauer der MD-Simulation von 50 Nanosekunden gemittelt. Reste 175-185: von S2 gebildeter Teil der Glycin_B-Bindungsstelle. Um Rest 50 und Rest 95: frei bewegliche Loops an der Oberfläche von NR1

Dies legt den Schluss nahe, dass die Umwandlung von antagonistischer in die antagonistische Rezeptorkonformation durch eine Konformationsänderung der Glycin_B-Bindungsstelle bei Bindung eines Agonisten eingeleitet wird. Diese befindet sich im Interface der beiden Domänen S1 und S2, welche zusammen die NR1-Untereinheit bilden. Die von Furukawa *et al.*³⁷ bei Bindung eines Agonisten in ihren Endpunkten identifizierte Bewegung der beiden Domänen zueinander konnte somit durch die funktionale Simulation des Bindungsereignisses bestätigt und genauer charakterisiert werden.

4 Zusammenfassung

Die Identifizierung neuartiger Verbindungsklassen für ein pharmakologisches Zielsystem ist eine fordernde Aufgabe für die frühe präklinische Forschung, insbesondere wenn bereits vorherige umfangreiche Studien durchgeführt und viele Leitstrukturserien gefunden wurden. In dieser Arbeit konnte gezeigt werden, dass Scaffold Hopping durch Methoden des Virtual Screenings auch für Systeme möglich ist, für die bereits eine Vielzahl von Referenzsubstanzen beschrieben ist und somit wenig freier chemischer Raum für Innovation zur Verfügung steht.

Als Beispielsystem wurde die Glycin_B-Bindungsstelle der NR1-Untereinheit des NMDA-Rezeptors betrachtet. Verschiedene zwei- und dreidimensionale Techniken des Virtual Screenings wurden einer umfangreichen retrospektiven Validierung unterworfen.

Zur Durchführung der prospektiven Virtual-Screening-Studie wurde eine automatisierte *in silico* Plattform entwickelt, die 8,9 Millionen käufliche Substanzen aus 46 Substanzkatalogen von 33 verschiedenen Anbietern sammelte, um etwa 5 Millionen unterschiedliche Moleküle in zweidimensionaler Darstellung aufzuarbeiten. Diese Menge an Substanzen stellt den größten Teil der zurzeit kommerziell verfügbaren chemischen Verbindungen, also den „verfügbaren chemischen Raum“ dar.

Anhand der retrospektiv validierten Virtual Screening Techniken konnten in einer prospektiven Suche 21 Glycin_B-Antagonisten mit neuartigen, d.h. für Glycin_B noch unbeschriebenen Scaffolds gefunden werden. Ausgehend von drei dieser Virtual Screening Hits wurden 53 weitere Verbindungen mit insgesamt fünf unterschiedlichen neuartigen Scaffolds und einem gemeinsamen Azomotiv identifiziert. Die Struktur-Wirkungsbeziehungen dieser fünf chemischen Serien wurden charakterisiert.

Das Ergebnis dieser Arbeit zeigt, dass es lohnend ist, alle vorhandenen Methoden auszuschöpfen, da sich die validierten Methoden komplementär zueinander verhielten und kein Virtual Screening Hit von mehr als einer Technik gefunden wurde.

Die Flexibilität von Proteinen als Antwort auf die Bindung unterschiedlicher Liganden stellt ein bislang ungelöstes chemieinformatisches Problem dar, welches auch grundlegende pharmakologische Bedeutung hat. So verursachen z.B. bei NMDA/Glycin_B agonistische Liganden eine Konformationsänderung des Rezeptors. Diese ruft dann eine direkte funktionale Antwort in Form der Öffnung des Ionenkanals hervor. Auch der Bindungsmodus der Antagonisten von

Glycin_B ist trotz Vorhandenseins von zwei Kristallstrukturen und mehreren Hundert zum Teil hochaffiner Referenzstrukturen zum großen Teil ungeklärt.

Im zweiten Teil dieser Arbeit wurde ein auf Moleküldynamiksimulationen basierendes Verfahren entwickelt, welches flexible Aminosäurereste im Rezeptor und damit induzierbare Bewegungen des Proteinrückgrates bestimmt. Die so identifizierten Reste wurden dann in einem erweiterten Verfahren des Induced-Fit-Dockings als explizit flexibel betrachtet. Hierdurch war die Berechnung verschiedener Bindungsmodi von Antagonisten möglich, die aufgrund ihrer Form und Größe nicht in die verfügbaren Kristallstrukturen von Glycin_B passten. Diese benötigten somit einen Induced-Fit-Effekt des Rezeptors, um eine Bindung einzugehen. Für die im ersten Teil dieser Arbeit identifizierten Azo-Liganden wurde auf Basis dieser Methode ein gemeinsamer Bindungsmodus vorgeschlagen. Ebenso konnte anhand der Methodik eine Aussage über die funktionale Auswirkung der Proteinflexibilität beim Übergang vom antagonistischen zum agonistischen Rezeptorzustand von Glycin_B getroffen werden.

Ein großes Problem aktueller Dockingverfahren ist die mangelnde Verfügbarkeit von Scoringfunktionen, welche die tatsächliche biologische Bindungsaffinität eines Liganden berechnen. Hier wurde ein Verfahren für das Zielsystem Glycin_B gezeigt, welches aufgrund der Berechnung des thermodynamischen Entropie- und Enthalpiegewinns durch Verdrängung von hydrophob eingeschlossenen Wasser aus der Bindungsstelle durch den Liganden eine Aussage über dessen zu erwartende Bindungsaffinität trifft. Dieses neuartige Scoring-system wurde auf die im Virtual Screening identifizierten Serie von Azo-Liganden angewandt und verfügte über eine im Vergleich zu klassischen Scoringfunktionen des Molecular Dockings verbesserte Vorhersagekraft der biologischen Bindungsaffinität.

5 Ausblick

In der vorliegenden Arbeit wurde eine Auswahl von Methoden des Virtual Screenings genutzt, um Glycin_B-Antagonisten mit neuartigen Molekülscaffolds zu identifizieren. Dieses Ergebnis zeigte, dass die Nutzung eines möglichst großen Portfolios vorhandener Methoden des Virtual Screenings notwendig ist, da sich alle Methoden komplementär zueinander verhielten und kein Virtual Screening Hit von mehr als einer Technik gefunden werden konnte. Die Methodenentwicklung im Bereich der Chemieinformatik und theoretischen/physikalischen Chemie ist noch nicht abgeschlossen, so dass fortwährend neue Ansätze zur Beschreibung von Protein-Ligand-Interaktion oder zur SAR-Analyse von Liganden und damit für die Suche nach neuartigen Wirkstoffen zur Verfügung stehen. Die kontinuierliche Validierung neuer und der Vergleich zu etablierten Methoden des Virtual Screenings wird somit auch in Zukunft ein Forschungsfeld bleiben.

Desweiteren wurde im Zuge dieser Arbeit deutlich, dass eine umfangreiche Validierung von Methoden – obwohl zeitlich aufwändig – eine essentielle Grundlage für das Verständnis der Leistungsfähigkeit der Methoden und damit ihrer prospektiven Anwendbarkeit darstellt. Eine solche Validierung entscheidet unter Umständen zwischen Erfolg und Misserfolg eines Virtual Screenings und somit sogar über den Erfolg eines präklinischen Forschungsprojektes. Die genaue Kenntnis der Anwendbarkeit von Methoden ermöglicht auch eine Priorisierung z.B. bei Ressourcenknappheit.

Im Verlauf dieser Arbeit wurde eine Serie von 53 bisher unbeschriebenen Azo-Verbindungen mit antagonistischer Wirkung auf die Glycin_B-Bindungsstelle des NMDA-Rezeptors identifiziert. Für eine mögliche weitere präklinische Entwicklung dieser Substanzen müsste als nächster Schritt überprüft werden, ob das Azomotiv bioisoster ersetzbar ist, da die Azogruppe gemeinhin als nicht *drug-like* und somit als nicht entwickelbar gilt. Aufgrund der vorgenommenen Analyse des Bindungsmodus ist eine funktionale Notwendigkeit des Azo-Motivs nicht zu erwarten. Weiterhin muss die Bindungsaffinität der Substanzen deutlich erhöht werden, damit sie als Leitstruktur (*Lead*) in Frage kommen. Da niedrig-nanomolare Affinität gemeinhin als Kriterium für einen klinischen Entwicklungskandidaten angesehen wird, die gefundenen Azo-Verbindungen jedoch eine niedrig-mikromolare Bindungsaffinität aufwiesen, muss die Tauglichkeit der Substanzen für eine weitere präklinische Entwicklung in Frage gestellt werden.

Desweiteren konnte die prinzipielle Tauglichkeit der Watermap-Scoringfunktion für die Berechnung der Bindungsaffinität einer kongenerischen Serie von Liganden gezeigt werden. Da eine solche „Wasserkarte“ der Bindungsstelle von der zugrundeliegenden Rezeptorkonformation abhängig ist, ist

eine Anwendung dieser Methode in großem Maßstab, z.B. für das Virtual Screening nicht möglich. Für jede Serie an Substanzen bzw. für jeden einzelnen Ligand müsste die passende Induced-Fit-Rezeptorkonformation und daraufhin die entsprechende Wasserkarte berechnet werden. Auch wenn dies schneller als eine komplette Simulation der freien Energie des Systems mit und ohne Ligand (und damit explizit die Berechnung von ΔG bis auf die Genauigkeit des zugrundeliegenden Kraftfeldes, z.B. anhand FEP-Berechnungen) ist, liegt diese Methode bei Millionen von Substanzen jenseits aktueller Rechenkapazitäten.

Alle in dieser Arbeit geschilderten Verfahren des Molecular Modelings basieren auf Kraftfeldern, also im Prinzip auf einer minimalistischen Reduktion der Komplexität von makromolekularen Systemen und ihrer Interaktionen auf die Prinzipien Newtonscher Mechanik. Auch wenn die Entwicklung neuer und genauerer Kraftfelder ständig voranschreitet, dürfte erst der Einzug quantenmechanischer Berechnungen auf ersten Prinzipien (*first principles*) einen „Quantensprung“ in Richtung Genauigkeit der simulierten Systeme liefern. Aber auch wenn es bereits erste Ansätze von quantenmechanisch polarisierten Dockings²⁰⁴ oder gemischt quantenmechanisch-molekularmechanischen (QM/MM) Simulationen²⁰⁵ gibt, so ist eine vollständige quantenmechanische Berechnung der Dynamik von Biomolekülen oder Protein-Ligand-Interaktionen in naher Zukunft noch nicht zu erwarten.

6 Literaturverzeichnis

1. Dingledine, R.; Borges, K.; Bowie, D.; Traynelis, S. F., The glutamate receptor ion channels. *Pharmacol. Rev.* **1999**, *51* (1), 7-61.
2. Davies, S.; Lodge, D., Evidence for involvement of N-methylaspartate receptors in "wind-up" of class 2 neurones in the dorsal horn of the rat. *Brain Res.* **1987**, *424* (2), 402-406.
3. Carlton, S. M.; Hargett, G. L.; Coggeshall, R. E., Localization and activation of glutamate receptors in unmyelinated axons of rat glabrous skin. *Neurosci. Lett.* **1995**, *197* (1), 25-28.
4. Carlton, S.; Zhou, S.; Coggeshall, R., Evidence for the interaction of glutamate and NK1 receptors in the periphery. *Brain Res.* **1998**, *790* (1-2), 160-169.
5. Greenamyre, J. T.; O'Brien, C. F., N-methyl-D-aspartate antagonists in the treatment of Parkinson's disease. *Arch. Neurol.* **1991**, *48* (9), 977-981.
6. Geddes, J. W.; Chang-Chui, H.; Cooper, S. M.; Lott, I. T.; Cotman, C. W., Density and distribution of NMDA receptors in the human hippocampus in Alzheimer's disease. *Brain Res.* **1986**, *399* (1), 156-161.
7. Danysz, W.; Parsons, C., The NMDA receptor antagonist memantine as a symptomatological and neuroprotective treatment for Alzheimer's disease: preclinical evidence. *Int. J. Geriatr. Psychiatry* **2003**, *18* (S1), S23-S32.
8. Olney, J., NMDA receptor hypofunction model of schizophrenia. *J. Psychiatr. Res.* **1999**, *33* (6), 523-533.
9. Nakanishi, S., Molecular diversity of glutamate receptors and implications for brain function. *Science* **1992**, *258* (5082), 597-603.
10. Kew, J. N. C.; Kemp, J. A., Ionotropic and metabotropic glutamate receptor structure and pharmacology. *Psychopharmacology (Berl)*. **2005**, *179* (1), 4-29.
11. Laube, B.; Kuhse, J.; Betz, H., Evidence for a Tetrameric Structure of Recombinant NMDA Receptors. *J. Neurosci.* **1998**, *18* (8), 2954-2961.
12. Furukawa, H.; Singh, S.; Mancusso, R.; Gouaux, E., Subunit arrangement and function in NMDA receptors. *Nature* **2003**, *438* (7065), 185-192.
13. Kleckner, N. W.; Dingledine, R., Requirement for Glycin in activation of NMDA-receptors expressed in *Xenopus* oocytes. *Science* **1988**, *241* (4867), 835-837.
14. Ishii, T.; Moriyoshi, K.; Sugihara, H.; Sakurada, K.; Kadotani, H.; Yokoi, M.; Akazawa, C.; Shigemoto, R.; Mizuno, N.; Masu, M., Molecular characterization of the family of the N-methyl-D-aspartate receptor subunits. *J. Biol. Chem.* **1993**, *268* (4), 2836-2843.
15. Danysz, W.; Parsons, C. G., GlycinB recognition site of NMDA receptors and its antagonists. *Amino Acids* **1998**, *14* (1-3), 205-206.

16. Danysz, W.; Parsons, C. G.; Bresink, I.; Quack, G., Glutamate in CNS Disorders. *Drug News and Perspectives* **1995**, *8* (5), 261-277.
17. Benveniste, M.; Mayer, M., Kinetic analysis of antagonist action at N-methyl-D-aspartic acid receptors. Two binding sites each for glutamate and Glycin. *Biophys. J.* **1991**, *59* (3), 560-573.
18. Clements, J.; Westbrook, G., Activation kinetics reveal the number of glutamate and Glycin binding sites on the N-methyl-d-aspartate receptor. *Neuron* **1991**, *7* (4), 605.
19. Laurie, D.; Seeburg, P., Ligand affinities at recombinant N-methyl-D-aspartate receptors depend on subunit composition. *Eur. J. Pharmacol.* **1994**, *268* (3), 335-345.
20. Nowak, L.; Bregestovski, P.; Ascher, P.; Herbert, A.; Prochiantz, A., Magnesium gates glutamate-activated channels in mouse central neurons. *Nature* **1984**, *307*, 462-465.
21. Danysz, W.; Parsons, C. G., Glycin and N-Methyl-D-Aspartate Receptors: Physiological Significance and Possible Therapeutic Applications. *Pharmacol. Rev.* **1998**, *50* (4), 597-664.
22. Forsythe, I.; Westbrook, G., Slow excitatory postsynaptic currents mediated by N-methyl-D-aspartate receptors on cultured mouse central neurones. *The Journal of Physiology* **1988**, *396* (1), 515-533.
23. Vicini, S.; Wang, J.; Li, J.; Zhu, W.; Wang, Y.; Luo, J.; Wolfe, B.; Grayson, D., Functional and pharmacological differences between recombinant N-methyl-D-aspartate receptors. *J. Neurophysiol.* **1998**, *79* (2), 555.
24. Lester, R.; Clements, J.; Westbrook, G.; Jahr, C., Channel kinetics determine the time course of NMDA receptor-mediated synaptic currents. *Nature* **1990**, *346* (6284), 565-567.
25. Simon, R. P.; Swan, J. H.; Griffiths, T.; Meldrum, B. S., Blockade of N-methyl-D-aspartate receptors may protect against ischemic damage in the brain. *Science* **1984**, *226* (4676), 850-852.
26. Gauthier, S.; Wirth, Y.; H. J. Möbius, Effects of memantine on behavioural symptoms in Alzheimer's disease patients: an analysis of the Neuropsychiatric Inventory (NPI) data of two randomised, controlled studies. *Int. J. Geriatr. Psychiatry* **2005**, *20* (5), 459-464.
27. McShane, R.; Areosa Sastre, A.; Minakaran, N. Memantine for dementia *Cochrane Database of Systematic Reviews* [Online], 2006.
<http://www.mrw.interscience.wiley.com/cochrane/clsysrev/articles/CD003154/frame.html>.
28. Danysz, W.; Parsons, C. G.; Karcz-Kubichal, M.; Schwaier, A.; Popik, P.; Wedzony, K.; Lazarewicz, J.; Quack, G., GlycinB antagonists as potential therapeutic agents. *Amino Acids* **1998**, *14* (1), 235-239.

29. Parsons, C., NMDA receptors as targets for drug action in neuropathic pain. *Eur. J. Pharmacol.* **2001**, 429 (1-3), 71-78.
30. Madden, D. R., The structure and function of glutamate receptor ion channels. *Nat Rev Neurosci* **2002**, 3 (2), 91-101.
31. Wo, Z. G.; Oswald, R. E., Unraveling the modular design of glutamate-gated ion channels. *Trends Neurosci.* **1995**, 18 (4), 161-168.
32. Parsons, C. G.; Zong, X.; Lux, H. D., Whole cell and single channel analysis of the kinetics of Glycin-sensitive N-methyl-D-aspartate receptor desensitization. *Br. J. Pharmacol.* **1993**, 109 (1), 213-221.
33. Gola, J.; Obrezanova, O.; Champness, E.; Segall, M., ADMET Property Prediction: The State of the Art and Current Challenges. *QSAR & Combinatorial Science* **2006**, 25 (12), 1172-1180.
34. Yu, H.; Adedoyin, A., ADME-Tox in drug discovery: integration of experimental and computational technologies. *Drug Discovery Today* **2003**, 8 (18), 852-861.
35. van de Waterbeemd, H.; Gifford, E., ADMET in silico modelling: towards prediction paradise? *Nat Rev Drug Discov* **2003**, 2 (3), 192-204.
36. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res.* **2000**, 28 (1), 235-242.
37. Furukawa, H.; Gouaux, E., Mechanisms of activation, inhibition and specificity: crystal structures of the NMDA receptor NR1 ligand-binding core. *EMBO J.* **2003**, 22 (12), 2873-2885.
38. Inanobe, A.; Furukawa, H.; Gouaux, E., Mechanism of Partial Agonist Action at the NR1 Subunit of NMDA Receptors. *Neuron* **2005**, 47 (1), 71-84.
39. Armstrong, N.; Gouaux, E., Mechanisms for Activation and Antagonism of an AMPA-Sensitive Glutamate Receptor Crystal Structures of the GluR2 Ligand Binding Core. *Neuron* **2000**, 28 (1), 165-181.
40. Armstrong, N.; Sun, Y.; Chen, G.; Gouaux, E., Structure of a glutamate-receptor ligand-binding core in complex with kainate. *Nature* **1998**, 395 (6705), 913-917.
41. Miyazaki, J.; Nakanishi, S.; Jingami, H., Expression and characterization of a Glycin-binding fragment of the N-methyl-D-aspartate receptor subunit NR1. *Biochem J.* **1999**, 340, 687-692.
42. Hirai, H.; Kirsch, J.; Laube, B.; Betz, H.; Kuhse, J., The Glycin binding site of the N-methyl-D-aspartate receptor subunit NR1: identification of novel determinants of co-agonist potentiation in the extracellular M3-M4 loop region. *PNAS* **1996**, 93 (12), 6031-6036.
43. Williams, K.; Chao, J.; Kashiwagi, K.; Masuko, T.; Igarashi, K., Activation of N-methyl-D-aspartate receptors by Glycin: role of an aspartate residue in the M3-M4 loop of the NR1 subunit. *Mol. Pharmacol.* **1996**, 50 (4), 701-708.

44. Lampinen, M.; Pentikäinen, O.; Johnson, M.; Keinänen, K., AMPA receptors and bacterial periplasmic amino acid-binding proteins share the ionic mechanism of ligand recognition. *EMBO J.* **1998**, *17*, 4704-4711.
45. McBain, C.; Kleckner, N.; Wyrick, S.; Dingledine, R., Structural requirements for activation of the Glycine coagonist site of N-methyl-D-aspartate receptors expressed in *Xenopus* oocytes. *Mol. Pharmacol.* **1989**, *36* (4), 556-565.
46. Tikhonova, I.; Baskin, I.; Palyulin, V.; Zefirov, N. In *A Spatial Model of the Glycine Site of the NR1 Subunit of NMDA-Receptor and Ligand Docking*, Springer: 2002; pp 67-70.
47. Kuryatov, A.; Laube, B.; Betz, H.; Kuhse, J., Mutational analysis of the Glycine-binding site of the NMDA receptor: structural similarity with bacterial amino acid-binding proteins. *Neuron* **1994**, *12* (6), 1291-300.
48. Sun, Y.; Olson, R.; Horning, M.; Armstrong, N.; Mayer, M.; Gouaux, E., Mechanism of glutamate receptor desensitization. *Nature* **2002**, *417*, 245-253.
49. Oprea, T. I., Current trends in lead discovery: Are we looking for the appropriate properties? *Mol. Divers.* **2000**, *5* (4), 199-208.
50. Schneider, G.; Neidhart, W.; Giller, T.; Schmid, G., "Scaffold-Hopping" by Topological Pharmacophore Search: A Contribution to Virtual Screening. *Angew. Chem. Int. Ed. Engl.* **1999**, *38* (19), 2894-2896.
51. Schneider, G.; Schneider, P.; Renner, S., Scaffold-Hopping: How Far Can You Jump? *QSAR Comb. Sci.* **2006**, *25* (12), 1162-1171.
52. Oprea, T., Chemical space navigation in lead discovery. *Curr. Opin. Chem. Biol.* **2002**, *6* (3), 384-389.
53. Dobson, C., Chemical space and biology. *Nature* **2004**, *432* (7019), 824-828.
54. Lipinski, C.; Hopkins, A., Navigating chemical space for biology and medicine. *Nature* **2004**, *432* (7019), 855-861.
55. Koch, M.; Schuffenhauer, A.; Scheck, M.; Wetzels, S.; Casaulta, M.; Odermatt, A.; Ertl, P.; Waldmann, H., Charting biologically relevant chemical space: A structural classification of natural products (SCONP). *PNAS* **2005**, *102* (48), 17272-17277.
56. Henkel, T.; Brunne, R. M.; Müller, H.; Reichel, F., Statistical Investigation into the Structural Complementarity of Natural Products and Synthetic Compounds. *Angew. Chem. Int. Ed.* **1999**, *38* (5), 643-647.
57. Feher, M.; Schmidt, J. M., Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (1), 218-227.
58. Hertzberg, R. P.; Pope, A. J., High-throughput screening: new technology for the 21st century. *Curr. Opin. Chem. Biol.* **2000**, *4* (4), 445-451.
59. Silverman, L., New assay technologies for high-throughput screening. *Curr. Opin. Chem. Biol.* **1998**, *2* (3), 397-403.

60. Gribbon, P.; Sewing, A., High-throughput drug discovery: what can we expect from HTS? *Drug Discovery Today* **2005**, *10* (1), 17-22.
61. Schneider, G., Trends in virtual combinatorial library design. *Curr. Med. Chem.* **2002**, *9* (23), 2095-2101.
62. Leach, A., The in silico world of virtual libraries. *Drug Discovery Today* **2000**, *5* (8), 326-336.
63. Van Drie, J., Approaches to virtual library design. *Drug Discovery Today* **1998**, *3* (6), 274-283.
64. Schneider, G.; Böhm, H.-J., Virtual screening and fast automated docking methods. *Drug Discovery Today* **2002**, *7*, 64-70.
65. Walters, W. P.; Stahl, M. T.; Murcko, M. A., Virtual screening - an overview. *Drug Discovery Today* **1998**, *3*, 160-178.
66. Langer, T.; Hoffmann, R. D., Virtual Screening An Effective Tool for Lead Structure Discovery. *Curr. Pharm. Des.* **2001**, *7*, 509-527.
67. Jenkins, J. L.; Kao, R. Y. T.; Shapiro, R., Virtual screening to enrich hit lists from high-throughput screening: A case study on small-molecule inhibitors of angiogenin. *Proteins* **2003**, *50* (1), 81-93.
68. Xu, H.; Agrafiotis, D. K., Retrospect and Prospect of Virtual Screening in Drug Discovery. *Curr. Top. Med. Chem.* **2002**, *2*, 1305-1320.
69. Chen; Beining; Harrison; Robert; Papadatos; George; Willett; Peter; Wood; David; Lewell; Xiao; Greenidge; Paulette; Stiefl; Nikolaus, Evaluation of machine-learning methods for ligand-based virtual screening. *J. Comput. Aided Mol. Des.* **2007**, *21* (1-3), 53-62.
70. Schneider, G.; Wrede, P., Artificial neural networks for computer-based molecular design. *Prog. Biophys. Mol. Biol.* **1998**, *70* (3), 175-222.
71. Byvatov, E.; Schneider, G., SVM-Based Feature Selection for Characterization of Focused Compound Collections. *J. Chem. Inf. Comput. Sci.* **2004**, *44* (3), 993-999.
72. Byvatov, E.; Schneider, G., Support vector machine applications in bioinformatics. *Appl. Bioinformatics* **2003**, *2* (2), 67-77.
73. Schneider, G.; Fechner, U., Computer-based de novo design of drug-like molecules. *Nat. Rev. Drug Discovery* **2005**, *4* (8), 649-663.
74. Snarey, M.; Terrett, N.; Willett, P.; Wilton, D., Comparison of algorithms for dissimilarity-based compound selection. *J. Mol. Graph. Model.* **1997**, *15* (6), 372-385.
75. Hert, J.; Willett, P.; Wilton, D. J.; Acklin, P.; Azzaoui, K.; Jacoby, E.; Schuffenhauer, A., Comparison of topological descriptors for similarity-based virtual screening using multiple bioactive reference structures. *Org. Biomol. Chem.* **2004**, *2* (22), 3256-3266.

76. Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D., Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57* (2), 225-242.
77. Lyne, P., Structure-based virtual screening: an overview. *Drug Discovery Today* **2002**, *7* (20), 1047-1055.
78. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J., Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discovery* **2004**, *3* (11), 935-949.
79. Willett, P., Searching for pharmacophoric patterns in databases of three-dimensional chemical structures. *J. Mol. Recognit.* **1995**, *8* (5), 290-303.
80. Sheridan, R. P.; Rusinko, A.; Nilakantan, R.; Venkataraghavan, R., Searching for pharmacophores in large coordinate data bases and its use in drug design. *PNAS* **1989**, *86* (20), 8165-8169.
81. Schneider, G.; Neidhart, W.; Giller, T.; Schmid, G., "Scaffold-Hopping" by Topological Pharmacophore Search: A Contribution to Virtual Screening. *Angew. Chem. Int. Ed.* **1999**, *38* (19), 2894-2896.
82. Mason, J. S.; Good, A. C.; Martin, E. J., 3-D Pharmacophores in Drug Discovery. *Curr. Pharm. Des.* **2001**, *7*, 567-597.
83. Emil Fischer, Ueber den Einfluss der Konfiguration auf die Wirkung der Enzyme III. *Berichte der deutschen chemischen Gesellschaft* **1895**, *28* (2), 1429-1438.
84. Frieder W. Lichtenthaler, 100 Years "Schlüssel-Schloss-Prinzip": What Made Emil Fischer Use this Analogy? *Angew. Chem. Int. Ed.* **1995**, *33* (23-24), 2364-2374.
85. Daniel E. Koshland Jr., The Key-Lock Theory and the Induced Fit Theory. *Angew. Chem. Int. Ed.* **1995**, *33* (23-24), 2375-2378.
86. Wilchek, M.; Bayer, E. A.; Livnah, O., Essentials of biorecognition: the (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction. *Immunol. Lett.* **2006**, *103* (1), 27-32.
87. Verkhivker, G.; Bouzida, D.; Gehlhaar, D. K.; Rejto, P. A.; Freer, S. T.; Rose, P. W., Complexity and simplicity of ligand-macromolecule interactions: the energy landscape perspective. *Curr. Opin. Struct. Biol.* **2002**, *12* (2), 197-203.
88. Ma, B.; Shatsky, M.; Wolfson, H. J.; Ruth Nussinov, Multiple diverse ligands binding at a single protein site: A matter of pre-existing populations. *Protein Sci.* **2002**, *11* (2), 184-197.
89. Kumar, S.; Ma, B.; Tsai, C.-J.; Sinha, N.; Nussinov, R., Folding and binding cascades: Dynamic landscapes and population shifts. *PRS* **2000**, *9* (01), 10-19.
90. Zavodszky, M. I.; Kuhn, L. A., Side-chain flexibility in protein-ligand binding: the minimal rotation hypothesis. *Protein Sci.* **2005**, *14* (4), 1104-1114.
91. Ma, B.; Kumar, S.; Tsai, C.-J.; Nussinov, R., Folding funnels and binding mechanisms. *Protein Eng.* **1999**, *12* (9), 713-720.

92. Philippopoulos, M.; Lim, C., Exploring the dynamic information content of a protein NMR structure: Comparison of a molecular dynamics simulation with the NMR and X-ray structures of Escherichia coli ribonuclease HI. *Proteins* **1999**, *36* (1), 87-110.
93. Snow, C. D.; Sorin, E. J.; Rhee, Y. M.; Pande, V. S., How well can simulation predict folding kinetics and thermodynamics? *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34* (1), 43-69.
94. Goh, C.-S.; Milburn, D.; Gerstein, M., Conformational changes associated with protein-protein interactions. *Curr. Opin. Struct. Biol.* **2004**, *14* (1), 104-109.
95. Lin, J.-H.; Perryman, A. L.; Schames, J. R.; McCammon, J. A., Computational Drug Design Accommodating Receptor Flexibility: The Relaxed Complex Scheme. *J. Am. Chem. Soc.* **2002**, *124* (20), 5632-5633.
96. Carlson, H., Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol.* **2002**, *6* (4), 447-452.
97. Knegt, R. M.; Kuntz, I. D.; Oshiro, C. M., Molecular docking to ensembles of protein structures. *J. Mol. Biol.* **1997**, *266* (2), 424-440.
98. Clauen, H.; Buning, C.; Rarey, M.; Lengauer, T., FlexE: efficient molecular docking considering protein structure variations. *J. Mol. Biol.* **2001**, *308* (2), 377-395.
99. Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S., A Critical Assessment of Docking Programs and Scoring Functions. *J. Med. Chem.* **2006**, *49* (20), 5912-5931.
100. Luque, I.; Freire, E., Structural stability of binding sites: Consequences for binding affinity and allosteric effects. *Proteins* **2000**, *41* (S4), 63-71.
101. Teague, S. J., Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discovery* **2003**, *2* (7), 527-541.
102. Olah, M.; Mracec, M.; Ostopovici, L.; Rad, R.; Bora, A.; Hadaruga, N.; Olah, I.; Banda, M.; Simon, Z.; Mracec, M.; Oprea, T. I., WOMBAT: World of Molecular Bioactivity. *Chemoinformatics in Drug Discovery* **2004**, 223-239.
103. Weininger, D., SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (1), 31-36.
104. Weininger, D.; Weininger, A.; Weininger, J., SMILES. 2. Algorithm for generation of unique SMILES notation. *J. Chem. Inf. Comput. Sci.* **1989**, *29* (2), 97-101.
105. Walters, P.; Stahl, M.; Murcko, M., Virtual screening - an overview. *Drug Discovery Today* **1998**, *3* (4), 160-178.

106. Hann, M.; Hudson, B.; Lewell, X.; Lively, R.; Miller, L.; Ramsden, N., Strategic Pooling of Compounds for High-Throughput Screening. *J. Chem. Inf. Comput. Sci.* **1999**, *39* (5), 897-902.
107. Lipinski, C.; Lombardo, F.; Dominy, B.; Feeney, P., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2001**, *46* (1-3), 3-26.
108. Lipinski, C., Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies* **2004**, *1* (4), 337-341.
109. Levy, J. B.; Taft, R. W.; Hammett, L. P., The Mechanism of the Acid-catalyzed Hydration of Olefins¹. *J. Am. Chem. Soc.* **1953**, *75* (5), 1253-1254.
110. Jorgensen, W.; Tirado-Rives, J., The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110* (6), 1657-1666.
111. Gentsch, W., Sun Grid Engine: Towards Creating a Compute Power Grid. In *Proceedings of the 1st International Symposium on Cluster Computing and the Grid*, IEEE Computer Society: 2001.
112. Willett, P.; Barnard, J. M.; Downs, G. M., Chemical Similarity Searching. *J. Chem. Inf. Comput. Sci.* **1998**, *38* (6), 983-996.
113. Jaccard, P., *Bulletin del la Société Vaudoisedes Sciences Naturelles* **1901**, *37*, 241-272.
114. Tanimoto, T., *IBM Internal Reports* **1957**.
115. Minkowski, H., Volumen und Oberflächen. *Mathematische Annalen* **1903**, *57*, 447-495.
116. Holliday, J. D.; Hu, C. Y.; Willett, P., Grouping of Coefficients for the Calculation of Inter-Molecular Similarity and Dissimilarity using 2D Fragment Bit-Strings. *Comb. Chem. High Throughput Screening* **2002**, *5*, 155-166.
117. James, C. A.; Weininger, D.; Delaney, J., Fingerprints - Screening and Similarity. *Daylight Theory Manual* **1997**.
118. Hanley, J.; McNeil, B., A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* **1983**, *148* (3), 839-843.
119. Zou, K. H.; O'Malley, A. J.; Mauri, L., Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models. *Circulation* **2007**, *115* (5), 654-657.
120. Matthews, B. W., Solvent content of protein crystals. *J. Mol. Biol.* **1968**, *33* (2), 491-497.
121. Morgan, H. L., The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. *J. Chem. Doc.* **1965**, *5* (2), 107-113.

122. Pearson, K., On lines and planes of closest fit to systems of points in space. *Philosophical Magazine* **1901**, 2 (6), 559-572.
123. Agrafiotis, D. K.; Lobanov, V. S., Nonlinear Mapping Networks. *J. Chem. Inf. Comput. Sci.* **2000**, 40 (6), 1356-1362.
124. Domine, D.; Devillers, J.; Chastrette, M.; Karcher, W., Non-linear mapping for structure-activity and structure-property modelling. *J. Chemom.* **1993**, 7 (4), 227-242.
125. Shoichet, B. K., Virtual screening of chemical libraries. *Nature* **2004**, 432 (7019), 862-865.
126. Bernardo, J.; Smith, A., Bayesian Theory. *Meas. Sci. Technol.* **2001**, 12 (2), 221-222.
127. Liu, Y., A comparative study on feature selection methods for drug discovery. *J. Chem. Inf. Comput. Sci.* **2004**, 44 (5), 1823-1828.
128. Vogt, M.; Bajorath, J., Bayesian Screening for Active Compounds in High-dimensional Chemical Spaces Combining Property Descriptors and Molecular Fingerprints. *Chem. Biol. Drug Des.* **2008**, 71 (1), 8-14.
129. Bender, A.; Mussa, H.; Glen, R., Screening for Dihydrofolate Reductase Inhibitors Using MOLPRINT 2D, a Fast Fragment-Based Method Employing the Naive Bayesian Classifier: Limitations of the Descriptor and the Importance of Balanced Chemistry in Training and Test Sets. *J. Biomol. Screen.* **2005**, 10 (7), 658-666.
130. Watson, P., Naive Bayes Classification Using 2D Pharmacophore Feature Triplet Vectors. *J. Chem. Inf. Model.* **2008**, 48 (1), 166-178.
131. Poschenrieder, H.; Stachel, H.; Hofner, G.; Mayer, P., Novel pyrrolinones as n-methyl-d-aspartate receptor antagonists. *Eur. J. Med. Chem.* **2005**, 40 (4), 391-400.
132. Guzikowski, A., 6,7,8,9-tetrahydro-3-hydroxy-1H-1-benzazepine-2,5-diones via a diels-alder reaction: antagonists with a non-planar hydrophobic region for NMDA receptor Glycin sites. *Bioorg. Med. Chem. Lett.* **1995**, 5 (22), 2747-2748.
133. Varano, F.; Catarzi, D.; Colotta, V.; Filacchioni, G.; Galli, A.; Costagli, C.; Carla, V., Synthesis and Biological Evaluation of a New Set of Pyrazolo[1,5-c]quinazoline-2-carboxylates as Novel Excitatory Amino Acid Antagonists. *J. Med. Chem.* **2002**, 45 (5), 1035-1044.
134. MacLeod, A. M.; Grimwood, S.; Barton, C.; Bristow, L.; Saywell, K. L.; Marshall, G. R.; Ball, R. G., Identification of 3,5-Dihydro-2-aryl-1H-pyrazolo[3,4-c]quinoline-1,4(2H)-diones as Novel High-Affinity Glycin Site N-Methyl-D-aspartate Antagonists. *J. Med. Chem.* **1995**, 38 (12), 2239-2243.
135. Baron, B. M.; Cregge, R. J.; Farr, R. A.; Friedrich, D.; Gross, R. S.; Harrison, B. L.; Janowick, D. A.; Matthews, D.; McCloskey, T. C.; Meikrantz, S.; Nyce, P. L.; Vaz, R.; Metz, W. A., CoMFA, Synthesis, and Pharmacological Evaluation of (E)-3-(2-Carboxy-2-arylvinyl)-4,6-dichloro-1H-indole-2-carboxylic Acids: 3-[2-(3-

- Aminophenyl)-2-carboxyvinyl]-4,6-dichloro-1H-indole-2-carboxylic Acid, a Potent Selective Glycin-Site NMDA Receptor Antagonist. *J. Med. Chem.* **2005**, *48* (4), 995-1018.
136. Renner, S.; Noeske, T.; Parsons, C. G.; Schneider, P.; Weil, T.; Schneider, G., New Allosteric Modulators of Metabotropic Glutamate Receptor 5 (mGluR5) Found by Ligand-Based Virtual Screening. *ChemBioChem* **2005**, *6* (4), 620-625.
137. Derksen, S.; Rau, O.; Schneider, P.; Schubert-Zsilavecz, M.; Schneider, G., Virtual Screening for PPAR Modulators Using a Probabilistic Neural Network. *ChemMedChem* **2006**, *1* (12), 1346-1350.
138. Noeske, T.; Gutcaits, A.; Parsons, Christopher G.; Weil, T., Allosteric Modulation of Family 3 GPCRs. *QSAR Comb. Sci.* **2006**, *25* (2), 134-146.
139. Renner, S.; Schneider, G., Scaffold-hopping potential of ligand-based similarity concepts. *ChemMedChem* **2006**, *1* (2), 181-185.
140. Hawkins, P. C. D.; Skillman, A. G.; Nicholls, A., Comparison of Shape-Matching and Docking as Virtual Screening Tools. *J. Med. Chem.* **2007**, *50* (1), 74-82.
141. Proschak, E.; Rupp, M.; Derksen, S.; Schneider, G., Shapelets: Possibilities and limitations of shape-based virtual screening. *J. Comput. Chem.* **2008**, *29* (1), 108-114.
142. Dannhardt, G.; von Gruchalla, M.; Kohl, B. K.; Parsons, C. G., A novel series of 2-carboxytetrahydroquinolines provides new insights into the eastern region of Glycin site NMDA antagonists. *Arch. Pharm. (Weinheim)*. **2000**, *333* (8), 267-274.
143. Zhang, L.; Zhou, W.; Li, D.-H., A descent modified Polak-Ribiere-Polyak conjugate gradient method and its global convergence. *IMA J. Numer. Anal.* **2006**, *26* (4), 629-640.
144. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T., Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112* (16), 6127-6129.
145. Pentikainen, U.; Settimo, L.; Johnson, M.; Pentikainen, O., Subtype selectivity and flexibility of ionotropic glutamate receptors upon antagonist ligand binding. *Org. Biomol. Chem.* **2006**, *4* (6), 1058-1070.
146. Kaye, S. L.; Sansom, M. S.; Biggin, P. C., Molecular dynamics simulations of the ligand-binding domain of an N-methyl-D-aspartate receptor. *J. Biol. Chem.* **2006**, *281* (18), 12736-12742.
147. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739-1749.

148. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47* (7), 1750-1759.
149. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrín, P. C.; Mainz, D. T., Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177-6196.
150. Hawkins, P.; Warren, G.; Skillman, A.; Nicholls, A., How to do an evaluation: pitfalls and traps. *J. Comput. Aided Mol. Des.* **2008**.
151. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrín, P. C.; Mainz, D. T., Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177-6196.
152. Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W., Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **1996**, *274* (5292), 1531-1534.
153. Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A., Site-directed ligand discovery. *PNAS* **2000**, *97* (17), 9367-9372.
154. Sherman, W.; Day, T.; Jacobson, M.; Friesner, R.; Farid, R., Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *J. Med. Chem.* **2006**, *49* (2), 534-553.
155. Sherman, W.; Beard, H. S.; Farid, R., Use of an Induced Fit Receptor Structure in Virtual Screening. *Chem. Biol. Drug Des.* **2006**, *67* (1), 83-84.
156. van Gunsteren, W. F.; Bakowies, D.; Baron, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Geerke, D. P.; Glättli, A.; Hünenberger, P. H.; Kastenholz, M. A.; Oostenbrink, C.; Schenk, M.; Trzesniak, D.; van der Vegt, N. F.; Yu, H. B., Biomolecular modeling: Goals, problems, perspectives. *Angew. Chem. Int. Ed. Engl.* **2006**, *45* (25), 4064-4092.
157. Karplus, M.; McCammon, J. A., Molecular dynamics simulations of biomolecules. *Nat. Struct. Mol. Biol.* **2002**, *9* (9), 646-652.
158. Adcock, S. A.; McCammon, J. A., Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev.* **2006**, *106* (5), 1589-1615.
159. Lennard-Jones, J. E., Cohesion. *Proc. Phys. Soc.* **1931**, (5), 461.
160. McCammon, J. A.; Gelin, B. R.; Karplus, M., Dynamics of folded proteins. *Nature* **1977**, *267* (5612), 585-590.
161. McCammon, J., *Molecular dynamics study of the bovine pancreatic trypsin inhibitor*. CECAM: Orsay, France, 1976.

162. McCammon, J.; Harvey, S. C., *Dynamics of proteins and nucelic acids*. Cambridge University Press: New York, 1987.
163. Duan, Y.; Kollman, P. A., Pathways to a Protein Folding Intermediate Observed in a 1-Microsecond Simulation in Aqueous Solution. *Science* **1998**, *282* (5389), 740-744.
164. Feig, M.; Brooks, C. L., Recent advances in the development and application of implicit solvent models in biomolecule simulations. *Curr. Opin. Struct. Biol.* **2004**, *14* (2), 217-224.
165. Levinthal, C., Are there pathways for protein folding. *J. Chim. Phys.* **1968**, *65* (1), 44-45.
166. de Groot, B. L.; Grubmuller, H., Water Permeation Across Biological Membranes: Mechanism and Dynamics of Aquaporin-1 and GlpF. *Science* **2001**, *294* (5550), 2353-2357.
167. Marelus, J.; Kolmodin, K.; Feierberg, I.; Aqvist, J., Q: a molecular dynamics program for free energy calculations and empirical valence bond simulations in biomolecular systems. *J. Mol. Graph. Model.* **1998**, *16* (4-6).
168. Nutt, D. R.; Smith, J. C., Molecular Dynamics Simulations of Proteins: Can the Explicit Water Model Be Varied? *J. Chem. Theory Comput.* **2007**, *3* (4), 1550-1560.
169. Hoffmann, D.; Kramer, B.; Washio, T.; Steinmetzer, T.; Rarey, M.; Lengauer, T., Two-Stage Method for Protein-Ligand Docking. *J. Med. Chem.* **1999**, *42* (21), 4422-4433.
170. Wang, J.; Kollman, P. A.; Kuntz, I. D., Flexible ligand docking: A multistep strategy approach. *Proteins* **1999**, *36* (1), 1-19.
171. Miranker, A.; Karplus, M., Functionality maps of binding sites: A multiple copy simultaneous search method. *Proteins* **1991**, *11* (1), 29-34.
172. Czerminski, R.; Elber, R., Computational studies of ligand diffusion in globins: I. Leghemoglobin. *Proteins* **1991**, *10* (1), 70-80.
173. Norberg, J.; Nilsson, L., Advances in biomolecular simulations: methodology and recent applications. *Q. Rev. Biophys.* **2003**, *36* (03), 257-306.
174. Park, H.; Lee, S., Free energy perturbation approach to the critical assessment of selective cyclooxygenase-2 inhibitors. *J. Comput. Aided Mol. Des.* **2005**, *19* (1), 17-31.
175. Rossi, K. A.; Weigelt, C. A.; Nayeem, A.; Krystek, S. R., Loopholes and missing links in protein modeling. *Protein Sci.* **2007**, *16* (9), 1999-2012.
176. Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E., Scalable algorithms for molecular dynamics simulations on commodity clusters. In *Proceedings of the 2006 ACM/IEEE conference on Supercomputing*, ACM: Tampa, Florida, 2006.

177. Evans, D. J.; Holian, B. L., The Nose--Hoover thermostat. *The Journal of Chemical Physics* **1985**, *83* (8), 4069-4074.
178. Martyna, G. J.; Tobias, D. J.; Klein, M. L., Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101* (5), 4177-4189.
179. Tuckerman, M.; Berne, B. J.; Martyna, G. J., Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **1992**, *97* (3), 1990-2001.
180. Sagui, C.; Darden, T. A., Molecular Dynamics Simulations of Biomolecules: Long-Range Electrostatic Effects. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28* (1), 155-179.
181. Koehl, P., Electrostatics calculations: latest methodological advances. *Curr. Opin. Struct. Biol.* **2006**.
182. Ciccotti, G.; Ferrario, M.; Ryckaert, J. P., Molecular dynamics of rigid systems in cartesian coordinates A general formulation. *Mol. Phys.* **1982**, *47*, 1253-1264.
183. Willis, B. T. M.; Pryor, A. W., *Thermal vibrations in crystallography*. Cambridge University Press: Cambridge, 1975.
184. Kuriyan, J.; Weis, W. I., Rigid Protein Motion as a Model for Crystallographic Temperature Factors. *PNAS* **1991**, *88* (7), 2773-2777.
185. Sternberg, M. J. E.; Grace, D. E. P.; Phillips, D. C., Dynamic information from protein crystallography: An analysis of temperature factors from refinement of the hen egg-white lysozyme structure. *J. Mol. Biol.* **1979**, *130* (3), 231-253.
186. Nienhaus, G. U.; Heinzl, J.; Huenges, E.; Parak, F., Protein crystal dynamics studied by time-resolved analysis of X-ray diffuse scattering. *Nature* **1989**, *338* (6217), 665-666.
187. Brooks, C. L.; Karplus, M.; Pettitt, B. M., Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics. In *Advances in Chemical Physics*, Wiley: New York, 1988.
188. Lazaridis, T., Inhomogeneous fluid approach to solvation thermodynamics. 1. Theory. *J. Phys. Chem. B-Cond. Phase* **1998**, *102* (18), 3531-3541.
189. Young, T.; Abel, R.; Kim, B.; Berne, B.; Friesner, R., Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *PNAS* **2007**, *104* (3), 808-813.
190. Zwanzig, R. W., High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22* (8), 1420-1426.
191. Lazaridis, T.; Karplus, M., Effective energy function for proteins in solution. *Proteins* **1999**, *35* (2), 133-152.
192. Sanschagrin, P. C.; Kuhn, L. A., Cluster analysis of consensus water sites in thrombin and trypsin shows conservation between serine proteases and contributions to ligand specificity. *Protein Sci.* **1998**, *7* (10), 2054-2064.

193. Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A., Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130* (9), 2817-2831.
194. Baron, B.; Siegel, B.; Harrison, B.; Gross, R.; Hawes, C.; Towers, P., [3H]MDL 105,519, a high-affinity radioligand for the N-methyl-D-aspartate receptor-associated Glycin recognition site. *J. Pharmacol. Exp. Ther.* **1996**, *279* (1), 62-68.
195. Englberger, W.; Przewosny, M.; Maul, C., Glutamate Receptors: 7.2 NMDA Receptors. In *Analgesics*, Helmut Buschmann; Christoph, T.; Maul, C.; Sundermann, B., Eds. 2005; pp 389-428.
196. Danysz, W.; Kozela, E.; Parsons, C.; Sladek, M.; Bauer, T.; Popik, P., Peripherally acting NMDA receptor/GlycinB site receptor antagonists inhibit morphine tolerance. *Neuropharmacology* **2005**, *48* (3), 360-371.
197. Parsons, C. G.; Danysz, W.; Quack, G., Memantine is a clinically well tolerated N-methyl--aspartate (NMDA) receptor antagonist--a review of preclinical data. *Neuropharmacology* **1999**, *38* (6), 735-767.
198. Wold, S.; Ruhe, A.; Wold, H.; Dunn, W. J., The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. *SIAM J. Scientific and Statistical Computing* **1984**, *5* (3), 735-743.
199. Wold, H., *Systems analysis by partial least squares*. Martinus Nijhoff: Boston, 1985.
200. Kulagowski, J. J.; Baker, R.; Curtis, N. R.; Mawer, I. M.; Moseley, A. M.; Ridgill, M. P.; Rowley, M.; Stansfield, I.; Leeson, P. D., 3'-(Arylmethyl)- and 3'-(Aryloxy)-3-phenyl-4-hydroxyquinolin-2(1H)-ones: Orally Active Antagonists of the Glycin Site on the NMDA Receptor. *J. Med. Chem.* **1994**, *37* (10), 1402-1405.
201. Hays, S. J.; Boxer, P. A.; Taylor, C. P.; Vartanian, M. G.; Robichaud, L. J.; Nielsen, E. Ø., N-sulfonyl derivatives of 6,7-dichloro 3,4-dihydro-3-oxo-quinoxalinecarboxylate as Glycin-site NMDA and AMPA antagonists. *Bioorg. Med. Chem. Lett.* **1993**, *3* (1), 77-80.
202. Jansen, M.; Dannhardt, G., Variations of acidic functions at position 2 and substituents at positions 4, 5 and 6 of the indole moiety and their effect on NMDA-Glycin site affinity. *Eur. J. Med. Chem.* **2003**, *38* (10), 855-865.
203. Rowley, M.; Leeson, P. D.; Stevenson, G. I.; Moseley, A. M.; Stansfield, I.; Sanderson, I.; Robinson, L.; Baker, R.; Kemp, J. A.; Marshall, G. R.; et al., 3-Acyl-4-hydroxyquinolin-2(1H)-ones. Systemically active anticonvulsants acting by antagonism at the Glycin site of the N-methyl-D-aspartate receptor complex. *J. Med. Chem.* **1993**, *36* (22), 3386-96.
204. Dean, P.; Wakelin, L., The docking manoeuvre at a drug receptor: a quantum mechanical study of intercalative attack of ethidium and its carboxylated derivative on a DNA fragment. *Phil. Trans. R. Soc. London, Ser. B* **1979**, 571-604.

