

**Proteinbasierte 3D-QSAR-Analysen
und virtuelles Screening an Inhibitoren der
Protein-Tyrosin-Phosphatase 1B**

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Dirk Claßen-Houben

aus Waldfeucht-Löcken

Düsseldorf 2005

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine Universität Düsseldorf

Referent: Prof. Dr. Hans-Dieter Höltje

Korreferent Prof. Dr. Wolfgang Sippl

Tag der mündlichen Prüfung: 28.01.2005

Diese Arbeit ist als elektronische Dissertation veröffentlicht.

<http://www.ub.uni-duesseldorf.de>

Die vorliegende Arbeit wurde unter Anleitung von Prof. Dr. Wolfgang Sippl und Prof. Dr. Hans-Dieter Höltje am Institut für Pharmazeutische Chemie der Heinrich-Heine-Universität Düsseldorf zwischen Januar 2000 und Dezember 2004 angefertigt.

Mein besonderer Dank geht an meine beiden Betreuer Prof. Dr. Wolfgang Sippl und Prof. Dr. Hans-Dieter Höltje für die umfangreiche Unterstützung und engagierte Betreuung meiner Arbeit, für die Diskussionsbereitschaft und die idealen Arbeitsbedingungen, die entscheidend zum Gelingen dieser Arbeit beigetragen haben.

Weiterhin möchte ich mich bei allen meinen Kolleginnen und Kollegen des Arbeitskreises für die vielfältige Unterstützung, ausgezeichnete Zusammenarbeit und Hilfsbereitschaft bedanken, besonders bei Edith Tot, Andrea Vollbrecht, Kirstin Jöhren, Andrea Hombrecher und Sonja Schlimme.

Peter Sippel bin ich zu Dank verpflichtet für die Hilfe bei Softwareinstallationen und das Aufspüren verloren gegangener Dateien.

Bei Christian Kirchhoff, Anja Sippel, Robin Ghosh und besonders Philipp Türk bedanke ich mich für die tolle Zusammenarbeit bei der Betreuung der Studenten im Praktikum Arzneibuchanalytik.

Andrea Hombrecher, Kirstin Jöhren und Karin Classen-Houben bin ich dankbar für die hilfreichen Korrekturen und Anmerkungen zum Manuskript dieser Arbeit.

Ein ganz besonderer Dank geht an meine Mutti für die Unterstützung während des Studiums und der Promotionszeit und meine Frau Karin für ihre Hilfe, Unterstützung, Geduld und Liebe.

1	EINLEITUNG	9
2	PROBLEMSTELLUNG UND ZIELSETZUNG	13
3	PHARMAKOLOGISCHE GRUNDLAGEN	17
3.1	SIGNALTRANSDUKTION DURCH KINASEN UND PHOSPHATASEN	18
3.2	PTP1B BEI DIABETES UND ÜBERGEWICHT	20
3.2.1	DIABETES.....	20
3.2.2	ÜBERGEWICHT	24
3.3	PROTEIN-TYROSIN-PHOSPHATASE 1B	27
3.3.1	EINTEILUNG DER PHOSPHATASEN.....	27
3.3.2	STRUKTURAUFKLÄRUNG DER PROTEIN-TYROSIN-PHOSPHATASE 1B	29
3.3.3	KATALYTISCHER MECHANISMUS	37
3.3.4	SUBSTRATSPEZIFITÄT, SELEKTIVITÄT	41
3.3.5	REGULATION DER PTP1B	44
4	METHODEN	47
4.1	BERECHNUNGSGRUNDLAGEN	48
4.1.1	QUANTENMECHANIK	48
4.1.2	MOLEKÜLMECHANIK.....	49
4.2	ALLGEMEINE METHODEN	51
4.2.1	GEOMETRIEOPTIMIERUNG	51
4.2.2	MOLEKÜLDYNAMIKSIMULATION	52
4.2.3	KONFORMATIONSANALYSE	52
4.2.4	CLUSTERANALYSE	54
4.2.5	PARTIALLADUNGEN.....	55
4.2.6	RÖNTGENKRISTALLSTRUKTUREN	56
4.3	SUPERPOSITIONIERUNG	57
4.3.1	FIT-ATOMS-ROUTINE (SYBYL) ZUR ÜBERLAGERUNG DER LIGANDEN.....	57
4.3.2	MULTIFIT-ROUTINE (SYBYL) ZUR ÜBERLAGERUNG DER LIGANDEN.....	58
4.3.3	DAS PROGRAMM FLEXS ZUR ÜBERLAGERUNG DER LIGANDEN	58
4.3.4	ÄHNLICHKEITSANALYSE.....	60
4.3.5	STRUKTUR-ALIGNMENT VON PROTEINEN	61
4.4	DOCKINGMETHODEN	62
4.4.1	FLEXX	62

4.4.2	AUTODOCK 2.4	63
4.4.3	GOLD	65
4.4.4	BEWERTUNG DER GEDOCKTEN KONFORMATIONEN	66
4.5	3D-QSAR-ANALYSE	68
4.5.1	MOLEKULARE WECHSELWIRKUNGSFELDER	68
4.5.2	MOLEKULARE INTERAKTIONSENERGIEN	69
4.5.3	PCA (PRINCIPLE COMPONENT ANALYSIS).....	70
4.5.4	PLS (PARTIAL LEAST SQUARE) ANALYSE.....	70
4.5.5	VALIDIERUNG DER MODELLE	71
4.5.6	VARIABLENREDUKTION	71
4.5.7	COMSIA (COMPARATIVE MOLECULAR SIMILARITY INDICES ANALYSIS)	72
5	DURCHFÜHRUNG UND ERGEBNISSE	73
5.1	DIE STRUKTUREN DER PTP1B.....	74
5.1.1	ANALYSE DER PTP1B-KRISTALLSTRUKTUREN.....	76
5.1.2	VERGLEICH MIT ANDEREN PHOSPHATASEN.....	82
5.1.3	UNTERSUCHUNG DES LADUNGSZUSTANDES	84
5.2	DIE UNTERSUCHTEN INHIBITOREN	87
5.2.1	BEKANNTE LIGANDEN AUS DER LITERATUR.....	87
5.2.2	SELEKTIVITÄT VON PTP1B INHIBITOREN	101
5.2.3	LIGANDEN DER 3D-QSAR-MODELLE	103
5.2.4	LIGANDEN ZUR DATENBANKSUCHE.....	109
5.3	LIGANDBASIERTE 3D-QSAR-MODELLE	113
5.3.1	KONFORMATIONSANALYSE	114
5.3.2	SUPERPOSITIONIERUNG	116
5.3.3	AUSWAHL DER GRID-SONDEN	118
5.3.4	3D-QSAR-MODELLE	120
5.4	DOCKINGBASIERTE 3D-QSAR-MODELLE.....	125
5.4.1	DOCKING IN DIE KRISTALLSTRUKTUR.....	126
5.4.2	GENERIERUNG VON ENZYMKONFORMATIONEN	127
5.4.3	OPTIMIERUNG DES PROGRAMMS AUTODOCK.....	133
5.4.4	PROTEINBASIERTE ÜBERLAGERUNGEN	137
5.4.5	INTERPRETATION DER PLS-FELDER DER 3D-QSAR-MODELLE	144
5.4.6	WEITERE GETESTETE QSAR-VERFAHREN	146

5.5	DATENBANKSUCHEN	148
5.5.1	VORGEHENSWEISE	149
5.5.2	WORLD DRUG INDEX UND MAYBRIDGE DATENBANK	152
5.5.3	EVOTEC OAI DATENBANK	155
5.5.4	EVOTEC OAI SUBSET	156
6	DISKUSSION	165
7	ZUSAMMENFASSUNG.....	175
8	LITERATUR.....	177
9	ANHANG	201
9.1	ABKÜRZUNGEN	202
9.1.1	AMINOSÄUREN	202
9.1.2	PROGRAMMABLAUFSYMBOLS	202
9.2	SKRIPTE	203
9.2.1	SKRIPT ZUR AUTOMATISIERUNG VON AUTODOCK.....	203
9.2.2	EINGABEDATEI FÜR DAS AUTOMATISIERTE AUTODOCK	212

1 Einleitung

Die Welt-Gesundheits-Organisation (WHO) stuft Typ 2 Diabetes als eine der häufigsten Ursachen für Tod und Invalidität im ersten Viertel des 21. Jahrhunderts ein. Von 1985 bis 1995 ist die Anzahl diagnostizierter Diabetiker weltweit von 30 Millionen auf 135 Millionen gestiegen, für den Zeitraum bis 2025 wird ein Anstieg auf 300 Millionen prognostiziert. Die Prävalenz für Deutschland liegt nach aktuellen Untersuchungen bei 7-8%, und man schätzt, dass bei der Altersgruppe der Menschen über 65 Jahre, deren Anteil an der Bevölkerung in den nächsten Jahrzehnten stark ansteigen wird, 40-45% an Typ 2 Diabetes oder der Vorstufe der verminderten Glucosetoleranz erkrankt sind [1][2].

Die Ursachen für diesen Anstieg an Diabeteserkrankungen in den Industrieländern sind die gestiegene Verfügbarkeit von Nahrungsmitteln in den letzten 50 Jahren, das damit verbundene Überangebot an hochkalorischen und fettreichen Lebensmittel sowie der so genannte sitzende Lebensstil.

Das Hauptmerkmal des Typ 2 Diabetes, die Insulinresistenz, steht in direktem Zusammenhang mit Übergewicht, Fettstoffwechselstörungen und Bluthochdruck. Durch den dauerhaft erhöhten Glucosespiegel im Blut entstehen oft schwerwiegende diabetische Folgeerkrankungen und Komplikationen wie Retinopathie, Neuropathie, Nephropathie, Atherosklerose und koronare Herzkrankheit.

Gleichzeitig mit der Prävalenz von Diabeteserkrankungen steigt der Anteil an Personen, die an Adipositas leiden, dem soweit gesteigerten Körpergewicht, dass als Konsequenz daraus Nachteile für die Gesundheit entstehen. So werden schon mehr als 30% der Bevölkerung als adipös eingestuft und weitere 35% haben mäßiges Übergewicht. Auch etwa 15% der Kinder und Jugendlichen sind zu dick. Übergewicht ist wiederum ein zusätzlicher Risikofaktor, der zu vielen lebensbedrohlichen Erkrankungen, einschließlich Typ-2-Diabetes, Herzinfarkt, Schlaganfall, sowie Brust- und Darmkrebs, beiträgt.

Die genauen Ursachen, die bei der Entstehung dieser Erkrankungen eine Rolle spielen, sind bisher nicht aufgeklärt, aber seit einigen Jahren weiß man, dass Tyrosin Phosphatasen dabei eine entscheidende Rolle spielen. Noch zu Beginn der neunziger Jahre des letzten Jahrhunderts sah man diese Enzyme als unspezifische Gegenspieler der an der zellulären Signalweiterleitung beteiligten

Kinasen an, doch in der letzten Zeit wurden für die einzelnen Phosphatasen besondere Informationswege und spezielle Angriffspunkte gefunden. Tyrosin-Phosphatasen sind als Schlüsselenzyme an der Regulation von Zellmetabolismus, Zelldifferenzierung, Wachstum, Proliferation und anderen Funktionen beteiligt [3].

Vor allem für die Protein-Tyrosin-Phosphatase 1B (PTP 1B) konnte durch einige Tiermodelle [4][5], Genanalysen [6][7] und Bindungsstudien [8] gezeigt werden, dass dieses Enzym an der Signaltransduktion des Insulins entscheidend beteiligt ist. Hemmung oder Ausschalten des Enzyms verhindert im Tierversuch nicht nur die Entstehung von Diabetes, sondern auch von Übergewicht. Eine Überexpression der PTP 1B führt zu einem gehäuften Auftreten von Insulinresistenz und Typ 2 Diabetes, was auf die Beteiligung einer genetischen Komponente hinweist [9].

Nach heutigem Behandlungsstandards können Diabetes und Adipositas zwar therapiert werden, eine Heilung ist derzeit jedoch nicht möglich. Die genannten erschreckenden statistischen Daten belegen die dringende Notwendigkeit für die Entwicklung neuer Wirkstoffe im Kampf gegen Typ-2-Diabetes und Adipositas.

Zurzeit sind in der Protein Data Bank (PDB) des *Research Collaboratory for Structural Bioinformatics* (RCSB) [10] 61 Kristallstrukturen der PTP 1B verzeichnet, und es sind unterschiedliche Klassen von Inhibitoren bekannt. Hieraus eröffnet sich die Möglichkeit, mit experimentellen und theoretischen struktur-basierten Methoden nach neuen, potenten und selektiven Inhibitoren für dieses Target zu suchen.

2 Problemstellung und Zielsetzung

Die Protein-Tyrosin-Phosphatase 1B (PTP1B) ist entscheidend an der Signaltransduktion des Insulins und des Leptins beteiligt und unterbricht nach der Bindung an den jeweiligen Rezeptor die Signale dieser beiden Botenstoffe innerhalb der Zelle. Die Blockade beider Signaltransduktionswege wird auch als ursächlich beteiligt an der Entstehung von Diabetes Typ-2 und Übergewicht angesehen. Substanzen, die die PTP1B hemmen und somit die Unterbrechung der Signalweiterleitung aufheben, zeigen in experimentellen Studien eine blutzuckersenkende Wirkung. Bisher gibt es aber noch Probleme, Inhibitoren zu entwickeln, die bei guter Aktivität auch geeignete pharmakokinetische Eigenschaften besitzen und sich für klinische Studien eignen. Somit besteht ein großes Interesse daran, neue PTP1B-Inhibitoren aus neuen Strukturklassen zu entwickeln.

Im ersten Teil der vorliegenden Arbeit sollten die PTP1B-Inhibitoren mit ligandbasierten theoretischen Methoden analysiert werden. Dazu wurden Inhibitoren aus einem einheitlichen Datensatz ausgewählt und in einer 3D-QSAR-Analyse untersucht, die die Wechselwirkungsmöglichkeiten der Liganden betrachtet und mit den biologischen Daten korreliert. Dazu war es notwendig die bioaktive Konformation der Liganden zu finden und geeignete Verfahren zur Überlagerung auszuwählen.

Anschließend sollte ein Vergleich dieser ligandbasierten Untersuchungen mit proteinbasierten Untersuchungen angestellt werden. Dazu mussten verschiedene Vorgehensweisen zum Docking getestet und zum Teil optimiert werden. Die bei den klassischen proteinbasierten Untersuchungen erfolgende Auswertung von Komplexenergien und Komplexgeometrien sollte hierbei aber nur als Vergleich dienen. Zusätzlich sollte die Auswertung der Dockingergebnisse mit 3D-QSAR Methoden erfolgen. Da hier die Proteinstrukturen für das Docking verwendet werden, bezeichnet man diese Vorgehensweise als rezeptor- oder proteinbasierte 3D-QSAR Methode [11][12][13][14].

Im zweiten Teil der Arbeit sollte, ausgehend von diesen rezeptorbasierten 3D-QSAR-Modellen, mit Hilfe des virtuellen Screenings in Moleküldatenbanken nach neuen Inhibitoren für die PTP1B gesucht werden. In einem ersten Schritt wird mit wenig rechenintensiven, ligandbasierten Methoden eine Vorauswahl

unter den zu Verfügung stehenden Strukturen getroffen. Der zweite Schritt besteht aus zeitaufwändigen Dockinganalysen und Vorhersagen mit den zuvor erstellten 3D-QSAR-Modellen, wodurch die interessanten Verbindungen ausgewählt werden.

3 Pharmakologische Grundlagen

3.1 Signaltransduktion durch Kinasen und Phosphatasen

Die Signaltransduktion, die Regulation zellulärer Prozesse als Antwort auf externe Signale, ist grundlegend für alle Prozesse des Lebens. Es sind zwar weit über hundert kovalente posttranslationelle Modifikationen von Proteinen bekannt, doch nur wenige, wie zum Beispiel Acetylierung, Methylierung und Phosphorylierung, sind reversibel und eignen sich somit zur Regulation von zellulären Prozessen. Schon um 1955 stellte man fest, dass Enzyme durch Protein Kinasen und Protein Phosphatasen reversibel phosphoryliert werden, wodurch sich ihre Aktivität ändert [15][16][17][18]. Der umkehrbare Prozess der zellulären Tyrosin-Phosphorylierung wird dabei nicht nur durch Protein-Tyrosin-Kinasen (PTKs), die die Übertragung des γ -Phosphates von Adenosintriphosphat (ATP) auf Tyrosinreste fördern, kontrolliert, sondern benötigt ebenso die gegensätzliche Enzymaktivität der Protein-Tyrosin-Phosphatasen (PTPs), die die Phosphatgruppen von Phosphotyrosin (pTyr) wieder abtrennen. Berücksichtigt man, dass etwa 30% aller intrazellulären Proteine Ziel für eine Phosphorylierung sind, so wird die große Bedeutung dieser Enzyme offensichtlich. Die Erkenntnis, dass der EGFR (Epidermal Growth Factor Receptor) eine Tyrosinkinase ist, führte zur Entdeckung einer großen Familie von PTKs, die durch Ligandbindung aktiviert werden, und zu der auch der Insulinrezeptor (IR) gehört.

Nahezu alle Rezeptoren für Peptidhormone und Wachstumsfaktoren verfügen über intrinsische PTK-Aktivität und die erste und entscheidende Aktion der meisten Signaltransduktionswege ist eine Tyrosin-Phosphorylierung. Rezeptoren für Zytokine, die diese Aktivität nicht besitzen, assoziieren mit intrazellulären PTKs.

Durch Signalkaskaden aus Proteinen, die an diese Rezeptoren binden, erfolgt die Weiterleitung der Signale innerhalb der Zelle. Die Addition oder Abspaltung eines Phosphatrestes kann hier die Enzymaktivität regulieren, die Stabilität von Proteinen beeinflussen, im Zellkern zur Aktivierung oder Unterdrückung der Transkription bestimmter Gene führen oder neue Motive für Protein-Protein-Interaktionen schaffen [19].

Die Bedeutung der PTPs wurde – im Vergleich zu den PTKs – lange Zeit unterschätzt, und weil man kein definiertes Substrat für spezifische Untersuchungen besaß, wurden diese Enzyme für konstitutive allgemeine Gegenspieler der PTKs gehalten, die deren Signale unspezifisch wieder abschalten, und als *Housekeeper* bezeichnet. Heute weiß man jedoch, dass die PTPs eine große Enzymfamilie ausmachen und den PTKs in struktureller Vielfalt und Komplexität vergleichbar sind.

Daraus resultieren für die PTPs analog wie für die PTKs mögliche Angriffspunkte zur Behandlung von Krankheiten wie Diabetes, Übergewicht, Krebs, Autoimmunkrankheiten, Entzündungen, Osteoporose, myelodysplastische Syndrom und myelogene Leukämie, sowie bakterielle Infektionen [20].

3.2 *PTP1B bei Diabetes und Übergewicht*

3.2.1 Diabetes

Aufgrund der Häufigkeit von Typ-2-Diabetes ist es nicht verwunderlich, dass beträchtliche Anstrengungen unternommen werden, um die Ursache dieser Krankheit aufzuklären, in der Hoffnung, neue Behandlungs- und Heilungsmöglichkeiten zu entwickeln. Das Peptidhormon Insulin spielt dabei eine wesentliche Rolle in der Wahrung der Glucosehomöostase und der Regulation des Metabolismus von Kohlenhydraten, Lipiden und Proteinen und die PTP1B wird seit ihrer Entdeckung mit der Insulinresistenz bei Übergewicht und Diabetes in Zusammenhang gebracht. In den letzten Jahren haben sich zwingende Beweise für diese Beteiligung angehäuft.

Der Insulinrezeptor (IR) ist der Prototyp der Rezeptor-Protein-Tyrosin-Kinasen (R-PTKs), die aus zwei extrazellulären peptidbindenden α -Untereinheiten und zwei membrandurchspannenden β -Untereinheiten bestehen. Die Bindung von Insulin, das bei steigenden Glucosekonzentrationen im Blut von den β -Zellen des Pankreas ausgeschüttet wird, verursacht eine Konformationsänderung des Rezeptors, wodurch die Kinase-Domäne im zytoplasmatischen Teil der β -Untereinheiten aktiviert wird. Beide β -Untereinheiten werden anschließend an mindestens sechs Tyrosinresten in drei getrennten Regionen autophosphoryliert (siehe Abbildung 3-1, Seite 21) [21].

Der somit aktivierte IR führt die Phosphorylierung von mehreren nachfolgenden Proteinsubstraten, einschließlich der Insulin Rezeptor Substrate (IRS) 1 bis 3 und anderer Verbindungsproteine (Gab1 und Shc) durch, wodurch die biologischen Effekte weitergeleitet werden [22][23].

Die IRS-Proteine stellen die Verbindung zwischen dem IR und der Phosphatidyl-Inositol-3-Kinase (PI3K) dar, die später im Signaltransduktionsweg aktiviert wird und dann Phosphatidyl-Inositol zum 3,4-bis-Phosphat und 3,4,5-tris-Phosphat umsetzt. Diese beiden Derivate des Inositols sind notwendig für die Aktivierung der PI3K-abhängigen Kinase 1 (PDK1), die schließlich die Proteinkinase AKT aktiviert.

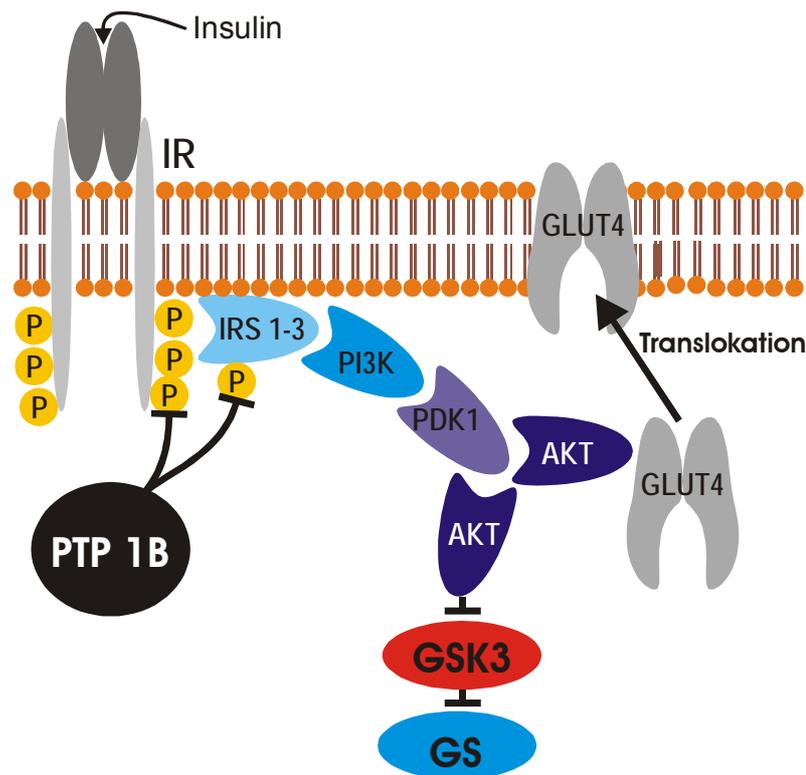


Abbildung 3-1: Signaltransduktion des Insulinrezeptors

AKT fördert die Translokalisierung des Glucosetransporters 4 (GLUT4), der in die Zellmembran wandert und dort für die Glucoseaufnahme in die Zelle sorgt. Zusätzlich wird die Glykogensynthase Kinase 3 (GSK3) durch AKT phosphoryliert und damit inaktiviert. Da die GSK3 die Glykogensynthese (GS) durch Phosphorylierung inaktiviert, wird diese Blockade der Signaltransduktion durch AKT aufgehoben und die Glykogensynthese stimuliert. [24][25][26]. Durch diese Mechanismen wird die Glucoseaufnahme in die Zellen gesteigert, der Glucoseabbau und die Glykogensynthese werden aktiviert und die Gluconeogenese und der Glykogenabbau werden gehemmt wodurch letztlich die Glucosespiegel im Blut sinken.

Ein Kennzeichen des Typ-2-Diabetes ist die Insulinresistenz, deren Ursache entweder in der reduzierten Kapazität oder der herabgesetzten Sensitivität zur Aufnahme und zum Metabolismus von Glucose zu suchen ist. Da die Unterbrechung des Insulin-Signals die Dephosphorylierung des IR und dessen Substrate erfordert, vermutet man, dass eine Dysregulation der PTPs auf Basis einer gesteigerten Aktivität oder Expression die Ursache für die abgeschwächten Insulineffekte und somit für die Insulinresistenz ist [27].

Die ersten Experimente, die auf eine Beteiligung der PTP1B an der Signaltransduktion des Insulins hinwiesen, zeigten, dass die Mikroinjektion der gereinigten katalytischen Domäne der PTP1B in *Xenopus* Oozyten die Insulinwirkung aufhebt [28][29].

Die PTP1B ist *in vitro* in der Lage, mit dem IR zu interagieren und die durch Autophosphorylierung entstandenen Tyrosinphosphate zu entfernen [30][31][32][33]. Ebenso kann PTP1B die IRS-Proteine dephosphorylieren und auch auf diesem Wege die Signaltransduktion des Insulins abschwächen oder ganz blockieren [34][35].

Wesentliche Beweise für die Rolle der PTP1B als der entscheidende negative Regulator bei Diabetes lieferten zwei Studien mit Mäusen, die keine PTP1B besitzen (PTP1B Knockout, PTP1B^{-/-}) [4][5]. Diese mutierten Mäuse weisen eine gesteigerte und verlängerte Phosphorylierung des IR und IRS-1 im Skelettmuskel und der Leber auf. Als Folge daraus besitzen sie eine verbesserte Insulinsensitivität in diesen Geweben. Der Blutspiegel für Glucose ist um etwa 13% niedriger, obwohl der Insulinspiegel nur halb so hoch ist wie bei den Kontrolltieren.

Zusätzlich sind die PTP1B^{-/-} Mäuse gegen Gewichtszuwachs geschützt und haben signifikant niedrigere Triglyzeridwerte, wenn sie einer fettreichen Diät ausgesetzt werden.

Da die PTP1B zumindest in Zellkulturen an der Abschwächung der Signaltransduktion von Wachstumsfaktoren wie IGF-IR, PDGFR und EDGFR beteiligt ist und viele weitere zelluläre Funktionen, wie zum Beispiel die von Integrin, Cadherin, Zytokinrezeptoren, Zellzyklus und Stressantwort reguliert [36]–[53], wurde angenommen, dass die Zerstörung des PTP1B-Gens zu nicht lebensfähigen Tieren oder zu einer erhöhten Anfälligkeit, Tumoren zu entwickeln, führt. Diese Befürchtungen haben sich jedoch nicht bestätigt: Die Tiere besitzen eine normale Entwicklung, eine übliche Lebensspanne und keine erhöhte Neigung zur Tumorbildung. Ein Grund dafür könnte darin begründet liegen, dass PTP1B nicht nur aktivierend sondern auch inhibitorisch auf die Kinase RAS einwirkt [54].