205. Cho, A. E.; Guallar, V.; Berne, B. J.; Friesner, R., Importance of accurate charges in molecular docking: Quantum mechanical/molecular mechanical (QM/MM) approach. *J. Comput. Chem.* **2005**, *26* (9), 915-931.
206. Acklin, P.; Allgeier, H.; Auberson, Y.; Bischoff, S.; Ofner, S.; Sauer, D.; Schmutz, M., 5-Aminomethylquinoxaline-2, 3-diones, part III: arylamide derivatives as highly potent and selective Glycin-site NMDA receptor antagonists. *Bioorg. Med. Chem. Lett.* **1998**, *8* (5), 493-498.
207. Auberson, Y.; Bischoff, S.; Moretti, R.; Schmutz, M.; Veenstra, S., 5-Aminomethylquinoxaline-2, 3-diones. Part I: A novel class of AMPA receptor antagonists. *Bioorg. Med. Chem. Lett.* **1998**, *8* (1), 65-70.
208. Jansen, M.; Potschka, H.; Brandt, C.; Loscher, W.; Dannhardt, G., Hydantoin-Substituted 4,6-Dichloroindole-2-carboxylic Acids as Ligands with High Affinity for the Glycin Binding Site of the NMDA Receptor. *J. Med. Chem.* **2003**, *46* (1), 64-73.
209. Katayama, S.; Ae, N.; Nagata, R., Synthesis of Tricyclic Indole-2-carboxylic Acids as Potent NMDA-Glycin Antagonists. *J. Org. Chem.* **2001**, *66* (10), 3474-3483.
210. Katayama, S.; Ae, N.; Kodo, T.; Masumoto, S.; Hourai, S.; Tamamura, C.; Tanaka, H.; Nagata, R., Tricyclic Indole-2-carboxylic Acids: Highly in Vivo Active and Selective Antagonists for the Glycin Binding Site of the NMDA Receptor. *J. Med. Chem.* **2003**, *46* (5), 691-701.
211. Micheli, F.; Di Fabio, R.; Baraldi, D.; Conti, N.; Cugola, A.; Gastaldi, P.; Giacobbe, S.; Marchioro, C.; Mugnaini, M.; Rossi, L.; Pecunioso, A.; Pentassuglia, G., Substituted indole-2-carboxylates as potent antagonists of the Glycin binding site associated with the NMDA receptor. *Arch. Pharm. Pharm. Med. Chem.* **1999**, *332* (8), 271-8.
212. Nagata, R.; Tanno, N.; Kodo, T.; Ae, N.; Yamaguchi, H.; Nishimura, T.; Antoku, F.; Tatsuno, T.; Kato, T., Tricyclic Quinoxalinediones: 5,6-Dihydro-1H-pyrrolo[1,2,3-de]quinoxaline-2,3-diones and 6,7-Dihydro-1H,5H-pyrido[1,2,3-de]quinoxaline-2,3-diones as Potent Antagonists for the Glycin Binding Site of the NMDA Receptor. *J. Med. Chem.* **1994**, *37* (23), 3956-3968.
213. Nagata, R.; Kodo, T.; Yamaguchi, H.; Tanno, N., Structure-activity relationships of tricyclic quinoxalinediones as potent antagonists for the Glycin binding site of the NMDA receptor 2. *Bioorg. Med. Chem. Lett.* **1995**, *5* (14), 1533-1536.
214. Catarzi, D.; Colotta, V.; Varano, F.; Filacchioni, G.; Galli, A.; Costagli, C.; Carla, V., Synthesis, Ionotropic Glutamate Receptor Binding Affinity, and Structure-Activity Relationships of a New Set of 4,5-Dihydro-8-heteroaryl-4-oxo-1,2,4-triazolo[1,5-a]quinoxaline-2-carboxylates Analogues of TQX-173. *J. Med. Chem.* **2001**, *44* (19), 3157-3165.
215. Brown, D. G.; Urbanek, R. A.; Bare, T. M.; McLaren, F. M.; Horchler, C. L.; Murphy, M.; Steelman, G. B.; Empfield, J. R.; Forst, J. M.; Herzog, K. J.; Xiao, W.; Dyroff, M. C.; Lee, C. M.; Trivedi, S.; Neilson, K. L.; Keith, R. A., Synthesis

- of 7-chloro-2,3-dihydro-2-[1-(pyridinyl)alkyl]-pyridazino[4,5-b]quinoline-1,4,10(5H)-triones as NMDA Glycin-site antagonists. *Bioorg. Med. Chem. Lett.* **2003**, *13* (20), 3553-6.
216. Cai, S. X.; Zhou, Z.-L.; Huang, J.-C.; Whittemore, E. R.; Egbuwoku, Z. O.; Lu, Y.; Hawkinson, J. E.; Woodward, R. M.; Weber, E.; Keana, J. F. W., Synthesis and Structure-Activity Relationships of 1,2,3,4-Tetrahydroquinoline-2,3,4-trione 3-Oximes: Novel and Highly Potent Antagonists for NMDA Receptor Glycin Site. *J. Med. Chem.* **1996**, *39* (17), 3248-3255.
217. Cai, S. X.; Zhou, Z.-L.; Huang, J.-C.; Whittemore, E. R.; Egbuwoku, Z. O.; Hawkinson, J. E.; Woodward, R. M.; Weber, E.; Keana, J. F. W., Structure-Activity Relationships of 4-Hydroxy-3-nitroquinolin-2(1H)-ones as Novel Antagonists at the Glycin Site of N-Methyl-d-aspartate Receptors. *J. Med. Chem.* **1996**, *39* (23), 4682-4686.
218. Cai, S. X.; Kher, S. M.; Zhou, Z.-L.; Ilyin, V.; Espitia, S. A.; Tran, M.; Hawkinson, J. E.; Woodward, R. M.; Weber, E.; Keana, J. F. W., Structure-Relationships of Alkyl- and Alkoxy-Substituted 1,4-Dihydroquinoxaline-2,3-diones: Potent and Systemically Active Antagonists for the Glycin Site of the NMDA Receptor. *J. Med. Chem.* **1997**, *40* (5), 730-738.
219. Conti, P.; De Amici, M.; Grazioso, G.; Roda, G.; Barberis Negra, F. F.; Nielsen, B.; Stensbol, T. B.; Madsen, U.; Brauner-Osborne, H.; Frydenvang, K.; De Sarro, G.; Toma, L.; De Micheli, C., Design, Synthesis, and Pharmacological Characterization of Novel, Potent NMDA Receptor Antagonists. *J. Med. Chem.* **2004**, *47* (27), 6740-6748.
220. Fray, M. J.; Bull, D. J.; Carr, C. L.; Gautier, E. C. L.; Mowbray, C. E.; Stobie, A., Structure-Activity Relationships of 1,4-Dihydro-(1H,4H)-quinoxaline-2,3-diones as N-Methyl-d-aspartate (Glycin Site) Receptor Antagonists. 1. Heterocyclic Substituted 5-Alkyl Derivatives. *J. Med. Chem.* **2001**, *44* (12), 1951-1962.
221. Zhou, Z. L.; Kher, S. M.; Cai, S. X.; Whittemore, E. R.; Espitia, S. A.; Hawkinson, J. E.; Tran, M.; Woodward, R. M.; Weber, E.; Keana, J. F., Synthesis and SAR of novel di- and trisubstituted 1,4-dihydroquinoxaline-2,3-diones related to licostinel (Acea 1021) as NMDA/Glycin site antagonists. *Bioorg. Med. Chem.* **2003**, *11* (8), 1769-80.
222. Colotta, V.; Catarzi, D.; Varano, F.; Calabri, F. R.; Filacchioni, G.; Costagli, C.; Galli, A., 3-Hydroxy-quinazoline-2,4-dione as a useful scaffold to obtain selective Gly/NMDA and AMPA receptor antagonists. *Bioorg. Med. Chem. Lett.* **2004**, *14* (9), 2345-2349.
223. Di Fabio, R.; Micheli, F.; Baraldi, D.; Bertani, B.; Conti, N.; Dal Forno, G.; Feriani, A.; Donati, D.; Marchioro, C.; Messeri, T.; Missio, A.; Pasquarello, A.; Pentassuglia, G.; Pizzi, D. A.; Provera, S.; Quaglia, A. M.; Sabbatini, F. M., Benzoazepine derivative as potent antagonists of the Glycin binding site associated to the NMDA receptor. *Il Farmaco* **2003**, *58* (9), 723-738.

224. Carling, R. W.; Leeson, P. D.; Moseley, A. M.; Baker, R.; Foster, A. C.; Grimwood, S.; Kemp, J. A.; Marshall, G. R., 2-Carboxytetrahydroquinolines. Conformational and stereochemical requirements for antagonism of the Glycin site on the N-methyl-D-aspartate (NMDA) receptor. *J. Med. Chem.* **1992**, *35* (11), 1942-1953.
225. Carling, R. W.; Leeson, P. D.; Moore, K. W.; Smith, J. D.; Moyes, C. R.; Mawer, I. M.; Thomas, S.; Chan, T.; Baker, R., 3-Nitro-3,4-dihydro-2(1H)-quinolones. Excitatory amino acid antagonists acting at Glycin-site NMDA and (RS)- α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors. *J. Med. Chem.* **1993**, *36* (22), 3397-3408.
226. Carling, R. W.; Leeson, P. D.; Moore, K. W.; Moyes, C. R.; Duncton, M.; Hudson, M. L.; Baker, R.; Foster, A. C.; Grimwood, S.; Kemp, J. A.; Marshall, G. R.; Tricklebank, M. D.; Saywell, K. L., 4-Substituted-3-phenylquinolin-2(1H)-ones: Acidic and Nonacidic Glycin Site N-Methyl-d-aspartate Antagonists with in Vivo Activity. *J. Med. Chem.* **1997**, *40* (5), 754-765.
227. Curtis, N. R.; Kulagowski, J. J.; Leeson, P. D.; Mawer, I. M.; Ridgill, M. P.; Rowley, M.; Grimwood, S.; Marshall, G. R., Synthesis and SAR of diiodotyrosine-derived Glycin-site N-methyl-D-aspartate receptor ligands. *Bioorg. Med. Chem. Lett.* **1996**, *6* (10), 1145-1150.
228. Leeson, P. D.; Carling, R. W.; Moore, K. W.; Moseley, A. M.; Smith, J. D.; Stevenson, G.; Chan, T.; Baker, R.; Foster, A. C., 4-Amido-2-carboxytetrahydroquinolines. Structure-activity relationships for antagonism at the Glycin site of the NMDA receptor. *J. Med. Chem.* **1992**, *35* (11), 1954-1968.
229. Leeson, P. D.; Baker, R.; Carling, R. W.; Curtis, N. R.; Moore, K. W.; Williams, B. J.; Foster, A. C.; Donald, A. E.; Kemp, J. A.; Marshall, G. R., Kynurenic acid derivatives. Structure-activity relationships for excitatory amino acid antagonism and identification of potent and selective antagonists at the Glycin site on the N-methyl-D-aspartate receptor. *J. Med. Chem.* **1991**, *34* (4), 1243-52.
230. Mawer, I.; Kulagowski, J.; Leeson, P.; Grimwood, S.; Marshall, G., Tetramic acids as novel Glycin site antagonists. *Bioorg. Med. Chem. Lett.* **1995**, *5* (22), 2643-2648.
231. Rowley, M.; Leeson, P.; Grimwood, S.; Marshall, G.; Saywell, K., 5, 6, 7, 8-Tetrahydroquinolones as antagonists at the Glycin site of the NMDA receptor. *Bioorg. Med. Chem. Lett.* **1995**, *5* (18), 2089-2092.
232. Ametamey, S. M.; Kocic, M.; Carrey-Remy, N.; Blauenstein, P.; Willmann, M.; Bischoff, S.; Schmutz, M.; Schubiger, P. A.; Auberson, Y. P., Synthesis, radiolabelling and biological characterization of (D)-7-iodo-N-(1-phosphonoethyl)-5-aminomethylquinoxaline-2,3-dione, a Glycin-binding site antagonist of NMDA receptors. *Bioorg. Med. Chem. Lett.* **2000**, *10* (1), 75-8.
233. Auberson, Y. P.; Acklin, P.; Allgeier, H.; Biollaz, M.; Bischoff, S.; Ofner, S.; Veenstra, S. J., 5-aminomethylquinoxaline-2,3-diones. Part II: N-aryl derivatives

- as novel NMDA/Glycin and AMPA antagonists. *Bioorg. Med. Chem. Lett.* **1998**, *8* (1), 71-74.
234. Zhou, Z. L.; Navratil, J. M.; Cai, S. X.; Whittemore, E. R.; Espitia, S. A.; Hawkinson, J. E.; Tran, M.; Woodward, R. M.; Weber, E.; Keana, J. F., Synthesis and SAR of 5-, 6-, 7- and 8-aza analogues of 3-aryl-4-hydroxyquinolin-2(1H)-one as NMDA/Glycin site antagonists. *Bioorg. Med. Chem.* **2001**, *9* (8), 2061-71.
235. Zheng, H.; Espitia, S.; Ilyin, V.; Hawkinson, J.; Woodward, R.; Weber, E.; Keana, J., 1-Hydroxy-1, 4-dihydroquinoxaline-2, 3-diones: Novel antagonists at NMDA receptor Glycin sites. *Bioorg. Med. Chem. Lett.* **1996**, *6* (4), 439-440.
236. Jackson, P.; Davenport, T.; Garcia, L.; McKinney, J.; Melville, M.; Harris, G.; Chapdelaine, M.; Damewood, J.; Pullan, L.; Goldstein, J., Synthesis and biological activity of a series of 4-aryl substituted benz [b] azepines: Antagonists at the strychnine-insensitive Glycin site. *Bioorg. Med. Chem. Lett.* **1995**, *5* (24), 3097-3100.
237. Guzikowski, A. P.; Hawkinson, J.; Weber, E.; Keana, J. F. W., 6,7,8,9-tetrahydro-3-hydroxy-1H-1-benzazepine-2,5-diones via a diels-alder reaction: antagonists with a non-planar hydrophobic region for NMDA receptor Glycin sites. *Bioorg. Med. Chem. Lett.* **1995**, *5* (22), 2747-2748.
238. Jimonet, P.; Audiau, F.; Aloup, J.; Barreau, M.; Blanchard, J.; Bohme, A.; Boireau, A.; Cheve, M.; Damour, D.; Doble, A., Synthesis and SAR of 2H-1, 2, 4-benzothiadiazine-1, 1-dioxide-3-carboxylic acid derivatives as novel potent Glycin antagonists of the NMDA receptor-channel complex. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2735-2735.
239. Moroni, F.; Alesiani, M.; Galli, A.; Mori, F.; Pecorari, R.; Carla, V.; Cherici, G.; Pellicciari, R., Thiokynurenates: a new group of antagonists of the Glycin modulatory site of the NMDA receptor. *Eur. J. Pharmacol.* **1991**, *199* (2), 227-32.
240. Salituro, F. G.; Tomlinson, R. C.; Baron, B. M.; Demeter, D. A.; Weintraub, H. J. R.; McDonald, I. A., Design, synthesis and molecular modeling of 3-acylamino-2-Carboxyindole NMDA receptor Glycin-site antagonists. *Bioorg. Med. Chem. Lett.* **1991**, *1* (9), 455-460.
241. Buchstaller, H.; Siebert, C.; Steinmetz, R.; Frank, I.; Berger, M.; Gottschlich, R.; Leibrock, J.; Krug, M.; Steinhilber, D.; Noes, C., Synthesis of Thieno [2, 3-b] Pyridinones Acting as Cytoprotectants and as Inhibitors of [3H] Glycin Binding to the N-Methyl-d-aspartate (NMDA) Receptor. *J. Med. Chem.* **2006**, *49* (3), 864-871.
242. Salituro, F. G.; Harrison, B. L.; Baron, B. M.; Nyce, P. L.; Stewart, K. T.; Kehne, J. H.; White, H. S.; McDonald, I. A., 3-(2-Carboxyindol-3-yl)propionic acid-based antagonists of the NMDA (N-methyl-D-aspartic acid) receptor associated Glycin binding site. *J. Med. Chem.* **1992**, *35* (10), 1791-1799.

243. Keana, J. F. W.; Kher, S. M.; Cai, S. X.; Dinsmore, C. M.; Glenn, A. G.; Guastella, J.; Huang, J.-C.; Ilyin, V.; Lu, Y., Synthesis and Structure-Activity Relationships of Substituted 1,4-Dihydroquinoxaline-2,3-diones: Antagonists of N-Methyl-D-aspartate (NMDA) Receptor Glycin Sites and Non-NMDA Glutamate Receptors. *J. Med. Chem.* **1995**, *38* (22), 4367-4379.
244. Guzikowski, A. P.; Cai, S. X.; Espitia, S. A.; Hawkinson, J. E.; Huettner, J. E.; Nogales, D. F.; Tran, M.; Woodward, R. M.; Weber, E.; Keana, J. F. W., Analogs of 3-Hydroxy-1H-1-benzazepine-2,5-dione: Structure-Activity Relationship at N-Methyl-d-aspartate Receptor Glycin Sites. *J. Med. Chem.* **1996**, *39* (23), 4643-4653.
245. Cai, S. X.; Kher, S. M.; Zhou, Z.-L.; Ilyin, V.; Espitia, S. A.; Tran, M.; Hawkinson, J. E.; Woodward, R. M.; Weber, E.; Keana, J. F. W., Structure-Activity Relationships of Alkyl- and Alkoxy-Substituted 1,4-Dihydroquinoxaline-2,3-diones: Potent and Systemically Active Antagonists for the Glycin Site of the NMDA Receptor. *J. Med. Chem.* **1997**, *40* (5), 730-738.
246. Balsamini, C.; Bedini, A.; Diamantini, G.; Spadoni, G.; Tontini, A.; Tarzia, G.; Di Fabio, R.; Feriani, A.; Reggiani, A.; Tedesco, G.; Valigi, R., (E)-3-(2-(N-Phenylcarbamoyl)vinyl)pyrrole-2-carboxylic Acid Derivatives. A Novel Class of Glycin Site Antagonists. *J. Med. Chem.* **1998**, *41* (6), 808-820.
247. Catarzi, D.; Colotta, V.; Varano, F.; Cecchi, L.; Filacchioni, G.; Galli, A.; Costagli, C., 4,5-Dihydro-1,2,4-triazolo[1,5-a]quinoxalin-4-ones: Excitatory Amino Acid Antagonists with Combined Glycin/NMDA and AMPA Receptor Affinity. *J. Med. Chem.* **1999**, *42* (13), 2478-2484.
248. Di Fabio, R.; Conti, N.; De Magistris, E.; Feriani, A.; Provera, S.; Sabbatini, F. M.; Reggiani, A.; Rovatti, L.; Barnaby, R. J., Substituted Analogues of GV150526 as Potent Glycin Binding Site Antagonists in Animal Models of Cerebral Ischemia. *J. Med. Chem.* **1999**, *42* (18), 3486-3493.
249. Kreimeyer, A.; Laube, B.; Sturgess, M.; Goeldner, M.; Foucaud, B., Evaluation and Biological Properties of Reactive Ligands for the Mapping of the Glycin Site on the N-Methyl-d-aspartate (NMDA) Receptor. *J. Med. Chem.* **1999**, *42* (21), 4394-4404.
250. Gray, N. M.; Dappen, M. S.; Cheng, B. K.; Cordi, A. A.; Biesterfeldt, J. P.; Hood, W. F.; Monahan, J. B., Novel indole-2-carboxylates as ligands for the strychnine-insensitive N-methyl-D-aspartate-linked Glycin receptor. *J. Med. Chem.* **1991**, *34* (4), 1283-1292.
251. Jimonet, P.; Ribeill, Y.; Bohme, G. A.; Boireau, A.; Cheve, Michel; Damour, D.; Doble, A.; Genevois-Borella, A.; Herman, F.; Imperato, A.; Le Guern, S.; Manfre, F.; Pratt, J.; Randle, J. C. R.; Stutzmann, J.-M.; Mignani, S., Indeno[1,2-b]pyrazin-2,3-diones: A New Class of Antagonists at the Glycin Site of the NMDA Receptor with Potent in Vivo Activity. *J. Med. Chem.* **2000**, *43* (12), 2371-2381.

252. Guzikowski, A. P.; Whittemore, E. R.; Woodward, R. M.; Weber, E.; Keana, J. F. W., Synthesis of Racemic 6,7,8,9-Tetrahydro-3-hydroxy-1H-1-benzazepine-2,5-diones as Antagonists of N-Methyl-d-aspartate (NMDA) and Alpha-Amino-3-hydroxy-5-methylisoxazole-4-propionic Acid (AMPA) Receptors. *J. Med. Chem.* **1997**, *40* (15), 2424-2429.
253. Buchstaller, H.-P.; Siebert, C. D.; Steinmetz, R.; Frank, I.; Berger, M. L.; Gottschlich, R.; Leibrock, J.; Krug, M.; Steinhilber, D.; Noe, C. R., Synthesis of Thieno[2,3-b]Pyridinones Acting as Cytoprotectants and as Inhibitors of [3H]Glycin Binding to the N-Methyl-d-aspartate (NMDA) Receptor. *J. Med. Chem.* **2006**, *49* (3), 864-871.
254. Calabri, F. R.; Colotta, V.; Catarzi, D.; Varano, F.; Lenzi, O.; Filacchioni, G.; Costagli, C.; Galli, A., Synthesis and pharmacological studies at the Gly/NMDA, AMPA and Kainate receptors of new oxazolo[4,5-c]quinolin-4-one derivatives bearing different substituents at position-2 and on the fused benzo ring. *Eur. J. Med. Chem.* **2005**, *40* (9), 897-907.

Danksagung

Die vorliegende Arbeit entstand in der Zeit von Februar 2007 bis April 2009 an der Professur für Chemie- und Bioinformatik unter Leitung von Prof. Dr. Gisbert Schneider. Für seine Bereitschaft, eine externe Doktorarbeit zu begleiten, die inspirierenden wissenschaftlichen Diskussionen und die Freiheit, meine eigenen Ideen verwirklichen zu können, sei meinem Doktorvater Gisbert Schneider herzlich gedankt. Seine konstruktive Kritik und Anregungen sowie seine persönliche Unterstützung haben maßgeblich zum Gelingen dieser Arbeit beigetragen.

Mein persönlicher Dank gilt Prof. Dr. Tanja Weil (Merz Pharmaceuticals, jetzt National University of Singapore), ohne deren tatkräftigen Beistand durch zahlreiche Anregungen und Diskussionen diese Arbeit nicht zustande gekommen wäre. Besonders danken möchte ich für ihr Vertrauen, mir bereits während meiner Promotionsphase in hohem Maße Verantwortung zu übertragen.

Frau Dr. Manuela López de la Paz danke ich sehr für die Betreuung und den immer hilfreichen persönlichen Beistand zu Beginn meiner Arbeit.

Den Kolleginnen und Kollegen der Abteilung Chemical R&D, insbesondere meiner Gruppe Drug Design (Dr. Manuela López de la Paz, Dr. Udo Meyer) sowie Dr. Sibylle Müller, Dr. Angela Bauer, Dr. Lutz Franke und Dr. Holger Kubas danke ich sehr für die immerwährende Unterstützung, spannende wissenschaftliche Diskussionen und vor allem für ein Arbeitsumfeld, in das ich herzlich aufgenommen wurde und in dem ich sehr gerne weiterhin tätig bin.

Den Kolleginnen und Kollegen der Abteilung In Vitro Screening bei Merz Pharmaceuticals, insbesondere Dr. Claudia Jatzke, Tanja Bauer, Christina Wollenburg und Dr. Meik Sladek danke ich für die Durchführung der Messungen zu den im Rahmen dieser Arbeit identifizierten Substanzen und die stetige Bereitschaft zur Erläuterung von experimentellen Fragestellungen.

Mein Dank gilt auch den Kolleginnen und Kollegen des Instituts für Organische Synthese in Riga (Lettland), insbesondere Dr. Valerjans Kauss und Dr. Ronalds Zemribo sowie Dr. Alexandrs Gutcaits für zahlreiche faszinierende Diskussionen zum Thema Glycin_B. Dr. Kristaps Jaudzems (IOS, jetzt Scripps Research Institute, La Jolla) danke ich für die Durchführung von STD-NMR-Experimenten zur Validierung möglicher Bindungsmodi meiner Liganden. Leider konnte dieses spannende Feld nicht in die vorliegende Arbeit mit einfließen.

Den Kolleginnen und Kollegen der Abteilungen Non-Clinical Safety, In Vivo Pharmacology, Biotechnology und Patent Affairs gilt mein Dank für vielerlei Unterstützung und interessante Blicke „über den Tellerrand“. Hier sei Dr. Nadine Zidek und Dr. Kai-Uwe Klein für spannende Diskussionen zum Thema vorklinische Substanzsicherheit, sowie Dr. Julia Hackenberg und Dr. Andreas Bane-

mann für die hilfreiche Unterstützung bei der Patenterstellung und Freigabe meiner Publikationen und nicht zuletzt dieser Arbeit gedankt.

Jörg Wilhelmi (Max-Planck-Institut für Polymerforschung, Mainz) danke ich für die gemeinsame Durchführung einer wissenschaftlichen Kooperation auf dem spannenden Feld der Biotechnologie, insbesondere des Einbaus chemisch modifizierter Aminosäuren in Proteine. Dr. Radu-Adrian Gropeanu sei für die Durchführung von ITC-Experimenten für einige meiner Substanzen gedankt.

Besonderer Dank gebührt Swetlana, Bettina, Lutz, Sina, Silke und Henrik für ihre Freundschaft und moralische Unterstützung, wenn es mal wieder notwendig wurde.

Meiner Familie, insbesondere meiner Frau Dr. Agnes Krüger, meiner Tochter Clara, meinen Eltern Mathilde und Hans-Jürgen und meinem Bruder Dr. Stefan Krüger sowie meinen Schwiegereltern Maria und Bernhard und meiner Schwägerin Stefanie gilt mein ganz persönlicher Dank. Ohne eure Unterstützung – während des Studiums, vor und während der Promotion - wäre diese Arbeit nicht entstanden. Ihr habt die Hoffnung nicht aufgegeben; ich hoffe, es hat sich gelohnt.

Lebenslauf

Björn Achim Krüger, geb. 10.01.1977 in Dillingen/Saar

verheiratet, eine Tochter

Akademischer Werdegang

- seit 02/2007 Anfertigung der Dissertation „Einfluss der Proteinflexibilität auf die Identifikation neuer Antagonisten für die Glycin_B-Bindungsstelle des NMDA-Rezeptors“
Prof. Dr. Gisbert Schneider, Professur für Chemie- und Bioinformatik
Goethe-Universität, Frankfurt am Main
- Durchführung in der Abteilung Chemical R&D,
Merz Pharmaceuticals GmbH, Frankfurt am Main
- 10/2000 – 12/2006 Studium der Bioinformatik (Diplom)
Diplomarbeit „Evaluierung von Fitnessfunktionen für das *de novo* Design“, Note 1.0
Prof. Dr. Gisbert Schneider, Professur für Chemie- und Bioinformatik
Johann Wolfgang Goethe-Universität, Frankfurt am Main
- Durchführung in der Abteilung Drug Design,
Sanofi-Aventis Deutschland GmbH, Frankfurt am Main
- Abschluss: Diplom-Bioinformatik**, Gesamtnote „sehr gut (1.1)“
- 09/1996 – 10/2000 Studium der Pharmazie
Johann Wolfgang Goethe-Universität Frankfurt am Main

Beruflicher Werdegang

- seit 01/2009 Section Head Drug Design
Chemical R&D, Merz Pharmaceuticals GmbH, Frankfurt am Main
- 04/2008 – 12/2008 Scientific Expert Drug Design,
Chemical R&D, Merz Pharmaceuticals GmbH, Frankfurt am Main

Schulische Ausbildung

- 1987 – 1996 Albert-Schweitzer-Gymnasium Dillingen/Saar
Allgemeine Hochschulreife (Note 1.4)

Stipendien

02/2007 – 04/2008 Stipendiat der Friedrich-Merz-Stiftung

Veröffentlichungen

- Krueger B.A.; Dietrich A.; Baringhaus K.-H.; Schneider G., Scaffold-hopping potential of fragment-based *de novo* design: The chances and limits of variation. *Combinatorial Chemistry and High-Throughput Screening* **2009**, 12 (4), 383-396.
- Krueger B.A.; Weil T.; Schneider G., Comparative Virtual Screening and QSAR Modeling for NMDA-Glycin_B Antagonists. *Journal of Computer-Aided Molecular Design* **2009**, submitted

Konferenzbeiträge

- Krüger B.; Schüller A.; Scheffer U.; Göbel M.; Schneider G., Seeking out novel lead candidates by *de novo* design and molecular library optimization. Poster, 3rd German Conference on Chemoinformatics, Goslar, Germany, 2007.

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides Statt, dass ich die vorgelegte Dissertation

„Einfluss der Proteinflexibilität auf die Identifikation
neuer Antagonisten für die Glycin_B-Bindungsstelle des NMDA-Rezeptors“

selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, dass aus Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche mit Angabe der betreffenden Schrift bezeichnet sind, nicht stattgefunden haben.

Frankfurt am Main, den 6. Juli 2009

.....
Björn Krüger

Anhang

A.1 Literatur-Referenzliganden

SMILES	Publikation	p/C_{50}
<chem>c1c(CNC(=O)c2ccccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>	Acklin <i>et al.</i> ²⁰⁶	7,22
<chem>c1c(CNC(=O)c2occc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,40
<chem>c1c(CNC(=O)c2sccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		8,00
<chem>c1c(CNC(=O)c2sccc2C)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,70
<chem>c1c(CNC(=O)c2sc(C)cc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,52
<chem>c1c(CNC(=O)c2ccsc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		8,00
<chem>c1c(CNC(=O)c2n(C)ccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,70
<chem>c1c(CNC(=O)c2ccnc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		6,55
<chem>c1c(CNC(=O)c2ccncc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		6,60
<chem>c1c(CNC(=O)c2ccncc2C)c(NC(=O)C(=O)N3)c3cc1Br</chem>		5,74
<chem>c1c(CNC(=O)c2ccc(C)nc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		6,96
<chem>c1c(CNC(=O)c2ccccc2O)c(NC(=O)C(=O)N3)c3cc1[N+][O-]=O</chem>		7,40
<chem>c1c(CNC(=O)c2cc(O)c(O)c2)c(NC(=O)C(=O)N3)c3cc1[N+][O-]=O</chem>		6,82
<chem>c1c(CNC(=O)c2cc(OC)c(OC)c2)c(NC(=O)C(=O)N3)c3cc1[N+][O-]=O</chem>		6,70
<chem>c1c(CNC(=O)c2ccccc2N)c(NC(=O)C(=O)N3)c3cc1[N+][O-]=O</chem>		6,77
<chem>c1c(CNC(=O)c2ccc(SN(O)O)cc2)c(NC(=O)C(=O)N3)c3cc1[N+][O-]=O</chem>		6,20
<chem>c1c(CNC(=O)Cc2ccsc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		8,00
<chem>c1c(CNC(=O)Cc2ccccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		8,00
<chem>c1c(CNC(=O)Cc2ccccc2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>		8,00
<chem>c1c(CNC(=O)Cc2ccccc2)c(NC(=O)C(=O)N3)c3cc1</chem>		7,05
<chem>c1c(CNC(=O)Cc2ccccc2)c(NC(=O)C(=O)N3)c3cc1[N+][O-]=O</chem>		7,40
<chem>c1c(CNC(=O)Cc2c(Cl)cccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		8,15
<chem>c1c(CNC(=O)Cc2cccc(Cl)c2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,40
<chem>c1c(CNC(=O)Cc2ccc(Cl)cc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		4,70
<chem>c1c(CNC(=O)Cc2cccc(OC)c2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,52
<chem>c1c(CNC(=O)Cc2c([N+][O-]=O)cccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,70
<chem>c1c(CNC(=O)Cc2cccc([N+][O-]=O)c2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,40
<chem>c1c(CNC(=O)Cc2ccc([N+][O-]=O)cc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		5,92
<chem>c1c(CNC(=O)Cc2ccc(OC)c(OC)c2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,30
<chem>c1c(CNC(=O)Cc2ccc3c(OCO3)c2)c(NC(=O)C(=O)N4)c4cc1Br</chem>		5,41
<chem>c1c(CNC(=O)Nc2ccccc2)c(NC(=O)C(=O)N3)c3cc1[N+][O-]=O</chem>		6,05
<chem>c1c(CNC(=O)Nc2ccccc2OC)c(NC(=O)C(=O)N3)c3cc1[N+][O-]=O</chem>		7,52
<chem>c1c(CNC(=O)Nc2ccc(OC)cc2)c(NC(=O)C(=O)N3)c3cc1[N+][O-]=O</chem>		7,52
<chem>c1c(CNC(=O)Nc2ccc(OC(F)F)cc2)c(NC(=O)C(=O)N3)c3cc1[N+][O-]=O</chem>		6,46
<chem>c1c(CNC(=O)C(C)c2ccccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		8,00
<chem>c1c(CNC(=O)C(OC(C)=O)c2ccccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,40

<chem>c1c(CNC(=O)C(O)c2ccccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		7,30
<chem>c1c(CNC(=O)C(=O)c2ccccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		5,40
<chem>c1c(CNC(=O)Oc2ccccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		5,82
<chem>c1c(CNC(=O)OCc2ccccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		6,96
<chem>c1c(CNC(=O)CCc2ccccc2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		6,07
<chem>c1c(CN2CCCC2C(O)=O)c(NC(=O)C(=O)N3)c3cc1Br</chem>	Auberson <i>et al.</i> ²⁰⁷	6,08
<chem>c1c(CN2CCC(C(O)=O)CC2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		5,41
<chem>c1c(CN2CCC(C(OC)=O)CC2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		6,23
<chem>c1c(CN2CCC(c3ccccc3)CC2)c(NC(=O)C(=O)N4)c4cc1[N+][[O-]]=O</chem>		6,48
<chem>c1c(CN2CC(C)CCC2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		5,52
<chem>c1c(CN2CCC2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		5,42
<chem>c1c(CN2[C@@H](C(O)=O)CC2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		5,70
<chem>c1c(CN2[C@@H](C(O)=O)[C@H](O)CC2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		5,82
<chem>c1c(CN2CC(O)CC2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		6,55
<chem>c1c(CNc2ccccc2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		5,85
<chem>c1c(CNc2c(F)cccc2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		6,30
<chem>c1c(CNc2cccc(F)c2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		6,14
<chem>c1c(CNc2ccc(F)cc2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		6,12
<chem>c1c(CNc2c(C(O)=O)cccc2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		7,30
<chem>c1c(CNc2scc(CC(OCC)=O)n2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		6,39
<chem>c1c(CNc2scc(CC(O)=O)n2)c(NC(=O)C(=O)N3)c3cc1[N+][[O-]]=O</chem>		6,92
<chem>c1c(CNc2scc(CC(OCC)=O)n2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		6,47
<chem>c1c(CNc2scc(CC(O)=O)n2)c(NC(=O)C(=O)N3)c3cc1Br</chem>		6,92
<chem>c1c(CNc2ccccc2C(O)=O)c(NC(=O)C(=O)N3)c3cc1Cl</chem>		7,52
<chem>c1c(Cl)c(c(\C=C(/c2ccc(F)cc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	Baron <i>et al.</i> ¹³⁵	7,45
<chem>c1c(Cl)c(c(\C=C(/c2ccc(Br)cc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		7,54
<chem>c1c(Cl)c(c(\C=C(\C(=O)Nc2ccccc2)c(C(O)=O)[nH]3)c3cc1Cl</chem>		7,82
<chem>c1c(Cl)c(c(\C=C(\C(=O)N2CCCC2)c(C(O)=O)[nH]3)c3cc1Cl</chem>		7,42
<chem>c1c(Cl)c(c(\C=C(/c2occc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		7,66
<chem>c1c(Cl)c(c(\C=C(/c2ccccc2)\C(N(C)C)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		7,17
<chem>c1c(Cl)c(c(\C=C(/c2sccc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		8,05
<chem>c1c(Cl)c(c(\C=C(/c2ccc(Cl)cc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		7,19
<chem>c1c(Cl)c(c(\C=C(/c2ccccc2)\C(NC)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		7,30
<chem>c1c(Cl)c(c(\C=C(/c2ccccc2)\C(Nc3ccccc3)=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		7,82
<chem>c1c(Cl)c(c(\C=C(/c2ccccc2)\C(NC3ccccc3)=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		7,51
<chem>c1c(Cl)c(c(\C=C(/c2cccc(cccc3)c23)\C(O)=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		6,84
<chem>c1c(Cl)c(c(\C=C(/c2ccc(C)cc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		6,98
<chem>c1c(Cl)c(c(\C=C(/c2ccc(C(F)F)cc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		7,05
<chem>c1c(Cl)c(c(\C=C(/c2ccc(N)cc2)\C#N)c(C(O)=O)[nH]3)c3cc1Cl</chem>		5,33
<chem>c1c(Cl)c(c(\C=C(/c2cccc([N+][[O-]]=O)c2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		7,68
<chem>c1c(Cl)c(c(\C=C(/c2ccc(OC)cc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		6,87

<chem>c1c(Cl)c(c(\C=C(/c2cnccc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	6,90
<chem>c1c(Cl)c(c(\C=C(/c2ccc(NC(=O)C)cc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	6,34
<chem>c1c(Cl)c(c(\C=C(/c2cc(OC)ccc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	6,93
<chem>c1c(Cl)c(c(\C=C(/C(O)=O)\c2cccc2)c(C(O)=O)[nH]3)c3cc1Cl</chem>	6,80
<chem>c1c(Cl)c(c(\C=C(/c2cccc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	7,96
<chem>c1c(Cl)c(c(\C=C(/c2cccc2)\C(NCCc3cccc3)=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>	6,65
<chem>c1c(Cl)c(c(\C=C(/c2ccncc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	6,82
<chem>c1c(Cl)c(c(\C=C(/c2c(Cl)cccc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	6,67
<chem>c1c(Cl)c(c(\C=C(/c2cc(N(C)C)ccc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	6,65
<chem>c1c(Cl)c(c(\C=C(/c2c(Cl)cc(Cl)cc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	6,94
<chem>c1c(Cl)c(c(\C=C(/c2cccc2)\C(N)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	7,70
<chem>c1c(Cl)c(c(\C=C(/c2ccc(NS(=O)(=O)C)cc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	6,78
<chem>c1c(Cl)c(c(\C=C(/c2cc(NC(=O)c3cccc3)ccc2)\C(O)=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>	6,29
<chem>c1c(Cl)c(c(\C=C(/c2nccc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	5,99
<chem>c1c(Cl)c(c(\C=C(/c2cc(NC(=O)OC)ccc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	6,94
<chem>c1c(Cl)c(c(\C=C(/c2cc(N)ccc2)\C(O)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>	7,61
<chem>c1(Cl)c([N+]([O-])=O)c(NC(=O)C(=O)N2)c2cc1Cl</chem>	7,70
<chem>c(c(cc1c2c(\C=C(/c3ccc3)\C(O)=O)c(C(O)=O)[nH]1)Cl)c2Cl</chem>	7,68
<chem>c(c(cc1c2c(\C=C(/c3ccc3)\C(O)=O)c(C(O)=O)[nH]1)Cl)c2Cl</chem>	7,85
<chem>c1cc(C(O)=C(c2cccc(Oc3cccc3)c2)C(=O)N4)c4cc1Cl</chem>	7,88
<chem>c1cc(nc(C(O)=O)[nH]2)c2cc1Cl</chem>	Dannhardt <i>et al.</i> ¹⁴² 5,00
<chem>c1(Cl)cc(nc(C(O)=O)[nH]2)c2cc1Cl</chem>	6,05
<chem>c1cc(nc(C(O)=O)[nH]2)c2cc1[N+]([O-])=O</chem>	4,42
<chem>c1cc(nc(S(=O)(O)=O)[nH]2)c2cc1F</chem>	3,40
<chem>c1cc(nc(S(=O)(O)=O)[nH]2)c2cc1Cl</chem>	3,35
<chem>c1c(Cl)c(nc(S(=O)(O)=O)[nH]2)c2cc1Cl</chem>	3,52
<chem>c1(Cl)cc(nc(S(=O)(O)=O)[nH]2)c2cc1Cl</chem>	3,65
<chem>c1cc(nc(S(=O)(O)=O)[nH]2)c2cc1Br</chem>	3,46
<chem>c1c(Br)c(nc(S(=O)(O)=O)[nH]2)c2cc1Br</chem>	3,50
<chem>c1cc(nc(S(=O)(O)=O)[nH]2)c2cc1C</chem>	3,40
<chem>c1(C)cc(nc(S(=O)(O)=O)[nH]2)c2cc1C</chem>	3,48
<chem>c1c(Cl)c([C@@H](NC(CCNC(=O)OC(C)(C)C)=O)C[C@@H](C(O)=O)N2)c2cc1Cl</chem>	5,18
<chem>c1c(Cl)c([C@@H](NC(CCNC(=O)OC(C)(C)C)=O)C[C@H](C(O)=O)N2)c2cc1Cl</chem>	6,39
<chem>c1c(Cl)c([C@@H](NC(CCCCNC(=O)OC(C)(C)C)=O)C[C@@H](C(O)=O)N2)c2cc1Cl</chem>	5,82
<chem>c1c(Cl)c([C@@H](NC(CCCCNC(=O)OC(C)(C)C)=O)C[C@H](C(O)=O)N2)c2cc1Cl</chem>	6,85
<chem>c1c(Cl)c([C@@H](NC(c2ccc(CNC(=O)OC(C)(C)C)cc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>	6,92
<chem>c1c(Cl)c([C@@H](NC(C2CN(C(=O)OC(C)(C)C)CCC2)=O)C[C@@H](C(O)=O)N3)c3cc1Cl</chem>	5,28
<chem>c1c(Cl)c([C@@H](NC(C2CN(C(=O)OC(C)(C)C)CCC2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>	6,33
<chem>c1c(Cl)c([C@@H](NC(Nc2cccc2)=O)C[C@@H](C(O)=O)N3)c3cc1Cl</chem>	8,00

<chem>c1c(Cl)c([C@@H](NC(Nc2ccccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		8,15
<chem>c1c(Cl)c([C@@H](NC(Nc2ccc(NC(=O)OC(C)(C)C)cc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,49
<chem>c1c(Cl)c([C@@H](NC(Nc2ccc(NC(=O)O[*](C)(C)C)cc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		8,10
<chem>c1c(Cl)c([C@@H](NC(Nc2ccc(CNC(=O)OC(C)(C)C)cc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,36
<chem>c1c(Cl)c([C@@H](NC(Nc2ccc(CNC(=O)OC(C)(C)C)cc2)=O)C[C@H](C(O)=O)N3)c3c1Cl</chem>		8,15
<chem>c1c(Cl)c([C@@H](NC(Nc2ccc(CCNC(=O)OC(C)(C)C)cc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,25
<chem>c1c(Cl)c([C@@H](NC(Nc2ccc(CCCNC(=O)OC(C)(C)C)cc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		4,14
<chem>c1c(Cl)c([C@@H](NC(Nc2ccc(N3CCN(C(OC(C)(C)C)=O)CC3)cc2)=O)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		5,91
<chem>c1c(Cl)c([C@@H](NC(Nc2ccc(N3CCN(C(OC(C)(C)C)=O)CC3)cc2)=O)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		8,22
<chem>c1c(Cl)c([C@@H](NC(Nc2ccc(CN3CCN(C(OC(C)(C)C)=O)CC3)cc2)=O)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		6,77
<chem>c1c(Cl)c([C@@H](NC(Nc2ccc(N3CCN(C(OC(C)(C)C)=O)CC3)cc2)=S)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		5,56
<chem>c1c(c(c(N2)cc1Cl)[C@@H](NC(Nc3ccc(CNC(OC(C)(C)C)=O)cc3)=S)C[C@H]2C(O)=O)Cl</chem>		5,92
<chem>c1c(c(c(N[C@H]2C(O)=O)cc1Cl)[C@H](NC(NS(=O)(=O)c3ccccc3)=S)C2)Cl</chem>		10,00
<chem>c1c(Cl)c(cc(C(O)=O)[nH]2)c2cc1Cl</chem>	Jansen <i>et al.</i> ²⁰²	5,54
<chem>c1c(Cl)c(cc(c2[nH]n2)[nH]3)c3cc1Cl</chem>		4,17
<chem>c1c(Cl)c(cc(C(Nc2[nH]n2)=O)[nH]3)c3cc1Cl</chem>		4,74
<chem>c1c(Cl)c(cc(C(NS(C)=O)=O)[nH]2)c2cc1Cl</chem>		4,13
<chem>c1c(Cl)c(cc(C(NS(C(F)F)=O)=O)[nH]2)c2cc1Cl</chem>		4,78
<chem>c1(C(C)C)cc(cc(C(O)=O)[nH]2)c2cc1</chem>		4,04
<chem>c1(C(C)C)cc(c(CN2CC(=O)N(c3ccccc3)C2=O)c(C(O)=O)[nH]4)c4cc1</chem>		4,82
<chem>c1c(Cl)c(c(CN2CC(=O)N(c3ccccc3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		7,66
<chem>c1cc(c(CN2CC(=O)N(c3ccccc3)C2=O)c(C(O)=O)[nH]4)c4cc1</chem>		5,54
<chem>c1(C(C)C)cc(c(CN2CC(=O)N(c3ccc(C)cc3)C2=O)c(C(O)=O)[nH]4)c4cc1</chem>		4,21
<chem>c1c(Cl)c(c(CN2CC(=O)N(c3ccc(C)cc3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		7,44
<chem>c1(C(C)C)cc(c(CN2CC(=O)N(c3ccc(OC)cc3)C2=O)c(C(O)=O)[nH]4)c4cc1</chem>		4,11
<chem>c1c(Cl)c(c(CN2CC(=O)N(c3ccc(OC)cc3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		7,00
<chem>c1c(Cl)c(c(CNCc2ccccc2)c(C(O)=O)[nH]3)c3cc1Cl</chem>	Jansen <i>et al.</i> ²⁰⁸	
<chem>c1c(Cl)c(c(CNC(=O)c2ccccc2)c(C(O)=O)[nH]3)c3cc1Cl</chem>		
<chem>c1c(Cl)c(c(CNS(=O)(=O)c2ccccc2)c(C(O)=O)[nH]3)c3cc1Cl</chem>		
<chem>c1c(Cl)c(c(CNC(Nc2ccccc2)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		
<chem>c1c(Cl)c(c(CNC(NS(=O)(=O)c2ccccc2)=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		
<chem>c1c(Cl)c(c(CN2CCN(c3ccccc3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c(Cl)c(c(CN2CC(=O)N(CC)C2=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		
<chem>c1c(Cl)c(c(CN2CC(=O)N(C(C)C)C2=O)c(C(O)=O)[nH]3)c3cc1Cl</chem>		
<chem>c1c(Cl)c(c(CN2CC(=O)N(C3CCCC3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		

<chem>c1c(Cl)c(c(CN2CC(=O)N(c3ccc(Cl)cc3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c(Cl)c(c(CN2CC(=O)N(c3cccc(Cl)c3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c(Cl)c(c(CN2CC(=O)N(c3cc(Cl)cc(Cl)c3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c(Cl)c(c(CN2CC(=O)N(c3ccc(F)cc3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c(Cl)c(c(CN2CC(=O)N(c3cccc(C)c3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c(Cl)c(c(CN2CC(=O)N(c3ccc(C(=O)C)cc3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c(Cl)c(c(CN2C(C)C(=O)N(c3cccc3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c(c1Cl)c(Cl)c(c(CN2C(C)C)C(=O)N(c3cccc3)C2=O)c(C(O)=O)[nH]4)c4c1</chem>		
<chem>c1c(Cl)c(c(CN2CC(=S)N(c3cccc3)C2=O)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c2c(c(C(CC(O)=O)CC2)c(C(O)=O)[nH]3)c3cc1Cl</chem>	Katayama <i>et al.</i> ²⁰⁹	
<chem>c1c2c(c([C@H](CC(Nc3cccc3)=O)CC2)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c2c(c([C@H](CC(Nc3c(OCC(O)=O)cc(CN)cc3)=O)CC2)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c2c(c(C(CC(Nc3c(OCC(O)=O)cc(CN)cc3)=O)CC2)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c2c(c([C@@H](CC(Nc3c(OCC(O)=O)cc(CN)cc3)=O)CC2)c(C(O)=O)[nH]4)c4cc1Cl</chem>	Katayama <i>et al.</i> ²¹⁰	7,48
<chem>c1c2c(c([C@@H](CC(Nc3cccc3)=O)CC2)c(C(O)=O)[nH]4)c4cc1Cl</chem>		7,72
<chem>c1c(cc2c(c([C@H](CC(Nc3c(CC(=O)O)cccc3)=O)CC4)c(C(O)=O)[nH]2)c14)Cl</chem>		8,12
<chem>c1c2c(c([C@H](CC(Nc3ccc(CN)cc3)=O)CC2)c(C(O)=O)[nH]4)c4cc1Cl</chem>		7,74
<chem>c1c2c(c([C@H](CC(Nc3c(CC(O)=O)cc(CN)cc3)=O)CC2)c(C(O)=O)[nH]4)c4cc1Cl</chem>		8,82
<chem>c1c2c(c([C@H](CC(Nc3c(O[C@H](C(O)=O)C)cc(CN)cc3)=O)CC2)c(C(O)=O)[nH]4)c4cc1Cl</chem>		8,57
<chem>c1c2c(c([C@H](CC(Nc3c(O[C@H](C(O)=O)C)cc(CN)cc3)=O)CC2)c(C(O)=O)[nH]4)c4cc1Cl</chem>		7,54
<chem>c1c(Cl)c(c(\C=C\C(=O)NC2CCCC2)c(C(O)=O)[nH]3)c3cc1Cl</chem>	Micheli <i>et al.</i> ²¹¹	
<chem>c1c(Cl)c(c(\C=C\C(=O)NC(CC(CC23)CC4C2)(C3)C4)c(C(O)=O)[nH]5)c5cc1Cl</chem>		
<chem>c1c(Cl)c(c(\C=C\C(=O)NCC(CC(CC23)CC4C2)(C3)C4)c(C(O)=O)[nH]5)c5cc1Cl</chem>		
<chem>c1c(Cl)c(c(\C=C\C(=O)NC(C(CC23)CC4C2)(C4)C3)c(C(O)=O)[nH]5)c5cc1Cl</chem>		
<chem>c1c(Cl)c(c(\C=C\C(=O)NC2CC2)c(C(O)=O)[nH]3)c3cc1Cl</chem>		
<chem>c1c(Cl)c(c(\C=C\C(=O)NCC2CC2)c(C(O)=O)[nH]3)c3cc1Cl</chem>		
<chem>c1c(Cl)c(c(\C=C\C(=O)NC(CN23)C(CC2)CC3)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c(Cl)c(c(\C=C\C(=O)NC2C(CC(CC23)CC4C3)C4)c(C(O)=O)[nH]5)c5cc1Cl</chem>		
<chem>C1(NC(=O)\C=C\C(c(Cl)c2)c3cc2Cl)c(C(O)=O)[nH]3)C(CC4)CC4C1</chem>		
<chem>C1(C(C)(CC2)C(C)(C)C2C1)NC(=O)\C=C\C(c(Cl)c3)c4cc3Cl)c(C(O)=O)[nH]4</chem>		
<chem>c1c(Cl)c(c(\C=C\C(=O)NC2CCCc(cccc3)c23)c(C(O)=O)[nH]4)c4cc1Cl</chem>		
<chem>c1c2c(N(CCC2)C(=O)C(=O)N3)c3cc1Br</chem>	Nagata <i>et al.</i> ²¹²	
<chem>c1c2c(N(C(C(O)=O)CC2)C(=O)C(=O)N3)c3cc1Br</chem>		
<chem>c1c2c(N(C(CC(O)=O)CC2)C(=O)C(=O)N3)c3cc1Br</chem>		
<chem>c1c2c(N(C(CCC(O)=O)CC2)C(=O)C(=O)N3)c3cc1Br</chem>		
<chem>c1c2c(N(C(C(=O)Nc3cccc3)CC2)C(=O)C(=O)N4)c4cc1Br</chem>		
<chem>c1c2c(N(C(CC(=O)Nc3cccc3)CC2)C(=O)C(=O)N4)c4cc1Br</chem>		
<chem>c1c2c(N(C(CCC(=O)Nc3cccc3)CC2)C(=O)C(=O)N4)c4cc1Br</chem>		
<chem>c1c2c(N(C(CNC(c3cccc3)=O)CC2)C(=O)C(=O)N4)c4cc1Br</chem>		