In Immunoprecipitations-Untersuchungen stellte sich heraus, dass katalytisch inaktive Mutanten der PTP1B in der Lage sind, den aktivierten IR fest zu binden [33][42], und man geht davon aus, dass daran die N-terminale Domäne der PTP1B mit den Aminosäuren Tyr152 und Tyr153 beteiligt ist [55]. Weitere Beweise für die Interaktion zwischen IR und PTP1B liefern kinetische und strukturelle Untersuchungen mit dem IR. In der Röntgenkristallstruktur des Komplexes aus PTP1B und dem Aktivierungssegment des IR kann man sehen, dass pTyr1162 (Phosphotyrosin) das aktive Zentrum besetzt und pTyr1163 in der daneben liegenden, zweiten Phosphotyrosin-Tasche bindet (PDB-Einträge 1G1F, 1G1G und 1G1H). [56]. Des Weiteren steigt die durch Insulin stimulierte Autophosphorylierung von IR und IRS-1 um den Faktor 2,2 bzw. 2,0, wenn die Zellen mit PTP1B-Antikörpern beladen werden [57]. Auch ein Peptid mit einem nicht hydrolysierbaren Phosphotyrosinanalogon, das die PTP1B inhibiert, macht die durch Überexpression verursachte Erniedrigung der Translokation von GLUT-4 rückgängig. Ebenso sind Thiazolidin-Derivate in der Lage, die Glucose induzierte Insulinresistenz, durch die Normalisierung der Aktivität der PTP1B, zu verbessern [59]. Zusätzliche Hinweise liefert die Behandlung von diabetische Mäuse mit dem PTP1B-ASO (antisense oligonucleotide), wodurch sich die Blutspiegel von Glucose und Insulin normalisieren und die Menge von PTP1B-Protein in der Leber und im Fettgewebe sich um 50% reduzieren. Genetische Untersuchungen an Menschen liefern weitere Anhaltspunkte für die Bedeutung der PTP1B bei Insulinresistenz. Die humane PTP1B ist auf dem Chromosom 20q13 lokalisiert, das mit den Merkmalen von Übergewicht, erhöhten Insulinspiegeln und Typ-2-Diabetes in Verbindung steht [60][61]. Für eine Genvariante, die zum Austausch der Aminosäure Pro387 zu Leu387 der PTP1B führt, erhöht sich das relatives Risiko an Diabetes zu erkranken um den Faktor 3,7 [62]. Verschiedene andere Varianten mit einer einzigen Mutation (single nucleotide polymorphism, SNP) im Gen der PTP1B führen zu einem Schutz vor verminderter Glucosetoleranz und Diabetes oder erhöhen das Diabetesrisiko. [63][64][65]. In zwei Populationen mit hoher Insulinresistenz beobachtet man eine gesteigerte Expression von mRNA (Boten-Ribonukleinsäure) der PTP1B im Skelettmuskel als Folge einer PTP1B-Variante, bei der in der Position 1484,

die selbst nicht in ein Protein übersetzt wird, ein zusätzliches Guanin eingefügt ist [7].

Wie Veröffentlichungen in Fachzeitschriften und Patentschriften belegen, arbeiten derzeit viele pharmazeutisch forschende Unternehmen an der Entwicklung von PTP1B-Inhibitoren, wie zum Beispiel Abbott, Astra Zeneca, Merk Frosst, Novartis, Novo Nordisk, Ontogen, Pharmacia & Upjohn, Serono, Takeda und Wyeth.

Während die meisten Firmen sich noch im Stadium der Forschung und Entwicklung befinden, wurden von Wyeth im Jahr 2000 klinische Studien mit dem Inhibitor Ertiprotafib begonnen. Diese Leitsubstanz ist in der Lage, die PTP1B effektiv zu inhibieren und bei diabetischen oder adipösen Mäusen die Plasmaspiegel von Insulin und Glucose zu normalisieren. Im Gegensatz zu anderen Verbindungen, die die Sensitivität für Insulin erhöhen wie zum Beispiel die bereits in der Therapie befindlichen Glitazone beruhen die *in-vivo*-Effekte von Ertiprotafib und anderen PTP1B Inhibitoren nicht auf einer Aktivierung von PPAR γ (peroxisome proliferator activated receptor) [66]. Unglücklicherweise mussten die klinischen Phase-II-Studien mit Ertiprotafib wegen unzureichender Effektivität und dosislimitierender Nebenwirkungen eingestellt werden.

Momentan werden als molekulare Ziele bei Diabetes [67][68] neben Inhibitoren der PTP1B [69] auch Glucagon-like-Peptid-1-Agonisten [70] und Dipeptidyl-peptidase-IV-Inhibitoren [71] die auch über das Glucagon-like-Peptid-1 wirken, sowie duale Agonisten des PPAR α/γ [72] untersucht.

3.2.2 Übergewicht

Leptin ist ein Hormon, das mit den Zytokinen verwandt ist und hauptsächlich im Fettgewebe produziert wird [73]. Sein wichtigstes Zielorgan ist der Hypothalamus, wo es an Leptinrezeptor (Ob-R1, Obesity-Rezeptor) bindet [74]. Die primäre Aufgabe des Leptins besteht darin, die Toleranz für Nahrungsfette zu

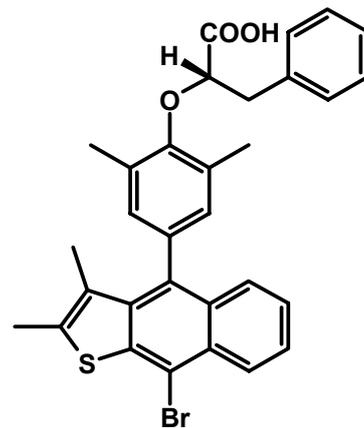


Abbildung 3-2: Ertiprotafib

erhöhen und die Fettsäure-Homöostase im Energiehaushalt mittels Fettspeicherung und Nahrungsaufnahme zu erhalten [75].

Der Ob-R1 (Obesity-Rezeptor, Leptinrezeptor) dimerisiert durch die extrazelluläre Bindung von Leptin und aktiviert die Ob-R1 assoziierte Kinase, JAK2 (Janus-Kinase 2). Die aktivierte JAK2 phosphoryliert den Ob-R1 an Tyr1138 (Abbildung 3-3), wodurch eine Bindungsstelle für die SH2-Domäne von STAT3 (signal transducers and activators of transcription 3) gebildet wird [76]. Wenn STAT3 an den Ob-R1 bindet, wird sie durch die JAK2 phosphoryliert, dimerisiert und wandert in den Zellkern, wo sie direkt die Transkription der Zielgene einleitet. Diese Gene kodieren unter anderem für ACC (Acetyl-Coenzym-A-Carboxylase), FAA (Fettsäure-Synthase), CPT-1 (Carnitin-Palmitoyl-Transferase 1) und ACO (Acetyl-Coenzym-A-Oxidase), welche für die Lipid-Homöostase wichtig sind [77]. Da die PTP1B wie in Abbildung 3-3 dargestellt auch die JAK2 dephosphoryliert und dadurch auch die Verarbeitung von Fettsäuren hemmt, könnte sie für Übergewicht mitverantwortlich sein [58].

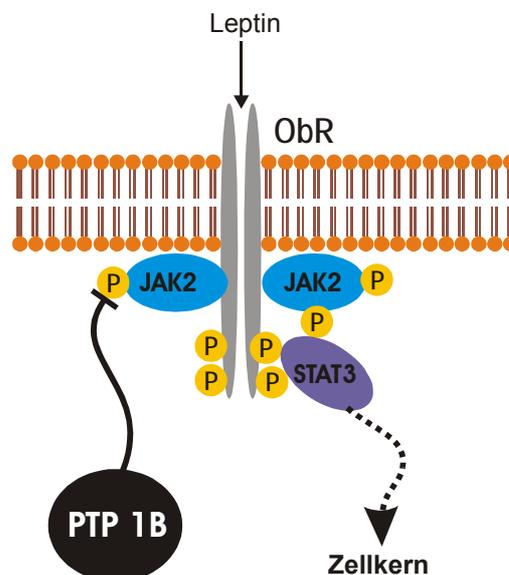


Abbildung 3-3: Funktion der PTP1B bei Adipositas

Beweise dafür liefern Mäuse, denen die PTP1B fehlt. Diese weisen bei einer fettreichen Diät niedrigere Leptin-Blutspiegel auf als Kontrolltiere und entwickeln kein Übergewicht [48]. Mäuse mit einer Mutation im Leptin-Gen, die normalerweise zu Gewichtszunahme führt, haben eine geringere Gewichtszunahme, sowie einen höheren Energieumsatz und Sauerstoffverbrauch und sprechen

stärker auf Leptin an, wenn bei ihnen zusätzlich das PTP1B-Gen ausgeschaltet wurde. Dies führt zu einem größeren Gewichtsverlust bei Futtermangel als bei den Vergleichstieren [5][49]. Leptin ist in der Lage ist, die Expression von PTP1B zu erhöhen [78], was zusätzlich die Verwobenheit der PTP1B in die Signaltransduktion des Ob-R1 bestätigt.

Wenn man bedenkt, dass Insulin die Speicherung von Glucose und Fetten fördert, und viele der bisher in der Therapie des Diabetes verwendeten Pharmaka (Sulfonylharnstoffe, Insulin, PPAR- γ -Agonisten) zu einer Gewichtszunahme führen, ist die Gewichtsreduktion durch Blockade der PTP1B umso erstaunlicher. Statt einer allgemeinen Erhöhung der Insulinwirkung führen PTP1B-Inhibitoren zu einer gewebsspezifischen Insulinsensitivität in Muskel und Leber, während im Fettgewebe die Insulinwirkung nicht gesteigert wird. Wäre die Insulinsensitivität auch im Fettgewebe erhöht, so wäre eine stärkere Fettspeicherung und somit Gewichtszunahme zu erwarten. Da dies nicht der Fall ist, kann man daraus folgern, dass die PTP1B im Fettgewebe eine andere Rolle spielen oder hier durch andere Phosphatasen kompensiert werden. [79].

3.3 Protein-Tyrosin-Phosphatase 1B

3.3.1 Einteilung der Phosphatasen

Protein Phosphatasen werden aufgrund ihrer Substratspezifität in zwei Hauptgruppen unterteilt: die Protein Phosphatasen (PPs), die spezifisch Phosphoester des Serins und Threonins hydrolysieren und die Tyrosin-Phosphatasen, die für Phosphotyrosin spezifisch sind. Die Tyrosin-Phosphatasen werden, wie in Abbildung 3-4 gezeigt, weiter in verschiedene Unterfamilien eingeteilt [80].

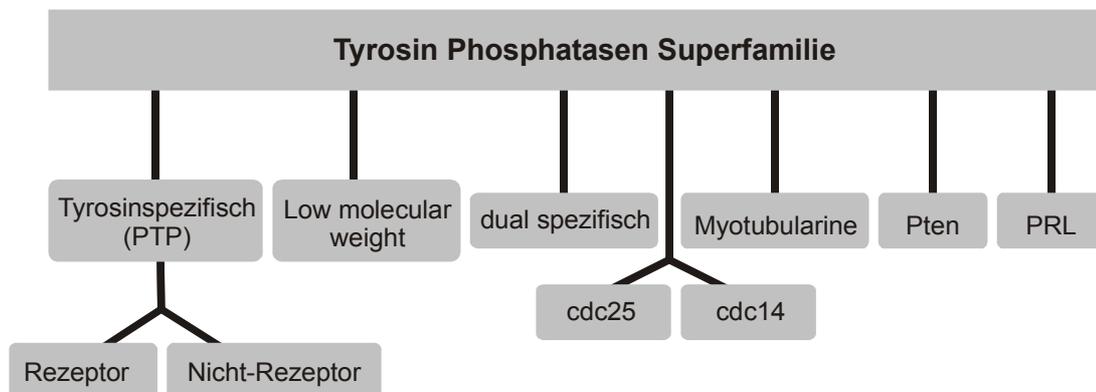


Abbildung 3-4: Einteilung der Phosphatasen in Familien

Obwohl beide Familien der PPs und PTPs die gleiche Reaktion, nämlich die Hydrolyse eines Phosphoesters, katalysieren, benutzen sie individuelle Mechanismen und haben, im Gegensatz zur großen Homologie innerhalb der Klassen, völlig unterschiedliche Strukturen. Die PPs gehören zu einer größeren Familie von Metallo-Phosphoesterasen, deren aktives Zentrum sich im Loopbereich eines β - α - β - α - β -Motivs befindet. Die beiden Metallatome, bei denen es sich um Eisen, Zink oder Mangan handeln kann, werden bei den PPs durch drei Histidine, zwei Aspartate und eine Asparaginsäure aus den Loopbereichen koordiniert, und sind sowohl für die katalytische Aktivität, als auch für die strukturelle Integrität notwendig [81][82][83]. Im Gegensatz zu den PPs, benötigen die PTPs keine Metallionen für die Katalyse der Reaktion. Stattdessen läuft die Hydrolyse über eine Zwischenstufe ab, bei der das Phosphat an das Cystein des aktiven Zentrums gebunden ist. Man unterteilt die PTPs in Rezeptor-PTPs und intrazelluläre PTPs, deren katalytische Domänen mit einer

Sequenzidentität von mehr als dreißig Prozent sehr ähnlich sind. Sie heben sich von den anderen Mitgliedern der Tyrosin-Phosphatase Superfamilie, trotz einer großen strukturellen und mechanistischen Ähnlichkeit durch eine Identität kleiner fünf Prozent ab.

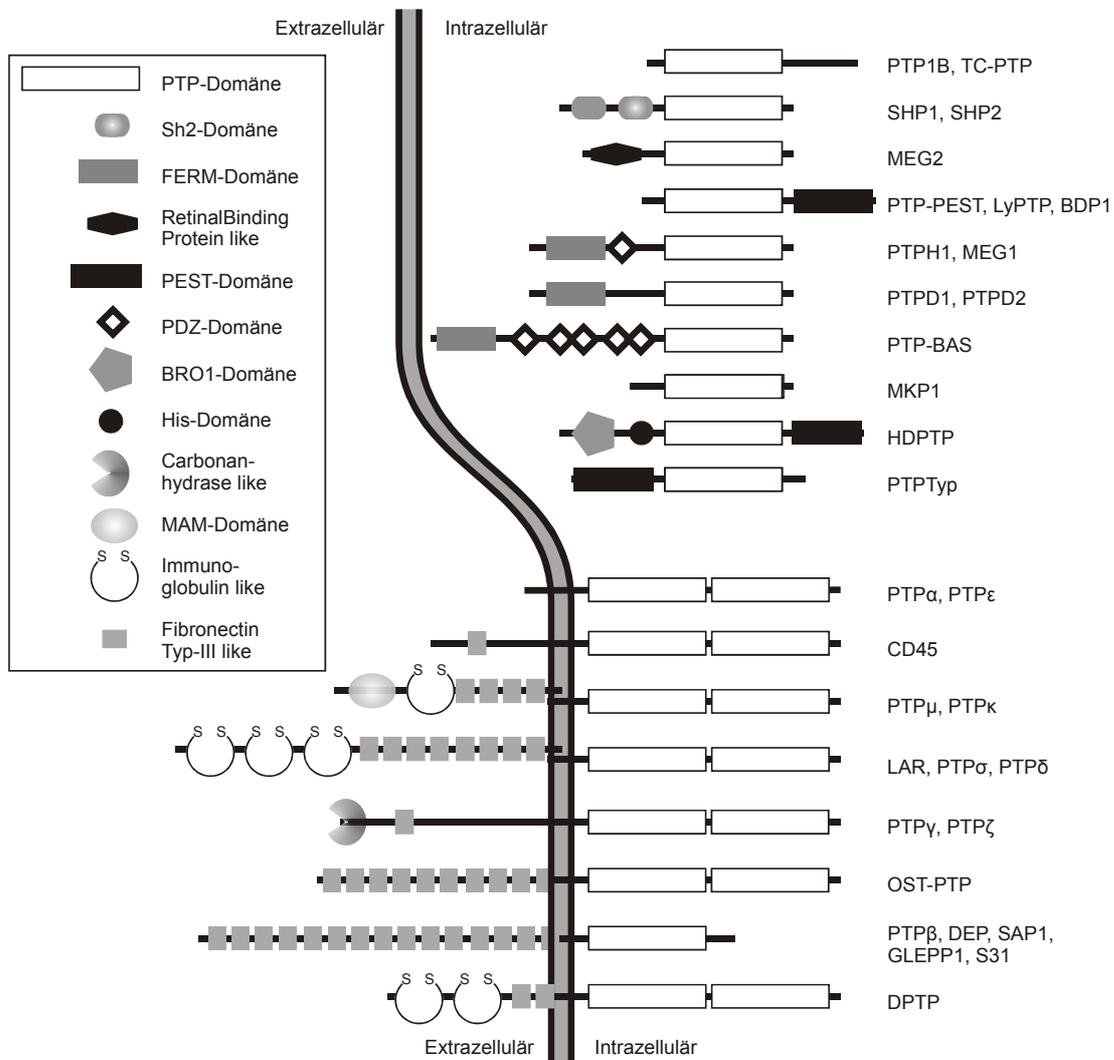


Abbildung 3-5: Vergleich von Protein-Tyrosin-Phosphatasen (schematisch)

Die Enzyme dieser Subfamilie, die aus etwa 40 Mitgliedern besteht, enthalten in der katalytischen Domäne (PTP-Domäne) von etwa 240 bis 250 Aminosäuren die für diese Enzymfamilie charakteristische, elf Aminosäuren lange, Sequenz [I/V]HCXAGXXR[S/T]G [84].

Sowohl am aminoterminalen als auch am carboxyterminalen Ende der katalytischen Domäne kann eine Vielzahl unterschiedlicher nichtkatalytischer, regulatorischer Domänen angehängt sein (Abbildung 3-5). Die Rezeptor-PTPs

bestehen aus einer die Membran durchspannenden α -Helix, die als transmembranäre Domäne bezeichnet wird, einer extrazellulären Domäne mit Fibronectin- oder Immunoglobulin-Repeats, die als Rezeptor dient, und meist zwei intrazellulären PTP-Domänen. Die Domäne, die näher an der Membran liegt, beinhaltet die katalytische Aktivität, die Funktion der zweiten PTP-Domäne ist noch nicht bekannt, man vermutet aber, dass eine intra- oder intermolekulare Regulation zwischen den beiden PTP-Domänen existiert. Intrazelluläre PTPs besitzen neben einer katalytischen Domäne unterschiedliche Bereiche, die Protein-Protein-Interaktionen ausüben können, wie die SH2 Domäne oder Bereiche, die das Enzym an eine Membran binden können, wie im Falle der PTP1B [85][86]. Aus der aufgeklärten Sequenz des humanen Genoms kann man entnehmen, dass der Mensch kodierende Gene für etwa 100 PTPs besitzt [87].

3.3.2 Strukturaufklärung der Protein-Tyrosin-Phosphatase 1B

Die PTPs wurden aufgrund ihrer Fähigkeit, tyrosinphosphorylierte Proteine zu dephosphorylieren, entdeckt. Als erste PTP wurde in den Jahren 1988 und 1989 aus der menschlichen Plazenta ein Enzym isoliert, dessen Sequenzaufklärung eine Homologie zu den intrazellulären Domänen der Leukozyten-Antigene CD45 und LAR, sowie einem kleinem Protein aus T-Zellen zeigte [88][89][90]. Das Enzym wurde als Protein-Tyrosin-Phosphatase 1B bezeichnet, da es bei der Reinigung und Charakterisierung aus dem zweiten Peak (B) der Säulenchromatographie des Plazenta-Lysates gefunden wurde. 1990 wurde eine cDNA (copy DNA, Rückübersetzung der mRNA in DNA) isoliert und sequenziert, die mit dem Enzym übereinstimmte, aber am C-Terminus 114 weitere, hauptsächlich hydrophobe Aminosäuren enthielt [91][92]. Diese Form der PTP1B wurde ebenfalls als monomeres Protein isoliert [93]. N-Terminal befinden sich die etwa 250 Aminosäuren der katalytischen Domäne, und die 35 Aminosäuren am C-terminalen Ende der Sequenz fungieren wie bei den Cytochromen als Membrananker. Dieser Anker besteht aus einer α -Helix von 20 lipophilen Aminosäuren, die von geladenen Aminosäuren begrenzt wird, und bindet die PTP 1B an das Endoplasmatische Retikulum (ER) [94]. Dazwischen

befinden sich zwei prolinreiche Domänen, von denen eine für die Protein-Protein-Interaktion mit SH3-Domänen (Src Homologie 3) notwendig ist. Diese Wechselwirkung mit SH3 ist nur von wenigen anderen Phosphatasen bekannt und läuft bei diesen über andere Mechanismen, während sie beim nächsten Verwandten, der ebenfalls an das ER gebundenen TCPTP, überhaupt nicht vorkommt [95][37]. Die Interaktionen mit dem IR finden mit einem Bereich in der katalytischen Domäne statt, der die Aminosäuren Tyr152 und Tyr153 enthält. Diese beiden Tyrosinreste können phosphoryliert werden und sind für die Wechselwirkung mit dem IR essentiell, wie in Mutationsstudien festgestellt wurde [32][55].

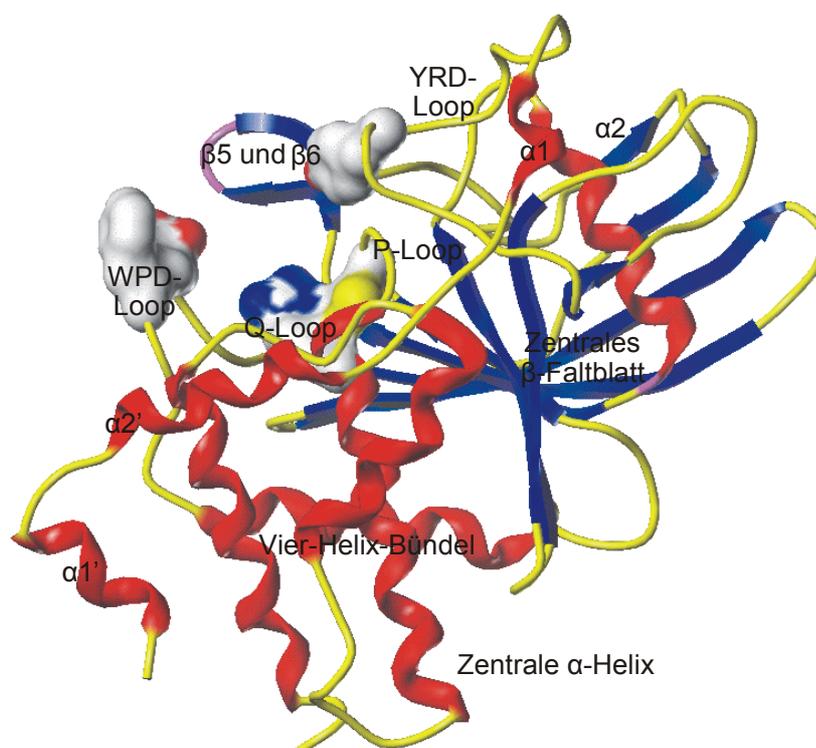


Abbildung 3-6: Röntgenkristallstruktur der PTP1B

Schon 1994 wurde die erste Röntgenkristallstruktur der verkürzten Form der PTP 1B veröffentlicht (PDB-Einträge 2HNP und 2HNQ) [96]. Die kristallisierte Struktur der 321 Aminosäuren der katalytischen Domäne der PTP 1B, von denen 278 aufgelöst sind, besteht aus einer einzigen globulären Domäne. In dieser aktiven Domäne oder PTP-Domäne ist die Peptidkette zu zwölf β -Faltblattsträngen und acht α -Helices gefaltet (Abbildung 3-6).

Motiv	Reste	Konservierung	Vermutete Aufgabe der Aminosäuren
1	40-46	NXX(K/R)NRY	pTyr-Erkennungs-Loop, Substratspezifität Asn44: Interaktion mit Asn68 Arg45: Elektrostatische Anziehung des Liganden Tyr46: Hydrophobe Interaktion mit pTyr
2	53-59	DXXR(V/I)XL	Konserviertes β -Faltblatt, Proteinoberfläche Arg56: Salzbrücke mit Asp65 Val/Ile57, Leu59: Hydrophober Proteinkern
3	65-70	DYINA(N/S)	Proteinkern Tyr66: Wasserstoffbrücke mit Asn44 Ile67, Ala69: Hydrophober Proteinkern Asn68: Wasserstoffbrücke mit Arg257
4	81-87	(F/Y)(I/V)A XQGP	Proteinkern, umgibt P-Loop Ile/Val82: Hydrophober Proteinkern Ala83: Interaktion mit P-Loop Gln85: Wasserstoffbrücke mit Wasser Gly86, Pro87: Interaktion mit P-Loop
5	91-101	TXDXFWX (M/L/V)XW(E/Q)	Konservierte α -Helix Asp94: Konservierte Subdomäne auf der Rückseite Phe95: π - π -Interaktion mit Trp96 Trp96: Wasserstoffbrücke mit Rückgrat von Tyr124 Met/Leu/Ile98: Hydrophober Proteinkern Trp100: Konservierte Subdomäne auf der Rückseite
6	107-111	(I/L/V)(V/I)MXT	Hydrophober Proteinkern Ile/Leu/Val107, Val/Ile108: Interaktion mit Trp96 Met109: Interaktion mit Trp125 Thr111: Interaktion mit P-Loop
7	120-126	KCXXYPW	Hydrophober Proteinkern Lys120: Ligandinduzierte Salzbrücke mit Asp181 Tyr124: Wasserstoffbrücke mit His214, Trp125: π - π -Interaktion mit Tyr124
8	176-185	(Y/F)XX WPDXGXP	WPD-Loop, Beweglicher Loop an der Proteinoberfläche Trp179: Zentrum des Loops Pro180: Wasserstoffbrücke mit Arg221 Asp181: Katalytisch wichtige Säure Gly183: Gelenk für die Bewegung
9	210-223	PXX(V/I)HCS AGXGR(T/S)G	P-Loop Sieben aufeinander folgende Rückgrat-NH-Gruppen interagieren mit den drei Sauerstoffen des Phosphats Pro210: Hydrophober Proteinkern His214: Senkt pKS-Wert von Cys215 Cys215: Nukleophiler Angriff an das Substrat Übernahme des Phosphat Ser216: Stabilisierung von Cys215 und Tyr46 Ala217: Unpolare Interaktion mit dem Substrat Gly218, Gly220: Interaktion mit dem Substrat Arg221: Stabilisierung des Übergangszustandes Thr/Ser222: Senkt pKS-Wert von Cys215
10	261-269	(V/I/L)QT XXQYXF	Q-Loop: Interaktion mit Wasser im aktiven Zentrum Gln262, Gln266: Wasserstoffbrücke mit Wasser Tyr267, Phe269: Konservierte α -Helix

Tabelle 3-1: Konservierte Motive in Tyrosinphosphatasen

Zehn der zwölf Stränge bilden ein stark verdrilltes, gemischtes β -Faltblatt, das die gesamte Phosphatase durchspannt. Der letzte Strang in der Primärsequenz, β 12, bildet den zentralen parallelen Teil des Faltblattes mit der Strangfolge β 3 (79-85), β 12 (211-214), β 4 (106-109) und β 11 (168-176), der von den antiparallelen Strängen β 2 (68-74) und β 1 (56-57) neben β 3 einerseits und β 10 (153-162), β 8-9 (140-142, 147-150) und β 7 (133-135) neben β 11 andererseits flankiert wird. Der Mittelteil des Faltblattes ist eingebettet zwischen den Helices α 2 (91-102) auf der einen Seite und den Helices α 3 (188-201) und α 4 (219-237) auf der anderen. Die Helices α 3 und α 4 bilden zusammen mit den Helices α 5 (246-254) und α 6 (264-279) ein 'Vier-Helix-Bündel'.

Ein weiteres β -Faltblatt wird aus den beiden kurzen Abschnitten β 5 (114-115), β 6 (118-119) und dem dazwischen liegenden Turn gebildet und liegt über dem großen β -Faltblatt. Am Ende des Faltblattes befindet sich die kurze Helix α 1 die den Anfang der konservierten PTP-Domäne markiert. In der Kristallstruktur befinden sich noch zwei weitere N-terminale Helices α 1' (6-12) und α 2' (16-26), die sich an der Oberfläche des Proteins angelagert haben, die aber nicht zur PTP-Domäne selbst gehören.

Die katalytische PTP-Domäne, die die Familie der PTPs definiert, besteht aus insgesamt zehn konservierten Motiven und sieben einzelnen konservierten Aminosäuren (in mindestens 80% der Sequenzen), die beim Sequenzvergleich gefunden wurden (siehe Tabelle 3-1). Die drei Motive M5, M6, und M7 bilden gemeinsam den dicht gepackten, hydrophoben Kern des Proteins mit entscheidenden π - π -Interaktionen der Aminosäuren Phe95, Trp96, Tyr124 und Trp125. Weitere, vor allem hydrophobe Interaktionen besitzen die Motive M3 und M4. Das Motiv M5 bildet zusätzlich eine Subdomäne auf der Rückseite des Enzyms, die dem aktiven Zentrum gegenüberliegt, die als Erkennungsstelle für Protein-Protein-Interaktionen dienen könnte.