c1c2c(N(C(CNC(Nc3ccccc3)=O)CC2)C(=O)C(=O)N4)c4cc1Br

c1c2c(N(C(CC(O)=O)C2)C(=O)C(=O)N3)c3cc1Br

c1c2c(N(C(C(=O)Nc3ccccc3)C2)C(=O)C(=O)N4)c4cc1Br

c1c2c(N(C(CC(=O)Nc3ccccc3)C2)C(=O)C(=O)N4)c4cc1Br

c1c2c(N(C(CNC(Nc3ccccc3)=O)C2)C(=O)C(=O)N4)c4cc1Br

c1c2c(N([C@@H](CC2)CC(O)=O)C(=O)C(=O)N3)c3cc1Br

c1c2c(N([C@@H](CC2)CC(Nc3ccccc3)=O)C(=O)C(=O)N4)c4cc1Br

c1c2c(N([C@H](CC2)CC(O)=O)C(=O)C(=O)N3)c3cc1Br

c1c2c(N([C@H](CC2)CC(Nc3ccccc3)=O)C(=O)C(=O)N4)c4cc1Br

c1c2c(N([C@@H](C[C@@H]2C)CC(O)=O)C(=O)C(=O)N3)c3cc1Br

Nagata *et al.*²¹³

c1c2c(N([C@@H](C[C@H]2C)CC(O)=O)C(=O)C(=O)N3)c3cc1Br

c1c2c(N([C@@](C)(CC2)CC(O)=O)C(=O)C(=O)N3)c3cc1Br

c1c2c(N([C@@H](CC2)C(C(O)=O)C)C(=O)C(=O)N3)c3cc1Br

c1c2c(N([C@@H](CC2)[C@H](C(O)=O)C)C(=O)C(=O)N3)c3cc1Br

c1c2c(N([C@@H](C[C@H](C)C2)CC(O)=O)C(=O)C(=O)N3)c3cc1Br

c1c2c(N([C@@H](C[C@@H](C)C2)CC(O)=O)C(=O)C(=O)N3)c3cc1Br

c1c2c(N([C@@H](C[C@@H](CC)C2)CC(O)=O)C(=O)C(=O)N3)c3cc1Br

c1c2c(N([C@@H](C[C@H](C)C)C2)CC(O)=O)C(=O)C(=O)N3)c3cc1Br

c1c2c(N([C@@H](C[C@@H]2C)CC(Nc3ccccc3)=O)C(=O)C(=O)N4)c4cc1Br

c1c2c(N([C@@H](C[C@H]2C)CC(Nc3ccccc3)=O)C(=O)C(=O)N4)c4cc1Br

c1c2c(N([C@@](C)(CC2)CC(Nc3ccccc3)=O)C(=O)C(=O)N4)c4cc1Br

c1c2c(N([C@@H](CC2)[C@H](C)C(Nc3ccccc3)=O)C(=O)C(=O)N4)c4cc1Br

c1c2c(N([C@@H](CC2)[C@H](C)C(Nc3ccccc3)=O)C(=O)C(=O)N4)c4cc1Br

c1c2c(N([C@@H](C[C@H](C)C2)CC(Nc3ccccc3)=O)C(=O)C(=O)N4)c4cc1Br

c1c2c(N([C@@H](C[C@@H](C)C2)CC(Nc3ccccc3)=O)C(=O)C(=O)N4)c4cc1Br

c1c2c(N([C@@H](C[C@@H](CC)C2)CC(Nc3ccccc3)=O)C(=O)C(=O)N4)c4cc1Br

c1(n(c2)cnn2)cc(n3c(nc(C(O)=O)n3)C(=O)N4)c4cc1Cl

Catarzi *et al.*²¹⁴

c1(n(c2)ccc2)cc(n3c(nc(C(OCC)=O)n3)C(=O)N4)c4cc1Cl

c1(n(c2)ccc2)cc(n3c(nc(C(O)=O)n3)C(=O)N4)c4cc1Cl

c1(n(c2)cc(C=O)c2)cc(n3c(nc(C(OCC)=O)n3)C(=O)N4)c4cc1Cl

c1(n(c2)cc(C=O)c2)cc(n3c(nc(C(O)=O)n3)C(=O)N4)c4cc1Cl

c1(n2nccc2)cc(n3c(nc(C(OCC)=O)n3)C(=O)N4)c4cc1Cl

c1(n2nccc2)cc(n3c(nc(C(O)=O)n3)C(=O)N4)c4cc1Cl

c1([N+][[O-]=O)cc(n2c(nc(C(OCC)=O)n2)C(=O)N3)c3cc1Cl

c1([N+][[O-]=O)cc(n2c(nc(C(O)=O)n2)C(=O)N3)c3cc1Cl

c1([N+][[O-]=O)cc(n2c(nc(C(O)=O)n2)C(=O)N3)c3cc1

c1(N)cc(n2c(nc(C(OCC)=O)n2)C(=O)N3)c3cc1Cl

c1(N)cc(n2c(nc(C(O)=O)n2)C(=O)N3)c3cc1Cl

c1cc(C(=O)C2=C(C(=O)NN(Cc3ccncc3)C2=O)N4)c4cc1Cl

Brown *et al.*²¹⁵

c1cc(C(=O)C2=C(C(=O)NN(Cc3cc3)ncc3)C2=O)N4)c4cc1Cl

c1cc(C(=O)C2=C(C(=O)NN([C@@H](C)c3cc3)ncc3)C2=O)N4)c4cc1Cl

<chem>c1cc(C(=O)C2=C(C(=O)NN(C(CC)c(cc3)ncc3)C2=O)N4)c4cc1Cl</chem>		
<chem>c1cc(C(=O)C2=C(C(=O)NN(C(C(C)C)c(cc3)ncc3)C2=O)N4)c4cc1Cl</chem>		
<chem>c1cc(C(=O)C2=C(C(=O)NN(C(C3CC3)c(cc4)ncc4)C2=O)N5)c5cc1Cl</chem>		
<chem>c1cc(C(=O)C2=C(C(=O)NN(Cc(c3)cncc3)C2=O)N4)c4cc1Cl</chem>		
<chem>c1cc(C(=O)C2=C(C(=O)NN(C(C)c(c3)cncc3)C2=O)N4)c4cc1Cl</chem>		
<chem>c1cc(C(=O)C2=C(C(=O)NN(C(c3ccncc3)C)C2=O)N4)c4cc1Cl</chem>		
<chem>c1cc(C(=O)C2=C(C(=O)NN(C(c3ccncc3)CC)C2=O)N4)c4cc1Cl</chem>		
<chem>c1cc(C(=O)C2=C(C(=O)NN(C(c3ccncc3)C4CC4)C2=O)N5)c5cc1Cl</chem>		
<chem>c1cc(C(=O)C2=C(C(=O)NN(C(c3ccncc3)CCC)C2=O)N4)c4cc1Cl</chem>		
<chem>c1cc(C(=O)C2=C(C(=O)NN(C(c3ccncc3)CCCC)C2=O)N4)c4cc1Cl</chem>		
<chem>c(c(c1)Cl)cc(C(=O)C2=C(C(=O)NN(C(c3ccncc3)C(C)CC)C2=O)N4)c14</chem>		
<chem>c1cc(C(=O)C2=C(C(=O)NN(Cc3cc[n+][[O-]]cc3)C2=O)N4)c4cc1Cl</chem>		
<chem>c1cc(c(NC(C(=O)NN([C@@H](C)c2cc[n+][[O-]]cc2)C3=O)=C34)cc1Cl)C4=O</chem>		
<hr/>		
<chem>c1cc(C(=O)\C(=N\O)\C(=O)N2)c2cc1</chem>	Cai et al. ²¹⁶	5,85
<chem>c1cc(C(=O)\C(=N\O)\C(=O)N2)c2cc1Cl</chem>		7,28
<chem>c1(Cl)cc(C(=O)\C(=N\O)\C(=O)N2)c2cc1</chem>		5,47
<chem>c1c(Cl)c(C(=O)\C(=N\O)\C(=O)N2)c2cc1</chem>		5,36
<chem>c1(Cl)c(Cl)c(C(=O)\C(=N\O)\C(=O)N2)c2cc1</chem>		6,77
<chem>c1c(Cl)c(C(=O)\C(=N\O)\C(=O)N2)c2cc1Cl</chem>		7,48
<chem>c1(Cl)cc(C(=O)\C(=N\O)\C(=O)N2)c2cc1Cl</chem>		7,92
<chem>c1(Cl)c(Cl)c(C(=O)\C(=N\O)\C(=O)N2)c2cc1Cl</chem>		8,15
<chem>c1c(C)c(C(=O)\C(=N\O)\C(=O)N2)c2cc1C</chem>		7,43
<chem>c1(F)cc(C(=O)\C(=N\O)\C(=O)N2)c2c(F)c1Cl</chem>		6,04
<chem>c1(F)c(F)c(C(=O)\C(=N\O)\C(=O)N2)c2c(F)c1F</chem>		5,48
<chem>c1cc(NC(=O)C(=O)N2)c2cc1</chem>		5,01
<chem>c1cc(NC(=O)C(=O)N2)c2cc1Cl</chem>		5,74
<chem>c1c(Cl)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		6,55
<chem>c1(Cl)cc(NC(=O)C(=O)N2)c2cc1Cl</chem>		6,89
<chem>c1(Cl)c(Cl)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		7,52
<chem>c1c(C)c(NC(=O)C(=O)N2)c2cc1C</chem>		5,68
<chem>c1(F)cc(NC(=O)C(=O)N2)c2c(F)c1Cl</chem>		6,20
<chem>c1(F)c(F)c(NC(=O)C(=O)N2)c2c(F)c1F</chem>		6,14
<hr/>		
<chem>c(ccc1c2C(O)=C([N+](=O)[O-])C(=O)N1)c2</chem>	Cai et al. ²¹⁷	4,47
<chem>c(c(cc1c2C(O)=C([N+](=O)[O-])C(=O)N1)Cl)c2</chem>		5,82
<chem>c(ccc1c2C(O)=C([N+](=O)[O-])C(=O)N1)c2Cl</chem>		4,21
<chem>c1c(Cl)cc2c(c1Cl)C(O)=C([N+](=O)[O-])C(=O)N2</chem>		6,66
<chem>c1(Cl)c(Cl)cc2c(c1)C(O)=C([N+](=O)[O-])C(=O)N2</chem>		6,40
<chem>c1c(C)cc2c(c1C)C(O)=C([N+](=O)[O-])C(=O)N2</chem>		6,54
<chem>c1(c(c(c2cc1Cl)C(O)=C([N+](=O)[O-])C(=O)N2)Cl)Cl</chem>		6,66
<chem>c1(cc(c2c(F)c1Cl)C(O)=C([N+](=O)[O-])C(=O)N2)F</chem>		4,40
<chem>c(F)(c(c1)Cl)cc(NC(=O)C(=O)N2)c12</chem>		6,20

<chem>c(F)(c(Cl)cc1c2C(=O)\C(=N\O)\C(=O)N1)c2</chem>		6,04
<chem>c1(C)cc(NC(=O)C(=O)N2)c2cc1C</chem>	Cai et al. ²¹⁸	8,23
<chem>c1(Cl)cc(NC(=O)C(=O)N2)c2cc1CC</chem>		5,85
<chem>c1(CC)cc(NC(=O)C(=O)N2)c2cc1CC</chem>		5,74
<chem>c1(OC)cc(NC(=O)C(=O)N2)c2cc1OC</chem>		4,77
<chem>c(cccc1)(cc(NC(=O)C(=O)N2)c2c3)c13</chem>		4,80
<chem>c(CCC1)(cc(NC(=O)C(=O)N2)c2c3)c13</chem>		4,82
<chem>c1c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1C</chem>		5,03
<chem>c1(OC)cc(NC(=O)C(=O)N2)c2cc1Br</chem>		6,00
<chem>c1(Cl)cc(NC(=O)C(=O)N2)c2cc1C</chem>		6,11
<chem>c1(C)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1C</chem>		7,54
<chem>c1(CC)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1CC</chem>		6,80
<chem>c(CCC1)(c([N+][O-])=O)c(NC(=O)C(=O)N2)c2c3)c13</chem>		5,51
<chem>c1(OC)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1Br</chem>		7,19
<chem>c1(Br)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1CC</chem>		7,09
<chem>c1(Cl)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1CC</chem>		7,54
<chem>c1(CC)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		6,89
<chem>c1(C)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1C#N</chem>		7,14
<chem>c1(C)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1F</chem>		7,02
<chem>c1(C)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1Br</chem>		8,06
<chem>c1(C)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		8,33
<chem>c1(Cl)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1C</chem>		7,35
<chem>c1c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1C#N</chem>		5,55
<chem>c1(C)c(N)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		6,82
<chem>c1(N)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1F</chem>		4,96
<chem>c1(OC)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1F</chem>		5,72
<chem>c1(OCC)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1F</chem>		5,44
<chem>c1(OCCCC)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1F</chem>		5,37
<chem>c1(OCCCCc2ccccc2)c([N+][O-])=O)c(NC(=O)C(=O)N3)c3cc1F</chem>		5,40
<chem>c1(OC)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		6,82
<chem>c1(SCC)c([N+][O-])=O)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		6,19
<chem>c1(OC)c(OC(=O)C)c(NC(=O)C(=O)N2)c2cc1OC</chem>		4,28
<chem>c1(OC)c(O)c(NC(=O)C(=O)N2)c2cc1OC</chem>		4,77
<chem>[C@H]1(C[C@H](N)C(O)=O)ON=C(C(O)=O)C1</chem>	Conti et al. ²¹⁹	
<chem>[C@H]1(C[C@@H](N)C(O)=O)ON=C(C(O)=O)C1</chem>		
<chem>C1(C)(C)ON=C(C(O)=O)[C@@]1([H])C[C@]([H])(N)C(O)=O</chem>		
<chem>c1n(Cc(c2Cl)c(NC(=O)C(=O)N3)c3cc2Cl)ccn1</chem>	Fray et al. ²²⁰	7,64
<chem>c1(Cl)c(Cn2cccn2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>		8,30
<chem>c1(Cl)c(Cn2cncn2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>		7,60
<chem>c1(Cl)c(Cn2ccnn2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>		9,00
<chem>c1(Cl)c(Cn(ch2)cn2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>		7,09

<chem>c1(Cl)c(Cn(c2C)ccn2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>	6,96
<chem>c1(Cl)c(Cn2cnc(C)n2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>	7,62
<chem>c1(Cl)c(Cn2c(C)ncn2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>	6,72
<chem>c1(Cl)c(Cn2cc(CN(CC)CC)nn2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>	8,22
<chem>c1(Cl)c(Cn2cc(CN3CCN(C)CC3)nn2)c(NC(=O)C(=O)N4)c4cc1Cl</chem>	8,12
<chem>c1(Cl)c(Cc2ncn(CCC)c2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>	8,36
<chem>c1(Cl)c(Cc2ncccc2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>	7,52
<chem>c1(Cl)c(Cc(n2)nnn2C)c(NC(=O)C(=O)N3)c3cc1Cl</chem>	7,11
<chem>c1n(C(CC)c(c2Cl)c(NC(=O)C(=O)N3)c3cc2Cl)ccn1</chem>	8,51
<chem>c1(Cl)c(C(CC)n2ccnn2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>	7,12
<chem>c1(Cl)c(C(CC)n(cn2)cn2)c(NC(=O)C(=O)N3)c3cc1Cl</chem>	8,59

<chem>C(\C(\C1=O)=N/O)(\C(\N1)=C/c2ccccc2)=O</chem>	Poschenrieder <i>et al.</i> ¹³¹	
<chem>C(\C(\C1=O)=N/O)(\C(\N1)=C(/c2ccccc2)\Br)=O</chem>		
<chem>C1(=O)\C(\C(=N/C/1=C\c2ccccc2)NCc3ccccc3)=N/O</chem>		
<chem>C1(\C(\C(=N/C/1=C\c2ccccc2)Nc3ccccc3)=N/O)=O</chem>		
<chem>C1(=O)\C(\C(=N/C/1=C\c2ccccc2)Nc3ccc(Cl)cc3)=N/O</chem>		
<chem>C(=O)(C([N+][O-])=O)=C1OC)\C(=C\c2ccccc2)\N1</chem>		
<chem>C1(=O)\C(=N/O)\C(Nc2ccccc2)=N/C/1=C(/c3ccccc3)\Br</chem>		
<chem>C(=O)(C([N+][O-])=O)=C1OC)\C(=C(/c2ccccc2)\Br)\N1</chem>		
<chem>C(=O)(C([N+][O-])=O)=C1Nc2ccccc2)\C(=C(/c3ccccc3)\Br)\N1</chem>		
<chem>C(O)(=C([N+][O-])=O)C1=O)\C(=C(/c2ccccc2)\Br)\N1</chem>		
<chem>C(Nc1ccccc1)(=C(N=O)C2=O)\C(=C\c3ccccc3)\N2</chem>		
<chem>C(Nc1ccccc1)(=C([N+][O-])=O)C2=O)\C(=C\c3ccccc3)\N2</chem>		

<chem>c1cc(c(cc(C(OCC)=O)n2)n2C(=O)N3)c3cc1</chem>	Varano <i>et al.</i> ¹³³	4,48
<chem>c1cc(c(cc(C(OCC)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		4,58
<chem>c1cc(c(cc(C(O)=O)n2)n2C(=O)N3)c3cc1</chem>		5,94
<chem>c1cc(c(cc(C(O)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		6,32
<chem>c1(Cl)cc(c(cc(C(O)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		6,80
<chem>c(c1)cc(c(c(C(O)OCC)c(C(OCC)=O)n2)n2C(=O)N3)c3c1</chem>		5,53
<chem>c(c(c1)Cl)cc(c(c(C(O)OCC)c(C(OCC)=O)n2)n2C(=O)N3)c13</chem>		5,47
<chem>c(Cl)(c(cc1c2c(c(C(O)OCC)c(C(OCC)=O)n3)n3C(=O)N1)Cl)c2</chem>		4,30
<chem>c1(Cl)cc(c(c(Br)c(C(OCC)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		5,15
<chem>c1cc(c(c(C(O)OCC)c(C(O)=O)n2)n2C(=O)N3)c3cc1</chem>		5,94
<chem>c1cc(c(c(C(O)OCC)c(C(O)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		6,07
<chem>c1(Cl)cc(c(c(C(O)OCC)c(C(O)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		6,66
<chem>c1cc(c(c(C(O)O)c(C(O)=O)n2)n2C(=O)N3)c3cc1</chem>		6,27
<chem>c1cc(c(c(C(O)O)c(C(O)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		7,05
<chem>c1(Cl)cc(c(c(C(O)O)c(C(O)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		7,23
<chem>c1cc(c(c(Cl)c(C(O)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		6,74
<chem>c1(Cl)cc(c(c(Cl)c(C(O)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		6,80
<chem>c1(Cl)cc(c(c(Br)c(C(O)=O)n2)n2C(=O)N3)c3cc1Cl</chem>		5,40

<chem>c1cc(c(c(C)C(C(O)=O)n2)n2C(=O)N3)c3cc1</chem>		6,11
<chem>c(c(c1)C)c(C(O)=O)c(NC(=O)C(=O)N2)c12</chem>	Zhou <i>et al.</i> ²²¹	4,89
<chem>c(c(cc1c2NC(=O)C(=O)N1)[N+](=O)[O-])c2C(O)=O</chem>		5,00
<chem>c(c(c1)Cl)c(C(O)=O)c(NC(=O)C(=O)N2)c12</chem>		5,19
<chem>c(c(cc1c2NC(=O)C(=O)N1)Cl)c2C(OC)=O</chem>		4,19
<chem>c1c(c(c(NC2=O)cc1Cl)NC2=O)C(Oc3ccccc3)=O</chem>		5,17
<chem>c(c(c1)Cl)c(C(N)=O)c(NC(=O)C(=O)N2)c12</chem>		5,51
<chem>c1c(cc2c(c1C(ON(C)C)=O)NC(=O)C(=O)N2)Cl</chem>		4,89
<chem>c1c(c(c(N2)cc1Cl)NC(=O)C2=O)C(ON(CC)CC)=O</chem>		4,62
<chem>c(CC)(c(cc1c2NC(=O)C(=O)N1)Cl)c2C=O</chem>		5,34
<chem>c(OC)(c(cc1c2NC(=O)C(=O)N1)Cl)c2C=O</chem>		5,54
<chem>c(c(cc1c2NC(=O)C(=O)N1)Cl)c2[N+](=O)[O-]</chem>		7,17
<chem>c1(Cl)c(C#N)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		7,49
<chem>c1(C)c(C#N)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		7,59
<chem>c1cc(C(=O)N(O)C(=O)N2)c2cc1Cl</chem>	Colotta <i>et al.</i> ²²²	
<chem>c1c(Cl)c(C(=O)N(O)C(=O)N2)c2cc1Cl</chem>		
<chem>c1(Cl)cc(C(=O)N(O)C(=O)N2)c2cc1</chem>		
<chem>c1([N+](=[O-])=O)cc(C(=O)N(O)C(=O)N2)c2cc1</chem>		
<chem>c1(N)cc(C(=O)N(O)C(=O)N2)c2cc1</chem>		
<chem>c1(n2cnnc2)cc(C(=O)N(O)C(=O)N3)c3cc1</chem>		
<chem>c1cc(\C=C\C(Nc2ccccc2)=O)\CC(O)C(=O)N3)c3cc1Cl</chem>	Di Fabio <i>et al.</i> ²²³	
<chem>c1cc(\C=C\C(Nc2ccccc2)=O)\CC(NS(C(F)(F)F)(=O)=O)C(=O)N3)c3cc1Cl</chem>		
<chem>c1cc(\C=C\C(Nc2ccccc2)=O)\CC(C(O)=O)C(=O)N3)c3cc1Cl</chem>		
<chem>c1cc(\C=C\C(Nc2ccccc2)=O)\CC(C(NO)=O)C(=O)N3)c3cc1Cl</chem>		
<chem>c1cc(\C=C\C(Nc2ccccc2)=O)\CC(C(N)=O)C(=O)N3)c3cc1Cl</chem>		
<chem>c1cc(\C=C\C(Nc2ccccc2)=O)\CC(C(Nc3ccccc3)=O)C(=O)N4)c4cc1Cl</chem>		
<chem>c1cc(\C=C\C(Nc2ccccc2)=O)\CC(C(N(C)C)=O)C(=O)N3)c3cc1Cl</chem>		
<chem>c1cc(\C=C\C(Nc2ccccc2)=O)\CC(C#N)C(=O)N3)c3cc1Cl</chem>		
<chem>c1cc(\C=C\C(Nc2ccccc2)=O)\CC(C(=O)NS(c3ccccc3)(=O)=O)C(=O)N4)c4cc1Cl</chem>		
<chem>c1cc(\C=C\C(Nc2ccccc2)=O)\CC(C(c3ccccc3)=O)C(=O)N4)c4cc1Cl</chem>		
<chem>c1cc(C(=O)C=C(C(O)=O)N2)c2cc1</chem>	Carling <i>et al.</i> ²²⁴	4,39
<chem>c1c(Cl)c(C(=O)C=C(C(O)=O)N2)c2cc1</chem>		5,28
<chem>c1cc(C(=O)C=C(C(O)=O)N2)c2cc1Cl</chem>		6,25
<chem>c1c(Cl)c(C(=O)C=C(C(O)=O)N2)c2cc1Cl</chem>		6,70
<chem>c1c(C)c(C(=O)C=C(C(O)=O)N2)c2cc1C</chem>		6,27
<chem>c1c(Br)c(C(=O)C=C(C(O)=O)N2)c2cc1Br</chem>		7,07
<chem>c1c(Cl)c(C(=O)C=C(C(O)=O)N2)c2cc1I</chem>		6,19
<chem>c1c(I)c(C(=O)C=C(C(O)=O)N2)c2cc1Cl</chem>		7,49
<chem>c1c(Cl)c(C(=O)CC(C(O)=O)N2)c2cc1</chem>		5,16
<chem>c1cc(C(=O)CC(C(O)=O)N2)c2cc1Cl</chem>		5,63
<chem>c1c(Cl)c(C(=O)CC(C(O)=O)N2)c2cc1Cl</chem>		5,64

c1c(C)c(C(=O)CC(C(O)=O)N2)c2cc1C	5,73
c1c(Br)c(C(=O)CC(C(O)=O)N2)c2cc1Br	5,89
c1c(Cl)c(C(=O)CC(C(O)=O)N2)c2cc1Cl	5,42
c1c(I)c(C(=O)CC(C(O)=O)N2)c2cc1Cl	6,44
c1c(C)c(C(=O)[C@H](C)[C@H](C(O)=O)N2)c2cc1C	4,86
c1c(Cl)c(CCC(C(O)=O)N2)c2cc1Cl	5,19
c1c(Cl)c([C@@H](C(OC)=O)C[C@@H](C(O)=O)N2)c2cc1	4,58
c1cc([C@@H](C(OC)=O)C[C@@H](C(O)=O)N2)c2cc1Cl	5,30
c1c(Cl)c([C@@H](C(OC)=O)C[C@@H](C(O)=O)N2)c2cc1Cl	5,92
c1c(Br)c([C@@H](C(OC)=O)C[C@@H](C(O)=O)N2)c2cc1Br	5,69
c1c(I)c([C@@H](C(OC)=O)C[C@@H](C(O)=O)N2)c2cc1Cl	5,38
c1c(Cl)c([C@@H](C(OC)=O)C[C@H](C(O)=O)N2)c2cc1	3,75
c1cc([C@@H](C(OC)=O)CC(C(O)=O)N2)c2cc1Cl	4,77
c1c(Cl)c([C@@H](C(OC)=O)C[C@H](C(O)=O)N2)c2cc1Cl	5,61
c1c(Br)c([C@@H](C(OC)=O)C[C@H](C(O)=O)N2)c2cc1Br	5,54
c1c(I)c([C@@H](C(OC)=O)C[C@H](C(O)=O)N2)c2cc1Cl	5,51
c1c(Cl)c([C@@H](C(O)=O)C[C@@H](C(O)=O)N2)c2cc1Cl	4,50
c1c(Cl)c([C@@H](C(O)=O)C[C@H](C(O)=O)N2)c2cc1Cl	4,70
c1c(Cl)c([C@@H](C(Cc2ccccc2)=O)C[C@@H](C(O)=O)N3)c3cc1Cl	5,63
c1c(Cl)c([C@@H](C(Cc2ccccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl	5,98
c1c(Cl)c([C@@H](C(N)=O)C[C@@H](C(O)=O)N2)c2cc1Cl	5,46
c1c(Cl)c([C@@H](C(N)=O)C[C@H](C(O)=O)N2)c2cc1Cl	6,14
c(c1Cl)c(Cl)c([C@@H](CC(=O)OC)C[C@H](C(O)=O)N2)c12	6,22
c1c(Cl)c([C@@H](Cc2onc(C)n2)C[C@H](C(O)=O)N3)c3cc1Cl	5,60
c1c(Cl)c([C@@H](CC(O)=O)C[C@H](C(O)=O)N2)c2cc1Cl	6,87
c1c(Cl)c([C@@H](CC(O)=O)C[C@@H](C(O)=O)N2)c2cc1Cl	4,82
<hr/>	
c1cc(C(CC(OC)=O)C([N+](=O)[O-])C(=O)N2)c2cc1Cl	Carling <i>et al.</i> ²²⁵ 5,88
c1cc(C(CC(OC)=O)C(C#N)C(=O)N2)c2cc1Cl	4,92
c1cc(C(CC(OC)=O)C([N+](=[O-])C(=O)N2)c2cc1Cl	4,03
c1cc(C(CC(OC)=O)[C@H]([N+](=O)[O-])C(=O)N2)c2cc1Cl	5,88
c1cc(C(CC(OC)=O)[C@@H]([N+](=O)[O-])C(=O)N2)c2cc1	4,54
c1cc(C(CC(OCC)=O)[C@@H]([N+](=O)[O-])C(=O)N2)c2cc1Cl	5,65
c1cc(C(c2onc(C)n2)[C@H]([N+](=O)[O-])C(=O)N3)c3cc1Cl	5,25
c1cc(C(CC(=O)N(C)c2ccccc2)[C@@H]([N+](=O)[O-])C(=O)N3)c3cc1Cl	5,74
c1cc(C(CC(=O)Nc2ccccc2)[C@H]([N+](=O)[O-])C(=O)N3)c3cc1Cl	5,13
c1cc(C(CC(=O)NC)[C@H]([N+](=O)[O-])C(=O)N2)c2cc1Cl	4,91
c1cc(C(CC(=O)N(C)C)[C@@H]([N+](=O)[O-])C(=O)N2)c2cc1Cl	5,43
c1cc(C[C@H]([N+](=O)[O-])C(=O)N2)c2cc1Cl	6,38
<hr/>	
c1cc(C(O)=C(c2ccccc2)C(N3)=O)c3cc1Cl	Carling <i>et al.</i> ²²⁶ 6,77
c1cc(C(N)=C(c2ccccc2)C(N3)=O)c3cc1Cl	5,17
c1cc(C(NC(C)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl	4,83