Das aktive Zentrum befindet sich am Ende einer 9 Å tiefen Spalte in der Oberfläche der Phosphatase und kann über das katalytisch aktive Cystein (Cys215) und das mitkristallisierte Wolframat identifiziert werden. Die Tiefe des aktiven Zentrums, das durch den YRD-Loop (37 bis 55, Loop 1 mit dem Motiv M1), den Loop 4 (86-90), den WPD-Loop (176-187, Loop 13 mit dem Motiv M8) und den

Q-Loop (255-263, Loop 17 mit dem Motiv M10) begrenzt wird (Abbildung 3-7), entspricht der Länge eines Phosphotyrosins. Die Seitenketten von Phosphoserin und Phosphothreonin hingegen sind zu kurz, so dass die Esterbindung nicht weit genug in das aktive Zentrum hineinragt, um gespalten zu werden.

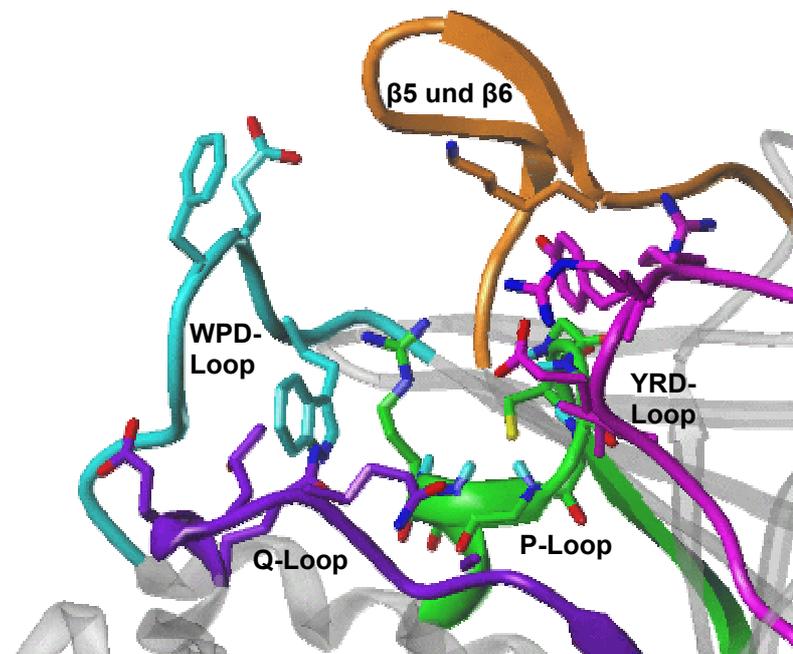
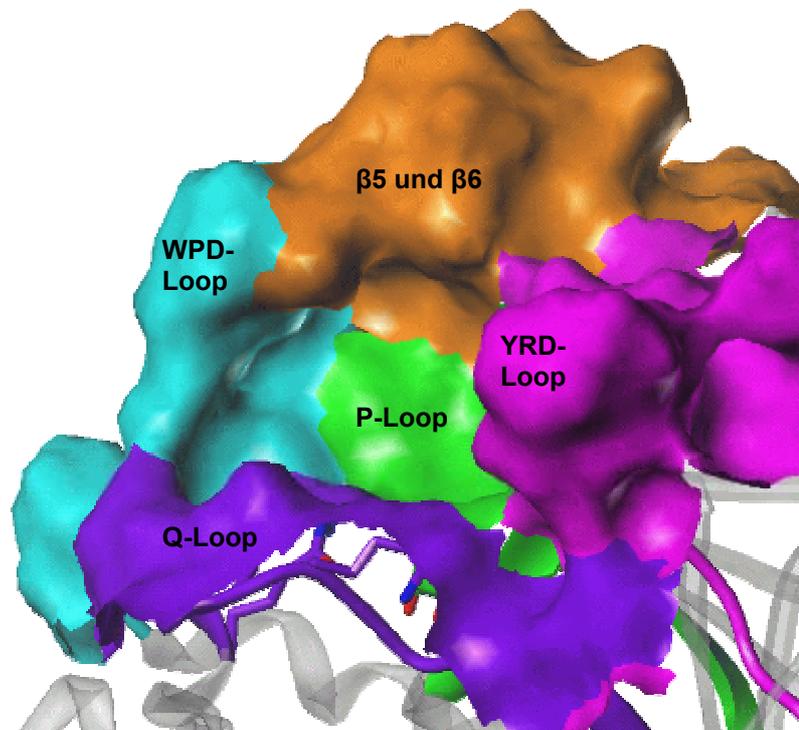


Abbildung 3-7: Aktives Zentrum der PTP1B

Vor allem der YRD-Loop spielt eine wichtige Rolle für die Aktivität und Selektivität des Enzyms indem er den Zugang begrenzt und die Tiefe der Bindungstasche festlegt. Der Q-Loop sorgt für die richtige Platzierung des Wassermoleküls, das das Intermediat spaltet [97] und das kleine β -Faltblatt befindet sich oberhalb des aktiven Zentrums. Der Boden der Vertiefung wird durch die Aminosäuren His214 bis Arg221 gebildet, die das Ende des zentralen β -Stranges β 12, den Loop L15 und den Anfang der Helix α 4 bilden und das Motiv M9 darstellen. Dieser Bereich ist die für Tyrosin Phosphatasen spezifische Sequenz, definiert die Familie der PTPs und wird deshalb oft als P-Loop, PTP-Loop oder PTP-Motiv bezeichnet. Er beinhaltet auch das aktive Cystein (Cys215) und das für den Ladungsausgleich bei der Substratbindung notwendige Arg221. Die Amidgruppen des Proteinrückgrats der Aminosäuren Ser216 bis Arg221 und die Guanidinium-Seitenkette von Arg221 bilden in der Kristallstruktur Wasserstoffbrücken zum Wolframat das als Phosphatanalogon mitkristallisiert wurde. Das konservierte His214 bildet Wasserstoffbrücken vom δ -Stickstoffatom zum Carbonylsauerstoff des Cys215 und vom ϵ -Stickstoffatom zur Hydroxylgruppe des Tyr124. Da eine direkte Interaktion mit dem Substrat nicht möglich ist, kann His214 keine Rolle bei der Katalyse spielen. Stattdessen fixiert es die Konformation von Cys215 und verstärkt die Polarisierung der Peptidbindung zwischen Cys215 und Ser216, so dass die Wasserstoffbrücke der Amidgruppe von Ser216 zum Phosphat verstärkt wird.

Schon im Jahr nach der ersten Strukturaufklärung wurde die erste Struktur einer inaktiven Mutante, bei der Cys215 im aktiven Zentrum durch ein Serin ersetzt ist (C215S), im Komplex mit verschiedenen Substraten, die aus dem EGFR stammen, sowie freiem pTyr veröffentlicht (PDB-Einträge 1PTT, 1PTU und 1PTV) [8].

Die Amidgruppen des Proteinrückgrats der Aminosäuren Ser216 bis Arg221 und die Guanidinium-Seitenkette von Arg221 bilden zusammen insgesamt sechs Wasserstoffbrücken und zwei Salzbrücken zu den drei terminale Sauerstoffatomen des Phosphotyrosins. Der YRD-Loop (Leu37 bis Ser55) mit dem YRD-Motiv (Tyr46, Arg47, Asp48) enthält mit dem Tyr46 eine der Aminosäuren, die den Aromaten des Phosphotyrosins bindet, und mit Arg47 und

Val49 zwei weitere an der Substratbindung beteiligte Reste. Die Aminosäure Phe182 liegt auf der anderen Seite dieses Aromaten, so dass eine Sandwich-Struktur entsteht (Abbildung 3-8).

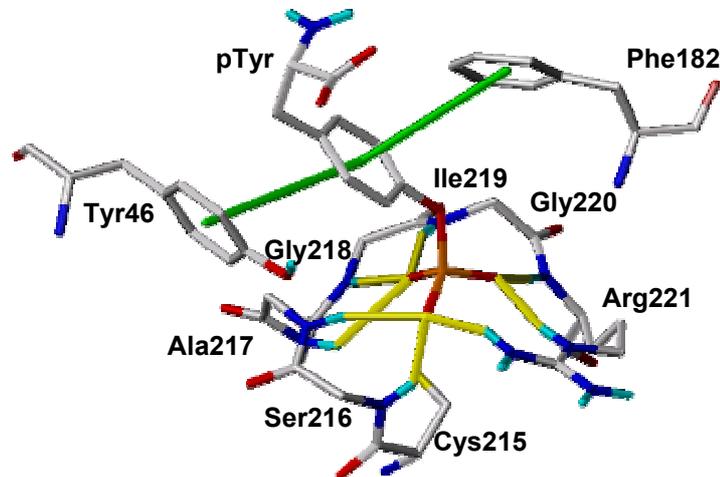


Abbildung 3-8: Bindung von Phosphotyrosin an das aktive Zentrum: Die Wasserstoffbrücken zum Phosphatrest sind in gelb dargestellt, die Sandwich-Anordnung der Aromaten durch die grünen Linien

Phe182 ist auf dem flexiblen WPD-Loop (Thr177 bis Ser187) lokalisiert, der seinen Namen von Trp179, Pro180 und Asp181 erhielt und in einer offenen und einer geschlossenen Konformation existiert. Die Atome des Proteinrückgrats bewegen sich um 5,5 Å, die Seitenketten von Asp181 und Phe182 um 8 bzw. 12 Å. Die Seitenkette von Asp181, die zuvor eine Salzbrücke zu Arg112 außerhalb des aktiven Zentrums ausbildete, geht jetzt eine mit Lys120 in der Nähe der pTyr-Bindungsstelle ein.

Wenn ein Substrat an das freie Enzym, in dem der WPD-Loop geöffnet ist bindet, wechselt der Loop in die geschlossene Konformation, wodurch Phe182 zur Bindung des Phosphotyrosins befähigt wird und Asp181 als Säure im aktiven Zentrum zur Verfügung steht (Abbildung 3-9).

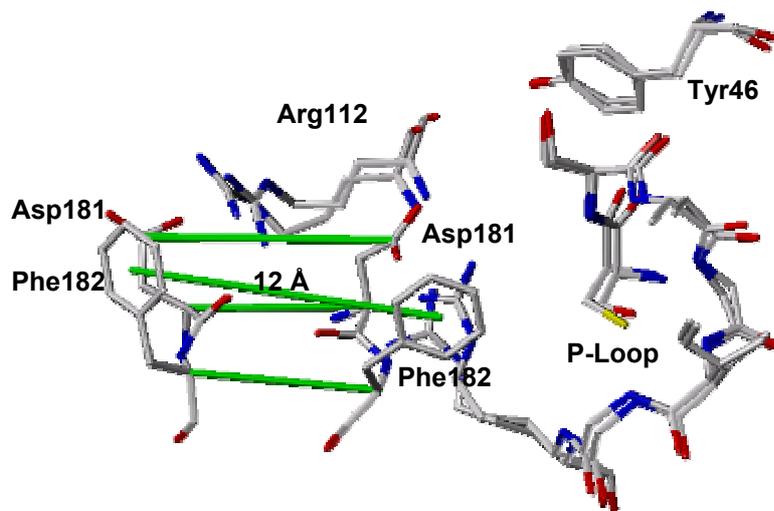


Abbildung 3-9: Bewegung des WPD-Loops

Zwei 1997 aufgeklärte Röntgenkristallstrukturen der PTP1B mit bis-(p-Phosphophenyl)methan (BPPM, siehe Abbildung 3-14, Seite 41, PDB-Eintrag 1AAX) und Phosphotyrosin (pTyr, PDB-Eintrag 1PTY) zeigen eine zweite Arylphosphat-Bindungsstelle in der Nähe des aktiven Zentrums. Die Elektronendichte für BPPM zeigte, dass aufgrund einer Überlappung nicht beide Arylphosphat-Bindungsstellen gleichzeitig besetzt sein können. In etwa 54% der Strukturen besetzt das Molekül die Bindungsstelle des aktiven Zentrums und bindet mit dem proximalen Phosphophenylrest analog pTyr im aktiven Zentrum, während die distale Phosphophenylgruppe mit Phe182 und über Wassermoleküle mit Gln266 interagiert. Zu 46% bindet BPPM in der zweiten Bindungstasche und dort mit der proximalen Phosphophenylgruppe an die Guanidinium-Gruppen der Aminosäuren Arg24 und Arg254 sowie über hydrophobe Wechselwirkungen an Ile219, Met258 und Gly259 während die distale Gruppe eine ähnliche Lage und vergleichbare Wechselwirkungen wie der distale Phosphophenylrest des ersten BPPM aufweist. In der Kristallstruktur mit zwei pTyr bindet das erste in der weiter oben beschriebenen Weise, und das zweite ähnlich dem proximalen Phosphophenylrest des zweiten BPPM [98]. Durch eine Kristallstruktur des Enzyms mit einem anderen peptidischen Inhibitor, der mit Arg47 im YRD-Motiv interagiert, wurde eine weitere Bindungsstelle für pTyr-Mimetika und ein alternativer Bindungsmodus für Biphosphonate gefunden (PDB-Eintrag 1LQF) [99].

Eine Konformation, die von den anderen Strukturen abweicht, besitzt die Röntgenkristallstruktur des Apoenzyms der PTP1B-Cys215Ser-Mutante (PDB-Eintrag 1I57), bei der vor allem der P-Loop eine völlig andere Anordnung einnimmt. Als Ursache hierfür könnte angeführt werden, dass der Übergang vom negativ geladenen Thiolat des Cysteins zu einem neutralen Alkohol des Serins diese starke Konformationsänderung notwendig macht, da die Gruppen, die das Thiolat stabilisieren, nun ihre Aufgabe verloren haben. Bei anderen Strukturen der Mutante wird diese negative Ladung von dem gebundenen Inhibitor oder Substrat beigesteuert, so dass hier keine Konformationsänderung im P-Loop zu beobachten ist [100].

Mittlerweile sind weitere hydrolytisch inaktive Mutationen der PTP1B bekannt, in denen zum Beispiel die Säure Asp181 gegen Alanin oder das für die Bindung eines wichtigen Wassermoleküls notwendige Gln262 gegen Alanin ausgetauscht wurde [101].

3.3.3 Katalytischer Mechanismus

Der Mechanismus der durch die Protein-Tyrosin-Phosphatase 1B katalysierten Hydrolyse des Phosphotyrosin-Esters wurde in den 1990er Jahren aufgeklärt und scheint bei allen PTPs gleich zu sein [97][102].

Dazu beigetragen haben besonders die in Abbildung 3-10 dargestellten Röntgenkristallstrukturen des freien Enzyms (PDB-Eintrag 2HPN) [103], des Phosphotyrosin-Enzym-Komplexes (PDB-Eintrag 1PTV) [8] und der intermediären Struktur, bei der das Phosphat kovalent als Cysteinyolphosphat an das Enzym gebunden ist (PDB-Eintrag 1A5Y) [104]. Als Übergangszustand für die Übertragung des Phosphates vom Substrat auf das Cystein wird das im aktiven Zentrum gebundene phosphatanaloge Vanadat angesehen, das in einer weiteren Kristallstruktur, pentavalent in trigonal-bipyramidaler Konformation vorliegt (PDB-Eintrag 2HPQ) [86][103][105].

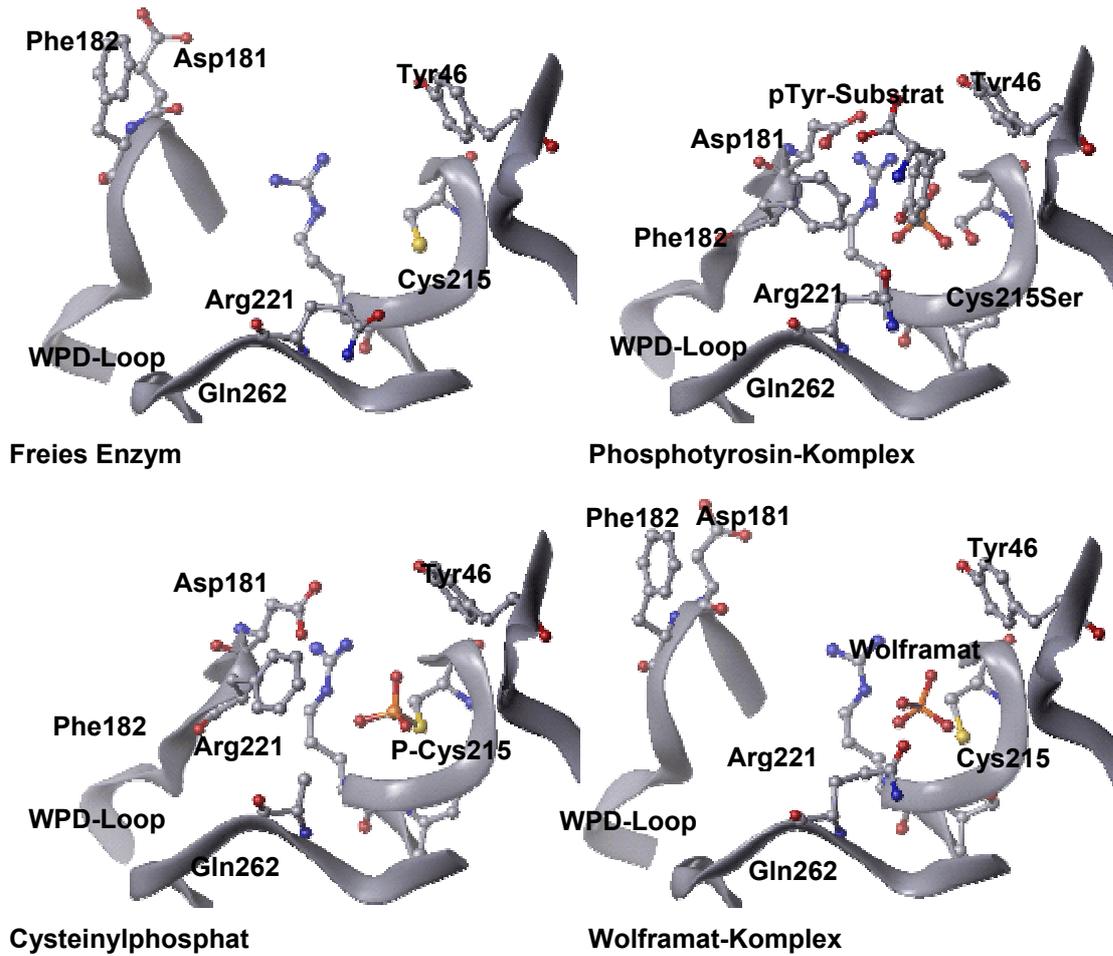


Abbildung 3-10: Dreidimensionale Struktur der einzelnen Reaktionsschritte

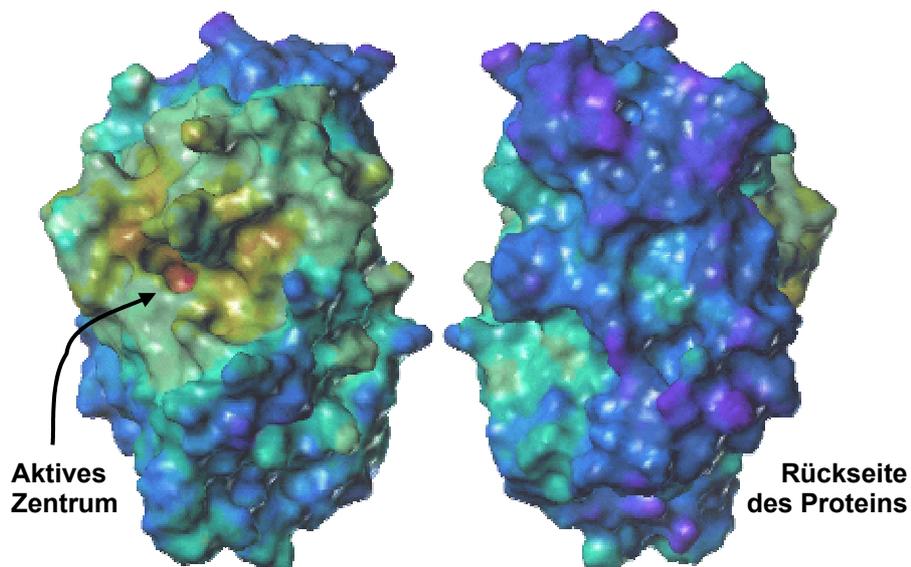


Abbildung 3-11: Elektrostatisches Potential des gesamten Proteins

Die hohe Affinität für negativ geladene Peptide ist auf den ersten Blick sehr verwunderlich, da die PTP1B insgesamt einen Überschuss von sechs negativen Ladungen besitzt. Betrachtet man allerdings die Ladungsverteilung, die in Abbildung 3-11 in Form des elektrostatischen Potentials auf der Oberfläche des Enzyms dargestellt ist, so fällt auf, dass die PTP1B in der Nähe des aktiven Zentrums ein positives Potential (orangerote Färbung) aufweist, während das restliche Protein überwiegend negativ geladen ist (blaue Färbung).

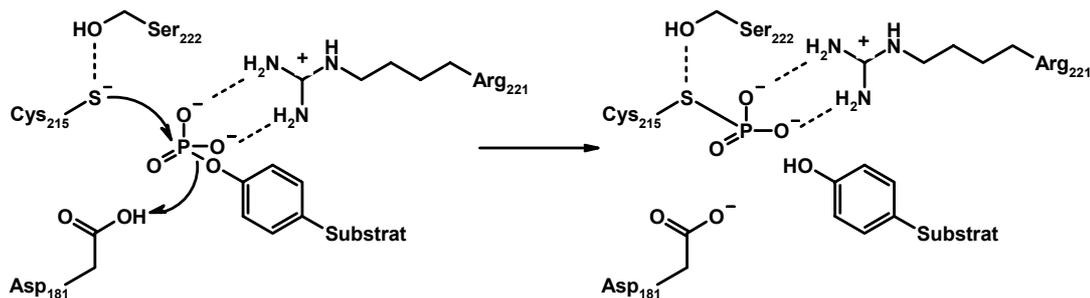


Abbildung 3-12: Bildung des Cysteinylphosphat

Nach der Bindung des Substrates übernimmt das deprotonierte nukleophile Cystein im aktiven Zentrum vom Tyrosin die Phosphatgruppe, wodurch eine kovalente Bindung zwischen dem Enzym und dem Phosphat entsteht. Wie in Abbildung 3-12 gezeigt, bildet dabei Arg221 zwei Salzbrücken zum Phosphatrest und ist sowohl für die Bindung des Substrates, als auch für die Stabilisierung des Übergangszustandes wichtig [106][107][108].

Die Freisetzung des Tyrosins, das durch die Säure des protonierten Asp181 stabilisiert wird, ist die treibende Kraft der Reaktion [107][109]. Dabei liegt das Tyrosin nicht deprotoniert als Phenolat vor, sondern die Protonierung des überbrückenden Sauerstoffatoms durch die Säure Asp181 erfolgt konzertiert mit der Übertragung des Phosphates auf Cystein [110].

Die Hydrolyse des Cysteinylphosphates durch ein Wassermolekül führt im geschwindigkeitsbestimmenden, zweiten Schritt zum nicht kovalent gebundenen Enzym-Phosphat-Komplex, von dem das anorganische Phosphat schließlich abdissoziiert. Das zur Hydrolyse benötigte Wassermolekül wird durch die Seitenkette von Gln262 fixiert, wie in Abbildung 3-13 gezeigt. Durch die Mutation dieser wichtigen Aminosäure zu Alanin war es möglich, die Reaktion auf diese Stelle anzuhalten, so dass auch das Cysteinylphosphat-

Intermediat und damit jeder einzelne Schritt der Reaktion visualisiert werden konnte (PDB-Eintrag 1A5Y) [104].

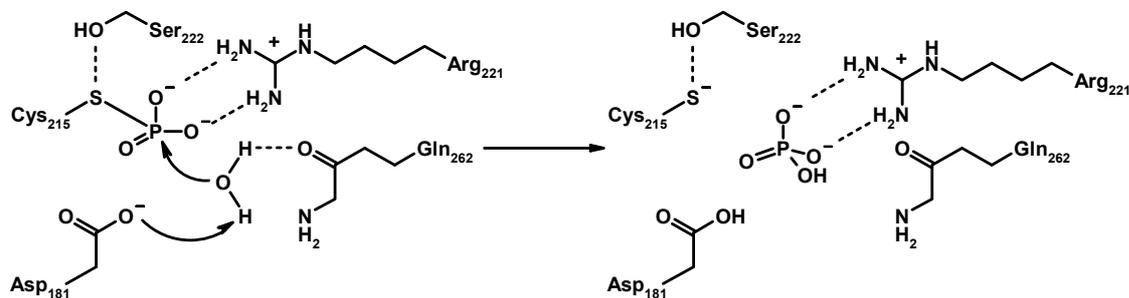


Abbildung 3-13: Spaltung des Cysteinylphosphat

Da verschiedene protonierbare Gruppen an der Reaktion beteiligt sind, ist die Reaktion pH-abhängig. So liegt der pK_S -Wert des Arylphosphates, das nach kinetischen Studien zweifach negativ geladen vorliegen muss, bei etwa 5,0 bis 5,5 für die zweite Deprotonierung. Die Aminosäure Asp181, die mit einem pK_S -Wert von etwa 4,9 deprotoniert vorliegen muss, ist für die Säurekatalyse der Reaktion zuständig. Die Aminosäure Glu115 muss deprotoniert sein, da sie eine Salzbrücke mit der Seitenkette von Arg221 im aktiven Zentrum bildet [111] [112]. Da die pK_S -Werte der beiden Seitenketten ebenfalls im Bereich von 5,0 bis 5,5 liegen, ergibt sich für das Enzym ein pH-Optimum von etwa 5,5 [113]. Weiterhin konnte gezeigt werden, dass der pK_S -Wert des aktiven Cys215 bei etwa 4,6 bis 5,6 liegt [112][114], also ungefähr 4 pK_S -Einheiten unter dem für Cystein in Proteinen üblichen Wert. Dieser niedrige pK_S -Wert scheint für den nukleophilen Angriff auf das Phosphoratom des Substrats absolut notwendig zu sein [115]. Im Gegensatz zu Cystein-Proteinasen, die das Thiolat-Anion durch eine Ion-Ion-Wechselwirkung mit einem Histidin, welches das Proton übernimmt, stabilisieren, findet bei den PTPs die Stabilisierung durch ein Netzwerk von Wasserstoffbrücken statt. Die Stickstoffatome des Proteinrückgrats des P-Loops sind zum Schwefel des Cysteins mit einer Entfernung von etwa 2,8 bis 3,3 Å im richtigen Abstand zur Bildung von Wasserstoffbrücken zu einem Schwefelatom [116].

Zusätzlich ist die Hydroxylgruppe von Ser222 als nächster Protonendonator für Cys215 vorhanden [117] und eine weitere Stabilisierung kann durch das ebenfalls konservierte His214 erfolgen [114]. Einen zusätzlichen wichtigen

Einfluss auf den pK_S -Wert des Cysteins besitzt die Sekundär- und Tertiärstruktur des aktiven Zentrums, in dem die Mikrodipole des Proteinrückgrats vom P-Loop sowie der Dipol der zentralen α -Helix in Richtung des Schwefelatoms zeigen [106] (siehe auch Abschnitt 5.1.3).

3.3.4 Substratspezifität, Selektivität

PTPs sind völlig spezifisch für Phosphotyrosin und somit nicht in der Lage, phosphoryliertes Serin oder Threonin in Peptiden zu dephosphorylieren. Die einzelnen PTPs besitzen zum Teil eine hohe Substratspezifität, wie K_m -Werte, die über 23 Größenordnungen reichen, eindrucksvoll belegen. Grundlage dieser Spezifität ist die Form des katalytischen Zentrums, das von seinem Eingang bis zum Boden, an dem sich das aktive Cystein befindet, 9 Å tief ist und somit einem Tyrosin exakt die richtige Größe bietet [118][119]. Als Substrate für kinetische Untersuchungen oder für die Testung von möglichen Inhibitoren dienen in erster Linie Phosphotyrosin (pTyr) und p-Nitrophenylphosphat (NPP) aber auch weitere substituierte Phosphophenole oder Phospho-Phenolphthalein (Abbildung 3-14). Mit Hilfe von Bis-Phosphophenylmethan (BPPM) wurden weitere Bindungsstellen für Arylphosphate gefunden.

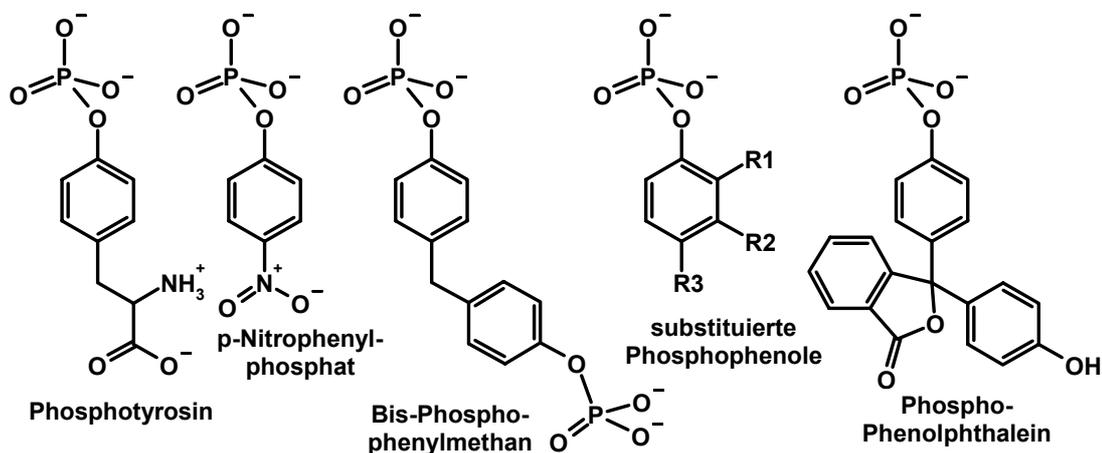


Abbildung 3-14: Substrate für PTP1B

Anhand kinetischer Untersuchungen von Mutanten des Enzyms wurde die Bedeutung der Aminosäuren Tyr46, Arg47 und Asp48 des YRD-Motivs, von Phe182 im WPD-Loop sowie der ebenfalls mit dem Tyrosin des Substrates interagierenden Aminosäure Gln262 für die Substratspezifität untersucht. Der

Aromat von Tyr46 ist wichtig für die Bindung des aromatischen Ringes vom Substrat und kann ohne Verlust an Spezifität gegen Phenylalanin ausgetauscht werden. Allerdings führt diese Mutation zu einem beschleunigten Abbau der PTP1B während die Mutation zum Alanin zu einem wesentlich schwächer aktiven Enzym führt. Arg47 spielt für die Hydrolyse keine Rolle, wie mit NPP als Substrat an Mutanten gezeigt wurde, ist aber für die Spezifität des Enzyms von großer Bedeutung, da es an die sauren Aminosäuren N-terminal vom pTyr bindet. Asp48 bindet hauptsächlich an die Amidgruppe vom pTyr des Substrates weshalb der Austausch gegen Alanin zwar zu einer geringeren Affinität für peptidische Substrate führt, aber keinen Einfluss auf die Spezifität hat. Phe182 zeigt nur einen geringen Einfluss auf die Spezifität, ist aber für die katalytische Aktivität wichtig, da es dazu beiträgt, die geschlossene Konformation des WPD-Loop zu stabilisieren und somit Asp181 in die richtige Position zu bringen. Gln262 interagiert zwar auch mit dem Substrat, ist aber besonderes bedeutsam, indem es das Wassermolekül bindet, das das Cysteinylphosphat im zweiten Reaktionsschritt hydrolysiert [120][121][122]. Eine weitere Analyse der Wechselwirkungen belegt nochmals, dass die Zerstörung der ionischen Interaktion zwischen Arg47 und den N-terminalen sauren Resten des Substrates die Interaktion zwischen Enzym und Substrat stark verringert. Das für die Katalyse wichtige Asp181 setzt hingegen die Affinität eher herab [123].

Bei einer Analyse der Komplexbildungsgeschwindigkeit und der Hydrolyserate des Aktivierungssegmentes des Insulinrezeptors mit der Sequenz R-D-I-Y-E-T-D-Y-Y-R-K-G-G-K-G-L-L, das als mono-, bis- und tris-Phosphat untersucht wurde, stellte man fest, dass die PTP1B eine hohe Präferenz für die bis- und tris-Phosphate aufweist. Vor allem ist die Sequenz Asp-pTyr-pTyr-Arg nicht nur für die Erkennung durch die PTP1B wichtig, sondern spielt auch eine entscheidende Rolle für die Selektivität gegenüber anderen Phosphatasen (PDB-Einträge 1G1F, 1G1G und 1G1H) [56].

Durch *Inverse-Alanine-Scanning* wurden jeweils die vier Positionen vor und hinter dem pTyr in der Peptidkette untersucht. Dazu wurde in dem Peptid Ac-A-A-A-A-pY-A-A-A-A-NH₂ jeweils ein Alanin gegen alle anderen Aminosäuren ausgetauscht und die Hydrolysegeschwindigkeit gemessen. In allen

Positionen des Substrates besteht eine Präferenz für saure Reste, in P-1 werden aber auch lipophile Aminosäuren wie Leucin, Isoleucin, Valin, Cystein und Threonin akzeptiert und am stärksten Phenylalanin bevorzugt. Die basischen Aminosäuren Arg und Leu sind in allen Positionen ungünstig und in P+1 ist besonders Pro unvorteilhaft. Als Ergebnis fand man das Konsensus-Peptid Ac-E-L-E-F-pY-M-D-Y-E-NH₂, das eine entsprechend hohe Hydrolyse-rate aufweist und ein modifiziertes Peptid D-A-D-BPA-pY-L-I-P-Q-Q-G vom EGFR (BPA: p-Benzoylphenylalanin) [124]. Kinetischen und röntgenkristallographischen Untersuchungen zufolge werden die wichtigsten Interaktionen durch das Peptidrückgrat des Substrates und pTyr selbst verursacht. Die Möglichkeit, sowohl aromatische als auch saure Reste in P-1 zu akzeptieren, basiert auf der Beweglichkeit der Aminosäure Arg47, welche zwei unterschiedliche Konformationen einnehmen kann und somit in der Lage ist, zwei Klassen von Substraten Wechselwirkungsmöglichkeiten zu bieten [125].

Der nächste Verwandte der PTP1B ist die TCPTP, die eine Sequenzidentität von 88,8% in der katalytischen Domäne besitzt. Trotz ihrer großen Sequenz-ähnlichkeit erkennen beide Phosphatasen unterschiedliche Substrate und haben verschiedene biologische Funktionen. Während *in vivo* PTP1B hauptsächlich den IR und den EGFR erkennt, dephosphoryliert die TCPTP neben dem EGFR andere tyrosinphosphorylierte Proteine aber nicht den IR. Einen weiteren Hinweis auf die unterschiedliche Funktion liefert die Tatsache, dass Mäuse, denen die PTP1B fehlt, im Allgemeinen gesund sind und eine Lebensspanne besitzen, die denen von Vergleichstieren entspricht [4][5], wohingegen Mäuse, denen die TCPTP fehlt, nach etwa zwei Wochen erste Krankheitssymptome entwickeln und nach fünf Wochen sterben [126]. *In vitro* kann allerdings auch die TCPTP den IR dephosphorylieren, so dass die Spezifität also eher aus der zellulären Lokalisierung zu resultieren scheint als aus der Selektivität des Enzyms selbst [33]. Besonders im so genannten YRD-Motiv mit den umgebenden Aminosäuren gibt es eine Übereinstimmung von 100% zwischen beiden Enzymen, und somit lassen sich Untersuchungen zur Kinetik von der TCPTP auf die PTP1B übertragen.

Ein genauer Vergleich der TCPTP mit der PTP1B zeigt, dass viele Eigenschaften der beiden Enzyme sehr ähnlich sind, wie etwa das pH-Optimum und die Aktivierungsenergie [127]. Seit Mai 2002 ist eine Kristallstruktur der TCPTP vorhanden (PDB-Eintrag 1L8K), in der zwar eine extrem große Strukturähnlichkeit zur PTP1B vorhanden ist, aber auch zwei Bereiche gefunden wurden, die sich entscheidend von der PTP1B unterscheiden. Die beiden Bereiche befinden sich in der Nähe des aktiven Zentrums und bestehen aus den Aminosäuren Cys32, Lys39 und Phe52 (His34, Glu41 und Tyr54 in TCPTP), sowie Ala17, Gln21 und Ala264 (Gln19, Leu23 und Pro262 in TCPTP) [128]. Ob die geringen Unterschiede zwischen den beiden Phosphatasen ausreichen, um selektive Inhibitoren zu synthetisieren, wird die zukünftige Wirkstoffforschung auf diesem Gebiet zeigen.

3.3.5 Regulation der PTP1B

Zurzeit sind mehrere Mechanismen bekannt, mit deren Hilfe die PTP1B-Aktivität reguliert wird. Die Aktivierung des IR führt zur Produktion von Wasserstoffperoxid (H_2O_2) und dadurch zu einer reversiblen Oxidation und Inaktivierung der PTP1B. Der Anteil an oxidiertem PTP1B ist in verschiedenen Geweben unterschiedlich, und man vermutet, dass ein erhöhter Anteil an nicht oxidiertem und somit aktiver PTP1B zur Insulinresistenz beiträgt [129][130][131]. Auch einige Inhibitoren, wirken über die Oxidation des aktiven Cysteins mit Hilfe von Wasserstoffperoxid [132]. Die Oxidation des Cysteins mit Wasserstoffperoxyd läuft wahrscheinlich über die Sulfensäure (Cys-SOH) und führt nachweislich zu einem Ringschluss zwischen dem Schwefel des Cysteins und den Rückgrat-Stickstoff von Ser216, wodurch die Konformation des P-Loops verändert wird, und mehrere Wasserstoffbrücken (zwischen Ser222 und Cys215 sowie Tyr46 und Ser215) gebrochen werden (PDB-Einträge 1OEM und 1OES). Bei höheren Konzentrationen des Oxidationsmittels entstehen die Sulfinensäure und die Sulfonsäure (Cys-SO₂H und Cys-SO₃H; PDB-Einträge 1OEO, 1OET, 1OEU und 1OEV) [133][134]. Neben Wasserstoffperoxid könnte auch Stickstoffmonoxid (NO) in der Lage sein, das Cystein des aktiven Zentrums zu oxidieren oder S-Nitrosothiole zu bilden [135].

Weiterhin kann die Enzymaktivität durch Anregung der entsprechenden Rezeptoren mit Insulin und IGF-1 (Insulin like Growth Factor 1) erhöht werden. [136]. Es wurde gezeigt, dass die PTP1B selbst durch den IR und andere Kinasen phosphoryliert werden kann. In Mutationsexperimenten wurden die Aminosäuren Tyr66, Tyr152 und Tyr153 als Phosphorylierungsstellen erkannt [32].

Auch für die Regulation der Expression der PTP1B sind Mechanismen und Regulatoren aufgeklärt worden. So führt die Überexpression des YB1 (Y-Box bindendes Protein 1) zu einer gesteigerten Synthese von PTP1B, die Ausschaltung von YB1 reduziert hingegen die Expression der PTP1B um 70%, beeinflusst aber nicht die Expression der TCPTP [137].

4 Methoden

4.1 Berechnungsgrundlagen

Noch vor wenigen Jahren war die Pharmaindustrie bei der Entwicklung neuer Wirkstoffe auf empirische Methoden beschränkt. Zehntausende unterschiedliche Substanzen wurden täglich experimentell auf ihre eventuelle Eignung zur Behandlung von Krankheiten getestet. Heute ist es durch Fortschritte in den Bio- und Computerwissenschaften möglich, vor der experimentellen Testung Berechnungen durchzuführen, die schon im Vorfeld die Anzahl möglicher Kandidaten aufgrund biologischer und biochemischer Erkenntnisse erheblich einschränken können.

In den letzten Jahrzehnten hat das exponentielle Wachstum biochemischer Daten auf molekularer Ebene dazu geführt, dass man vom rein empirischen Vorgehen zum rationalen und systematischen Design übergegangen ist. Zur Verarbeitung großer Mengen biologischer oder chemischer Daten und zur anschaulichen und interpretierbaren Darstellung, ist der Einsatz von Computern unumgänglich. Computer sind jedoch nicht in der Lage, Entscheidungen selbständig und auf dem Kontext basierend zu treffen.

Die in dieser Arbeit verwendeten computergestützten Methoden werden im Folgenden kurz erläutert.

4.1.1 Quantenmechanik

Man unterscheidet beim Molecular Modelling zwei grundsätzlich verschiedene Ansätze zur Berechnung von Molekülgeometrien und Moleküleigenschaften, die Quantenmechanik und die Molekülmechanik. Sie unterscheiden sich durch den Grad der notwendigen Näherungen und im benötigten Rechenaufwand.

Quantenmechanische Berechnungen werden weiterhin in *ab-initio*-Rechnungen und semiempirische Berechnungen unterteilt. *Ab-initio*-Rechnungen basieren nicht auf abgeleiteten Parametern, sondern nur auf Fundamentalkonstanten wie zum Beispiel der Elementarladung, der Masse des Elektrons und dem Planckschen Wirkungsquantum. Die Schrödinger-Gleichung, die Grundlage der Quantenchemie, lässt sich allerdings nur für das Wasserstoffatom exakt lösen, weshalb für größere Moleküle Näherungen entwickelt wurden. Ein bekanntes

Näherungsverfahren ist das von Hartree und Fock entwickelte 'Self Consistent Field' (SCF). Die Atomkerne werden hierbei als ortsfest angesehen und getrennt von den Elektronen berechnet, die Vielelektronen-Wellenfunktion wird in einzelne Eielektronen-Wellenfunktionen aufgeteilt und die Molekülorbitale werden als Linearkombination der einzelnen Atomorbitale betrachtet.

Um die Rechnungen bei größeren Molekülen weiter zu verkürzen, werden so genannte semi-empirische Verfahren verwendet, bei denen zusätzliche Vereinfachungen angenommen werden. So werden zum Beispiel nur die Orbitale der Valenzelektronen berücksichtigt und viele Parameter, die sonst aufwendig aus Integralen berechnet werden müssten, werden auf empirischen Daten basierend abgeschätzt.

4.1.2 Molekülmechanik

Um aber auch dynamische Effekte über einen bestimmten Zeitraum betrachten oder große Moleküle wie Proteine simulieren zu können, reichen die Abstraktionen der semi-empirischen Verfahren noch nicht aus. Man greift daher zur Molekülmechanik, die den Energiegehalt von Molekülen durch die Gesetze der klassischen Mechanik beschreibt.

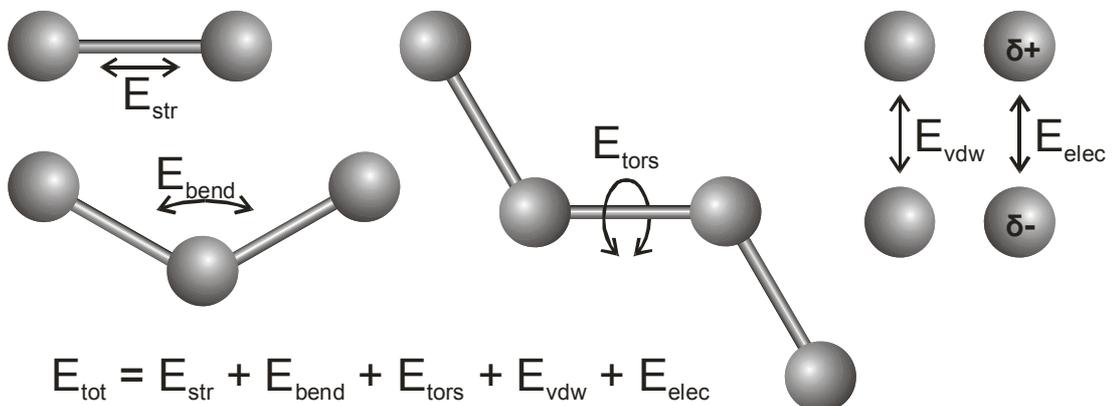


Abbildung 4-1: Kraftfeldterme

Die Energie des Moleküls, dessen Atome dabei als über Federn miteinander verbundene Massenpunkte betrachtet werden, wird als Funktion der Atomkoordinaten berechnet. Die meisten Kraftfelder verwenden Energieterme für Bindungslänge (E_{str}), Bindungswinkel (E_{bend}) und Torsionswinkel (E_{tors}), sowie zwei

Terme für Wechselwirkungen durch den Raum, die Dispersionsenergie (E_{vdw}) und die elektrostatische Energie (E_{elec}) (Abbildung 4-1).

In dieser Arbeit wurde zur Geometrieoptimierung kleiner Moleküle das Tripos-Kraftfeld [138] in der Software Sybyl [139] verwendet, das die qualitativ gleichwertige Berechnung von kleinen organischen Molekülen und Biopolymeren erlaubt. Zur Moleküldynamiksimulation von Enzymen und Enzym-Inhibitor-Komplexen wurden das CVFF (consistent valence force field) [140] des Programmpaketes Insight/Discover [141] und das Amber-Kraftfeld im Programm PrGen verwendet, die ursprünglich für Proteine entworfen wurden [142][143], inzwischen aber auch auf andere organische Verbindungen erweitert worden sind.

4.2 Allgemeine Methoden

4.2.1 Geometrieoptimierung

Nach dem Bau von Molekülen in einem Modelling-Programm liegen diese häufig in einer energetisch ungünstigen Konformation vor. Auch die Strukturen, die aus einer Kristallstruktur-Datenbank eingelesen wurden, befinden sich, durch Packungseffekte und die Auflösung der Struktur bedingt, nicht immer in einem lokalen Energieminimum. Mit Hilfe einer Geometrieoptimierung, auch Energieminimierung genannt, sucht man auf der Energiehyperfläche das nächste lokale Minimum. Man muss hierbei berücksichtigen, dass die resultierende Konformation von den Parametern des verwendeten Kraftfeldes abhängen kann. Unter den zur Geometrieoptimierung verwendbaren Methoden haben sich vor allem *Steepest Descents* und *Conjugate Gradient* etabliert. Diese beiden gradientenbasierten Verfahren benutzen zur Ermittlung der Suchrichtung die erste Ableitung der Energiefunktion und gelangen so über viele Iterationsschritte, in denen die Atompositionen optimiert werden, immer weiter in Richtung eines lokalen Energieminimums. Wenn die Energiedifferenz zwischen zwei Iterationen einen bestimmten Betrag unterschreitet oder eine festgelegte Anzahl an Schritten erreicht ist, wird die Optimierung beendet. Die Methode *Conjugate Gradient* berechnet zusätzlich die zweite Ableitung der Energiefunktion und speichert die Werte der vorhergehenden Iterationen, um die Suche so zu optimieren. Da dies aber mit einer Erhöhung der Rechenzeit erkauft wird, verwendet man zu Beginn einer Geometrieoptimierung meist das schnellere *Steepest Descents*. In der Nähe des Energieminimums wechselt man dann zu *Conjugate Gradient*, weil die Methode effektiver ist, da man sich dem Minimum schneller nähert und Rückschritte vermeidet.

Bei dieser Arbeit wurden zur Berechnung von Proteinen und Protein-Ligand-Komplexen im Programm Insight/Discover beide Methoden nacheinander angewendet, das Programm PrGen enthält nur die Algorithmen für *Conjugate Gradient*. Für die Liganden wurden die Berechnungen in Sybyl mit der Powell-Methode, einem abgewandelten *Conjugate Gradient* Verfahren durchgeführt.

4.2.2 Moleküldynamiksimulation

Auch Moleküldynamiksimulationen (MDS) basieren auf den Parametern eines Kraftfeldes, berücksichtigen aber zusätzlich die Bewegung und Flexibilität der betrachteten Moleküle in Abhängigkeit von der Zeit. Bei der Simulation von Bewegungen sollte der Zeitschritt etwa um den Faktor 10 kleiner sein als die Periode der schnellsten zu simulierenden Schwingung. Je nachdem, ob Schwingungen einzelner Bindungen (Periode 10^{-14} Sekunden) oder die Bewegung von Proteinseitenketten oder Loops (Periode 10^{-8} Sekunden) simuliert werden sollen, sind unterschiedliche Zeitschritte notwendig. Auch die Simulationstemperatur spielt eine Rolle für die Geschwindigkeit, mit der die Prozesse ablaufen. Dies wird zum Beispiel beim *Simulated Annealing* und bei Hochtemperaturdynamiken ausgenutzt, die verwendet werden, um unterschiedliche Konformationen für Moleküle zu generieren. Die MDS für diese Arbeit wurden in Insight/Discover mit dem CVFF durchgeführt, um Startgeometrien des Enzyms für das Docking zu erzeugen und die gedockten Komplexe zu relaxieren.

4.2.3 Konformationsanalyse

Nur selten sind organische Moleküle rigide und können nur in einer einzigen Konformation vorkommen. Durch eine Geometrieoptimierung gelangt man jedoch nur in das nächste Minimum auf der Energiehyperfläche. Um aber auch andere Energieminima, also weitere Konformationen, in denen das Molekül häufig vorliegt, sowie das globale Minimum und gegebenenfalls die bioaktive Konformation zu finden, bedient man sich der Konformationsanalyse. Man unterscheidet die systematische Analyse von den zufallsbasierten Verfahren der 'Monte-Carlo Simulation' und des 'Simulated Annealing' (Tabelle 4-1).

Bei einer systematischen Konformationsanalyse werden die Torsionswinkel der Einfachbindungen in festgelegten Schritten verändert, die Energie der entstandenen Konformationen berechnet und solche mit einer hohen Energie ausgeschlossen. Von Nachteil ist, dass je nach Inkrement für den Drehwinkel und Anzahl der drehbaren Bindungen eine große Anzahl Konformere entsteht und Ringsysteme als starr betrachtet werden müssen. Vorteilhaft ist aber, dass

der Konformationsraum vollständig analysiert werden kann. Bei einer *Monte-Carlo Simulation* werden einzelne Atome nach dem Zufallsprinzip verschoben und die entstehenden Konformere relaxiert, beim *Simulated Annealing* wird eine Moleküldynamiksimulation bei hoher Temperatur durchgeführt und die Temperatur dann schrittweise bis zum Nullpunkt reduziert. Beide Methoden erlauben die Analyse von Ringsystemen. Zur Abdeckung des gesamten Konformationsraumes sind allerdings viele Simulationen notwendig, wodurch einzelne Konformationen sehr häufig gefunden werden, wodurch hohe Redundanz entsteht. Sowohl bei der systematischen Analyse als auch bei den statistischen Verfahren ist es erforderlich, die Geometrie der gewonnenen Konformere zu optimieren und doppelt vorkommende Strukturen auszusortieren. Da die in der vorliegenden Arbeit betrachteten PTP 1B Inhibitoren nur aromatische Ringsysteme enthalten, die keine Flexibilität besitzen, wurde für diese Verbindungen das Verfahren der systematischen Konformationsanalyse angewendet. Für das aktive Zentrum des Proteins wurde eine Dynamiksimulation bei erhöhter Temperatur durchgeführt, um verschiedene Konformationen zu generieren. Damit die erhaltenen Konformationen des Proteins nicht wieder in das Energieminimum der Kristallstruktur zurück fallen, wurden diese nicht weiter optimiert.

	Monte-Carlo Simulation	Simulated Annealing	Systematische Analyse
Veränderung der Atompositionen	zufällige Verschiebung	Dynamiksimulation bei höherer Temperatur	systematische Veränderung der Torsionswinkel
Ringsysteme	Flexibel	Flexibel	Starr
Abdeckung des Konformationsraumes	unvollständig	unvollständig	vollständig
Vermeidung ungünstiger Konformationen	Energieminimierung	Senken der Simulationstemperatur bis zum Nullpunkt	Energieminimierung
Rechenaufwand abhängig von	Anzahl der generierten Konformationen	Anzahl der generierten Konformationen	Anzahl der betrachteten Torsionswinkel

Tabelle 4-1: Methoden zur Konformationsanalyse

4.2.4 Clusteranalyse

Von Proteinstrukturen gibt es oft unterschiedliche Konformere, die aus verschiedenen Röntgenkristall- oder NMR-Strukturen stammen oder durch eine Moleküldynamiksimulation gewonnen wurden. Um die Konformere in Gruppen einzuteilen, die sich nur durch kleine Veränderungen in der Konformation unterscheiden, bedient man sich der Clusteranalyse. Mit dem Programm NMRClust [144] werden die Konformere eines Ensembles zuerst überlagert und für alle paarweisen Überlagerungen die RMS-Werte (root mean square, Wurzel aus der Summe der Abstandsquadrate definierter Atome) berechnet. Mittels dieser RMS-Matrix werden die beiden ähnlichsten Strukturen beziehungsweise Cluster zu einem neuen Cluster zusammengefasst und die Streuung aller RMS-Werte eines Clusters berechnet. Dies wird solange wiederholt, bis nur noch ein Cluster mit allen Konformeren übrig bleibt. Aus den Mittelwerten der Streuung und der Anzahl der Cluster wird ein Grenzwert berechnet, der dann zur endgültigen Einteilung in sinnvolle Cluster dient. Für jeden Cluster wird als repräsentative Struktur diejenige ausgewählt, die dem Mittelpunkt des Clusters am nächsten liegt. Strukturen, die in keinen der Cluster hineinpassen, werden als Ausreißer betrachtet. Dieses Verfahren wurde zur Analyse der Röntgenkristallstrukturen der PTP1B und zur Auswertung der Moleküldynamiksimulationen des Enzyms herangezogen.

Alternativ dazu wurde zum Clustern der Proteinstrukturen auch eine Hauptkomponentenanalyse (PCA, Principle Component Analysis, siehe Abschnitt 4.5.3) von Interaktionsfeldern (siehe Abschnitt 4.5) angewandt. Zu diesem Zweck wurden die dreidimensionalen Proteinstrukturen nach einem Sequenz-Alignment überlagert (Abschnitt 4.3.5). Für das aktive Zentrum wurden GRID-Felder (siehe Abschnitt 4.5.1) berechnet, die mittels einer Hauptkomponentenanalyse (siehe Abschnitt 4.5.3) analysiert wurden. Das Verfahren wird im Kapitel 5.1 ausführlich behandelt und mit der Clusteranalyse von RMS-Werten verglichen.

4.2.5 Partiaalladungen

Alle Atome, die in einem Kraftfeld berechnet werden sollen, benötigen elektrostatische Punktladungen, die mit quantenmechanischen, semi-empirischen oder empirischen Methoden berechnet werden können. Die quantenmechanischen Ladungen haben den Nachteil, dass sie aufwändig zu berechnen sind, weshalb sie sich hauptsächlich zur Parametrisierung von Kraftfeldern und empirischen Methoden eignen. Der Nachteil der semi-empirischen Verfahren liegt darin, dass sie bei Biopolymeren und stickstoffhaltigen Verbindungen oft ungenaue oder falsche Werte liefern. Deshalb wurden empirische Ladungen nach der Gasteiger-Hückel-Methode für die PTP1B Inhibitoren berechnet. Diese setzt sich zusammen aus der Methode nach *Hückel* zur Berechnung der π -Ladungen und der Methode nach *Gasteiger* und *Marsili*, mit der σ -Ladungen ermittelt werden. Insgesamt wird so für jedes Atom die Partiaalladung unter Berücksichtigung der Bindungspartner und der freien Valenzen berechnet. Für die Proteinstrukturen wurden Ladungen nach *Kollman* verwendet, die aus einer Liste mit Partiaalladungen für die Aminosäuren eingelesen werden. Dabei wurden je nach Notwendigkeit die Ladungen für das *All-Atom*-Modell oder das *United-Atom*-Modell, bei dem lipophile Wasserstoffatome dem entsprechenden Kohlenstoff zugerechnet werden, verwendet.