<chem>c1cc(C(NS(C)(=O)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,57
<chem>c1cc(C(NS(c2ccccc2)(=O)=O)=C(c3ccccc3)C(N4)=O)c4cc1Cl</chem>	5,92
<chem>c1cc(C(NCc2ccccc2)=C(c3ccccc3)C(N4)=O)c4cc1Cl</chem>	5,70
<chem>c1cc(C(NCCN(C)C)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	5,49
<chem>c1cc(C(NCCCN(C)C)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	5,38
<chem>c1cc(C(NC(C(=O)O)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,24
<chem>c1cc(C(NC(C(=O)OC)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	5,57
<chem>c1cc(C(NC(C(=O)NCCN(C)C)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	4,60
<chem>c1cc(C(NC(CC(O)=O)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	5,96
<chem>c1cc(C(OCC(OC)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	5,49
<chem>c1cc(C(OCC(O)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,96
<chem>c1cc(C(OCC(N(C)C)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	4,94
<chem>c1cc(C(OCC(N)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,38
<chem>c1cc(C(OCC#N)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,49
<chem>c1cc(c2cc1Cl)C(OCc3ccccc3)=C(c4ccccc4)C(N2)=O</chem>	5,55
<chem>c1cc(C(OCCN(C)C)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	4,96
<chem>c1cc(C(OCC(=O)C)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	5,77
<chem>c1cc(C(OC\C(=N/O)\C)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,11
<chem>c1cc(C(C(O)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,05
<chem>c1cc(C(CC(=O)OC)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	5,96
<chem>c1cc(C(CC(=O)O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,80
<chem>c1cc(C(CCC(OC)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,30
<chem>c1cc(C(CCC(O)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,77
<chem>c1cc(C(CCCc2onc(C)n2)=C(c3ccccc3)C(N4)=O)c4cc1Cl</chem>	5,46
<chem>c1cc(C(CCC(N)=O)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,72
<chem>c1cc(C(CCC#N)=C(c2ccccc2)C(N3)=O)c3cc1Cl</chem>	6,12
<chem>c1cc(C(CCc2[nH]nnc2)=C(c3ccccc3)C(N4)=O)c4cc1Cl</chem>	6,62
<chem>c1cc(C(N)=C(c2cccc(Oc3ccccc3)c2)C(N4)=O)c4cc1Cl</chem>	6,38
<chem>c1cc(C(OCC#N)=C(c2cccc(Oc3ccccc3)c2)C(N4)=O)c4cc1Cl</chem>	5,92
<chem>c1cc(C(OCC(O)=O)=C(c2cccc(Oc3ccccc3)c2)C(N4)=O)c4cc1Cl</chem>	7,62
<chem>c1(O)ccc(CCOc2c(l)cc(C[C@H](C(O)=O)N)cc2l)cc1</chem>	Curtis <i>et al.</i> ²²⁷ 6,37
<chem>c1ccc(CCOc2c(l)cc(C[C@H](C(O)=O)N)cc2l)cc1</chem>	6,29
<chem>c1(OC)ccc(CCOc2c(l)cc(C[C@H](C(O)=O)N)cc2l)cc1</chem>	6,27
<chem>c1ccc(COc2c(l)cc(C[C@H](C(O)=O)N)cc2l)cc1</chem>	6,10
<chem>c1(O)ccc(Oc2c(l)cc(C[C@H](C(O)=O)N)cc2l)cc1</chem>	5,39
<chem>c1(l)cc(Oc2c(l)cc(C[C@@H](N)C(=O)O)cc2l)ccc1O</chem>	5,82
<chem>c1(l)cc(Oc2c(l)cc(C[C@@H](N)C(=O)O)cc2l)cc(l)c1O</chem>	6,48
<chem>C(Cc1cccc(O)c1)Oc2c(l)cc(C[C@@H](N)C(=O)O)cc2l</chem>	6,46
<chem>C(Cc1cncc1)Oc2c(l)cc(C[C@@H](N)C(=O)O)cc2l</chem>	6,34
<chem>C(C)Oc1c(l)cc(C[C@@H](N)C(=O)O)cc1l</chem>	5,82
<chem>C(Cc1cccc1)Oc2ccc(C[C@@H](N)C(=O)O)cc2l</chem>	5,14

<chem>c1c(Cl)c(C(CCC(O)=O)C(C(O)=O)=N2)c2cc1Cl</chem>	Kulagowski <i>et al.</i> ²⁰⁰	7,16
<chem>c1cc(C(O)=C(c2cccc(Cc3cccc3)c2)C(N4)=O)c4cc1Cl</chem>		8,39
<chem>c1cc(C(O)=C(c2cccc(Cc3ccc(OC)cc3)c2)C(N4)=O)c4cc1Cl</chem>		8,35
<chem>c1cc(C(O)=C(c2cccc(Cc3ccc(OCOC)cc3)c2)C(N4)=O)c4cc1Cl</chem>		8,66
<chem>c1c(cc2c(c1)C(O)=C(c3cccc(Cc4csc4)c3)C(N2)=O)Cl</chem>		8,64
<chem>c1cc(C(O)=C(c2cccc(Oc3ccsc3)c2)C(N4)=O)c4cc1Cl</chem>		8,85
<chem>c1c(Cl)c([C@@H](NC(C)=O)C[C@H](C(O)=O)N2)c2cc1Cl</chem>	Leeson <i>et al.</i> ²²⁸	6,05
<chem>c1c(Cl)c([C@@H](NC(CCC)=O)C[C@H](C(O)=O)N2)c2cc1Cl</chem>		6,43
<chem>c1c(Cl)c([C@@H](NC(C2CCCC2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,77
<chem>c1c(Cl)c([C@@H](NC(C2CCCC2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,96
<chem>c1c(Cl)c([C@@H](NC(c2ccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,82
<chem>c1c(Cl)c([C@@H](NC(C2NC=CC=C2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,20
<chem>c1c(Cl)c([C@@H](NC(Cc2ccsc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,52
<chem>c1c(Cl)c([C@@H](NC(Cc2sc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,49
<chem>c1c(Cl)c([C@@H](NC(Cc2ccc(cc2)N)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,74
<chem>c1c(Cl)c([C@@H](NC(Cc2ccc(cc2)O)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,52
<chem>c1c(Cl)c([C@@H](NC(Cc2ccc(cc2)C)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,47
<chem>c1c(Cl)c([C@@H](NC(Cc2ccc(cc2)OC)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,41
<chem>c1c(Cl)c([C@@H](NC(Cc2ccc(cc2)Cl)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,28
<chem>c1c(Cl)c([C@@H](NC(Cc2cccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,12
<chem>c1c(Cl)c([C@@H](NC(CCCc2cccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,77
<chem>c1c(Cl)c([C@@H](NC(C(c2cccc2)c3cccc3)=O)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		7,21
<chem>c1c(Cl)c([C@@H](NC(C2c(c3c(c24)cccc4)cccc3)=O)C[C@H](C(O)=O)N5)c5cc1Cl</chem>		7,60
<chem>c1c(Cl)c(CC[C@H](C(O)=O)N2)c2cc1Cl</chem>		5,19
<chem>c1c(Cl)c([C@H](N)C[C@H](C(O)=O)N2)c2cc1Cl</chem>		4,85
<chem>c1c(Cl)c([C@@H](N)C[C@H](C(O)=O)N2)c2cc1Cl</chem>		4,95
<chem>c1c(Cl)c([C@H](NC(OC(C)(C)C)=O)C[C@H](C(O)=O)N2)c2cc1Cl</chem>		4,94
<chem>c1c(Cl)c([C@@H](NC(OC(C)(C)C)=O)C[C@H](C(O)=O)N2)c2cc1Cl</chem>		6,03
<chem>c1c(Cl)c([C@H](NC(c2cccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		5,08
<chem>c1c(Cl)c([C@@H](NC(c2cccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,00
<chem>c(cc1c2[C@@H](NC(C(N)Cc3cccc3)=O)C[C@H](C(O)=O)N1)Cl)c2Cl</chem>		5,15
<chem>c1c(Cl)c([C@@H](NC([C@H](N)Cc2cccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		5,98
<chem>c1c(Cl)c([C@@H](NC(=O)Cc2cccc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,85
<chem>c1c(Cl)c([C@@H](NC(=O)NC2CCCC2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,55
<chem>c1c(Cl)c([C@@H](NC(=O)N(C)c2cccc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,35
<chem>c1c(Cl)c([C@@H](NC(=O)N(c2cccc2)c3cccc3)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		7,72
<chem>c1c(Cl)c([C@@H](NC(=O)NCc2cccc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,52
<chem>c1c(Cl)c([C@@H](NC(=O)Nc2c(c3ccc2)cccc3)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		7,96
<chem>c1c(Cl)c([C@@H](NC(=O)Nc2ccc(C)cc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		8,30
<chem>c1c(Cl)c([C@@H](NC(=O)N(C)c2cccc(C)c2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,74
<chem>c1c(Cl)c([C@@H](NC(=O)N(C)c2cccc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		8,14

<chem>c1c(Cl)c([C@@H](NS(=O)(=O)c2ccccc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		5,28
<chem>c1c(Cl)c([C@@H](NCc2ccccc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		4,97
<chem>c1c(Cl)c([C@@H](NCSNc2ccccc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,05
<chem>c1c(Cl)c([C@@H](N1C(Nc2ccccc2)=N1C#N)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,24
<chem>c1c(Cl)c([C@@H](NCC(Nc2ccccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		6,38
<chem>c1c(Cl)c([C@@H](C(NC2ccccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		5,77
<chem>c1c(Cl)c([C@@H](NC(=O)OCc2ccccc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,10
<chem>c1c(Cl)c([C@@H](N2C(=O)CCC2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,34
<chem>c1c(Cl)c([C@@H](N2Cc(cccc3)c3C2=O)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		7,57
<chem>c1c(Cl)c([C@@H](N2C(=O)c(cccc3)c3C2=O)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		5,72
<chem>c1c(Cl)c([C@@H](N2C(=O)Cc(cccc3)c3C2)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		7,14
<chem>c1c(Cl)c([C@@H](N2C(=O)N(c3ccccc3)CC2)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		8,22
<chem>c1c(Cl)c([C@@H](N2C(=O)Nc(cccc3)c3C2)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		6,29
<chem>c1c(Cl)c([C@@H](NC(N2c(cccc3)c3CC2)=O)C[C@H](C(O)=O)N4)c4cc1Cl</chem>		6,66
<chem>c1c(Cl)c([C@@H](NC(Cc2ccccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		8,00
<chem>c1c(Cl)c([C@@H](NC(Cc2ccccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,38
<chem>c1c(Cl)c([C@@H](NC(Cc2ccccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,17
<chem>c1c(Cl)c([C@@H](NC(Nc2ccccc2)=O)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		8,22
<chem>c1cc(C(C(=O)N(c2ccccc2)N3)=C3C(=O)N4)c4cc1</chem>	MacLeod <i>et al.</i> ¹³⁴	6,87
<chem>c1cc(C(C(=O)N(c2c(Cl)cccc2)N3)=C3C(=O)N4)c4cc1</chem>		6,21
<chem>c1cc(C(C(=O)N(c2cc(Cl)ccc2)N3)=C3C(=O)N4)c4cc1</chem>		6,79
<chem>c1cc(C(C(=O)N(c2ccc(Cl)cc2)N3)=C3C(=O)N4)c4cc1</chem>		6,86
<chem>c1cc(C(C(=O)N(c2ccc(OC)cc2)N3)=C3C(=O)N4)c4cc1</chem>		7,06
<chem>c1cc(C(C(=O)N(c2ccccc2)N3)=C3C(=O)N4)c4cc1Cl</chem>		7,80
<chem>c1cc(C(C(=O)N(c2ccc(OC)cc2)N3)=C3C(=O)N4)c4cc1Cl</chem>		8,48
<chem>c1cc(C(C(=O)N(Cc2ccccc2)N3)=C3C(=O)N4)c4cc1</chem>		6,25
<chem>c1cc(C(C(=O)N(C)N2)=C2C(=O)N3)c3cc1</chem>		5,69
<chem>c1c(Cl)c(C(=O)C=C(C(O)=O)N2)c2cc1Cl</chem>	Leeson <i>et al.</i> ²²⁹	6,20
<chem>c1c(Cl)c(C(=O)C=C(C(O)=O)N2)c2cc1</chem>		5,68
<chem>c1c(Cl)c(C(=O)C=C(C(O)=O)N2)c2cc1</chem>		5,32
<chem>C1([N+](=O)[O-])C=C(NC(=O)C(=O)N2)C2=CC1[N+](=O)[O-]</chem>		5,35
<chem>c1cc(S(=O)(=O)C(C(=O)OC)C(=O)N2)c2cc1Cl</chem>	Rowley <i>et al.</i> ²⁰³	5,41
<chem>c(c(c1)Cl)cc(C(=O)C(S(=O)OC)C(=O)N2)c12</chem>		4,81
<chem>c1cc(C(=O)C(C#N)C(=O)N2)c2cc1Cl</chem>		5,41
<chem>c1cc(C(=O)C(C(=O)OCC)C(=O)N2)c2cc1Cl</chem>		4,78
<chem>c1cc(C(=O)C(C(=O)Oc2ccccc2)C(=O)N3)c3cc1Cl</chem>		5,50
<chem>c1cc(C(=O)C(C(=O)OC)C(=O)N2)c2cc1Cl</chem>		5,19
<chem>c1cc(C(O)=C(C(=O)OCC=C)C(=O)N2)c2cc1Cl</chem>		5,19
<chem>c1cc(C(O)=C(C(=O)OCC#C)C(=O)N2)c2cc1Cl</chem>		5,05
<chem>c1cc(C(O)=C(C(=O)OCc2cc(O)ccc2)C(=O)N3)c3cc1Cl</chem>		5,21
<chem>c1cc(C(O)=C(C(=O)OCCc2cccc(O)c2)C(=O)N3)c3cc1Cl</chem>		5,74

<chem>c1cc(C(O)=C(C(=O)OCCc2cc(O)ccc2)C(=O)N3)c3cc1Cl</chem>		5,58
<chem>c1cc(C(O)=C(C(=O)OCCc2cc(O)ccc2)C(=O)N3)c3cc1Cl</chem>		5,58
<chem>c1cc(C(O)=C(C(=O)OCCc2cccc2O)C(=O)N3)c3cc1Cl</chem>		5,15
<chem>c1cc(C(O)=C(C(=O)OCCc2ccc(O)cc2)C(=O)N3)c3cc1Cl</chem>		5,53
<chem>c1cc(C(O)=C(C(=O)OCCc2cccn2)C(=O)N3)c3cc1Cl</chem>		4,56
<chem>c1cc(C(O)=C(C(=O)OCCc2sccc2)C(=O)N3)c3cc1Cl</chem>		5,52
<chem>c1cc(C(O)=C(C(=O)OCCc2c[nH]c(cccc3)c23)C(=O)N4)c4cc1Cl</chem>		6,37
<chem>c1cc(C(O)=C(C(=O)OCC#Cc2cc(OC)ccc2)C(=O)N3)c3cc1Cl</chem>		6,57
<chem>c1cc(C(O)=C(C(=O)OCC#Cc2cc(O)ccc2)C(=O)N3)c3cc1Cl</chem>		6,77
<chem>c1cc(C(O)=C(C(=O)OCC)C(=O)N2)c2cc1</chem>		4,78
<chem>c(c1C#N)cc(C(O)=C(C(=O)OCC)C(=O)N2)c2c1</chem>		4,59
<chem>c(c(cc1c2C(O)=C(C(=O)OCC)C(=O)N1)C(F)(F)F)c2</chem>		4,91
<chem>c(c(N(=O)O)c1)cc(C(O)=C(C(=O)OCC)C(=O)N2)c12</chem>		4,60
<chem>c(c1Cl)c(l)c(C(O)=C(C(=O)OCC)C(=O)N2)c2c1</chem>		5,79
<chem>c1(N(=O)O)c(cc2c(c1)C(O)=C(C(=O)OCC)C(=O)N2)N(=O)O</chem>		4,98
<chem>c1cc(C(O)=C(C(=O)c2csc2)C(=O)N3)c3cc1Cl</chem>		5,71
<chem>c1cc(C(O)=C(C(=O)c2occc2)C(=O)N3)c3cc1Cl</chem>		5,62
<chem>c1cc(C(O)=C(C(=O)c2ccnc2)C(=O)N3)c3cc1Cl</chem>		5,16
<chem>c1cc(C(O)=C(C(=O)Cc2cccc2)C(=O)N3)c3cc1Cl</chem>		6,03
<chem>c1cc(C(O)=C(C(=O)Cc2sccc2)C(=O)N3)c3cc1Cl</chem>		5,97
<chem>c1cc(C(O)=C(C(=O)C2CC2)C(=O)N3)c3cc1Cl</chem>		6,38
<chem>c1cc(C(O)=C(C(=O)O)C(=O)N2)c2cc1Cl</chem>		5,24
<chem>c1cc(C(O)=C(C(=O)NO)C(=O)N2)c2cc1Cl</chem>		5,20
<chem>c1cc(C(O)=C(C(=O)NOC)C(=O)N2)c2cc1Cl</chem>		5,51
<chem>c1cc(C(O)=C(C(=O)N(O)CCc2cccc2)C(=O)N3)c3cc1Cl</chem>		5,03
<chem>c1cc(C(O)=C(C(=O)N(C)OC)C(=O)N2)c2cc1Cl</chem>		4,33
<chem>C(O)=(C(c1cccc1)C2=O)\C(=C\CCC)\N2</chem>	Mawer <i>et al.</i> ²³⁰	4,34
<chem>C(O)=(C(c1cccc1)C2=O)\C(=C\c3cccc3)\N2</chem>		4,62
<chem>C(O)=(C(c1cccc1)C2=O)\C(=C\CCCC)\N2</chem>		4,80
<chem>C1(\C=C(N2)/C(O)=C(c3cccc(C)c3)C2=O)CC1</chem>		5,02
<chem>C(O)=(C(c1cccc(Oc2cccc2)c1)C3=O)\C(=C\CCC)\N3</chem>		6,15
<chem>C(O)=(C(c1cccc(OC(C)=C)c1)C2=O)\C(=C\CCC)\N2</chem>		5,52
<chem>C1(C(O)=C(c2cccc2)C(=O)N3)=C3CC(C)CC1</chem>	Rowley <i>et al.</i> ²³¹	5,46
<chem>C1(C(O)=C(c2cccc(Oc3cccc3)c2)C(=O)N4)=C4CC(C)CC1</chem>		7,24
<chem>c1(C(O)=C(c2cccc(Oc3cccc3)c2)C(=O)N4)c4cc(C)cc1</chem>		7,70
<chem>c1c(Cl)c([C@@H](C(=O)Nc2cccc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>	Carling <i>et al.</i> ²²⁵	7,72
<chem>c1c(Cl)c([C@@H](N(C)C(=O)Nc2cccc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		8,05
<chem>c1c(Cl)c([C@@H](NC(=O)Cc2ccc(CN)cc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,72
<chem>c1c(Cl)c([C@@H](NC(=O)Cc2ccc(CC(C(O)=O)N)cc2)C[C@H](C(O)=O)N3)c3cc1Cl</chem>		7,49
<chem>c1c(CN[C@@H](P(O)(O)=O)C)c(NC(=O)C(=O)N2)c2cc1F</chem>	Ametamey <i>et al.</i> ²³²	8,05
<chem>c1c(CN[C@@H](P(O)(O)=O)C)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		8,22

<chem>c1c(CN[C@@H](P(O)(O)=O)C)C(=O)C(=O)N2)c2cc1Br</chem>		8,30
<chem>c1c(CN[C@@H](P(O)(O)=O)C)C(=O)C(=O)N2)c2cc1C#N</chem>		7,12
<chem>c1c(CN[C@@H](P(O)(O)=O)C)C(=O)C(=O)N2)c2cc1C(F)(F)F</chem>		8,10
<chem>c1c(CN[C@@H](P(O)(O)=O)C)C(=O)C(=O)N2)c2cc1[N+](=[O-])=O</chem>		7,49
<chem>c1c(CN[C@@H](P(O)(O)=O)C)C(=O)C(=O)N2)c2cc1</chem>		8,10
<chem>c1(N2CCOCC2)cc(N(CP(O)(O)=O)C(=O)C(=O)N3)c3cc1C(F)(F)F</chem>	Auberson <i>et al.</i> ²³³	5,48
<chem>c1c(CNCCC(O)=O)C(=O)C(=O)N2)c2cc1[N+](=[O-])=O</chem>		
<chem>c1c(CNCC(O)=O)C(=O)C(=O)N2)c2cc1[N+](=[O-])=O</chem>		6,80
<chem>c1c(CN[C@@H](C)C(O)=O)C(=O)C(=O)N2)c2cc1[N+](=[O-])=O</chem>		6,51
<chem>c1c(CN[C@H](C)C(O)=O)C(=O)C(=O)N2)c2cc1[N+](=[O-])=O</chem>		6,42
<chem>c1c(CNCC(O)=O)C(=O)C(=O)N2)c2cc1Br</chem>		7,40
<chem>c1c(CN[C@H](C)C(O)=O)C(=O)C(=O)N2)c2cc1Br</chem>		6,92
<chem>c1c(CN[C@@H](C)C(O)=O)C(=O)C(=O)N2)c2cc1Br</chem>		5,66
<chem>c1c(CSCC(O)=O)C(=O)C(=O)N2)c2cc1Br</chem>		6,92
<chem>c1c(COCC(O)=O)C(=O)C(=O)N2)c2cc1Br</chem>		6,96
<chem>c1c(COCC(O)=O)C(=O)C(=O)N2)c2cc1[N+](=[O-])=O</chem>		6,51
<chem>c1c(COC(C)C(O)=O)C(=O)C(=O)N2)c2cc1[N+](=[O-])=O</chem>		5,89
<chem>c1c(CNCP(O)(O)=O)C(=O)C(=O)N2)c2cc1[N+](=[O-])=O</chem>		6,00
<chem>c1c(CNCP(O)(O)=O)C(=O)C(=O)N2)c2cc1Br</chem>		7,00
<chem>c1c(CN[C@@H](C)P(O)(O)=O)C(=O)C(=O)N2)c2cc1Br</chem>		8,22
<chem>c1c(CNCP(O)(O)=O)C(=O)C(=O)N2)c2cc1C(F)(F)F</chem>		6,52
<chem>c1c(CN[C@@H](C)P(O)(O)=O)C(=O)C(=O)N2)c2cc1C(F)(F)F</chem>		8,22
<chem>c1c(CN(CC)[C@H](C)P(O)(O)=O)C(=O)C(=O)N2)c2cc1C(F)(F)F</chem>		5,40
<chem>c1nc(C(O)=C(c2ccccc2)C(=O)N3)c3cc1Cl</chem>	Zhou <i>et al.</i> ²³⁴	5,48
<chem>c1nc(C(O)=C(c2ccc(Cl)cc2)C(=O)N3)c3cc1Cl</chem>		4,89
<chem>c1nc(C(O)=C(c2cc(Cl)ccc2)C(=O)N3)c3cc1Cl</chem>		5,52
<chem>c1nc(C(O)=C(c2c(Cl)ccc2)C(=O)N3)c3cc1Cl</chem>		4,62
<chem>c1nc(C(O)=C(c2cc(OC)ccc2)C(=O)N3)c3cc1Cl</chem>		5,54
<chem>c1nc(C(O)=C(c2cc(C)ccc2)C(=O)N3)c3cc1Cl</chem>		5,74
<chem>c1nc(C(O)=C(c2cc([N+](=[O-])=O)ccc2)C(=O)N3)c3cc1Cl</chem>		5,13
<chem>c1nc(C(O)=C(c2cc(Oc3ccccc3)ccc2)C(=O)N4)c4cc1Cl</chem>		6,96
<chem>c1nc(C(O)=C(c2ccsc2)C(=O)N3)c3cc1Cl</chem>		5,37
<chem>c1nc(C(O)=C(C2=SCC=C2)C(=O)N3)c3cc1Cl</chem>		5,01
<chem>n1cc(C(O)=C(c2cc(Oc3ccccc3)ccc2)C(=O)N4)c4cc1</chem>		4,70
<chem>c1cc(C(O)=C(c2cc(Oc3ccccc3)ccc2)C(=O)N4)c4cn1</chem>		5,27
<chem>c1cc(C(O)=C(c2cc(Oc3ccccc3)ccc2)C(=O)N4)c4nc1</chem>		4,51
<chem>c1(Cl)cc(C(O)=C(c2cc(Oc3ccccc3)ccc2)C(=O)N4)c4nc1</chem>		4,72
<chem>c(c(Cl)c1)nc(C(O)=C(N(O)=O)C(=O)N2)c12</chem>		4,54
<chem>c(c(Cl)c1)cc(C(O)=C(N(O)=O)C(=O)N2)c12</chem>		5,82
<chem>c(c(Cl)cc1c2C(O)=C(N(O)=O)C(=O)N1)c2Cl</chem>		6,66
<chem>c1(Cl)cc(NC(=O)C(=O)N2O)c2cc1Cl</chem>	Zheng <i>et al.</i> ²³⁵	7,09