Da einige Aminosäuren in Proteinen die Möglichkeit besitzen, Protonen mit dem Medium auszutauschen und in verschiedenen Protonierungszuständen vorkommen können, ist es wichtig die pK_S -Werte und somit die Gesamtladungen dieser titrierbaren Aminosäuren für den vorliegenden pH-Wert zu bestimmen. Zu diesem Zweck wurde das Programm UHBD (University of Houston Brownian Dynamics Program) [145][146][147] verwendet, das den Ladungszustand jeder Aminosäure für den pH-Wert von 0 bis 14 berechnet und daraus den pK_S -Wert bestimmt. Das Programm benutzt dazu folgenden Ansatz: der Unterschied im Protonierungsgrad einer bestimmten protonierbaren Gruppe im Protein und in einer Modellaminosäure entspricht dem Unterschiede in der elektrostatischen Arbeit, die notwendig ist, um die entsprechende Gruppe im Protein und im Modell von der unprotonierten in die protonierte Form zu überführen, also ihre Gesamtladung zu verändern. Somit können die pK_S -Werte der titrierbaren

Gruppen im Protein aus den korrespondierenden pK_S -Werten der Modellamino-säuren bestimmt werden, indem die elektrostatische Arbeit berücksichtigt wird. Diese setzt sich aus dem Term für die Protonierung der Gruppe im Protein, wenn alle anderen Gruppen neutral sind, und der Interaktion aller titrierbaren Gruppen zusammen. Es werden also die Energiedifferenzen $\Delta G_{\text{Protein}}$ und ΔG_{Model} berechnet, aus denen dann die Differenz $\Delta\Delta G$ berechnet wird, die zum Unterschied im pK_S -Wert zwischen Modell und Protein proportional ist.

4.2.6 Röntgenkristallstrukturen

Die in der Arbeit verwendeten Röntgenkristallstrukturen der PTP1B und anderer Tyrosinphosphatasen stammen aus der Protein Data Bank des *Research Collaboratory for Structural Bioinformatics* (RCSB) [10].

4.3 Superpositionierung

Eine bei Molecular Modelling Untersuchungen häufig notwendige Aufgabe ist die Überlagerung von mehr oder weniger ähnlichen Molekülen, auch Superpositionierung oder Alignment genannt. Sie wird bei Proteinen ebenso angewandt wie bei kleinen organischen Molekülen, zum Beispiel Rezeptoragonisten und Rezeptorantagonisten oder Enzyminhibitoren, wo sie als ligandbasierte Methode zur Generierung von Pharmakophor-Modellen dienen kann. Eine sinnvolle Überlagerung bildet sowohl für weitergehende Untersuchungen von Liganden, wie zum Beispiel das 3D-QSAR, als auch zum Vergleich unterschiedlicher Proteinkonformationen oder verschiedener Proteine eine wichtige Grundlage. Man unterscheidet zwischen einer flexiblen und einer rigiden Überlagerung, wobei für Proteine und Liganden unterschiedliche Methoden angewendet werden. Die für Liganden benutzten Methoden sind in Tabelle 4-2 zusammengefasst und im Folgenden näher erläutert.

	Fit-Atoms	Multifit	FlexS
Flexibilität	Templat und zu überlagernde Struktur starr	Templat und zu überlagernde Struktur flexibel	Templat starr, zu überlagernde Struktur flexibel
Überlagerung	einzelner Atome	einzelner Atome	Interaktionssphären
Verfahren	kleinster RMS-Wert	Geometrieoptimierung mit Zusatzbedingungen	Inkrementeller Aufbau der Struktur unter Berücksichtigung der Templatstruktur
Resultat	eine Überlagerung	eine Überlagerung	Ensemble von Strukturvorschlägen mit Bewertung

Tabelle 4-2: Methoden zur Superpositionierung von kleinen Molekülen

4.3.1 Fit-Atoms-Routine (Sybyl) zur Überlagerung der Liganden

Bei dieser Methode werden mehrere Atompaare aus zwei Molekülen und somit die Moleküle selbst rigide überlagert und der RMS-Wert für die überlagerten Atompaare berechnet. Nachteilig ist bei diesem Verfahren, dass jeweils nur zwei Moleküle betrachtet werden können und die zu überlagernden Atome genau spezifiziert werden müssen. Bei letzterem Punkt kann zwar das

Definieren von Pseudo-Atomen (zum Beispiel das Zentrum eines Aromaten) hilfreich sein, da aber keine Flexibilität ermöglicht wird, sind andere Verfahren oft besser geeignet um eine gute Überlagerung zu erzeugen.

4.3.2 Multifit-Routine (Sybyl) zur Überlagerung der Liganden

Die Multifit-Routine bietet die Möglichkeit, mehr als zwei Moleküle gleichzeitig zu überlagern und betrachtet diese nicht mehr als rigide sondern als flexibel. Es findet eine parallele Geometrieoptimierung für mehrere Moleküle unter Berücksichtigung von Nebenbedingungen, die definierte Atome aus den unterschiedlichen Molekülen mit einer spezifizierten Kraft aufeinander hält, statt. Da es bei den betrachteten PTP1B Inhibitoren nicht sinnvoll erscheint, bestimmte Atome zu überlagern, sondern lipophile Bereiche als Ganzes zu betrachten, ist auch dieses Verfahren nicht die Methode der Wahl. Auch eine Nachoptimierung der mittels FlexS (siehe Abschnitt 4.3.3) gewonnenen Superpositionierungen wurde aus diesem Grund nicht durchgeführt.

4.3.3 Das Programm FlexS zur Überlagerung der Liganden

Das Programm FlexS [148] verfolgt einen ganz anderen Ansatz, sowohl hinsichtlich der Vorgehensweise zum Aufbau der Überlagerung, als auch bei der Methode zur Auswahl der zu überlagernden Molekülteile. Anstatt einzelne Atome aufeinander zu legen, werden Interaktionssphären für Wasserstoffbrücken-Akzeptoren und -Donoren, hydrophile und hydrophobe Wechselwirkungsbereiche sowie das Volumen und die Partialladungen berücksichtigt. Eine vorherige Bestimmung der zu überlagernden Atome oder Interaktionssphären ist nicht notwendig.

Als Basis liegt ein Referenzmolekül rigide vor und ein weiteres Molekül wird flexibel auf diesen Referenzliganden überlagert. Der zu überlagernde Ligand wird an allen Einfachbindungen, an denen kein Wasserstoff beteiligt ist in rigide Teilstrukturen zerlegt, wobei Ringsysteme nicht gespalten werden und somit nur starr betrachtet werden. Sollen bei der Untersuchung auch Ringsysteme flexibel sein, ist es möglich, mit entsprechenden Programmen, vorher alle energetisch günstigen Konformationen für den Ring zu berechnen. Für die in

dieser Arbeit analysierten Liganden, die nur rigide, aromatische Ringe enthalten, war dies allerdings nicht erforderlich.

Zuerst wird für ein oder mehrere ausgewählte Startfragmente nach korrespondierenden Eigenschaften im Referenzmolekül gesucht. Von jeder dieser Startpositionen wird anschließend das Molekül weiter aufgebaut, indem weitere Fragmente unter Berücksichtigung verschiedener Torsionswinkel angehängt werden. Hierbei werden nur bestimmte Torsionswinkel berücksichtigt, die in einer Datenbank abgelegt sind, da nur wenige Geometrien für ein bestimmtes Fragment Strukturen mit niedrigem Energiegehalt liefern. Die verwendeten Torsionswinkel stammen aus Analysen von Kristallstrukturen der Cambridge Structural Database [149] und repräsentieren die dort häufig vorkommenden Torsionswinkel [150]. Durch Anknüpfen weiterer Fragmente wird das Molekül vollständig aufgebaut und damit der Konformationsraum analysiert. Jede Zwischenstruktur wird mittels einer Scoringfunktion [148] hinsichtlich der Interaktionssphären und Volumenüberlappungen bewertet und gegebenenfalls schon während des Aufbaus wieder verworfen. Als Endergebnis erhält man ein Ensemble von Strukturvorschlägen für die Überlagerung mit den dazugehörigen Bewertungen.

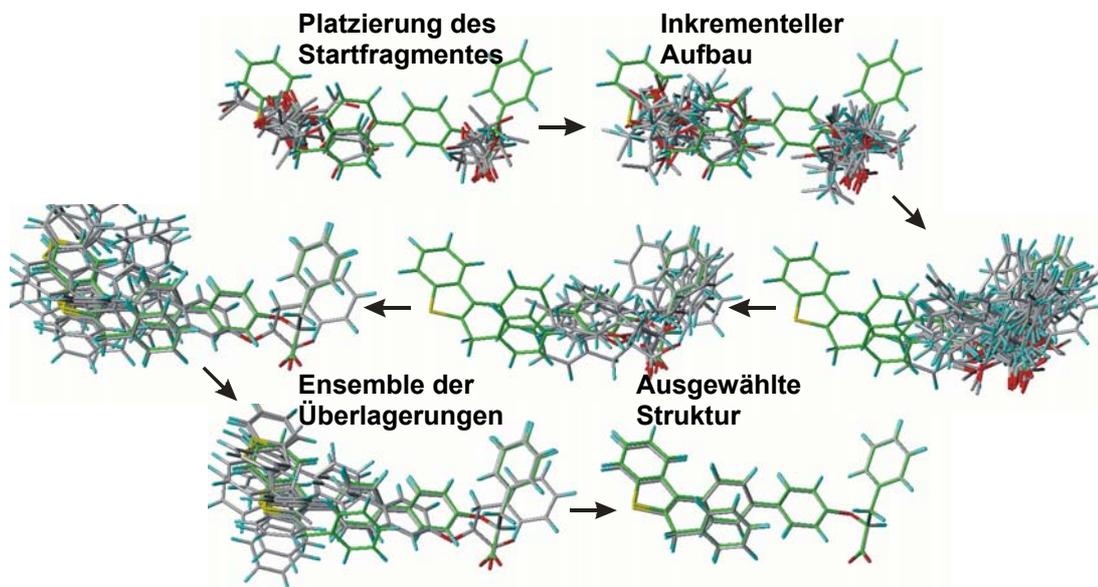


Abbildung 4-2: Überlagerung mittels FlexS

Abbildung 4-2 zeigt im ersten Schritt die Platzierung des Startfragmentes (Carboxylatgruppe mit einer Methylengruppe) für das zwei Bereiche im Templat

gefunden werden. Als nächstes werden eine weitere Methylengruppe und dann der Phenolsauerstoff angehängt (zweite Struktur), wodurch die Anzahl der möglichen Konformationen weiter steigt. Durch den weiteren Aufbau mit den Phenylringen des Biphenyls und der Seitenkette werden die Überlagerungen der Carbonsäure auf das Benzothiophen und viele weitere Konformationen ausgeschlossen. Nach dem Anhängen des Benzothiophen und der Benzylgruppe bleiben mehrere vollständige Überlagerungen übrig (Abbildung 4-2 unten links), von denen die mit dem höchsten Score im letzten Teil der Abbildung 4-2 dargestellt ist.

Für die weitere Optimierung enthält das Programm noch die RigFit-Routine, die es ermöglicht, die Überlagerung der bereits überlagerten Liganden bezüglich ihrer Interaktionspunkte nochmals mit einer rigiden Methode weiter zu optimieren.

4.3.4 Ähnlichkeitsanalyse

Das Programm FeatureTrees [151] ist in der Lage, kleine Moleküle miteinander zu vergleichen. Wegen der hohen Geschwindigkeit von mehr als fünfzig Vergleichen in einer Sekunde ist es geeignet, ganze Datenbanken innerhalb einiger Stunden zu durchsuchen.

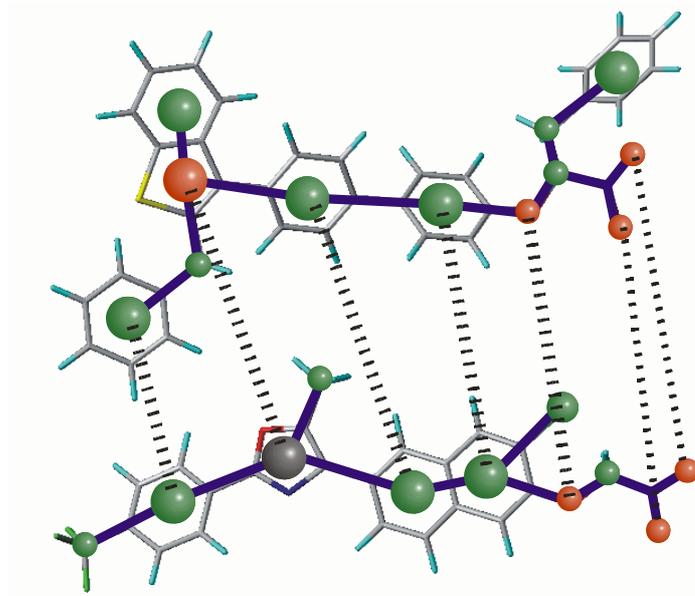


Abbildung 4-3: Strukturvergleich mit FeatureTrees

Die 3D-Struktur eines Moleküls wird dazu in einen Eigenschaftsbaum (Feature Tree) übersetzt, der die Eigenschaften und funktionellen Gruppen an den Knoten des Baumes repräsentiert (Abbildung 4-3). Der topologische Vergleich der Knotenpunkte zweier Moleküle erfolgt unabhängig von den Konformationen und liefert einen Ähnlichkeitswert zwischen Null und Eins. Da nicht die dreidimensionalen Moleküle selbst sondern nur die eindimensionalen Eigenschaftsbäume betrachtet werden, stellt das Verfahren keine Überlagerung im herkömmlichen Sinne dar.

4.3.5 Struktur-Alignment von Proteinen

Die beiden Modellierungspakete Insight/Discover und Sybyl enthalten ein Modul zur Analyse und zum Modellieren von 3D-Strukturen von Proteinen, mit denen eine Überlagerung verschiedener Proteinstrukturen realisiert werden kann. Bei unterschiedlichen Proteinen wird zuerst ein Sequenz-Alignment erzeugt, anschließend werden die homologen Bereiche basierend auf dem Sequenz-Alignment und den α -Kohlenstoffatomen der Proteine superpositioniert. Mit Insight/Discover ist ein so genanntes *Multiple Sequence Alignment*, ein simultanes Alignment für mehrere Strukturen durchführbar, wobei die Sequenzen zunächst alle gemeinsam miteinander verglichen werden und danach überlagert werden. In Sybyl hingegen können zwar auch mehrere Strukturen überlagert werden, dabei wird jedoch jeweils nur eine Sequenz mit einem Templat verglichen und dann das dazugehörige Protein auf die Struktur dieses Templats überlagert.

4.4 Dockingmethoden

Unter Docking versteht man das automatische Einpassen eines Liganden in eine Bindetasche eines Proteins mit einem speziellen Dockingprogramm, und die anschließende Bewertung der gefundenen Komplexe. Bei diesen automatischen Methoden ist ebenso wie beim Superpositionieren zwischen Algorithmen zu unterscheiden, die den Liganden nur rigide verarbeiten und solchen, bei denen der Ligand flexibel ist, was mittlerweile bei allen aktuellen Programmen der Fall ist (Tabelle 4-3). Flexibilität erhöht die benötigte Rechenkapazität in Abhängigkeit der Freiheitsgrade, deshalb verfügen erst wenige Dockingprogramme über die Möglichkeit, auch für das Protein eine meist sehr eingeschränkte Flexibilität zu erlauben. Diese Flexibilität kann dadurch simuliert werden, dass für das Docking mehrere Konformationen des Proteins verwendet werden, was jedoch die Dauer der Berechnungen vervielfacht.

	FlexX	Autodock	Gold
Flexibilität	Ligand flexibel, Protein starr	Ligand flexibel, Protein starr	Ligand flexibel, Protein teilweise flexibel
Docking des Liganden	in Interaktionssphären des Proteins	in Interaktionsfelder des Proteins	direkt in ein Atommodell des Proteins
Verfahren	Inkrementeller Aufbau der Struktur	Monte-Carlo Simulation und Simulated Annealing	Genetischer Algorithmus
Resultat	Ensemble von Strukturvorschlägen mit Bewertung	Ensemble von Strukturvorschlägen mit Bewertung	Ensemble von Strukturvorschlägen mit Bewertung

Tabelle 4-3: Dockingmethoden

4.4.1 FlexX

Das Programm FlexX [152] generiert basierend auf den Koordinaten und Atomtypen von Protein und Ligand und der Definition der Bindungstasche Dockinglösungen mit einem dem Superpositionierungsprogramm FlexS ähnlichen, inkrementbasierten Algorithmus. Wie bei FlexS werden die Liganden fragmentiert, Basisfragmente platziert, die weiteren Fragmente angehängt und die Zwischen- und Endergebnisse mit einer Scoringfunktion bewertet. Die Interaktionssphären stammen aus Analysen der in der PDB (Protein Data Bank) [10] enthaltenen Protein-Ligand-Komplexe. Für das Basisfragment wird versucht

solche Lösungen zu finden, die drei Interaktionen mit dem Rezeptor ermöglichen. Die weiteren Fragmente dürfen nicht mit dem Protein oder anderen Teilen des Liganden überlappen und sollten noch weitere Interaktionsmöglichkeiten hinzufügen. Da das Programm in erster Linie Wasserstoffbrückenbindungen berücksichtigt, bei den betrachteten Inhibitoren aber hauptsächlich apolare Wechselwirkungen zum Enzym auftreten, konnte es trotz des Geschwindigkeitsvorteils gegenüber Autodock (siehe Abschnitt 4.4.2) nicht für die Untersuchung der PTP1B Komplexe verwendet werden.

4.4.2 Autodock 2.4

Mittels einer grundsätzlich anderen Methode arbeitet das Programm Autodock [153]. Die in der Arbeit angewandte Version 2.4 verwendet einen *Monte-Carlo Simulated-Annealing* Algorithmus (MC-SA) zur Konformationssuche.

Zuerst werden für das Protein Affinitätsgitter, so genannte GRID-Felder (siehe 4.5.1, Seite 68), die sich aus den Wechselwirkungsenergien zwischen einem Atom und dem Protein ergeben, in einem dreidimensionalen Gitter berechnet. Diese Berechnung wird für alle in den Liganden enthaltenen Atomtypen durchgeführt. Zusätzlich wird für die elektrostatischen Wechselwirkungen mittels einer positiven Punktladung ein weiteres Gitter berechnet. Daran anschließend wird das eigentliche Docking durchgeführt, bei dem die Liganden flexibel in diese Affinitätsgitter gedockt werden, wobei die Gesamtaffinität die Summe der Wechselwirkungsenergien für die einzelnen Gitter beträgt. Die Affinität eines Atoms zum Protein ergibt sich aus der Interaktion mit den acht umgebenden Gitterpunkten. Der Ligand bewegt sich frei innerhalb der Gitter, die das starre Protein darstellen. Bei der Berechnung werden die Freiheitsgrade des Liganden bei jedem Simulationsschritt einer zufälligen kleinen Veränderung unterzogen (Monte-Carlo Methode), wodurch sich eine neue Geometrie ergibt, für die die Interaktionsenergie ausgerechnet wird (Abbildung 4-4). Diese neue Struktur wird sofort akzeptiert, wenn die Interaktionsenergie geringer ist, eine Lösung mit höherer Energie wird nur dann angenommen, wenn die Energie einen bestimmten Wert, der von der aktuellen Simulationstemperatur abhängt, nicht überschreitet.

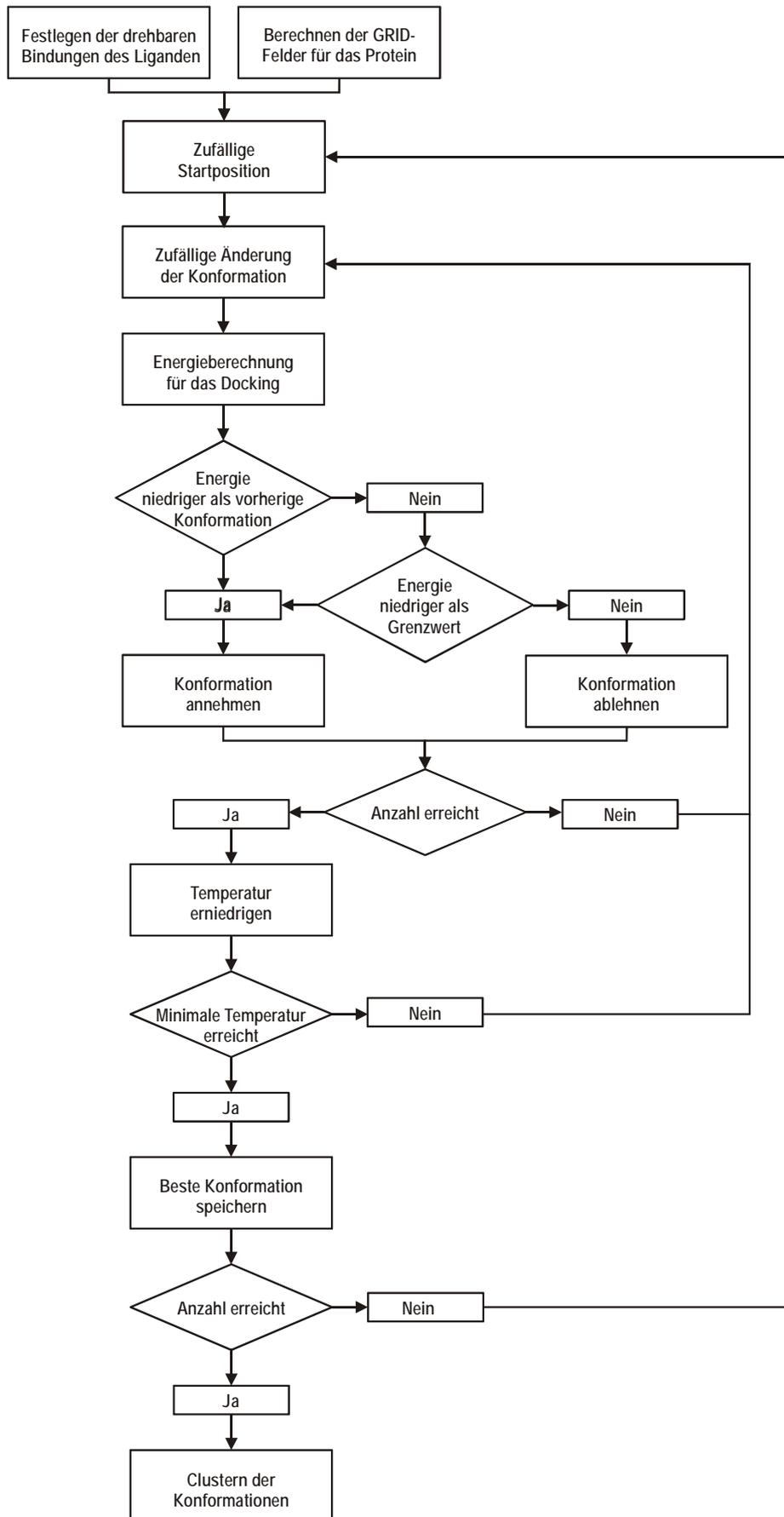


Abbildung 4-4: Ablauf des Dockings mit dem MC-SA in Autodock

Dieser Schritt wird bei konstanter Temperatur solange wiederholt, bis eine vorgegebene Anzahl von akzeptierten oder abgelehnten Lösungen erreicht wird. Dann wird von der günstigsten Konformation ausgehend ein neuer Zyklus mit erniedrigter Temperatur gestartet (Simulated Annealing), bis eine bestimmte Anzahl Zyklen durchlaufen wurde.

Die letzte Konformation wird als Ergebnis eines Durchlaufes gespeichert und ein neuer Durchlauf mit der gleichen Ausgangstemperatur und einer neuen Zufallskonformation gestartet. Die Ergebnisse werden gesammelt bis eine festgelegte Menge an Durchläufen erreicht wird.

Für die Lösungen der einzelnen Durchläufe werden die RMS-Werte berechnet und nach einem vorher festgelegten Maximalwert in Cluster eingeteilt. Aus jedem Cluster wird die Struktur mit der günstigsten Interaktionsenergie zusammen mit der Anzahl der darin enthaltenen Strukturen und der mittleren und der niedrigsten Interaktionsenergie des Clusters ausgegeben. Durch die Interaktionsfelder werden auch Wechselwirkungen zwischen apolaren Molekülfragmenten adäquat berücksichtigt, die in Kraftfeldern durch die Van-der-Waals-Wechselwirkungen beschrieben werden.

Zusätzlich wirkt sich auch die Unabhängigkeit der Ergebnisse von einer korrekten Platzierung eines Startfragmentes positiv aus. Allerdings sind die Rechenzeiten in Abhängigkeit von der Anzahl an Schritten, Zyklen und Durchläufen wesentlich länger als mit FlexX.

4.4.3 Gold

Gold (Genetic Optimization for Ligand Docking) [154][155] ist eine Software, die für die Suche beim Docking einen genetischen Algorithmus (GA) verwendet und es zulässt, ausgewählte Wasserstoffatome des aktiven Zentrums flexibel zu betrachten. Die Torsionswinkel und die Wasserstoffbrücken zwischen Ligand und Rezeptor, für die zu Beginn zufällige Werte ausgewählt werden, werden hierbei in so genannten Genen kodiert. Die Auswahl besitzt einen Bias in Richtung der günstigsten Mitglieder der Gesamtpopulation und die ausgewählten Gene werden mutiert (kleine Veränderungen) oder miteinander

kombiniert. In vielen tausenden Zyklen von Mutationen, Kombinationen und jeweils anschließender Selektion werden die Liganden immer besser gedockt. Dabei werden für die Bewertung (Fitness) in erster Linie Wasserstoffbrücken zwischen Protein und Ligand und auch die innere Energie des Liganden berücksichtigt. Nach mehreren Läufen wird dann der Komplex mit der höchsten Fitness als Ergebnis festgehalten.

4.4.4 Bewertung der gedockten Konformationen

Die Bewertung der mittels Docking generierten Komplexe erfolgt zwar im Dockingprogramm selbst, zusätzlich ist es aber möglich die Komplexe mit weiteren, externen Programmen zu bewerten. In der Arbeit wurde hierzu das Programm X-Score [156] verwendet, das drei unterschiedliche Scoringfunktionen benutzt. Das Programm berechnet jeweils Energieterme für die Van-der-Waals-Interaktionen und die Wasserstoffbrückenbindungen zwischen Ligand und Protein, die Deformation der beteiligten Moleküle und den hydrophoben Effekt, also der Änderung der Entropie durch die Desolvatisierung. Für die ersten drei Terme werden in allen Scoringfunktionen die gleichen Formeln zur Berechnung verwendet, für den hydrophoben Effekt besitzen die drei unterschiedlichen Scoringfunktionen verschiedene Berechnungsgrundlagen. Dabei nutzen die ersten beiden Funktionen den Zusammenhang zwischen dem hydrophoben Effekt auf der einen Seite und der hydrophoben Oberfläche oder den hydrophoben Kontakten zwischen Ligand und Protein auf der anderen Seite. Die dritte Funktion benötigt zur Berechnung zusätzlich den logP-Wert, der mit dem XLogP2-Algorithmus [157] berechnet wird. Die einzelnen Terme der Scoringfunktionen wurden anhand von Protein-Ligand-Komplexen aus der PDB an die Wechselwirkungskonstanten angepasst und mit einem separaten Testsatz validiert. Neben den einzelnen Scores der drei Scoringfunktionen, berechnet das Programm einen Durchschnittsscore (Consensus Score) sowie den logP-Wert.

Daneben wurde für das Screening mit FlexX auch die Software C-Score verwendet, die vier verschiedene Scoringfunktionen enthält: G_Score [155], D_Score [158], PMF_Score [159] und ChemScore [160].

Alle vier Scoringfunktionen betrachten verschiedene Aspekte der Protein-Ligand-Wechselwirkungen und liefern je nach Art der Interaktion unterschiedlich gute Ergebnisse. G_Score betrachtet neben Wasserstoffbrücken die Komplexenergie und die interne Energie des Liganden, während D_Score nur die Ladung und Van-der-Waals-Interaktionen zwischen dem Protein und dem Liganden zur Berechnung verwendet. Für den PMF_Score wiederum wurden basierend auf einer Analyse von Protein-Ligand-Komplexen aus der Protein Data Bank die freien Energien (Helmholtz free energies) für interagierende Protein-Ligand Atompaaire (Potential of Mean Force, PMF) berechnet, die für das Scoring herangezogen werden. ChemScore schließlich enthält Terme für Wasserstoffbrücken, Interaktionen mit Metallatomen, lipophile Kontakte und die Entropie von drehbaren Bindungen.