<chem>c1(Cl)c(N)c(NC(=O)C(=O)N2O)c2cc1Cl</chem>		5,74
<chem>c1(Cl)c(Br)c(NC(=O)C(=O)N2O)c2cc1Cl</chem>		5,57
<chem>c1c(Cl)c(NC(=O)C(=O)N2O)c2cc1C(F)(F)F</chem>		4,60
<chem>c1(Cl)c(Cl)c(NC(=O)C(=O)N2O)c2cc1Cl</chem>		5,60
<chem>c1c(Cl)c(NC(=O)C(=O)N2O)c2cc1Cl</chem>		4,72
<chem>c1c(Br)c(NC(=O)C(=O)N2O)c2cc1C(F)(F)F</chem>		4,27
<chem>c1c(Br)c(NC(=O)C(=O)N2O)c2cc1Br</chem>		4,70
<chem>c1(Cl)c(N)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		7,03
<chem>c1(Cl)c(Br)c(NC(=O)C(=O)N2)c2cc1Cl</chem>		6,68
<chem>c1c(Cl)c(NC(=O)C(=O)N2)c2cc1C(F)(F)F</chem>		5,68
<chem>c1cc(C(=O)C(c2cccc2)=C(O)C(=O)N3)c3cc1Cl</chem>	Jackson <i>et al.</i> ²³⁶	6,42
<chem>c1cc(C(=O)C(c2ccc(OC)cc2)=C(O)C(=O)N3)c3cc1Cl</chem>		6,24
<chem>c1cc(C(=O)C(c2ccc(O)cc2)=C(O)C(=O)N3)c3cc1Cl</chem>		6,24
<chem>c1cc(C(=O)C(C2NC=CC=C2)=C(O)C(=O)N3)c3cc1Cl</chem>		6,72
<chem>c1cc(C(=O)C(c2ccc2)=C(O)C(=O)N3)c3cc1Cl</chem>		6,68
<chem>c1cc(C(=O)C(c2sccc2)=C(O)C(=O)N3)c3cc1Cl</chem>		6,20
<chem>c1cc(C(=O)C(c2ccsc2)=C(O)C(=O)N3)c3cc1Cl</chem>		6,23
<chem>C1CC(C(=O)C=C(O)C(=O)N2)=C2CC1</chem>	Guzikowski <i>et al.</i> ²³⁷	5,49
<chem>C1(C)CC(C(=O)C=C(O)C(=O)N2)=C2CC1C</chem>		6,64
<chem>C1(Cl)CC(C(=O)C=C(O)C(=O)N2)=C2CC1Cl</chem>		6,88
<chem>c1c(Cl)c(S(=O)(=O)N(C)C(C(=O)CC)N2)c2cc1Cl</chem>	Jimonet <i>et al.</i> ²³⁸	6,40
<chem>c1c(Cl)c(S(=O)(=O)N(c2cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		5,59
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,32
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,05
<chem>c1c(Cl)c(S(=O)(=O)N(CCCc2cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,48
<chem>c1c(Cl)c(S(=O)(=O)N(CCCCc2cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,11
<chem>c1c(Cl)c(S(=O)(=O)N(CCCCCc2cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,44
<chem>c1c(Cl)c(S(=O)(=O)N(CC(O)=O)C(C(=O)O)N2)c2cc1Cl</chem>		5,02
<chem>c1c(Cl)c(S(=O)(=O)N(CN(C)C)C(C(=O)O)N2)c2cc1Cl</chem>		5,44
<chem>c1c(Cl)c(S(=O)(=O)N(CC(=O)c2cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,66
<chem>c1c(Cl)c(S(=O)(=O)N(CCC(=O)c2cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,39
<chem>c1c(Cl)c(S(=O)(=O)N(CCS2cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,42
<chem>c1c(Cl)c(S(=O)(=O)N(CCS(=O)(=O)c2cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,15
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2c(C)cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,10
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2c(OC)cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,68
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2c(C(F)(F)F)cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,00
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2c(F)cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,59
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2c(Cl)cccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,32
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(C)ccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,85
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(OC)ccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,92
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(C(F)(F)F)ccc2)C(C(=O)O)N3)c3cc1Cl</chem>		7,74

<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(F)ccc2)C(C=O)O)N3c3cc1Cl</chem>		7,89
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(Cl)ccc2)C(C=O)O)N3c3cc1Cl</chem>		7,96
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(Br)ccc2)C(C=O)O)N3c3cc1Cl</chem>		8,10
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(I)ccc2)C(C=O)O)N3c3cc1Cl</chem>		8,10
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(C#N)ccc2)C(C=O)O)N3c3cc1Cl</chem>		7,77
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc([N+][O-])ccc2)C(C=O)O)N3c3cc1Cl</chem>		7,92
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(OC(F)F)ccc2)C(C=O)O)N3c3cc1Cl</chem>		7,70
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(Oc3ccccc3)ccc2)C(C=O)O)N4c4cc1Cl</chem>		8,05
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(N)ccc2)C(C=O)O)N3c3cc1Cl</chem>		7,54
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2cc(c3ccccc3)ccc2)C(C=O)O)N4c4cc1Cl</chem>		7,70
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2ccc(C)cc2)C(C=O)O)N3c3cc1Cl</chem>		7,35
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2ccc(OC)cc2)C(C=O)O)N3c3cc1Cl</chem>		7,40
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2ccc(C(F)F)cc2)C(C=O)O)N3c3cc1Cl</chem>		7,04
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2ccc(F)cc2)C(C=O)O)N3c3cc1Cl</chem>		7,36
<chem>c1c(Cl)c(S(=O)(=O)N(Cc2ccc(Cl)cc2)C(C=O)O)N3c3cc1Cl</chem>		7,70
<chem>c1(Cl)cc(N=C(C(=O)NS(=O)(C=O)C(=O)N2)c2cc1Cl</chem>	Hays <i>et al.</i> ²⁰¹	6,44
<chem>c1(Cl)cc(N=C(C(=O)NS(=O)(c2ccccc2)=O)C(=O)N3)c3cc1Cl</chem>		6,92
<chem>c1c(F)c(c(O)cc(C(=O)O)n2)c2cc1</chem>	Moroni <i>et al.</i> ²³⁹	4,30
<chem>c1c(Cl)c(c(O)cc(C(=O)O)n2)c2cc1</chem>		4,52
<chem>c1(Cl)cc(c(O)cc(C(=O)O)n2)c2cc1</chem>		2,70
<chem>c1(Cl)cc(c(S)cc(C(=O)O)n2)c2cc1</chem>		3,00
<chem>c1cc(c(O)cc(C(=O)O)n2)c2cc1Cl</chem>		5,15
<chem>c1cc(c(S)cc(C(=O)O)n2)c2cc1Cl</chem>		5,30
<chem>c1cc(c(S)cc(C(=O)O)n2)c2cc1C(F)F</chem>		3,74
<chem>c1cc(c(O)cc(C(=O)O)n2)c2cc1OC</chem>		2,70
<chem>c1cc(c(S)cc(C(=O)O)n2)c2cc1OC</chem>		3,60
<chem>c1cc(c(O)cc(C(=O)O)n2)c2c(Cl)c1</chem>		2,70
<chem>c1c(Cl)c(c(O)cc(C(=O)O)n2)c2cc1Cl</chem>		4,82
<chem>c1c(Cl)c(c(S)cc(C(=O)O)n2)c2cc1Cl</chem>		5,00
<chem>c1c(Cl)c(c(NC(=O)c2ccccc2)c(C(=O)O)[nH]3)c3cc1Cl</chem>	Salituro <i>et al.</i> ²⁴⁰	4,77
<chem>c1c(Cl)c(c(N(C)C(=O)c2ccccc2)c(C(=O)O)[nH]3)c3cc1Cl</chem>		6,57
<chem>c1c(Cl)c(c(N(CC(O)=O)C(=O)c2ccccc2)c(C(=O)O)[nH]3)c3cc1Cl</chem>		6,40
<chem>c1c(Cl)c(c(NC(C(=O)O)=O)c(C(=O)O)[nH]2)c2cc1Cl</chem>		6,77
<chem>c1c(Cl)c(c(N(C)C(C(=O)O)=O)c(C(=O)O)[nH]2)c2cc1Cl</chem>		5,02
<chem>c1cc(c(NC(c2ccccc2)=O)c(C(=O)O)[nH]3)c3cc1Cl</chem>		5,52
<chem>c1cc(c(N(C)C(c2ccccc2)=O)c(C(=O)O)[nH]3)c3cc1Cl</chem>		6,96
<chem>c1(Cl)cc(N(C)C(=O)C(=O)N2)c2cc1Cl</chem>		4,77
<chem>c1(Cl)cc(N(CC)C(=O)C(=O)N2)c2cc1Cl</chem>		4,02
<chem>c1(Cl)cc(N2C(=NCC2)C(=O)N3)c3cc1Cl</chem>		5,22
<chem>c1(Cl)cc(n(cnn2)c2C(=O)N3)c3cc1Cl</chem>		5,15
<chem>c1(Cl)cc(n(nnn2)c2C(=O)N3)c3cc1Cl</chem>		4,11

N([H])(C(C(c1cccc1)=C2O)=O)c3c2cc(s3)C	Buchstaller <i>et al.</i> ²⁴¹	5,02
N([H])(C(C(c1cccs1)=C2O)=O)c(c2cc3C)s3		4,64
N(C)(C(C(c1cccc1)=C2O)=O)c(sc(c3)CC)c23		5,51
N([H])(C(C(c1cccc1)=C2O)=O)c(sc(c3C)C)c23		5,90
N([H])(C(C(c1cccc1)=C2O)=O)c3c2c(C)cs3		5,08
N([H])(C(C(c1cccs1)=C2O)=O)c(c2c(c3)C)s3		4,34
N([H])(C(C(c1cccc1)=C2O)=O)c(scc3CC)c23		6,28
N([H])(C(C(=C1O)c2cccc2)=O)c(scc3C(C)C)c13		5,82
N([H])(C(=O)C(=C1O)c2cccc2)c(scc3C4CC4)c13		5,32
N([H])(C(=O)C(=C1O)c2cccc2)c(scc3C(CC)C)c13		6,03
N([H])(C(C(c1cccs1)=C2O)=O)c(sc3c4CCCC3)c24		4,64
N(C(C(=C1O)c2cccc2)=O)([H])c(sc3c4CCCC3)c14		5,37
N(C1=O)([H])c(sc(Br)c2)c2C(O)=C1c3cccc3		5,75
N(C(C(=C1O)c2cccc2)=O)([H])c(sc(Br)c3C)c13		6,41
N(C(C(c1cccc1)=C2O)=O)([H])c(sc(Br)c3CC)c23		7,98
N(C(C(=C1O)c2cccc2)=O)([H])c(sc(Cl)c3C)c13		6,57
N(C(C(c1cccc1)=C2O)=O)([H])c(sc(Cl)c3CC)c23		8,05
N(C(C(=C1O)c2cccc2)=O)([H])c(sc(Cl)c3C(C)C)c13		8,18
N(C(C(=C1O)c2cccc2)=S)([H])c(sc(Cl)c3C(C)C)c13		6,05
N([H])(C(=O)C(c1cccc1)=C2O)c(sc(Cl)c3C4CC4)c23		7,24
N(C(C(=C1O)c(c2)cccc2)=O)([H])c(sc(Cl)c3C(C)CC)c13		8,24
N(C(C(=C1O)c2cccc2)=O)([H])c(sc(Cl)c3CCN)c13		5,34
N([H])(C(=O)C(c(c1)cccc1O)c2cccc2)=C3O)c(c34)scc4		6,43
N([H])(C(=O)C(c(cc1O)c2cccc2)ccc1)=C3O)c(c34)scc4C		6,48
N1([H])C(=O)C(c(cc2O)c3cccc3)ccc2)=C(c4c1scc4CC)O		6,75
N1([H])C(=O)C(c(cc(c2)O)c3cccc3)cc2)=C(c4c1scc4COC)O		7,44
N1(c(c2C(O)=C(c(cc(c3)O)c4cccc4)cc3)C1=O)scc2C(C)CC)[H]		7,02
N1([H])C(=O)C(=C(c2c1sc(CCC3)c23)O)c(cc(cc4)O)c5cccc5)c4		6,93
N1([H])C(=O)C(=C(c2c1sc3c2CCCC3)O)c4cc(ccc4)O)c5cccc5		6,41
N1([H])C(=O)C(=C(c2c1sc3c2SCCC3)O)c4cc(ccc4)O)c5cccc5		7,43
N1([H])C(=O)C(c(cc(c2)O)c3cccc3)cc2)=C(c4c1sc(CN)c4C)O		5,70
N([H])(C(=O)C(c(cc1O)c2cccc2)ccc1)=C3O)c(c34)sc(Cl)c4		7,78
N1([H])C(=O)C(c(cc2O)c3cccc3)ccc2)=C(c4c1sc(Cl)c4C)O		7,64
N1([H])C(=O)C(c(cc(c2)O)c3cccc3)cc2)=C(c4c1sc(Cl)c4CC)O		8,95
N1(c(c2C(O)=C(c(cc(c3)O)c4cccc4)cc3)C1=O)sc(Cl)c2C(C)C)[H]		8,84
N1(c(c2C(O)=C(c(cc(c3)O)c4cccc4)cc3)C1=S)sc(Cl)c2C(C)C)[H]		6,86
N1(c(c2C(O)=C(c(cc(c3)O)c4cccc4)cc3)C1=O)sc(Cl)c2CCC)[H]		8,39
N1(c(c2C(O)=C(c(cc(c3)O)c4cccc4)cc3)C1=O)sc(Cl)c2C5CC5)[H]		8,02
N1(c(c2C(O)=C(c(cc(c3)O)c4cccc4)cc3)C1=O)sc(Cl)c2COC)[H]		8,80
N1(c(c2C(O)=C(c(cc(cc3)O)c4cccc4)c3)C1=O)sc(Cl)c2C(CC)C)[H]		8,70
N1(c(c2C(O)=C(c(cc(cc3)O)c4cccc4)c3)C1=O)sc(Cl)c2C(C)(C)C)[H]		8,48

<chem>N(C1=O)(c(c2C(O)=C1c3cc(ccc3)Oc4cccc4)sc(Cl)c2CCCC)[H]</chem>		7,10
<chem>N1([H])C(=O)C(c(cc(c2)Oc3cccc3)cc2)=C(c4c1sc(Cl)c4CO)O</chem>		7,16
<chem>N1(c(c2C(O)=C(c(cc(c3)Oc4cccc4)cc3)C1=O)sc(Cl)c2CCN)[H]</chem>		4,78
<chem>c(ccc(N1)c2C(C=CC3)C3C1C(=O)O)(Cl)c2</chem>	Grünenthal, WO 03/013530 A2	6,52
<chem>c(c(Cl)cc(NC1C(=O)[O-])c2C(C=CC3)C13)c2Cl</chem>	Grünenthal, WO 03/013530 A2	6,46
<chem>c1(C(O)=O)[nH]c(c2c1CCC(O)=O)cccc2</chem>	Salituro <i>et al.</i> ²⁴²	4,57
<chem>c1(CCC(O)=O)c2c([nH]c1C(O)=O)cccc2Cl</chem>		5,22
<chem>c1(C(O)=O)[nH]c(c2c1CCC(O)=O)ccc(Cl)c2</chem>		4,64
<chem>c1(C(O)=O)c(CCC(O)=O)c(c2[nH]1)ccc(Cl)c2</chem>		5,70
<chem>c1(CCC(O)=O)c2c([nH]c1C(O)=O)cc(Cl)cc2Cl</chem>		6,77
<chem>n(C)(c(cc(Cl)cc1Cl)c1c2CCC(O)=O)c2C(O)=O</chem>		4,35
<chem>c1(CCC(O)=O)c2c([nH]c1C(O)=O)cccc2F</chem>		4,77
<chem>c1(C(O)=O)c(CCC(O)=O)c(c2[nH]1)ccc(F)c2</chem>		5,30
<chem>c1(CCC(O)=O)c2c([nH]c1C(O)=O)cc(F)cc2F</chem>		5,40
<chem>c1(C(O)=O)c(CCC(O)=O)c(c2[nH]1)ccc(OC)c2</chem>		4,89
<chem>c1(CCC(O)=O)c2c([nH]c1C(O)=O)cccc2N(=O)=O</chem>		4,12
<chem>c1(C(O)=O)c(CCC(O)=O)c(c2[nH]1)ccc(c2)N(=O)=O</chem>		5,60
<chem>c1(CCC(O)=O)c2c([nH]c1C(O)=O)cccc2C(F)(F)F</chem>		4,89
<chem>c1(C(O)=O)c(CCC(O)=O)c(c(c2)[nH]1)ccc2C(F)(F)F</chem>		5,10
<chem>c1(C(O)=O)[nH]c(c2c1SCC(O)=O)cccc2</chem>		4,92
<chem>c1(C(O)=O)[nH]c(c2c1SCC(O)=O)cc(cc2Cl)Cl</chem>		7,00
<chem>c1(C(O)=O)cc(c2[nH]1)cccc2</chem>		3,97
<chem>n1c(cccc2)c2c(cc1C(O)=O)OCC(O)=O</chem>		4,60
<chem>c12c(cc(Cl)cc1Cl)N[C@@H](C[C@@H]2C(=O)OCc3cccc3)C(O)=O</chem>	Carling <i>et al.</i> ²²⁴	5,63
<chem>c12c(cc(Cl)cc1Cl)N[C@H](C[C@@H]2C(=O)OCc3cccc3)C(O)=O</chem>		5,98
<chem>c1([C@H](C[C@@H](N2)C(O)=O)NC(=O)[C@@H](N)Cc3cccc3)c2cc(Cl)cc1Cl</chem>		5,98
<chem>c1([C@H](C[C@@H](N2)C(O)=O)NC(c(cc3)ccn3)=O)c2cc(Cl)cc1Cl</chem>		6,20
<chem>c1([C@@H](NC(=O)Nc2cccc(C)c2)C[C@@H](N3)C(O)=O)c3cc(Cl)cc1Cl</chem>		7,74
<chem>c1([C@@H](NC(=O)Nc2cccc2C)C[C@@H](N3)C(O)=O)c3cc(Cl)cc1Cl</chem>		8,14
<chem>c1([C@@H](NC(=S)Nc2cccc2)C[C@@H](N3)C(O)=O)c3cc(Cl)cc1Cl</chem>		7,05
<chem>c1([C@H](C[C@@H](N2)C(O)=O)N\C(\Nc3cccc3)=N/C#N)c2cc(Cl)cc1Cl</chem>		6,24
<chem>c12c(cc(c(Cl)c1N(=O)=O)Cl)NC(C(=O)N2)=O</chem>	Keana <i>et al.</i> ²⁴³	
<chem>c12c(cc(c(Br)c1N(=O)=O)Br)NC(C(=O)N2)=O</chem>		
<chem>c12c(cc(c(Cl)c1N(=O)=O)Br)NC(C(=O)N2)=O</chem>		
<chem>c12c(cc(c(Br)c1N(=O)=O)Cl)NC(C(=O)N2)=O</chem>		
<chem>c1(N(=O)=O)c(Br)c(cc(NC(C(=O)N2)=O)c12)C(F)(F)F</chem>		
<chem>c1(N(=O)=O)c(Cl)c(cc(NC(C(=O)N2)=O)c12)C(F)(F)F</chem>		
<chem>c12c(cc(c(F)c1N(=O)=O)Br)NC(C(=O)N2)=O</chem>		
<chem>c1(N(=O)=O)c(c(Cl)cc(NC(C(=O)N2)=O)c12)N(=O)=O</chem>		
<chem>c12c(cc(c(Br)c1N)Br)NC(C(=O)N2)=O</chem>		

<chem>c12c(cc(c(F)c1N(=O)=O)Cl)NC(C(=O)N2)=O</chem>		
<chem>c12c(cc(c(Br)c1N(=O)=O)F)NC(C(=O)N2)=O</chem>		
<chem>c(c1)(C(F)(F)F)c(Cl)c(c2c1NC(=O)C(N2)=O)Cl</chem>		
<chem>c12c(cc(c(Cl)c1C(F)(F)F)Cl)NC(C(=O)N2)=O</chem>		
<chem>N(c1c2cc(Br)c(Br)c1)C(=O)C(N2)=O</chem>		
<chem>c1(N(=O)=O)c(Cl)c2c(cc1C(F)(F)F)NC(C(=O)N2)=O</chem>		
<chem>c12c(cc(c(Cl)c1NC(C(=O)Cl)NC(C(=O)N2)=O</chem>		
<chem>c1(cc(Br)cc(c12)NC(C(=O)N2)=O)C(F)(F)F</chem>		
<chem>c(c(Br)cc(c12)NC(C(=O)N1)=O)(c2)C(F)(F)F</chem>		
<chem>c(c(Cl)cc(c12)NC(C(=O)N1)=O)(c2)C(F)(F)F</chem>		
<chem>c12c(c(F)c(c(F)c1N(=O)=O)Cl)NC(=O)C(=O)N2</chem>		
<chem>c1(c(Cl)cc(c2c1)NC(C(=O)N2)=O)N(=O)=O</chem>		
<chem>c12c(cc(c(F)c1N(=O)=O)F)NC(C(=O)N2)=O</chem>		
<chem>c1(cc(cc(NC(C(=O)N2)=O)c12)C(F)(F)F)N(=O)=O</chem>		
<chem>N(c1c2cc(Br)c(Cl)c1)C(=O)C(N2)=O</chem>		
<chem>c1(cc(cc(c12)NC(C(=O)N2)=O)C(F)(F)F)C(F)(F)F</chem>		
<chem>c12c(ccc(Cl)c1N(=O)=O)NC(C(=O)N2)=O</chem>		
<chem>c1(cc(Cl)cc(NC(C(=O)N2)=O)c12)N(=O)=O</chem>		
<chem>c1(cc(Cl)cc(c12)NC(C(=O)N2)=O)C(F)(F)F</chem>		
<chem>N1C(=O)C(=O)Nc(cc(Br)cc2Br)c12</chem>		
<chem>N1C(=O)C(=O)Nc(cc2C(F)(F)F)c1c(Br)c2</chem>		
<chem>c12c(c(Cl)c(c(Cl)c1N(=O)=O)Cl)NC(=O)C(=O)N2</chem>		
<chem>c12c(ccc(Br)c1N(=O)=O)NC(C(=O)N2)=O</chem>		
<chem>c12c(cc(F)c(F)c1Cl)NC(C(=O)N2)=O</chem>		
<chem>N1c2c(ccc(Br)c2)NC(=O)C1=O</chem>		
<chem>N(=O)(=O)c(ccc(c12)NC(C(=O)N1)=O)c2</chem>		
<chem>c(c(F)cc(c12)NC(C(=O)N1)=O)(c2)C(F)(F)F</chem>		
<chem>c1(cc(F)cc(c12)NC(C(=O)N2)=O)C(F)(F)F</chem>		
<chem>N(c1c2cc(F)c(F)c1)C(=O)C(N2)=O</chem>		
<chem>c12c(cc(F)c(F)c1F)NC(C(=O)N2)=O</chem>		
<chem>c12c(cccc1N(=O)=O)NC(C(=O)N2)=O</chem>		
<hr/>		
<chem>N1c(cccc2)c2C(C=C(O)C1=O)=O</chem>	Guzikowski <i>et al.</i> ²⁴⁴	6,08
<chem>N1c(cc(C)cc2)c2C(C=C(O)C1=O)=O</chem>		6,89
<chem>N1C(=O)C(O)=CC(=O)c(ccc(Cl)c2)c12</chem>		7,89
<chem>N1C(=O)C(O)=CC(=O)c(ccc(Br)c2)c12</chem>		7,10
<chem>N1C(=O)C(O)=CC(=O)c(ccc(F)c2)c12</chem>		6,31
<chem>N(=O)(=O)c(ccc1c2C(=O)C=C(C(=O)N1)O)c2</chem>		4,92
<chem>c1(C(=O)C=C(C(=O)N2)O)c2cc(C)c(C)c1</chem>		6,62
<chem>C(=O)(c1c2cccc1C)C=C(O)C(N2)=O</chem>		4,89
<chem>C(=O)(c1c2cc(C)cc1C)C=C(O)C(N2)=O</chem>		6,52
<chem>c12c(cc(C)c(c1C)N(=O)=O)NC(C(O)=CC2=O)=O</chem>		5,66

<chem>N1c(cc(CC)cc2)c2C(C=C(O)C1=O)=O</chem>		5,29
<chem>C1(N(=O)=O)=C(O)c2c(cccc2)NC1=O</chem>		4,46
<chem>C1(N(=O)=O)=C(O)c(c2NC1=O)ccc(Cl)c2</chem>		5,82
<chem>C1(N(=O)=O)=C(O)c2c(cccc2Cl)NC1=O</chem>		4,21
<chem>C1(N(=O)=O)=C(O)c2c(cc(cc2Cl)Cl)NC1=O</chem>		6,66
<chem>C1(N(=O)=O)=C(O)c2c(cc(c(Cl)c2)Cl)NC1=O</chem>		6,40
<chem>C1(N(=O)=O)=C(O)c2c(cc(cc2C)C)NC1=O</chem>		6,54
<chem>C(N(=O)=O)(C(=O)Nc(c12)cc(c(Cl)c1Cl)Cl)=C2O</chem>		6,66
<chem>C(N(=O)=O)(C(=O)Nc1c2cc(c(Cl)c1F)F)=C2O</chem>		4,40
<chem>c12c(cc(cc1N(=O)=O)C)NC(C(=O)N2)=O</chem>	Cai et al. ²⁴⁵	5,03
<chem>c12c(cc(c(C)c1N(=O)=O)C)NC(C(=O)N2)=O</chem>		7,54
<chem>c12c(cc(c(CC)c1N(=O)=O)CC)NC(C(=O)N2)=O</chem>		6,80
<chem>c12c(cc(c3c1N(=O)=O)CCC3)NC(C(=O)N2)=O</chem>		5,51
<chem>c12c(cc(c(OC)c1N(=O)=O)Br)NC(C(=O)N2)=O</chem>		7,19
<chem>c12c(cc(c(Br)c1N(=O)=O)CC)NC(C(=O)N2)=O</chem>		7,09
<chem>c12c(cc(c(Cl)c1N(=O)=O)CC)NC(C(=O)N2)=O</chem>		7,54
<chem>c12c(cc(c(CC)c1N(=O)=O)Cl)NC(C(=O)N2)=O</chem>		6,89
<chem>c12c(cc(c(C)c1N(=O)=O)C#N)NC(C(=O)N2)=O</chem>		7,14
<chem>c12c(cc(c(C)c1N(=O)=O)F)NC(C(=O)N2)=O</chem>		7,02
<chem>c12c(cc(c(C)c1N(=O)=O)Br)NC(C(=O)N2)=O</chem>		8,06
<chem>c12c(cc(c(C)c1N(=O)=O)Cl)NC(C(=O)N2)=O</chem>		8,33
<chem>c12c(cc(c(Cl)c1N(=O)=O)C)NC(C(=O)N2)=O</chem>		7,35
<chem>c12c(cc(cc1N(=O)=O)C#N)NC(C(=O)N2)=O</chem>		5,55
<chem>c12c(cc(c(N)c1N(=O)=O)F)NC(C(=O)N2)=O</chem>		4,96
<chem>c12c(cc(c(OC)c1N(=O)=O)F)NC(C(=O)N2)=O</chem>		5,72
<chem>c12c(cc(c(OCC)c1N(=O)=O)F)NC(C(=O)N2)=O</chem>		5,44
<chem>c12c(cc(c(OCCCC)c1N(=O)=O)F)NC(C(=O)N2)=O</chem>		5,37
<chem>c12c(cc(c(OCCCCc3ccccc3)c1N(=O)=O)F)NC(C(=O)N2)=O</chem>		5,40
<chem>c12c(cc(c(OC)c1N(=O)=O)Cl)NC(C(=O)N2)=O</chem>		6,82
<chem>c12c(cc(c(SCC)c1N(=O)=O)Cl)NC(C(=O)N2)=O</chem>		6,19
<chem>C1(c2ccccc2)=C(OC\ C\ C)=N\ O)c(c3NC1=O)ccc(Cl)c3</chem>		6,11
<chem>c1(CCC(=O)Nc2ccccc2)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl</chem>		
<chem>c1(\C=C\ C(O)=O)c2c([nH]c1C(O)=O)cc(Cl)cc2Cl</chem>		
<chem>c1(\C=C\ C(O)=O)OC(C)(C)c2c([nH]c1C(O)=O)cc(Cl)cc2Cl</chem>		
<chem>c1(\C=C\ S(=O)(=O)Nc2ccccc2)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl</chem>		
<chem>c1(\C=C\ C(O)=O)NCc2ccccc2)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl</chem>		
<chem>c1(\C=C\ C(O)=O)Nc(cc2)ccc2NC(N)=O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl</chem>		
<chem>c1(\C=C\ C(O)=O)Nc2ccccc(N)c2)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl</chem>		
<chem>c1(\C=C\ C(O)=O)Nc2ccccc2N)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl</chem>		
<chem>c1(\C=C\ C(O)=O)Nc(cc2)ccc2N)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl</chem>		
<chem>c1(\C=C\ C(O)=O)Nc(cc2)ccc2O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl</chem>		

c1(\C=C\C(=O)Nc2cccc(O)c2)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc2cccc2O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc2cccc2N(=O)=O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc2ccc(c(OC)c2)OC)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc2ccc(cc2C)OC)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc2ccc(cc2N(=O)=O)F)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(ccc2N(=O)=O)c2F)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc2ccc(cc2F)F)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2C(O)=O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(ccc2N(=O)=O)c2O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2N(C)C)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2N(=O)=O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2F)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2OCC)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc2cccc2C(C)C)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(ccc2N(=O)=O)Cl)c2)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2N(CC)CC)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2C(F)(F)F)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2Nc3cccc3)c4c([nH]c1C(O)=O)cc(Cl)cc4Cl