4.5 3D-QSAR-Analyse

Zur Erstellung von 3D-QSAR-Modellen müssen von den betrachteten Molekülen unter Berücksichtigung ihrer Konformation und Orientierung im dreidimensionalen Raum Eigenschaften berechnet werden. Dies kann in Form von Interaktionsenergien (siehe Abschnitt 4.5.2), die mit der *COMBINE-Methode*, oder Wechselwirkungsfeldern (Abschnitt 4.5.1) die mit dem Programm GRID [161] berechnet werden, geschehen. Für die statistische Analyse dieser Wechselwirkungsfelder und Interaktionsenergien in Bezug auf die biologische Aktivität der einzelnen Substanzen wurde die Software GOLPE (Generating Optimal Linear PLS Estimations) verwendet [162]. Mit diesem Programm sind vergleichende Untersuchungen molekularer Wechselwirkungsfelder möglich ((CoMFA) Comparative Molecular Field Analysis) [163] und es können neben den im Namen erwähnten PLS-Modellen (Partial Least Square) auch PCA-Modelle (Principle Component Analysis) erstellt werden, die mit verschiedenen Methoden wie *Leave one out* (LOO) oder *Random-Groups* validiert werden können. Aus den Wechselwirkungsfeldern resultieren riesige Datenmengen, die mittels verschiedener Verfahren zur Variablenreduktion reduziert werden. Als Eingabe benötigt das Programm die Wechselwirkungsenergien wie sie GRID oder COMBINE liefern (x-Variablen) und die biologische Aktivität oder Affinität der untersuchten Verbindungen (y-Variablen).

4.5.1 Molekulare Wechselwirkungsfelder

Um die Eigenschaften und Wechselwirkungsmöglichkeiten von Molekülen berechnen und dreidimensional visualisieren zu können, wurde das Programm GRID [161] verwendet. Es bedient sich hierzu so genannter molekularer Sonden, auch GRID-Sonden genannt, die die physikochemischen Eigenschaften von Atomtypen beziehungsweise kleinen Atom- und funktionellen Gruppen simulieren. Das zu untersuchende Molekül, von einer kleinen organischen Verbindung bis zum Protein, wird in einen Gitterkäfig gelegt, dessen Größe automatisch (das Molekül umschließend) oder manuell festgelegt werden kann (Abbildung 4-5). Für die Sonde mit den gewünschten Eigenschaften, werden nun an jedem Knotenpunkt des Gitters die Wechselwirkungsenergien berech-

net, die sich aus der Van-der-Waals-Wechselwirkung, der elektrostatischen Energie und einer Wasserstoffbrücken-Bindungsenergie zusammensetzen.

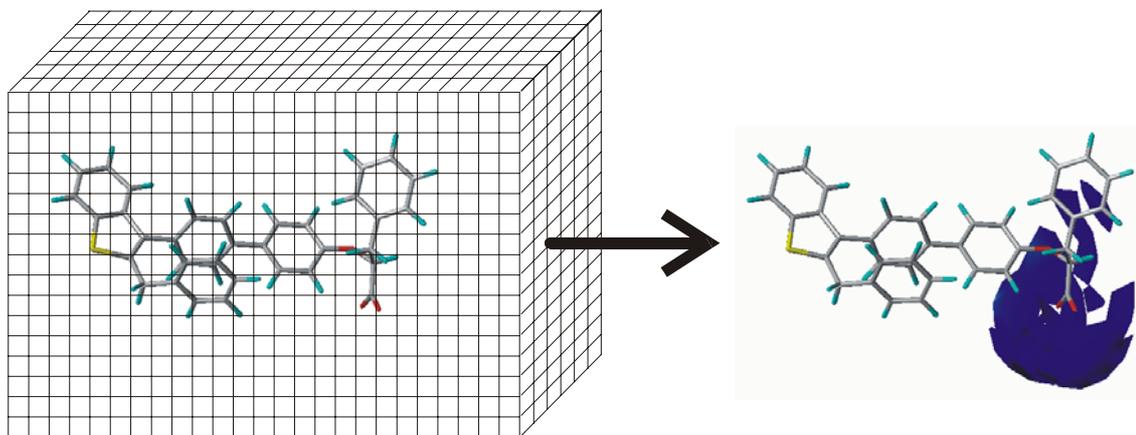


Abbildung 4-5: Molekulare Wechselwirkungsfelder

Die Wechselwirkungsfelder lassen sich als Isokonturlinien für einen bestimmten Betrag in einer Molecular Modelling Software darstellen. Sowohl die für ein Protein als auch die für einen Liganden berechneten Felder liefern oft aussagekräftige Hinweise auf mögliche Protein-Ligand-Interaktionen. So kann man zum Beispiel anhand der GRID-Felder der Bindungstasche eines Proteins bestimmen, welche Eigenschaften die bindenden Liganden aufweisen müssen. Ein visueller Vergleich der Felder für unterschiedliche Moleküle ist oft sehr schwierig und bei einer großen Anzahl an Liganden nicht mehr praktikabel. Mit Hilfe des Programms GOLPE (siehe Abschnitt 4.5) kann aber eine statistische Auswertung der Energiewerte stattfinden.

4.5.2 Molekulare Interaktionsenergien

Wenn die Bindungstasche des Proteins bekannt ist und die Liganden gedockt wurden, lassen sich für eine 3D-QSAR-Analyse neben den Wechselwirkungsfeldern auch die Interaktionsenergien zwischen dem Liganden und den einzelnen Aminosäuren des Proteins in GOLPE einlesen. Für diese als COMBINE [164] bezeichnete Methode werden mit dem Programm Discover nach einer Moleküldynamiksimulation oder Geometrieoptimierung die Wechselwirkungen zwischen einzelnen Resten des Proteins und dem Liganden im CVFF berechnet. Auch diese Energiewerte können mit dem Programm GOLPE (siehe 4.5) statistisch ausgewertet werden.

4.5.3 PCA (Principle Component Analysis)

Bei einer 3D-QSAR Untersuchung besteht der Datensatz im Allgemeinen aus etwa zwanzig bis hundert Verbindungen und jeweils mehreren tausend Wechselwirkungsenergien, aus denen quantitative Beziehungen zwischen Struktur und Wirkung aufzustellen sind. Dabei kann man aus der Betrachtung der Rohdaten der Wechselwirkungsenergien keine Trends oder Muster in der Datenmenge erkennen. Durch eine Hauptkomponentenanalyse (PCA) werden die Informationen aus der Matrix der Wechselwirkungsenergien zu Vektoren zusammengefasst und als Produkt zweier kleinerer Matrices betrachtet. In den so genannten Hauptkomponenten werden die Variablen zusammengefasst, die die Unterschiede am besten erklären, wobei die biologische Aktivität der Verbindungen unberücksichtigt bleibt. Die einzelnen Hauptkomponenten sind jeweils orthogonal zueinander und die erste Hauptkomponente erklärt immer einen größeren Anteil der Varianz als die folgende. Mit einer graphischen Analyse der Hauptkomponenten lassen sich Gruppen ähnlicher Verbindungen finden.

Die Hauptkomponentenanalyse wurde nicht nur durchgeführt, um die Liganden zu analysieren und die GRID-Sonde mit der größten Varianz bezüglich des Datensatzes zu finden, sondern auch, um die verschiedenen Kristallstrukturen der Protein-Tyrosin-Phosphatase 1B untereinander und mit anderen Tyrosin Phosphatasen zu vergleichen. Dazu wurden die Proteinstrukturen zuerst auf eine ausgewählte Templatstruktur überlagert. Anschließend wurden mit dem Programm GRID Interaktionsfelder für alle Proteinstrukturen berechnet. Diese Berechnungen wurden auf die Umgebung des aktiven Zentrums beschränkt. Die Interaktionsfelder wurden mit dem Programm GOLPE mit Hilfe einer PCA ausgewertet.

4.5.4 PLS (Partial Least Square) Analyse

Die PLS-Analyse ist eine Regressionsanalyse, mit deren Hilfe man versucht, den Wert einer abhängigen Variablen (biologische Aktivität, y-Variable) aus den Werten für die unabhängigen Variablen (Interaktionsenergien, x-Variablen) möglichst genau zu berechnen. Die Variablen werden unter besonderer Berück-

sichtigung der y-Variable analog zur PCA organisiert. Mit dem Korrelationskoeffizienten r^2 wird die Qualität des Modells beurteilt, wobei ein hoher Wert ein Modell beschreibt, bei dem die Wechselwirkungsenergien mit der biologischen Aktivität korrelieren.

4.5.5 Validierung der Modelle

Da der r^2 -Wert mit der Anzahl der Hauptkomponenten steigt, ohne dass sich die Aussagekraft des Modells verbessert, bedient man sich zum Testen der Modelle verschiedener Methoden der Kreuzvalidierung. Bei der LOO-Methode wird aus dem Datensatz eine Verbindung herausgenommen, und deren Aktivität mit Hilfe eines Modells aus den verbleibenden Verbindungen vorhergesagt. Dies wird für alle Verbindungen einmal durchgeführt und der Korrelationskoeffizient für die Vorhersage q^2 berechnet, der bei einem Modell, das in der Lage ist, die biologische Aktivität aus den Wechselwirkungsenergien vorherzusagen, nahe bei 1,0 liegen sollte. Ein weiteres Qualitätsmaß ist der SDEP-Wert (Standardabweichung der Fehler der Vorhersagen), bei dem ein Modell mit guten Voraussagen möglichst kleine Werte erreicht. Ein anspruchsvolleres Verfahren zur Kreuzvalidierung, die *Random-Group*-Methode wurde ebenfalls zur Beurteilung herangezogen, wobei jeweils nicht nur eine Verbindung aus dem Modell herausgelassen wird sondern eine bestimmte Anzahl, zum Beispiel zwanzig Prozent aller Verbindungen (L20%O). Diese Methode liefert immer etwas schlechtere, dafür aber aussagekräftigere q^2 - und SDEP-Werte, da der interne Testdatensatz wesentlich größer ist.

4.5.6 Variablenreduktion

Viele der Interaktionsenergien, die als x-Variablen in ein PLS-Modell einfließen, sind für die Unterschiede in der biologischen Aktivität nicht relevant oder bringen nur redundante Informationen ein. Deshalb gibt es in GOLPE verschiedene Verfahren zur Reduktion der Anzahl an Variablen.

Mit der vergleichsweise schnellen Methode 'D-optimal-Selection' werden aus einem bereits aufgestellten PLS-Modell die Variablen entfernt, die den kleinsten Anteil an den Vektoren des Modells haben, also die geringste Information

einbringen. Der Anteil Variablen, die entfernt werden sollen – üblich sind zwischen dreißig und fünfzig Prozent – wird vorgegeben. Anschließend wird ein neues PLS-Modell aufgestellt, und die *D-optimal-Selection* gegebenenfalls wiederholt.

Mit der rechenintensiven Methode des fraktioniert faktoriellen Designs (FFD) werden viele PLS-Modelle aufgebaut, wobei jeweils einige Variablen ausgelassen werden, und mit einer Kreuzvalidierung auf ihre Aussagekraft hin überprüft werden. Da nicht alle Kombinationen von Variablen geprüft werden, spricht man vom fraktionierten faktoriellen Design, im Gegensatz zum (vollständigen) faktoriellen Design, bei dem alle Möglichkeiten analysiert werden. Um die Signifikanz zu bewerten, werden zusätzlich Zufallsvariablen eingestreut. Schließlich werden dann nur noch die Variablen in das endgültige Modell eingeschlossen, die die Vorhersagekraft positiv beeinflussen.

Bevor aber einzelne Variablen mittels FFD gewichtet und ausgeschlossen werden, ist es sinnvoll, die Variablen zu gruppieren. Dazu bietet GOLPE die Methode *Smart Region Definition* [165] an, bei der Variablen aufgrund ihrer statistischen Ähnlichkeit und ihrer räumlichen Nähe in Gruppen zusammengefasst werden, die dann dem FFD unterzogen werden.

4.5.7 CoMSIA (Comparative Molecular Similarity Indices Analysis)

Ausgehend vom Verfahren der CoMFA wurde CoMSIA [166] abgeleitet. Bei dieser Methode werden die Interaktionsfelder nicht mit Hilfe der üblichen Lennard-Jones-Potentiale sondern mittels Gauss-Funktionen berechnet. Dies hat den Vorteil, dass Energiewerte der Felder kontinuierlich sind, da diese nicht an der Oberfläche der Moleküle mit einem Grenzwert belegt werden müssen, damit die Werte nicht ins Unendliche steigen. Die Felder werden für das sterische, elektrostatische und hydrophobe Potential berechnet und um die Potentiale für Wasserstoffbrücken erweitert. Die mit der CoMSIA erhaltenen Ergebnisse sind im Allgemeinen mit den Ergebnissen von CoMFA oder GRID/GOLPE vergleichbar, aber die Felder sind leichter zu interpretieren als die CoMFA-Felder und wurden deshalb auch in dieser Arbeit verwendet.

5 Durchführung und Ergebnisse

5.1 Die Strukturen der PTP1B

Zum Zeitpunkt des Beginns der Arbeit befanden sich in der Proteindatenbank 25 Einträge für Röntgenkristallstrukturen der PTP1B, die in den folgenden Untersuchungen zur Analyse der Proteinkonformationen als Trainingssatz für die Hauptkomponentenanalyse verwendet wurden. Die Anzahl der Strukturen ist mittlerweile auf 51 angewachsen, und die hinzugekommenen Strukturen wurden als Testsatz verwendet. Tabelle 5-1 gibt eine thematisch gegliederte Übersicht über die bisher veröffentlichten Strukturen. Die Auflösung der einzelnen Kristallstrukturen, die zwischen 284 und 321 Aminosäuren der N-terminalen, katalytisch aktiven Domäne enthalten, liegt zwischen 3,1 und 1,6 Å. Alle weisen das gleiche Faltungsmuster auf, und die C-terminale, membranbindende Domäne ist erwartungsgemäß in keiner der Strukturen vorhanden. Die Unterschiede bestehen in den mitkristallisierten Liganden, Mutationen einzelner Aminosäuren und der Oxidation des aktiven Cysteins. Der WPD-Loop des Enzyms existiert in der offenen und der geschlossenen Konformation (Tabelle 5-1).

PDB-ID	Mutation	Konformation	Besonderheiten / Liganden	Quelle
2HNP		offen	freies Enzym	[103]
2HNQ		offen	Wolframat	[103]
1PTV	C215S	geschlossen	pTyr	[8]
1PTY	C215S	geschlossen	jeweils ein BPPM in beiden Bindungsstelle	[98]
1AAX	C215S	geschlossen	jeweils ein pTyr in beiden Bindungsstelle	[98]
1A5Y		geschlossen	Cysteinylphosphat Intermediat	[104]
1I57	C215S		freies Enzym	[100]
1PA1	C215D	geschlossen	freies Enzym	[167]
1JF7		offen	freies Enzym	[168]
1OEM		offen	Oxidation des aktiven Cysteins zur Sulfenylamid	[133]
1OEO		offen	Oxidation des aktiven Cysteins zur Sulfonsäure	[133]
1OES		offen	Oxidation des aktiven Cysteins zur Sulfensäure	[134]
1OEV		offen	Oxidation des aktiven Cysteins zur Sulfonsäure	[134]
1OET		offen	Oxidation des aktiven Cysteins zur Sulfensäure	[134]
1OEU		offen	Oxidation des aktiven Cysteins zur Sulfensäure	[134]
1PTU	C215S	geschlossen	EGFR-Peptid, D-A-D-E-pY-L- NH ₂	[8]
1PTT	C215S	geschlossen	EGFR-Peptid, Ac-D-E-pY-L-NH ₂	[8]
1EEN	C215S	geschlossen	EGFR-Peptid, Ac-D-A-D-Bpa-pY-L-I-P-Q-Q-G	[125]
1EEO	C215S	geschlossen	Consensus-Peptid Ac-I-E-L-E-F-pY-M-D-Y-E-NH ₂	[125]
1G1F	C215A	geschlossen	IR-Peptid, R-D-I-pY-E-T-D-pY-pY-R-K	[56]
1G1G	C215A	geschlossen	IR-Peptid, E-T-D-Y-pY-R-K-G-G-K-G-L-L	[56]
1G1H	C215A	geschlossen	IR-Peptid, E-T-D-pY-pY-R-K-G-G-K-G-L-L	[56]

Tabelle 5-1: Kristallstrukturen der PTP1B

PDB-ID	Mutation	Konformation	Besonderheiten / Liganden	Quelle
1BZJ		geschlossen	DFMP-Derivat	[169]
1BZC		geschlossen	DFMP-Derivat	[169]
1KAV		geschlossen	DFMP-Derivat	[170]
1KAK		geschlossen	DFMP-Derivat	[170]
1Q6J		geschlossen	DFMP-Derivat	[171]
1Q6M		geschlossen	DFMP-Derivat	[171]
1Q6N		geschlossen	DFMP-Derivat	[171]
1Q6P		geschlossen	DFMP-Derivat	[171]
1Q6S		geschlossen	DFMP-Derivat	[171]
1Q6T		geschlossen	DFMP-Derivat	[171]
1C83		geschlossen	OBA-Derivat	[172]
1ECV		geschlossen	OBA-Derivat	[172]
1C84		geschlossen	OBA-Derivat	[172]
1C85		geschlossen	OBA-Derivat	[172]
1C88		geschlossen	OBA-Derivat	[173]
1C87		geschlossen	OBA-Derivat	[173]
1C86		geschlossen	OBA-Derivat	[173]
1GFY		geschlossen	OBA-Derivat	[174]
1L8G		geschlossen	OBA-Derivat	[175]
1NO6		offen	OBA-Derivat	[176]
1NL9		offen	OBA-Derivat in der zweiten Bindungsstelle	[176]
1NNY		offen	OBA-Derivat in der zweiten Bindungsstelle	[176]
1ONY		offen	OBA-Derivat in der zweiten Bindungsstelle	[177]
1ONZ		offen	OBA-Derivat in der zweiten Bindungsstelle	[177]
1Q1M		geschlossen	OBA-Derivat in der zweiten Bindungsstelle	[178]
1NZ7		offen	OBA-Derivat in der zweiten Bindungsstelle	[179]
1PH0		offen	OBA-Derivat mit nur einer Carbonsäure	[180]
1PYN		offen	OBA-Derivat in der zweiten Bindungsstelle	[181]
1QXK		geschlossen	OBA-Derivat mit nur einer Carbonsäure	[182]
1BZH		offen	zyklisches Peptidomimetikum	[169]
1G7F		offen	Peptidomimetikum in beiden Bindungsstellen	[183]
1G7G		geschlossen	Peptidomimetikum in beiden Bindungsstellen	[183]
1JF7		geschlossen	Peptidomimetikum in beiden Bindungsstellen	[168]
1LQF		geschlossen	Peptidomimetikum in beiden Bindungsstellen	[99]
1N6W		geschlossen	Peptidomimetikum in beiden Bindungsstellen	[184]
1PXH		geschlossen	Peptidomimetika, zweite Bindungsstelle	[184]
1NWL	R47C	offen	Inhibitor kovalent an Cys47 gebunden	[185]
1NWE	R47C	geschlossen	Inhibitor kovalent an Cys47 gebunden	[185]
1T48			Inhibitor in der allosterischen Bindungsstelle	[186]
1T49			Inhibitor in der allosterischen Bindungsstelle	[186]
1T4J			Inhibitor in der allosterischen Bindungsstelle	[186]

Tabelle 5-1: Kristallstrukturen der PTP1B (Fortsetzung)

Die beiden ersten, gleichzeitig veröffentlichten Strukturen stammen vom freien Enzym, das ohne einen Liganden kristallisiert wurde und einer PTP1B, die ein Wolframat im aktiven Zentrum enthält. Die nachfolgend veröffentlichten Strukturen stammen von inaktiven Punktmutationen des Enzyms und haben ein oder zwei Moleküle Phosphotyrosin (pTyr) oder Bis-Phosphophenylmethan (BPPM) als Substrat im aktiven Zentrum gebunden. Diese liegen im Gegensatz zu den beiden ersten Strukturen in einer Konformation mit geschlossenem WPD-Loop

vor. Weitere Strukturen des freien Enzyms, die später gewonnen wurden, enthalten weitere Mutationen. Der nächste Teil von Tabelle 5-1 zählt verschiedene Oxidationsstufen des aktiven Cysteins auf, die unter natürlichen oder künstlichen Bedingungen gewonnen wurden. Darauf folgen einige Strukturen des inaktivierten Enzyms mit peptidischen Substraten, die Abschnitte aus vermuteten Zielorten der PTP1B wie dem *Epidermal Growth Factor Receptor* (EGFR) und dem Insulinrezeptor (IR) entsprechen. Für die Substrate wird allgemein angenommen, dass sie an die offene Konformation der PTP1B binden, die sich daraufhin schließt, so dass diese Kristallstrukturen einen geschlossenen WPD-Loop besitzen.

Der größte Teil der in Tabelle 5-1 aufgeführten Kristallstrukturen enthält Inhibitoren, die sich in verschiedene Klassen einteilen lassen: Derivate von Difluormethylenphosphonsäure (DMFP) oder Oxylaminobenzoesäure (OBA) und peptidische Inhibitoren). Die Inhibitoren selbst werden in Abschnitt 5.2 im Einzelnen genauer behandelt. Es fällt auf, dass die meisten Inhibitoren in der geschlossenen Konformation binden, aber einige der neueren OBA-Derivate die offene Konformation bevorzugen. Man nimmt an, dass auch bei den Liganden, die in der geschlossenen Konformation des Enzyms kristallisiert wurden, die Bindung an das Enzym mit geöffnetem WPD-Loop stattfindet und sich das Enzym erst daraufhin schließt, dieses Schließen des Enzyms bei den neueren Inhibitoren aber nicht mehr möglich ist.

Relativ neu sind auch die Inhibitoren, die über eine Disulfid-Brücke an eine in der Nähe des aktiven Zentrums befindliche Aminosäure kovalent gebunden wurden. Über diese künstlich erzeugte Bindung zwischen dem Enzym und möglichen Liganden konnten neue pTyr-Mimetika identifiziert werden.

5.1.1 Analyse der PTP1B-Kristallstrukturen

Die anhand des Proteinrückgrates überlagerten Kristallstrukturen der PTP1B wurden mit dem Programm GRID mittels einer Wasser-Sonde abgetastet und die daraus resultierenden Wechselwirkungs-Felder im Programm GOLPE auf Gemeinsamkeiten und Unterschiede hin untersucht. Anhand der auf den GRID-Feldern basierenden ersten Hauptkomponente kann man problemlos zwischen

den Strukturen mit offener und denen mit geschlossener Konformation differenzieren (siehe Abbildung 5-1), und auch die Strukturen, bei denen ein Inhibitor in der offenen Konformation der Phosphatase gebunden ist, heben sich von den Strukturen des freien Enzyms ab (Abbildung 5-3). Auf den Cluster, der sich etwa in der Mitte der Abbildung befindet und nur Strukturen des Testsatzes enthält, wird später genauer eingegangen.

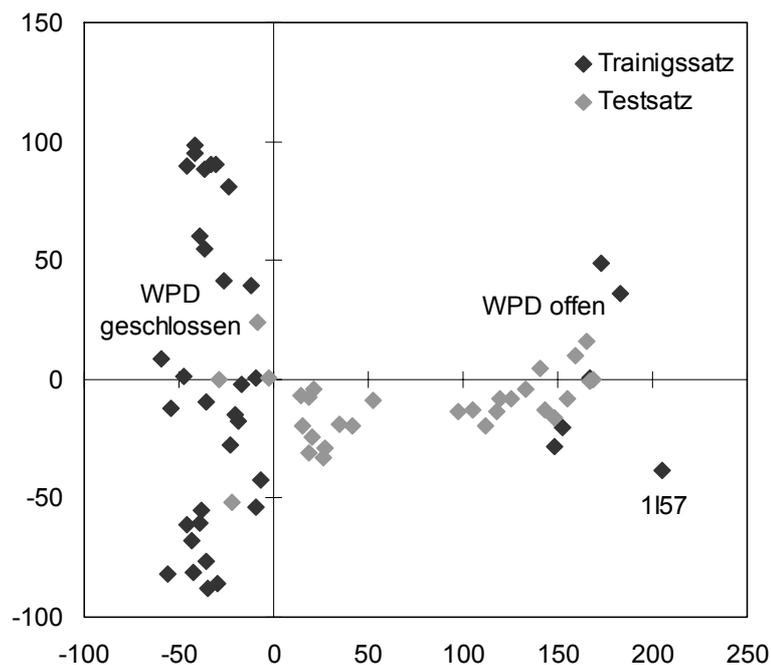


Abbildung 5-1: Vergleich der PTP1B-Strukturen mittels PCA; Trainings- und Testsatz

Die Struktur 1157, die zwar auch einen geöffneten WPD-Loop besitzt, in der aber der P-Loop eine andere Konformation einnimmt, liegt in der PCA in einem geringen Abstand zu den Verbindungen mit offener Konformation. Vermutlich stellt diese Struktur aber ein Artefakt dar, das aus der Mutation des Cysteins und den Kristallisationsbedingungen resultiert und weder physiologisch noch für theoretische Untersuchungen relevant ist. Alle neuen Kristallstrukturen, die erst im Laufe des Jahres 2003 veröffentlicht wurden, konnten auf der Basis der Wechselwirkungsfelder korrekt vorhergesagt und erfolgreich in die oben beschriebenen Gruppen eingeordnet werden.

In Abbildung 5-2 (Seite 78) sind die Bewegungen, die beim öffnen oder schließen des WPD-Loops ablaufen müssen, anhand von zwei Konformationen dargestellt.

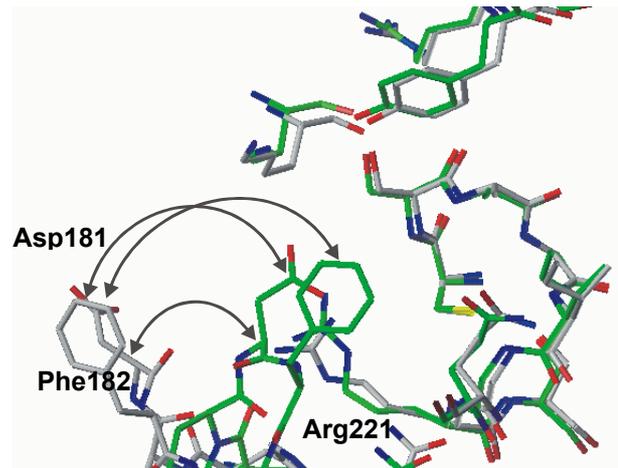


Abbildung 5-2: Vergleich der Aktiven Zentren mit offenem (1BZH, weiß) und geschlossenem (1ECV, grün) WPD-Loop

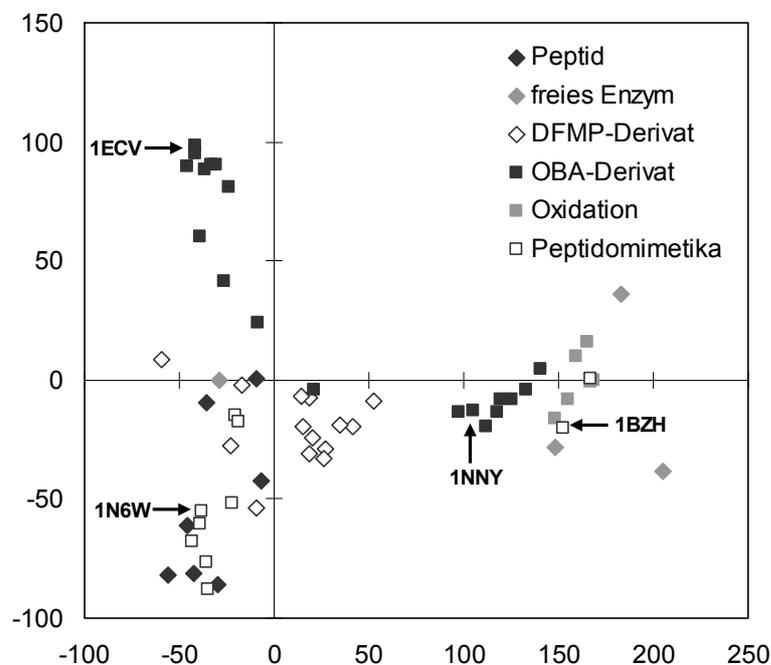


Abbildung 5-3: Einteilung der PTP1B-Strukturen in Subgruppen

Innerhalb der Gruppen von geschlossenem und geöffnetem WPD-Loop sind die Strukturen aber relativ ähnlich, anhand der zweiten Hauptkomponente lassen sich aber in erster Linie die Kristallstrukturen mit geschlossenem WPD-Loop nochmals in Untergruppen einteilen. Eine weitere Untersuchung der Interaktionsfelder zeigt, dass diese von der Art des gebundenen Liganden, der bei den Berechnungen selbst nicht berücksichtigt wurde, abhängig sind (Abbildung 5-3). Die gebundenen peptidischen Substrate unterscheiden sich von den Peptidomimetika und den DFMP-Derivaten sowohl in der PCA als auch in ihrer räum-

lichen Struktur nur geringfügig. Die Kristallstrukturen, die in der PCA-Analyse in den gleichen Bereichen wie diese beiden Strukturen angeordnet wurden, besitzen auch sehr ähnliche Konformationen, andere Komplexe unterscheiden sich aber deutlich. Vergleicht man zum Beispiel wie in der Abbildung 5-4 die Struktur 1ECV, in der ein OBA-Derivat bindet, mit der Struktur 1N6W, in der ein Peptidomimetikum bindet, so erkennt man, dass hier die Seitenkette von Asp181 eine andere Konformation einnimmt.

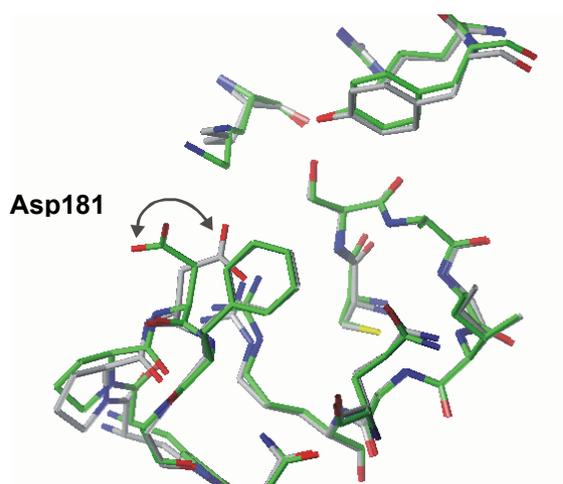


Abbildung 5-4: Vergleich von zwei Strukturen mit geschlossenem WPD-Loop; 1ECV (weiß) und 1N6W (grün)

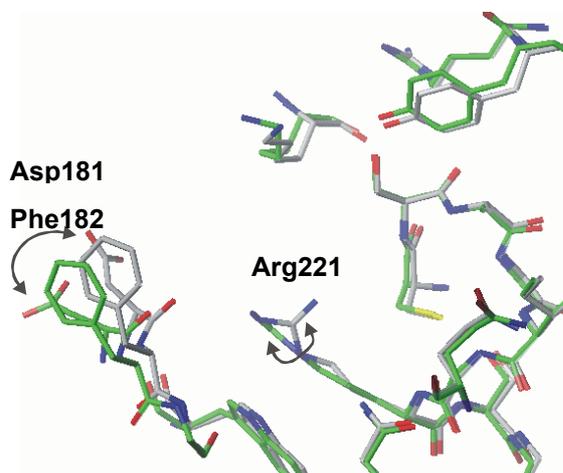


Abbildung 5-5: Vergleich von zwei Strukturen mit offenem WPD-Loop; 1BZH (weiß) und 1NNY (grün)

Auch bei offenem WPD-Loop lassen sich solche Unterschiede finden (Abbildung 5-5). So ist etwa zwischen den Strukturen 1BZH und 1NNY der WPD-Loop geringfügig verschoben, so dass die Asp181 und Phe182 jeweils eine unterschiedliche Lage besitzen. Die Enzymkonformationen, die OBA-Derivate

als Liganden enthalten, besitzen ähnliche GRID-Felder, lassen sich aber in zwei Gruppen unterteilen, wobei die zweite die offene Konformation und die später entwickelten Verbindungen enthält. DFMP-Derivate binden den Röntgenkristallanalysen zufolge sowohl an die offene als auch an die geschlossene Konformation der PTP1B. Die Proteinkonformationen dieser neueren Kristallstrukturen werden in der PCA-Analyse zwar im Bereich der geschlossenen Konformation angesiedelt, befinden sich aber näher an den offenen Strukturen, was mit den experimentellen Befunden über die Bindung der DFMP-Derivate an die PTP1B übereinstimmt [100]. Die Unterschiede in der PCA-Analyse der verschiedenen Kristallstrukturen des ligandfreien Enzyms kommen vor allem durch die Mutation einzelner Aminosäuren zustande.

Die PCA der Interaktionsfelder hat gegenüber den üblichen Analysemethoden, die nur auf einem Clustering der RMS-Werte der überlagerten Proteinstrukturen basieren, verschiedene Vorteile. Zum einen ist es häufig schwierig, die RMS-Werte zu interpretieren, da auch unterschiedliche Konformationen ähnliche RMS-Werte im Vergleich zu einer Templatstruktur besitzen können. Zum anderen ist es mit der PCA möglich, die Strukturen nicht nur in Cluster einzuteilen, sondern die einzelnen Strukturen auch innerhalb der Cluster und die Cluster zueinander anzuordnen. Zusätzlich bietet die PCA eine sehr einfache graphische Darstellung der Ergebnisse. Ein wichtiger Vorteil der PCA ist die Möglichkeit, auch Strukturen mit unterschiedlichen Sequenzen zu vergleichen, was bei anderen Methoden nicht möglich ist, wenn die Seitenketten in die Betrachtung mit einbezogen werden sollen oder wenn die Sequenzen sich erheblich unterscheiden. Außerdem konnte die Analyse so auf die Umgebung des aktiven Zentrums beschränkt werden.

Zum Vergleich wurde auch ein Clustering der Strukturen mit dem Programm NMRClust durchgeführt. Dabei fallen mehr als 60 % aller Strukturen in einem großen Cluster zusammen, der sowohl das Cysteinylphosphat Intermediat als auch Strukturen mit Inhibitoren aus verschiedenen Klassen wie DFMP-Derivate, und OBA-Derivate aber auch peptidische Liganden und die Struktur mit dem pTyr-Substrat enthält. In der PCA erstreckt sich dieser Cluster über den gesamten Bereich der geschlossenen Konformationen, und anhand der RMS-

Werte ist nicht nachvollziehbar, welche Strukturen innerhalb dieses Clusters besonders ähnlich sind (Abbildung 5-6). In drei kleinen Clustern ist jeweils nur eine Gruppe von Strukturen zu finden. Einer enthält einen Teil der DFMP-Derivate, ein anderer die Strukturen mit dem IR-Peptid. Die Cluster befinden sich zwischen den Strukturen des großen, ersten Clusters und sind noch mit einigen Ausreißern, die etwa 15 % der Strukturen ausmachen vermischt. Ein Teil der Ausreißer kann auch nach der PCA als Ausreißer betrachtet werden, man könnte sie aber auch zu einem Cluster mit Strukturen der offenen Konformation zusammenfassen. Der dritte kleine Cluster enthält die anderen beiden Strukturen der offenen Konformation. Im Unterschied zu der PCA ist es anhand der RMS-Werte nicht möglich, die unterschiedlichen Cluster zueinander in Beziehung zu setzen oder zu bestimmen in der Nähe welcher Cluster die Ausreißer liegen.

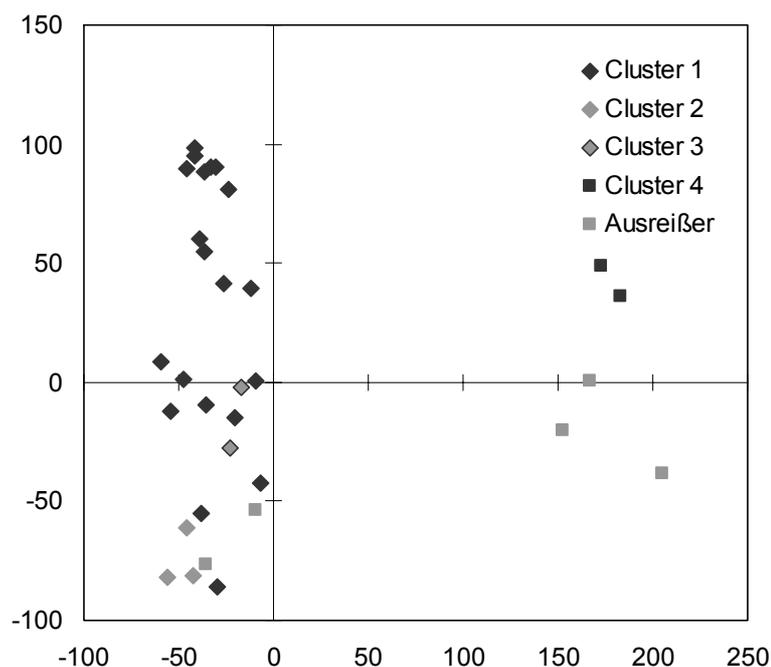


Abbildung 5-6: Vergleich der Ergebnisse aus NMRClust mit der PCA

5.1.2 Vergleich mit anderen Phosphatasen

Neben der PTP1B sind eine Reihe weiterer PTPs bekannt, und für einige von ihnen gibt es bereits eine Röntgenkristallstruktur (Tabelle 5-2). Die Strukturen dieser Phosphatasen wurden ebenfalls durch eine Hauptkomponentenanalyse der GRID-Felder mit denen der PTP1B verglichen. Diese konnten aufgrund ihrer großen Homologie mittels eines sequenzbasierten Alignments auf die bereits superpositionierten Strukturen der PTP1B überlagert werden.

PDB-ID	Enzym	Quelle
1frp	SHP1	[187]
1gwz	SHP1	[188]
1jln	PTP-SL	[189]
1i8k	TCPTP	[128]
1rpm	PTP μ	[190]
1yfo	PTP α	[191]
2shp	SHP2	[192]

Tabelle 5-2: Kristallstrukturen verwandter PTPs

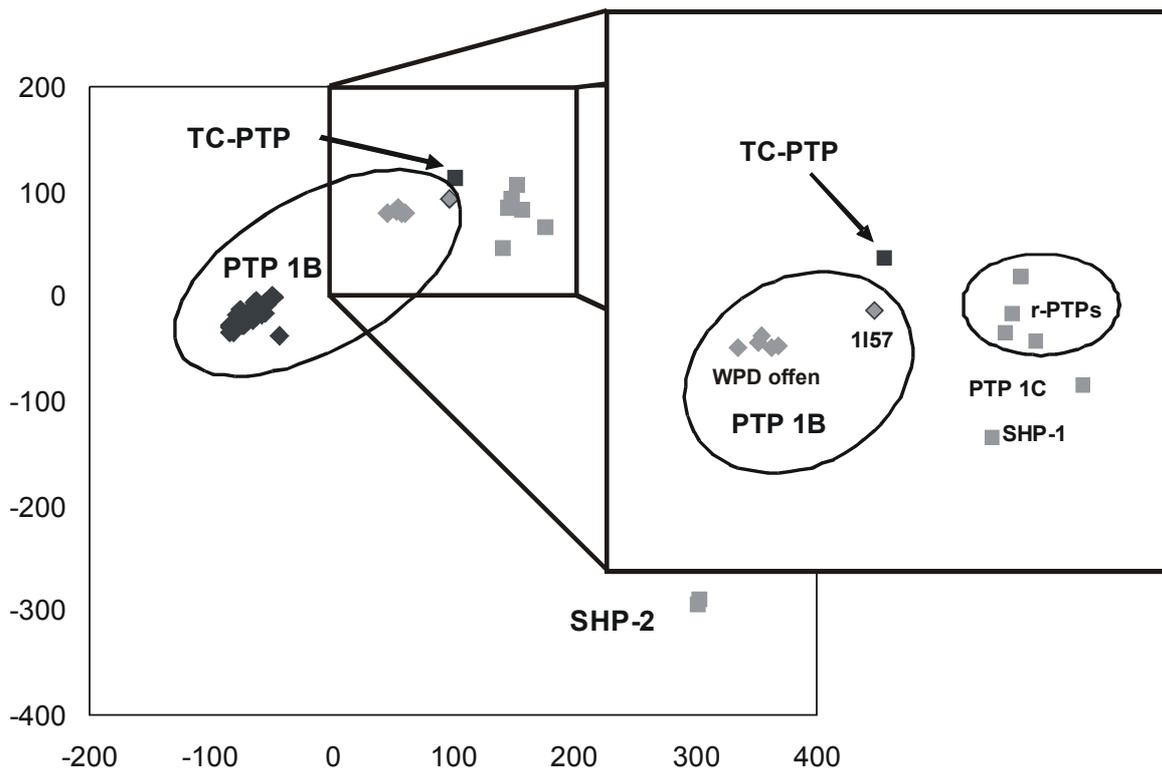


Abbildung 5-7: Vergleich der PTP1B mit anderen PTPs mittels PCA

Wie Abbildung 5-7 zeigt, lassen sich auch bei diesem Vergleich die Strukturen der PTP1B mit offenem und geschlossenem WPD-Loop voneinander unterscheiden. Die PTP1B ist von anderen Phosphatasen klar abgegrenzt, wobei die PTP-Domänen verschiedene Rezeptor-PTPs und die intrazellulären Phosphatasen PTP1C und SHP1 relativ nahe bei der PTP1B zu finden sind und nur die strukturell sehr unterschiedliche Phosphatase SHP2 einen großen Abstand zu den anderen Kristallstrukturen besitzt. Die Tatsache, dass die anderen Phosphatasen näher bei den Strukturen mit geschlossener Konformation zu finden sind, legt nahe, dass diese ebenfalls einen geschlossenen WPD-Loop besitzen, was durch eine visuelle Untersuchung der Strukturen bestätigt wurde.

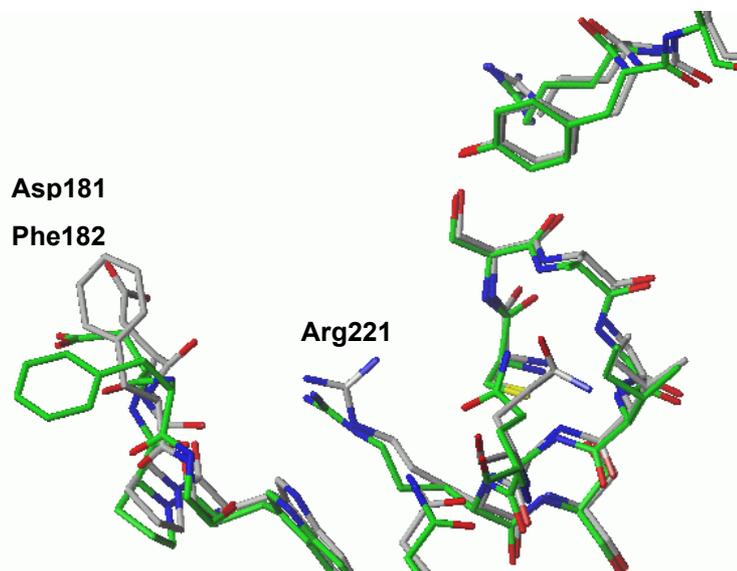


Abbildung 5-8: Vergleich des aktiven Zentrums von PTP1B (weiß) mit TCPTP (grün)

In der Hauptkomponentenanalyse wird ebenfalls die große Ähnlichkeit von PTP1B und TCPTP deutlich (Abbildung 5-7). Dieser nächste Nachbar der PTP1B hat im Diagramm mit den ersten beiden Hauptkomponenten der GRID-Untersuchung einen geringeren Abstand zu den PTP1B-Strukturen mit offenem WPD-Loop als die PTP1B Strukturen mit geschlossenem WPD-Loop. Die kristallisierte Struktur der TCPTP befindet sich wie viele der PTP1B-Strukturen in der offenen Konformation (Abbildung 5-8) und weist im aktiven Zentrum keine nennenswerten Unterschiede zur PTP1B auf. Die Untersuchungen der Wechselwirkungsfelder mittels der Hauptkomponentenanalyse, die sehr deutlich die Unterschiede und Gemeinsamkeiten von Enzymen aufzeigen, wären noch

aussagekräftiger, wenn auch für die anderen Phosphatasen mehrere Kristallstrukturen, die mit unterschiedlichen Bedingungen oder Liganden gewonnen wurden, zur Verfügung stehen würden.

5.1.3 Untersuchung des Ladungszustandes

Da der Ladungszustand der Aminosäuren, die an der enzymatischen Katalyse beteiligt sind, für die Reaktionsgeschwindigkeit mitentscheidend ist, wurden für alle titrierbaren Aminosäuren die Ladungen in Abhängigkeit vom pH-Wert berechnet. Die mit dem Programm UHBD [145] durchgeführten Berechnungen, deren Ergebnisse in Tabelle 5-3 für die wichtigsten Aminosäuren aufgeführt sind, lieferten für mehrere Aminosäuren pK_S -Werte, die von den üblichen Werten für die entsprechenden Seitenketten zum Teil erheblich abweichen.

Aminosäure	pK_S (normal)	pK_S (berechnet)	stabilisiert durch	Lokalisation
Tyr46	9,6	6,2	Lys120	Aktives Zentrum
Glu97	4,4	11,0	Glu101	
Glu101	4,4	8,3	Glu97	
Asp181	4,0	1,5	Lys120	Aktives Zentrum
Cys215	8,3	2,1	P-Loop	Aktives Zentrum

Aminosäure	pK_B (normal)	pK_B (berechnet)	stabilisiert durch	Lokalisation
His214	7,7	11,4	Tyr124	Aktives Zentrum
Arg257	2,0	4,8	Ser70	zweite pTyr-Tasche

Tabelle 5-3: Abweichende pK_S -Werte einzelner Aminosäuren

Neben den beiden Glutamaten Glu97 und Glu101, die aufgrund ihrer räumlichen Nähe nicht beide negativ geladen vorliegen können und somit gegenseitig ihre pK_S -Werte erhöhen, werden die pK_S -Werte verschiedener Aminosäuren durch die Proteinumgebung erniedrigt. Darunter befinden sich Arg257 aus der zweiten pTyr-Bindungsstelle und vier Aminosäuren, die im aktiven Zentrum liegen (Tyr46, Asp181, His214 und Cys215).

Die niedrigeren pK_S -Werte kommen durch eine Stabilisierung der deprotonierten Form durch basische Aminosäuren oder wie im Falle des Cys215 durch die Amidgruppen des Proteinerückgrats zustande. Aus den ermittelten pK_S -Werten kann man schließen, dass das katalytische Cystein deprotoniert vorliegt. In Abbildung 5-9 und Tabelle 5-4 sind die sieben möglichen Wasserstoffbrücken

dargestellt und beschrieben, von denen aber aufgrund der Koordinationszahl des Schwefels maximal fünf gleichzeitig vorhanden sein können. Die Länge der Wasserstoffbrückenbindung, also der Abstand zwischen dem Wasserstoffatom und dem Schwefelatom ist durch den größeren Atomradius des Schwefel größer als bei Wasserstoffbrücken, an denen andere Akzeptoren beteiligt sind.

Aminosäure	Gruppe	Abstand	Winkel
Ser216	Amid	2,3 Å	138°
Gly218	Amid	2,5 Å	128°
Ile219	Amid	2,8 Å	130°
Arg221	Amid	3,1 Å	122°
Ser222	Amid	2,9 Å	157°
Arg221	Guanidinium	3,3 Å	147°
Ser222	Hydroxyl	2,6 Å	166°

Tabelle 5-4: Wasserstoffbrücken zum negativ geladenen Schwefelatom von Cys215

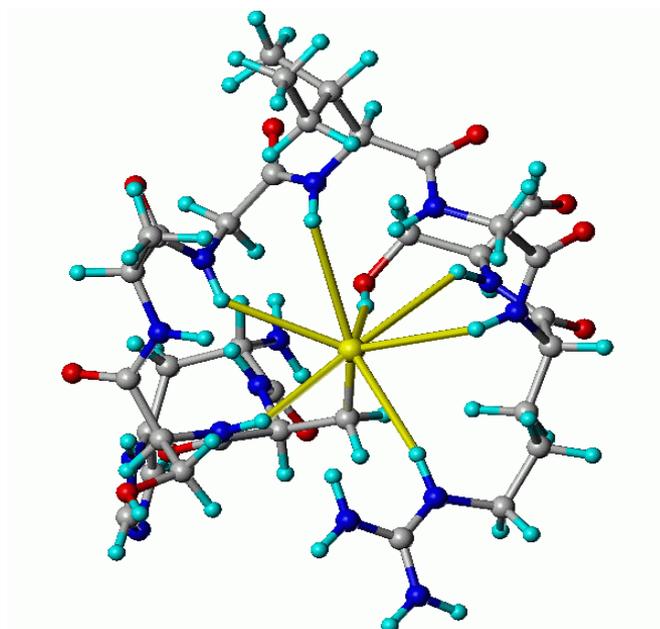


Abbildung 5-9: Stabilisierung der negativen Ladung von Cys215 durch Wasserstoffbrücken (gelb) mit dem P-Loop

In Abbildung 5-10 zeigt der magentafarbene Pfeil das Dipolmoment, das durch die zentrale α -Helix erzeugt wird. Das Dipolmoment verläuft in einer α -Helix parallel zur Achse der Helix. Durch die Ausrichtung der Helix direkt auf den Schwefel von Cys215 im aktiven Zentrum, kann eine zusätzliche Stabilisierung der negativen Ladung durch das Dipolmoment erfolgen.

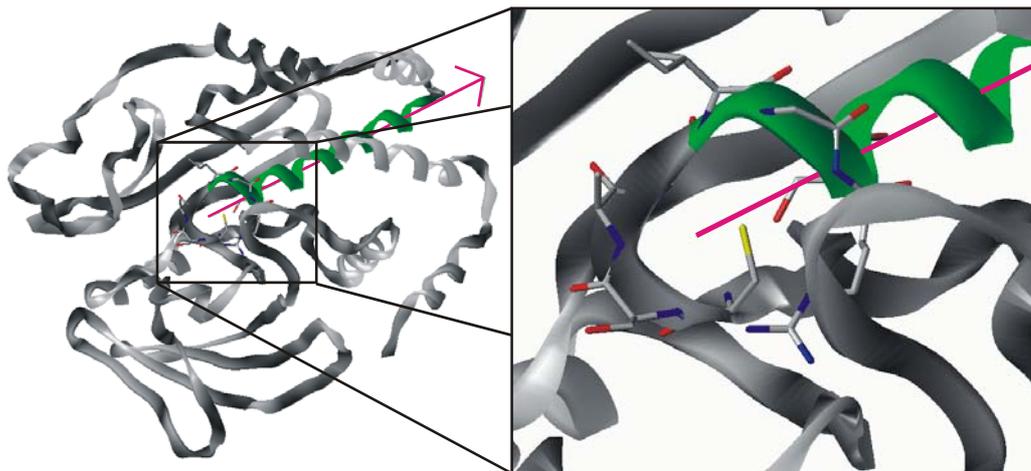


Abbildung 5-10: Stabilisierung der negativen Ladung von Cys215 durch das Dipolmoment der zentralen Helix

Bei den anderen Aminosäuren des aktiven Zentrums und der zweiten Bindungstasche, deren berechnete pK_S -Werte erheblich von den üblichen Werten für die entsprechende Aminosäure abweichen (Tabelle 5-3), führt die Abweichung nicht zu einer Änderung des Protonierungszustandes so dass Tyr46 und Arg257 protoniert und Asp181 deprotoniert vorliegen, sowie His214 in der neutralen Form. Die Protonierung der Aminosäuren außerhalb des aktiven Zentrums wurde nicht analysiert, da diese in den weiteren Berechnungen keine Rolle spielen.

5.2 Die untersuchten Inhibitoren

5.2.1 Bekannte Liganden aus der Literatur

Die PTP1B-Inhibitoren lassen sich in die folgenden acht Klassen einteilen:

- Vanadium-Verbindungen,
- Naturstoffe als Inhibitoren,
- Irreversible, kovalente Inhibitoren,
- Peptidische Inhibitoren,
- Nicht peptidische Inhibitoren
- Inhibitoren mit zwei Phosphonatgruppen,
- Peptidische Inhibitoren ohne Phosphonatgruppe und
- Nicht peptidische Inhibitoren ohne Phosphonatgruppe.

Aufgrund großer struktureller Unterschiede ist es nicht möglich, alle diese Inhibitoren in einer gemeinsamen 3D-QSAR-Analyse zu betrachten. Die in unterschiedlichen Laboren entwickelten und getesteten Inhibitoren wurden außerdem mit unterschiedlichen Testsystemen analysiert, wodurch die Messergebnisse nicht vergleichbar und damit für QSAR-Analysen nicht geeignet sind. Außerdem wurden jeweils verschiedene Werte über ihre Hemmwirkung (IC_{50} , K_i , Ergebnisse aus Tierversuchen) publiziert. Im Folgenden werden die Struktur-Wirkungs-Beziehungen der verschiedenen Klassen von PTP1B-Inhibitoren kurz beschrieben.

5.2.1.1 Vanadium-Verbindungen

Vanadat (VO_4^{3-}) ist ein unspezifischer Phosphataseinhibitor [193], der vermutlich als Analogon des Übergangszustandes fungiert und die PTP1B mit einem K_i -Wert von $0,38 \mu M$ hemmt. Somit können Vanadiumverbindungen als Insulinmimetika verwendet werden [194][195]. Die Bindung von Vanadat im aktiven Zentrum von Phosphatasen wurde durch eine Röntgenkristallstruktur der Yersinia-PTP bestätigt [86][105]. Da anorganisches Vanadat ein unspezifischer Inhibitor ist, schlecht resorbiert wird und bei längerer Anwendung toxische Eigenschaften besitzt, wurden organische Vanadiumverbindungen auf die Verwendbarkeit als Inhibitor der PTP1B untersucht. Zwei Verbindungen, Bis-

(N,N-dimethylhydroxamido)hydroxyoxovanadat [196][197] und Bis(maltolato)-oxovanadium(IV) [198] besitzen eine größere Resorptionsrate und senken stärker den Blutzuckerspiegel im diabetischen Tiermodell als anorganisches Vanadium. Dabei scheint aber trotzdem anorganisches Vanadat, das aus den beiden metallorganischen Verbindungen freigesetzt wird, die aktive Substanz zu sein [199].

Komplexe aus Vanadium(V) und Wasserstoffperoxid, so genannte Pervanadate, oxidieren irreversibel das katalytische Cystein der PTP1B zur Cysteinsäure, erhöhen die Menge an phosphorylierten Proteinen und wirken als potentere Insulinmimetika im Vergleich zu reinem Vanadat [195][200].

5.2.1.2 Naturstoffe als Inhibitoren

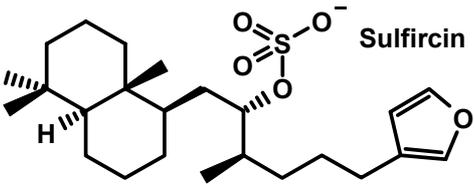
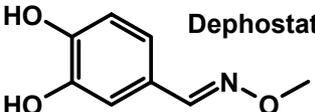
Verbindung	Literatur	Aktivität gegen PTP1B
 <p>Sulfircin</p>	[201]	IC ₅₀ = 29,8 µM
 <p>Dephostatin</p>	[202]	IC ₅₀ = 3 µM

Tabelle 5-5: Naturstoffe als Inhibitoren

Sulfircin (Tabelle 5-5) aus einem Wasser-Schwamm und dessen synthetische Analoga hemmen neben anderen PTPs auch die PTP1B (IC₅₀ = 29,8 µM) [201]. Auch das aus einer Streptomyces-Art isolierte Dephostatin (Tabelle 5-5) und vor allem das Analogon Et-3,4-dephostatin hemmt die PTP1B (IC₅₀ = 580 nM) aber nicht die Phosphatasen CD45 und LAR. Es wurde ein nitrosaminfreies Derivat hergestellt, das die potentiellen mutagenen und karzinogenen Effekte von Nitrosaminen ausschließt. Es besitzt etwa fünfmal größere IC₅₀-Werte für PTP1B und SH-PTP-1 als die Ausgangsverbindung und zeigt in Tierversuchen antidiabetische Aktivität [202][203][204].

Aus der asiatischen Pflanze *Broussonetia papyrifera* (Papiermaulbeerbaum) wurden mehrere Flavonoide isoliert, die die PTP1B in mikromolaren Konzentrationen hemmen [205].