C1(c2ccc(cc2)O)=C(O)c3c(cc(cc3Cl)Cl)NC1=O

Balsamini *et al.*²⁴⁶

7,89

c1(C(O)=O)[nH]c(c(CC)c1\C=C\C(=O)Nc2cccc2)C
c1(C(O)=O)[nH]c(c(C)C1\C=C\C(=O)Nc2cccc2)CC
c1(C(O)=O)[nH]c(c(C)C1\C=C\C(=O)Nc2cccc2)C
c1(C(O)=O)[nH]c(c2c1\C=C\C(=O)Nc3cccc3)CCC2
c1(C(O)=O)[nH]c(c2c1\C=C\C(=O)Nc3cccc3)CCCC2
c1(\C=C\C(=O)Nc2cccc2)c(Br)c([nH]c1C(O)=O)C
c1(\C=C\C(=O)Nc2cccc2)c(Cl)c([nH]c1C(O)=O)C
c1(\C=C\C(=O)Nc2cccc2)c(C(C)C)c([nH]c1C(O)=O)C
c1(\C=C\C(=O)Nc2cccc2)c(C)c([nH]c1C(O)=O)C(C)C
c1(\C=C\C(=O)Nc2cccc2)c(OCC)c([nH]c1C(O)=O)C
c1(\C=C\C(=O)Nc2cccc2)c(OCCc3cccc3)c([nH]c1C(O)=O)C
c1(C(O)=O)[nH]c(cc1\C=C\C(=O)Nc2cccc2)C
c1(c(Br)c[nH]c1C(O)=O)\C=C\C(=O)Nc2cccc2
c1(\C=C\C(=O)Nc2cccc2)c(Br)c([nH]c1C(O)=O)Br
c1(\C=C\C(=O)Nc2cccc2)c(I)c([nH]c1C(O)=O)I
c1(\C=C\C(=O)Nc2cccc2)c(I)c([nH]c1C(O)=O)C
c1(\C=C\C(=O)Nc2cccc2)c(C(C)=O)c([nH]c1C(O)=O)C

c12n(c3c(ccc3)NC1=O)nc(C(OCC)=O)n2

Catarzi *et al.*²⁴⁷

c12n(c(c3NC1=O)ccc(Cl)c3)nc(C(OCC)=O)n2

c12n(c3c(ccc(Cl)c3)NC1=O)nc(C(OCC)=O)n2

c1(C(=O)Nc(c23)cc(cc2Cl)Cl)n3nc(C(OCC)=O)n1

c1(C=O)Nc(c23)cc(c(Cl)c2)Cl)n3nc(C(OCC)=O)n1
c12n(c(c(c3)NC1=O)ccc3C(F)(F)F)nc(C(OCC)=O)n2
c12n(c(c3NC1=O)ccc(Br)c3)nc(C(OCC)=O)n2
c12n(c3c(cccc3)NC1=O)nc(C(O)=O)n2
c12n(c(c3NC1=O)ccc(Cl)c3)nc(C(O)=O)n2
c12n(c3c(ccc(Cl)c3)NC1=O)nc(C(O)=O)n2
c1(C=O)Nc(c23)cc(cc2Cl)Cl)n3nc(C(O)=O)n1
c1(C=O)Nc(c23)cc(c(Cl)c2)Cl)n3nc(C(O)=O)n1
c12n(c(c(c3)NC1=O)ccc3C(F)(F)F)nc(C(O)=O)n2
c12n(c(c3NC1=O)ccc(c3)N(=O)=O)nc(C(O)=O)n2
c12n(c(c3NC1=O)ccc(Br)c3)nc(C(O)=O)n2
c1(C=O)Nc(c23)ccc(c2)N(=O)=O)n3nc(C(O)=O)n1
c1(C=O)Nc(c2c3)cc(c3N(=O)=O)Cl)n2nc(C(O)=O)n1
c1(C=O)N2)n(c(c23)cc(cc3N(=O)=O)Cl)nc(C(O)=O)n1
c12n(c(c3N(CC)C1=O)ccc(Cl)c3)nc(C(O)=O)n2
c12n(c(c3NC1=O)ccc(Cl)c3)nc(CO)n2
c12n(c(c3NC1=O)ccc(Cl)c3)nc(CC#N)n2
c12n(c(c3NC1=O)ccc(Cl)c3)nc(CC(OCC)=O)n2
c12n(c(c3NC1=O)ccc(Cl)c3)nc(CC(O)=O)n2
c1(n2)n(c(c3NC1=O)ccc(Cl)c3)nc2C(NO)=O
c1(n2)n(c(c3NC1=O)ccc(Cl)c3)nc2C(NN)=O
c12n(c(c3NC1=O)ccc(Cl)c3)nc(C(N)=O)n2
c12n(c(c3NC1=O)ccc(Cl)c3)nc(C(=O)Nc4[nH]nnn4)n2
c12n(c(c3NC1=O)ccc(Cl)c3)nc(C(NCCO)=O)n2
c12n(c(c3NC1=O)ccc(Cl)c3)nc(C(=O)NCC(OC)=O)n2
c12n(c(c3NC1=O)ccc(Cl)c3)nc(C(=O)NCC(O)=O)n2

c1(\C=C\C(=O)Nc(cc2)ccc2CCNC(N)=O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2CNC(NCC)=O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2CNC(NCC)=S)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2CNC(=O)NCC(O)=O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2CNC(=O)NC3CC3)c4c([nH]c1C(O)=O)cc(Cl)cc4Cl
c1(\C=C\C(=O)Nc(cc2)ccc2CNC(=O)NC(CC3)CCO3)c4c([nH]c1C(O)=O)cc(Cl)cc4Cl
c1(\C=C\C(=O)Nc(cc2)ccc2CNC(=O)Nc3ccccc3)c4c([nH]c1C(O)=O)cc(Cl)cc4Cl
c1(\C=C\C(=O)Nc(cc2)ccc2CNC(=O)Nc(cc3)ccc3OC)c4c([nH]c1C(O)=O)cc(Cl)cc4Cl
c1(\C=C\C(=O)Nc(cc2)ccc2CNC(=O)Nc(cccn3)c3)c4c([nH]c1C(O)=O)cc(Cl)cc4Cl
c1(\C=C\C(=O)Nc(cc2)ccc2[C@@H](C)NC(N)=O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl
c1(\C=C\C(=O)Nc(cc2)ccc2[C@H](C)NC(N)=O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl

Di Fabio *et al.*²⁴⁸

C1(c2cccc(Oc3ccccc3)c2)=C(O)c(c4NC1=O)ccc(c4)N=C=S
C1(c2cccc(Oc3ccccc3)c2)=C(O)c(c4NC1=O)ccc(c4)NC(CCl)=O
C1(c2cccc(Oc3ccccc3)c2)=C(OC(CCl)=O)c(c4NC1=O)ccc(c4)NC(CCl)=O
C1(c2cccc(Oc3ccccc3)c2)=C(OC(=O)c(cc4)ccc4S(F)(=O)=O)c(c5NC1=O)ccc(Cl)c5

Kreimeyer *et al.*²⁴⁹

<chem>C1(c2cccc(c2)N=C=S)=C(O)c(c3NC1=O)ccc(Cl)c3</chem>		
<chem>C1(c2cccc(c2)NC(CCl)=O)=C(O)c(c3NC1=O)ccc(Cl)c3</chem>		
<chem>C1(c2cccc(c2)NCCS(F)(=O)=O)=C(O)c(c3NC1=O)ccc(Cl)c3</chem>		
<chem>C1(c2cccc(c2)Oc3cccc3NC(CCl)=O)=C(O)c(c4NC1=O)ccc(Cl)c4</chem>		
<chem>C1(c2cccc(c2)Oc3cccc3NCCS(F)(=O)=O)=C(O)c(c4NC1=O)ccc(Cl)c4</chem>		
<chem>C1(c2cccc(c2)Oc3cccc(c3)NC(CCl)=O)=C(O)c(c4NC1=O)ccc(Cl)c4</chem>		
<chem>C1(c2cccc(Oc3cccc3)c2)=C(NC(CCl)=O)c(c4NC1=O)ccc(Cl)c4</chem>		
<hr/>		
<chem>c12c(cc(c(Cl)c1Cl)Cl)NC(=CC2=O)C(O)=O</chem>	Leeson <i>et al.</i> ²²⁹	5,62
<chem>C(C(O)=O)=(C1)Nc(c2C1=O)cccc2I</chem>		6,20
<chem>C1(C(O)=O)=CC(=O)c(c2N1)ccc(Br)c2</chem>		6,07
<chem>C1(C(O)=O)=CC(=O)c(c2N1)ccc(I)c2</chem>		5,44
<chem>C1(C(O)=O)=CC(=O)c2c(cc(cc2)C)N1</chem>		5,74
<chem>c1(c(cc(C)cc1Br)N2)C(=O)C=C2C(O)=O</chem>		6,09
<chem>c1(c(cc(CC)cc1Br)N2)C(=O)C=C2C(O)=O</chem>		5,49
<chem>C(C(O)=O)=(C1)Nc(c2C1=O)cc(cc2C)Br</chem>		5,96
<chem>C(C(O)=O)=(C1)Nc(c2C1=O)cc(cc2CC)Br</chem>		7,18
<chem>C(C(O)=O)=(C1)Nc(c2C1=O)cc(cc2C)I</chem>		5,92
<chem>C(C(O)=O)=(C1)Nc(c2C1=O)cc(cc2CC)Cl</chem>		6,59
<chem>c1(c(cc(CC)cc1Cl)N2)C(=O)C=C2C(O)=O</chem>		5,31
<chem>c1(c(OC)cc(n2)C(O)=O)c2cc(Cl)cc1Cl</chem>		4,00
<chem>C(O)(CNc(cc(Cl)cc1Cl)c1)=O</chem>		4,31
<chem>C(=O)(Nc1cc(Cl)cc(Cl)c1)C(O)=O</chem>		5,17
<chem>N1c2c(cccc2Cl)NC(=O)C1=O</chem>		4,25
<chem>c1(cc2c(cc1N(=O)=O)NC(C(=O)N2)=O)N(=O)=O</chem>		5,35
<chem>N1(C)c2c(cc(C)c(C)c2)NC(=O)C1=O</chem>		5,18
<hr/>		
<chem>c1(C(O)=O)cc2c(ccc(F)c2)[nH]1</chem>	Gray <i>et al.</i> ²⁵⁰	
<chem>c1(C(O)=O)cc(c2[nH]1)ccc(Cl)c2</chem>		
<chem>c1(C(O)=O)[nH]c(c2c1CC(O)=O)cccc2</chem>		
<chem>c1(C(O)=O)[nH]c(c2c1CC(O)=O)ccc(Cl)c2</chem>		
<chem>c1(C(O)=O)[nH]c(c2c1CC(O)=O)ccc(Br)c2</chem>		
<chem>c1(C(O)=O)[nH]c(c2c1CC(O)=O)ccc(F)c2</chem>		
<chem>c1(C(O)=O)[nH]c2c(ccc(Cl)c2)c1CC(OC)=O</chem>		
<chem>c1(C(O)=O)[nH]c2c(ccc(Cl)c2)c1CC(OCCC)=O</chem>		
<chem>c1(C(O)=O)[nH]c2c(ccc(Cl)c2)c1CC(N)=O</chem>		
<chem>c1(C(O)=O)[nH]c2c(ccc(Cl)c2)c1CC(NC)=O</chem>		
<chem>c1(C(O)=O)[nH]c2c(ccc(Cl)c2)c1CC(NCCO)=O</chem>		
<chem>c1(C(O)=O)[nH]c2c(ccc(Cl)c2)c1CC(NCC)=O</chem>		
<chem>c1(C(O)=O)[nH]c2c(ccc(Cl)c2)c1CC(NCCC)=O</chem>		
<chem>c1(C(O)=O)[nH]c2c(ccc(Cl)c2)c1CC(NCCc3cccc3)=O</chem>		
<chem>c1(C(O)=O)c(CCO)c(c2[nH]1)ccc(Cl)c2</chem>		
<chem>c1(C(O)=O)c(CCCl)c(c2[nH]1)ccc(Cl)c2</chem>		

<chem>c1(C(O)=O)c(CCOCC)c(c2[nH]1)ccc(Cl)c2</chem>		
<chem>c1(C(O)=O)[nH]c(c2c1C=O)cccc2</chem>		
<chem>c1(C(O)=O)[nH]c(c2c1CC(C)=O)cccc2</chem>		
<chem>n12c(c(c(cc(c3N(=O)=O)Cl)NC1=O)c3)cc(C(OCC)=O)n2</chem>		
<chem>n12c(c(c(cc(c3N(=O)=O)Cl)NC1=O)c3)cc(C(O)=O)n2</chem>		
<chem>n12c(c3c(cc(c(Cl)c3)Cl)NC1=O)cc(C(OC)=O)n2</chem>		
<chem>n12c(c(c(cc(c3n(cnn4)c4)Cl)NC1=O)c3)cc(C(O)=O)n2</chem>		
<chem>n12c(c(c(cc(c3n4ccc(C=O)c4)Cl)NC1=O)c3)cc(C(OCC)=O)n2</chem>		
<chem>n12c(c(c(cc(c3n4ccc(C=O)c4)Cl)NC1=O)c3)cc(C(O)=O)n2</chem>		
<chem>N12C(c3c(NC1=O)cccc3)CC(C(O)=O)=N2</chem>		
<chem>N12C(CC(C(OC)=O)=N1)c(ccc(c3)Cl)c3NC2=O</chem>		
<chem>N12C(CC(C(O)=O)=N1)c(ccc(c3)Cl)c3NC2=O</chem>		
<chem>N12C(CC(C(OC)=O)=N1)c3c(NC2=O)cc(Cl)c(Cl)c3</chem>		
<chem>N12C(CC(C(O)=O)=N1)c3c(NC2=O)cc(Cl)c(Cl)c3</chem>		
<hr/>		
<chem>c1(C(O)=O)cc(cc(NC(C(=O)N2)=O)c12)N(=O)=O</chem>	Zhou <i>et al.</i> ²²¹	5,00
<chem>c1(cc(Cl)cc(c12)NC(C(=O)N2)=O)C(c3cccc3)=O</chem>		5,17
<chem>c1(cc(Cl)cc(c12)NC(C(=O)N2)=O)C(=O)N(C)C</chem>		4,89
<chem>c1(cc(Cl)cc(c12)NC(C(=O)N2)=O)C(=O)N(CC)CC</chem>		4,62
<chem>c12c(cc(c(C)c1C(OC)=O)Br)NC(C(=O)N2)=O</chem>		6,40
<chem>c12c(cc(c(CCC3=O)c13)Cl)NC(C(=O)N2)=O</chem>		5,34
<chem>c12c(COC1=O)c(Cl)cc(NC(C(=O)N3)=O)c23</chem>		5,54
<chem>N(c1c2cc(Cl)cc1C#N)C(=O)C(N2)=O</chem>		7,17
<hr/>		
<chem>N1(c2cccc2)C(=O)CN(Cc3c4c([nH]c3C(O)=O)cc(Cl)cc4Cl)C1=S</chem>	Jansen <i>et al.</i> ²⁰⁸	
<chem>C1(C2=O)=C(Nc3c(ccc(c3)Cl)C1=O)C(NN2C(c(cc4)ccn4)CC(C)C)=O</chem>	Brown <i>et al.</i> ²¹⁵	
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc2cccc2)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2N)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2NC(C)=O)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2NC(COC)=O)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2NS(C(=O)=O)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>S(N)(=O)(=O)Nc(cc1)ccc1NC(=O)\C=C(/CC2C(O)=O)\c3c(N2)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2CC(N)=O)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2CC#N)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2CC(O)=O)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2CNC(N)=O)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2CNC(C)=O)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2CNC(C(C)C)=O)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2\C=C\C(N)=O)\c3c(N1)cc(Cl)cc3Cl</chem>		
<chem>C(/CC1C(O)=O)(=C/C(=O)Nc(cc2)ccc2\C=C\C#N)\c3c(N1)cc(Cl)cc3Cl</chem>		
<hr/>		
<chem>C12=C(c3c(cccc3)C1(C)CC(O)=O)NC(=O)C(=O)N2</chem>	Jimonet <i>et al.</i> ²⁵¹	7,68
<chem>C12=C(c(c3C1(C)CC(O)=O)cccc3Cl)NC(=O)C(=O)N2</chem>		7,10
<chem>C12=C(c(c3C1(C)CC(O)=O)cc(cc3Cl)Cl)NC(=O)C(=O)N2</chem>		7,62

<chem>C12=C(NC(=O)C(=O)N1)c3c(cc(c(Cl)c3)Cl)C2(C)CC(O)=O</chem>		6,85
<chem>C1(C(C)(CC(O)=O)c(c23)ccc(Cl)c2Cl)=C3NC(=O)C(=O)N1</chem>		5,59
<chem>C12=C(c(c3C1(C)CC(O)=O)cc(cc3)C)NC(=O)C(=O)N2</chem>		7,92
<chem>C12=C(c(c3C1(C)CC(O)=O)cc(cc3)F)NC(=O)C(=O)N2</chem>		7,82
<chem>C12=C(c(c3C1(C)CC(O)=O)cc(cc3)OC(F)(F)F)NC(=O)C(=O)N2</chem>		5,37
<chem>C12=C(c(c3C1(CCC)CC(O)=O)cc(cc3)Cl)NC(=O)C(=O)N2</chem>		7,68
<chem>C1(C[C@H](Cl)[C@H](C2)Cl)=C2C(=O)C=C(O)C(=O)N1</chem>	Guzikowski <i>et al.</i> ²⁵²	6,89
<chem>c1(\C=C(/c(cccc2N(=O)=O)c2)\C(O)=O)c3c([nH]c1C(O)=O)cc(Cl)cc3Cl</chem>	Baron <i>et al.</i> ¹³⁵	7,68
<chem>c1(ccs2)c2NC(C(c3ccccc3)=C1O)=O</chem>	Buchstaller <i>et al.</i> ²⁵³	
<chem>C(c1ccccc1)(C(=O)Nc(c2c3)cs3)=C2O</chem>		
<chem>C1(c2ccccc2)=C(O)c3c(ccs3)NC1=O</chem>		
<chem>C1(c2ccccc2)=C(O)c3c(cccc3)NC1=O</chem>		
<chem>c1(cc(s2)CC)c2NC(=O)C(c3ccccc3)=C1O</chem>		
<chem>c1(c(CCCS2)c2s3)c3NC(=O)C(c4cccc(Oc5ccccc5)c4)=C1O</chem>		
<chem>N1(O)C(=O)Nc(c2C1=O)cc(c(Cl)c2)Cl</chem>	Colotta <i>et al.</i> ²²²	
<chem>N1(O)C(=O)Nc(c(c2)C1=O)cc(c2N(=O)=O)Cl</chem>		
<chem>N1(O)C(=O)Nc(c2C1=O)cc(c(N)c2)Cl</chem>		
<chem>N1(c2ccccc2)C(=O)c3c(NC1=O)cc(Cl)cc3Cl</chem>		
<chem>C(=N\O)/C1=O\C(NC2ccccc2)=N/C/1=C1c3ccccc3</chem>	Poschenrieder <i>et al.</i> ¹³¹	
<chem>C1(Nc2ccccc2)=N\C(\C(=O)/C/1=N/O)=C/c3ccccc3</chem>		
<chem>C1(Nc(cc2)ccc2Cl)=N\C(\C(=O)/C/1=N/O)=C/c3ccccc3</chem>		
<chem>C1(N(=O)=O)=C(OC)N\C(=C/c2ccccc2)\C1=O</chem>		
<chem>C1(=C(/Br)\c2ccccc2)/C(=O)\C(\C(Nc3ccccc3)=N1)=N\O</chem>		
<chem>C(=C(\Br)/c1ccccc1)/(NC(=C2N(=O)=O)OC)\C2=O</chem>		
<chem>C(=C(\Br)/c1ccccc1)/(NC(=C2N(=O)=O)Nc3ccccc3)\C2=O</chem>		
<chem>C(=C(\Br)/c1ccccc1)/(N2)\C(O)=C(C2=O)N(=O)=O</chem>		
<chem>C1(N(=O)=O)=C(Nc2ccccc2)\C(\NC1=O)=C1c3ccccc3</chem>		
<chem>c12c(oc(S)n1)c(c3NC2=O)ccc(Cl)c3</chem>	Calabri <i>et al.</i> ²⁵⁴	
<chem>c1(C(=O)Nc(c23)cc(cc2Cl)Cl)c3oc(S)n1</chem>		
<chem>c1(C(=O)Nc(c23)cc(c(Cl)c2)Cl)c3oc(S)n1</chem>		
<chem>c12c(oc(S)n1)c(c(c3)NC2=O)ccc3C(F)(F)F</chem>		
<chem>C12=C(OC(=O)N1)c(c3NC2=O)ccc(Cl)c3</chem>		
<chem>C1(C(=O)Nc(c23)cc(cc2Cl)Cl)=C3OC(=O)N1</chem>		
<chem>C1(C(=O)Nc(c23)cc(c(Cl)c2)Cl)=C3OC(=O)N1</chem>		
<chem>C12=C(OC(=O)N1)c(c(c3)NC2=O)ccc3C(F)(F)F</chem>		
<chem>C1(NC(C)=O)=C(O)c(c2NC1=O)ccc(Cl)c2</chem>		
<chem>C(NC(C)=O)(C(=O)Nc(c12)cc(cc1Cl)Cl)=C2O</chem>		
<chem>C1(NC(C)=O)=C(O)c2c(cc(c(Cl)c2)Cl)NC1=O</chem>		

A.2 Hann-Filter

```

*/ Hann, Hudson, Lewell, Lively, Miller, Ramsden
* Strategic Pooling of Compounds for High-Throughput Screening
* J. Chem. Inf. Comput. Sci. 1999, 39, 897-902
*/

```

```

[Br,Cl,I][CX4;CH,CH2]
[S,C](=[O,S])[F,Br,Cl,I]
O=CN=[N+]=[N-]
COS(=O)O[C,c]
COS(=O)(=O)[C,c]
C(=O)OC(=O)
OO
C(=O)Oc1c(F)c(F)c(F)c(F)c1(F)
C(=O)Oc1ccc(N(=O)=O)cc1
C(=O)Oonn
N=C=[S,O]
OS(=O)(=O)C(F)(F)F
P(=S)(S)S
NP(=O)(N)N
cN=[N+]=[N-]
C(=O)C[N+,n+]
[N;R0][N;R0]C(=O)
[C+,Cl+,I+,P+,S+]
C=P
[Cl]C([C&R0])=N
[N&D2](=O)
[P,S][Cl,Br,F,I]
N=C=N
[N+]#[C-]
C(=O)N(C(=O))OC(=O)
N#CC[OH]
S(=O)(=O)C#N
P(OCC)(OCC)(=O)C#N
[N;R0]=[N;R0]C#N
[N;R0]=[N,R0]CC=O
[CD2;R0][CD2;R0][CD2;R0][CD2;R0][CD2;R0][CD2;R0][CD2;R0]
[C,S,P](=O)[OH].[C,S,P](=O)[OH].[C,S,P](=O)[OH].[C,S,P](=O)[OH]
[O;R1][C;R1][C;R1][O;R1][C;R1][C;R1][O;R1]
SS
[SH]
C1[O,S,N]C1
c([OH])c([OH])c([OH])
c([OH])c([OH])cc([OH])
N=NC(=S)N
SC#N
cC[N+]
C[O,S;R0][C;R0](=S)
N[CH2]C#N
C1(=O)OCC1
P(=O)([OH])OP(=O)[OH]
N1CCC1=O
O=C1[#6]~[#6]C(=O)[#6]~[#6]1
C=CC=CC=CC=C
O1CCCCC1OC2CCC3CCCCC3C2
O=C1NCC2CCCCC21
O=C1CCCC(N1)=O
O1CCCCC1C2CCCCO2
[OH]c1cc([OH])cc2=[O+]C(=C([OH])Cc21)c3cc([OH])c([OH])
C12OCCC(O1)CC2

```