5.2.1.3 Irreversible und kovalent bindende Inhibitoren

Über das Cystein des aktiven Zentrums reagieren PTPs mit elektrophilen Verbindungen wie α -halogenierten Ketonen und Michael-Akzeptoren (Tabelle 5-6) zu kovalenten Addukten [206][207][208][209].

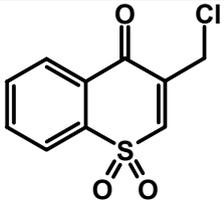
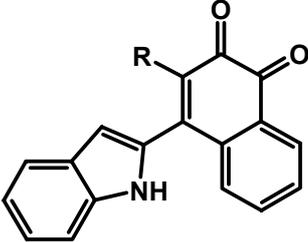
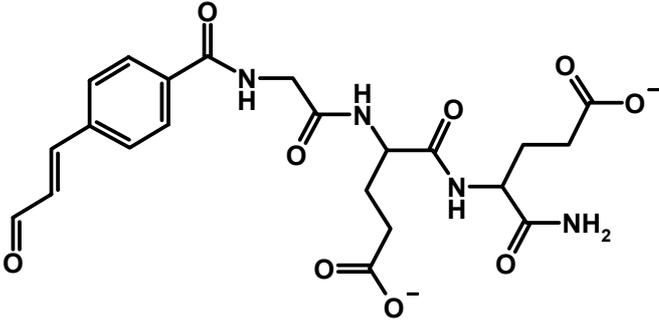
Verbindung	Literatur	Aktivität gegen PTP1B
	[207]	$K_i = 3,5 \mu\text{M}$
	[208]	$IC_{50} = 270 \text{ nM}$
	[209]	$K_i = 5,42 \mu\text{M}$

Tabelle 5-6: Irreversible und kovalente Inhibitoren

Auch über die Oxidation des aktiven Cysteins ist eine irreversible Inhibition der Phosphatasen möglich. Neben den schon unter 5.2.1.1 erwähnten Pervanadaten sind hierzu auch S-Nitrosothiole [210] und Peroxynitrit (ONOO^-) [211] in der Lage.

5.2.1.4 Peptidische Inhibitoren

Frühe Untersuchungen zum Design von PTP1B Inhibitoren hatten zum Ziel, Phosphotyrosin (pTyr) in hoch affinen peptidischen Substraten durch nicht hydrolysierbare pTyr-Mimetika zu ersetzen. Dieser Ansatz führte zum Einbau von Phosphonomethyl-D/L-phenylalanin (PMP) in verschiedene Peptide, die der Autophosphorylierungs-Stelle der IR-Kinase (1155 bis 1165) und des EGFR

analog sind. Aus dem daraus abgeleiteten Hexapeptid Ac-D-A-D-E-Pmp-L-NH₂ wurde das α,α -difluoro Analogon, Phosphonodifluoromethyl-L-phenylalanin (F₂Pmp), mit einem K_i-Wert von 100 nM, entwickelt [212][213].

Mittels Screening von Peptidbibliotheken wurden neben anderen hoch affinen Inhibitoren, das Peptid E-E-D-E-F₂Pmp-M mit einem K_i-Wert von 7,2 nM identifiziert [214]. Durch anhängen einer Pentadecansäure an das Peptid F₂Pmp-E- F₂Pmp-NH₂ konnte ein zellpermeabler Inhibitor entwickelt werden, der die Insulinsensitivität erhöht und zusätzlich insulinmimetisch wirkt [215].

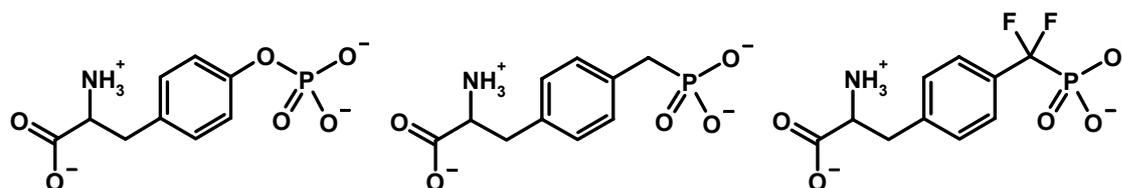


Abbildung 5-11: Phosphotyrosin (pTyr) und pTyr-Mimetika Pmp und F₂Pmp

5.2.1.5 Nicht peptidische Inhibitoren

Die Suche nach nicht peptidischen Inhibitoren begann bei relativ einfachen Phenylderivaten, die verschiedene Phosphonatgruppen enthielten, aber nur eine sehr geringe Aktivität besaßen oder inaktiv waren. Nur einige Naphthylverbindungen mit Phosphonodifluoromethyl als Phosphonatgruppe hatten eine nennenswerte Aktivität (Tabelle 5-7) [216][217][218].

Im Gegensatz zu den fluorierten Naphthylphosphonaten, besitzen nur wenige der α -fluorierten Phenylphosphonate eine nennenswerte Aktivität. Von den untersuchten meta-substituierten Verbindungen sind nur diejenigen aktiv, die Kation- π -Wechselwirkungen mit Lys116 und Lys120 eingehen können (Tabelle 5-7) [219]. Bei den para-substituierten Verbindungen konnte die Aktivität durch die Einführung weiterer saurer Gruppen in die Arylphosphonate, die mit der basischen Seitenkette von Arg47 nahe dem aktiven Zentrum interagieren sollten, erhöht werden [220]. Mit Hilfe von Erweiterungen an den relativ kleinen Substraten konnte auch eine weitere Bindungsstelle in der Nähe von Arg47 adressiert werden (Tabelle 5-7) [221]. Die folgende Entwicklung dieser Verbindungen führte zu Substanzen mit einer Selektivität von 7,2 für PTP1B gegenüber TCPTP bei einem IC₅₀-Wert von 5,0 nM (PDB-Einträge 1Q6J, 1Q6M, 1Q6N, 1Q6P, 1Q6S und 1Q6T) [171][222][223][224][225][227].

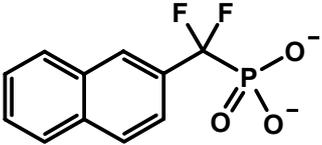
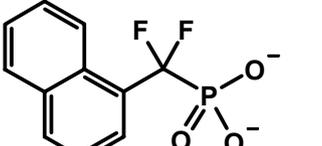
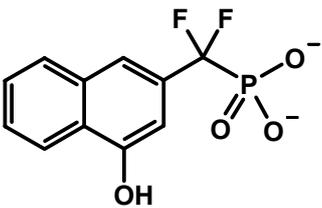
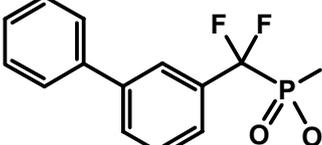
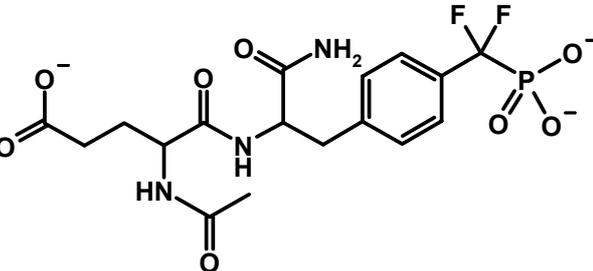
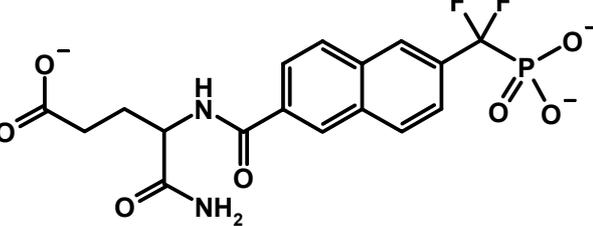
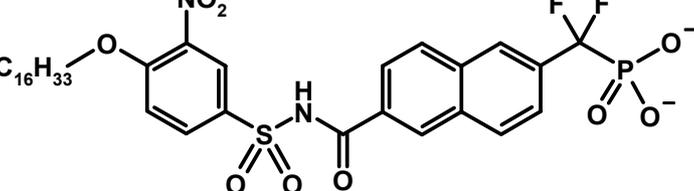
Verbindung	Literatur	Aktivität gegen PTP1B
	[216]	91 % Inhibition bei 200 μM
	[216]	88 % Inhibition bei 200 μM $K_i = 179 \pm 25 \mu\text{M}$
	[218]	$K_i = 93 \pm 8 \mu\text{M}$
	[219]	94 % Inhibition bei 500 μM
	[220]	$K_i = 2,0 \pm 0,14 \mu\text{M}$
	[220]	$K_i = 12,4 \pm 0,9 \mu\text{M}$
	[221]	$\text{IC}_{50} = 0,35 \mu\text{M}$

Tabelle 5-7: Nicht peptidische Inhibitoren

5.2.1.6 Inhibitoren mit zwei Phosphonatgruppen

Neben der katalytischen Arylphosphat-Bindungsstelle besitzt die PTP1B eine zweite niedrig affine Bindungsstelle für Arylphosphate, die zwar keine katalytische Aktivität besitzt aber für die effiziente Erkennung des Substrates notwendig ist [8][124][125].

Diese wurde durch die Kokristallisation und Röntgenstrukturanalyse von katalytisch inaktiver PTP1B mit Bis-(Phosphophenyl)methan (BPPM, $K_m = 16 \mu\text{M}$) oder pTyr ($K_m = 53 \text{ mM}$) in gesättigter Lösung entdeckt [98].

Auch schon vor der Entdeckung der zweiten Arylphosphat-Bindungsstelle hatte man festgestellt, dass bestimmte Verbindungen wie Fluoresceindiphosphat und BPPM, die zwei Phosphatgruppen enthalten, hinsichtlich der K_m -Werte bessere Substrate für die PTP1B sind als ihre analogen Monophosphate [229]. Es wurde eine Serie von Arylderivaten entwickelt, die mit zwei Difluormethylenphosphonsäure-Gruppen (DFMP) substituiert sind (Tabelle 5-8) [219][230].

Molecular Modelling Untersuchungen an den Verbindungen zeigen zwar, dass der Abstand zwischen den Phosphonatgruppen bei einem Spacer aus vier Methylengruppen groß genug ist, um beide Bindungsstellen zu besetzen, aber die IC_{50} -Werte und die Kristallstrukturen der unterschiedlichen Verbindungen belegen, dass dies nicht der Fall ist [170][231][233] (PDB-Einträge 1KAK und 1KAV).

Der aus dem Aktivierungs-Loop des IR stammende Peptidabschnitt E-T-D-pY-pY-R-K-G-G-K-G-G-L (Aminosäuren 1156-1171) besetzt hingegen beide Bindungsstellen und besitzt eine 70-fach größere Affinität zur inaktiven PTP1B-D181A als das analoge Peptid, das nur das zweite pTyr enthält. Während pTyr1162 im katalytischen Zentrum bindet, besetzt pTyr1163 die zweite Phosphotyrosin-Bindungsstelle mit den wichtigen Aminosäuren Arg24 und Arg254 [56]. Um auszuschließen, dass bei dem ebenfalls aktiven Tripeptid Glu-F2Pmp-F2Pmp-NH₂ nur die Gesamtladung für die Aktivität verantwortlich ist, wurde auch das Peptid F2Pmp-Glu-F2Pmp-NH₂ mit vertauschten Aminosäuren getestet. Der zehnfach höhere K_i -Wert zeigt, dass die Sequenz und nicht die Ladung des Peptides ausschlaggebend ist [234].

Daneben wurden etliche weitere peptidische und nicht peptidische Verbindungen, die zwei Phosphonatgruppen enthalten getestet [99][124][184][214][232][233][235][236][237][238] von denen nur die aktivsten Verbindungen am Ende von Tabelle 5-8 abgebildet sind.

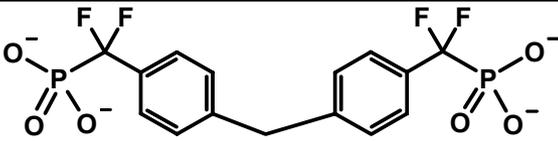
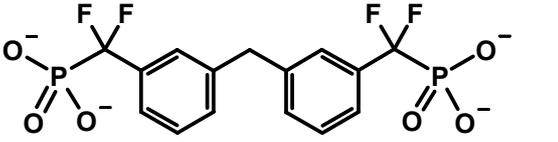
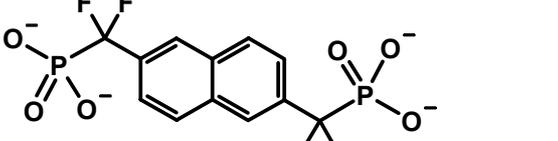
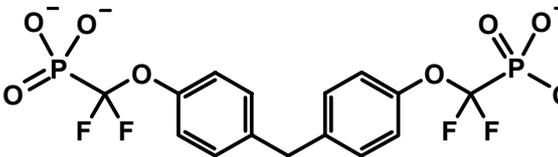
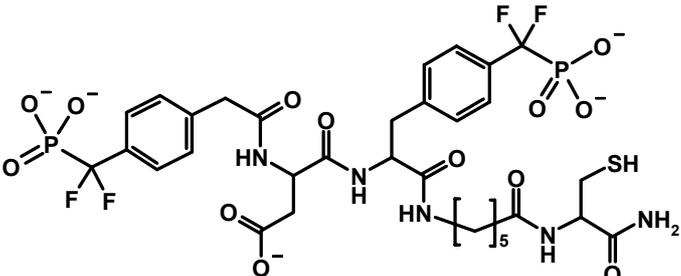
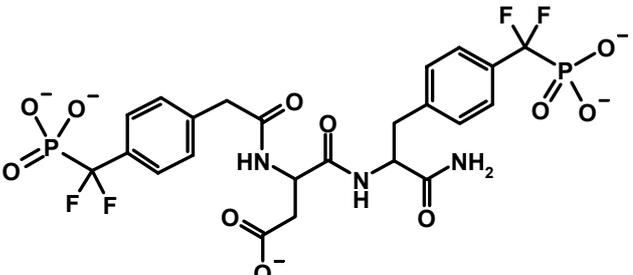
Verbindung	Literatur	Aktivität gegen PTP1B
	[219]	n 1 $IC_{50} = 23 \pm 2,0 \mu M$ 2 $6,8 \pm 0,2 \mu M$ 3 $5,8 \pm 0,3 \mu M$ 4 $4,8 \pm 0,4 \mu M$
	[219]	$IC_{50} = 45 \pm 1 \mu M$
	[230]	$IC_{50} = 26 \pm 3 \mu M$
	[230]	$IC_{50} = 29 \pm 4 \mu M$
	[231]	$K_i = 47 \pm 5 \mu M$
	[184]	$K_i = 0,0018 \mu M$
	[232]	$K_i = 0,0024 \mu M$

Tabelle 5-8: Inhibitoren mit zwei Phosphonatgruppen

5.2.1.7 Peptidische Inhibitoren ohne Phosphatgruppe

Die α,α -Difluoromethylenphosphat-Gruppe ist zwar ein sehr effektives Mimetikum für die Phosphatgruppe und wurde dazu vielfach bei der Entwicklung von Inhibitoren der PTP1B verwendet, sie führt aber zu einer schlechten Bioverfügbarkeit der Substanzen. Deshalb wurde versucht, Verbindungen zu entwickeln, die keine stark geladene Phosphatgruppe mehr enthalten.

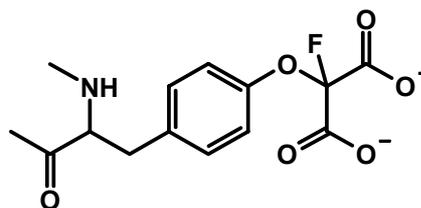


Abbildung 5-12: FOMT

Dass eine Malonylgruppe wie in 4'-O-(2-Malonyl)-L-tyrosin (OMT) oder eine Fluor-Malonylgruppe (FOMT) als nicht hydrolysierbarer Ersatz für die Phosphatgruppe dienen kann, wurde mit den in Tabelle 5-9 dargestellten linearen und zyklischen Peptiden belegt [217][239][240][241].

Verbindung	Literatur
	[239]
	[241]
	[241]
	[241]

Tabelle 5-9: Peptidische, FOMT basierte Inhibitoren

Die Röntgenkristallstruktur des Komplexes mit dem zyklischen Peptid (PDB-Eintrag 1BZH), in der das Enzym die offene Konformation für den WPD-Loop einnimmt, zeigt, dass für die Bindung eines Liganden das Schließen des WPD-Loops nicht notwendig ist. Durch die zweite Carboxylatgruppe des sperrigen Malonylrestes wird ein Wassermolekül aus dem aktiven Zentrum verdrängt, und es werden zusätzliche Wasserstoffbrücken zwischen Ligand und Rezeptor gebildet [169].

Die Sulfatgruppe (Sulfotyrosin, sTyr, sY) stellt ein weiteres Mimetikum für die Phosphatgruppe dar, und wurde im Dodecapeptid (T-R-I-sY-E-T-D-sY-sY-R-K), das der Sequenz 1132-1153 des IR entspricht, und einer Vielzahl kürzerer sTyr enthaltender Peptide untersucht wurde [242][243][244][245]. In Analogie zu den F2Pmp enthaltenden Peptiden wurden deshalb auch vergleichbare Verbindungen mit Sulfonodifluoromethyl-L-phenylalanin (F2Smp) ihre Aktivität hin untersucht. Diese sind aber weniger wirksam als die entsprechenden F2Pmp-Derivate [223].

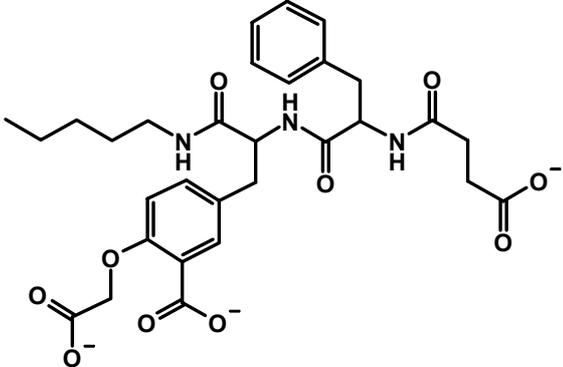
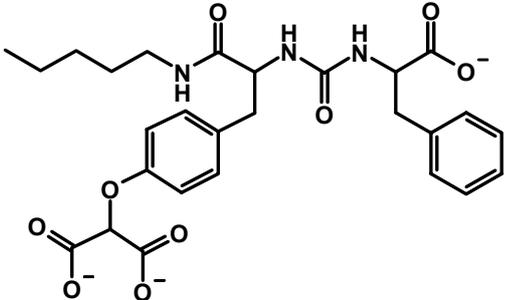
Verbindung	Literatur	Aktivität gegen PTP1B
	[183]	$K_i = 0,25 \mu\text{M}$
	[183]	$K_i = 3,4 \mu\text{M}$

Tabelle 5-10: Peptidomimetische Inhibitoren

Daneben wurden viele weitere funktionelle Gruppen als nicht hydrolysierbare pTyr-Mimetika in Peptide eingebaut und besitzen eine geringe oder mäßige Aktivität [246][247][248][249][250][251][252]. Besonders aktiv waren vor allem die Salicylatgruppe [168] und 3-Carboxy-4-(O-carboxymethyl)-tyrosin (Tabelle 5-10) das in einem peptidomimetischen Inhibitor als PTP1B-Komplex kristallisiert werden konnte (PDB-Einträge 1G7G, 1G7F und 1JF7) [183].

5.2.1.8 Nicht peptidische Inhibitoren ohne Phosphatgruppe

Die größten Anstrengungen auf dem Gebiet der Entwicklung von PTP1B-Inhibitoren wurde im Bereich der nicht peptidischen, nicht phosphathaltigen Inhibitoren unternommen, mit dem Ziel, oral verfügbare, hoch aktive Arzneistoffe zur Behandlung von Typ-2-Diabetes und Adipositas zu erhalten. Das Spektrum der publizierten Substanzen reicht von relativ einfachen Naphthyl- und Biphenylverbindungen mit α -fluorierten Phosphatmimetika, wie die α,α -Difluorotetrazolat-Gruppe oder die α,α -Difluorocarboxylat-Gruppe [223][226][253] über die großen Gruppen der Oxylaminobenzoessäurederivate und den vom Urikosurikum Benzbromaron oder den PPAR γ -Agonisten abgeleiteten Verbindungen zu einer Reihe von Substanzen, die aus dem virtuellem Screening oder dem High-Throughput-Screening stammen.

Das relativ kleine Molekül 2-(Oxylamino)benzoessäure und seine heterozyklische Derivate, die in der geschlossenen Konformation des Enzyms binden, bilden mit der o-Carboxylatgruppe eine Salzbrücke zu Lys120 und Wasserstoffbrücken zu Asp181 und Tyr46. Außerdem werden Wechselwirkungen zwischen dem Oxylaminorest und dem P-Loop (His214 bis Ser222) beobachtet (PDB-Einträge 1C83 bis 1C88 und 1ECV) [172]. Nach der Optimierung dieser Leitstruktur erhielt man selektivere Substanzen, die an die Aminosäure Gly259 binden. Da diese Aminosäure bei anderen PTPs durch eine größere Seitenkette ersetzt ist, ist die Bindung dort nicht möglich (PDB-Eintrag 1L8G) [122][175]. Im Folgenden entstanden Inhibitoren mit einem K_i von 18 nM und relativ niedrig affine Verbindungen, die aber eine 23-fachen Selektivität für PTP1B gegenüber TCPTP besitzen (PDB-Einträge 1NL9, 1NNY, 1NO6, 1NZ7, 1ONY, 1ONZ und 1PH0, 1PYN, 1QXK) [176][177][178][179][180][181][182][254].

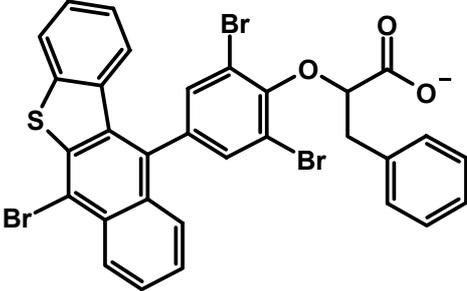
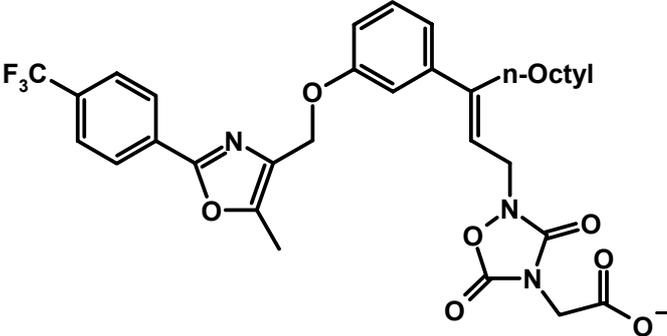
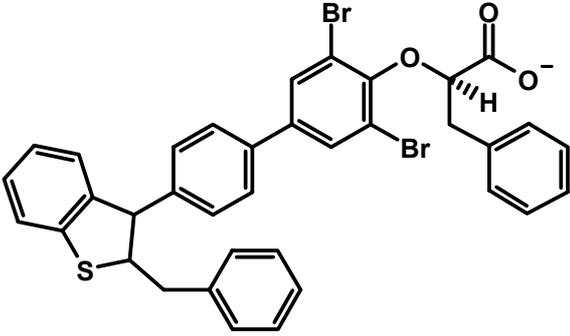
Verbindung	Literatur	Aktivität gegen PTP1B
	[255]	IC ₅₀ = 83 nM
	[257]	IC ₅₀ = 12 nM
	[258]	IC ₅₀ = 25 nM

Tabelle 5-11: Inhibitoren der Firma Wyeth-Ayerst Research

Ausgehend von dem Urikosurikum Benzbromaron, das eine gute Resorption und Pharmakokinetik besitzt, wurden Arylobenzo[b]naphthothiophene, Arylobenzo[b]naphthofurane und Thiacyclopenta[b]fluorene mit einem Oxoacetatrest als potente PTP1B-Inhibitoren entwickelt. Ein Beispiel für diese Verbindungen ist in Tabelle 5-11 dargestellt [255][256]. Innerhalb einer großen Gruppe von Azolidindionen, die in der gleichen Arbeitsgruppe synthetisiert und auf ihre PTP1B-Inhibition untersucht wurden, hatten die aktivsten Verbindungen IC₅₀-Werte zwischen 120 und 300 nM (Tabelle 5-11) [257]. Ein großer Datensatz mit Benzofuran- und Benzothiophen-Biphenylen, die IC₅₀-Werte bis zu etwa 25 nM

besitzen [258], diente als Gegenstand der vorliegenden Untersuchungen und wird im Abschnitt 5.2.3 noch genauer besprochen. Daneben wurde eine Reihe von 1,2,4-Oxadiazolidin-3,5-dionen synthetisiert, für die keine Ergebnisse aus *in vitro*-Untersuchungen sondern nur die Aktivitäten in Tierversuch veröffentlicht wurden, so dass diese Daten nicht in die QSAR-Analysen aufgenommen werden können [259].

Unter Verwendung von Docking Programmen wurden in einem theoretischen Ansatz aus einer Datenbank mit 150.000 kommerziell erhältlichen Substanzen 2.000 Verbindungen herausgefiltert, die nach elektrostatischen und geometrischen Eigenschaften komplementär zum aktiven Zentrum der PTP1B sind. Aus diesen 2.000 am besten bewerteten Verbindungen wurden unter Berücksichtigung von chemischer Diversität, vorhergesagtem Bindungsmodus, Löslichkeit, Stabilität, kommerzieller Verfügbarkeit und Beschaffungskosten 25 Substanzen ausgewählt. Viele dieser Verbindungen besitzen wie die in Tabelle 5-12 Ki-Werte zwischen 20 μM und 100 μM [260].

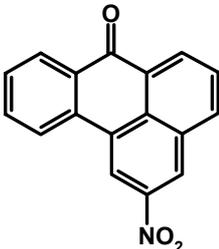
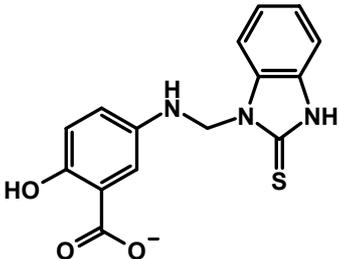
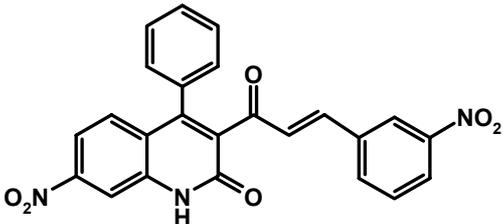
Verbindung	Literatur	Aktivität gegen PTP1B
	[260]	$K_i = 21 \pm 5$
	[260]	$K_i = 61 \pm 3$
	[260]	$K_i = 54 \pm 6$

Tabelle 5-12: Inhibitoren aus virtuellem Screening

Ebenfalls mit Hilfe des Dockings wurden aus einer Datenbank mit 235.000 Verbindungen 365 Substanzen ausgewählt und als PTP1B-Inhibitoren getestet. 127 Verbindungen hemmten die PTP1B mit einem IC_{50} kleiner als 100 μM , die aktivste Verbindung, deren Struktur aber nicht veröffentlicht ist, mit einem IC_{50} Wert von 1,7 μM . Für 25 Verbindungen wurden detaillierte kinetische Untersuchungen durchgeführt, die zeigten, dass nur vier dieser Substanzen als klassische kompetitive Inhibitoren fungieren, während die anderen entweder zeitabhängige kompetitive oder nicht kompetitive Inhibitoren sind. [261].

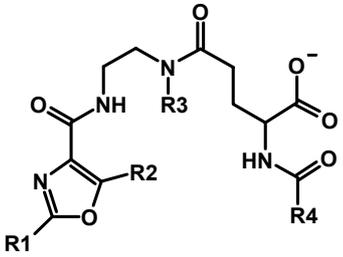
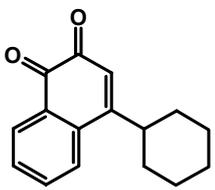
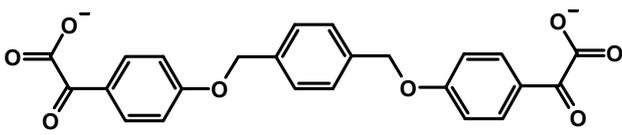
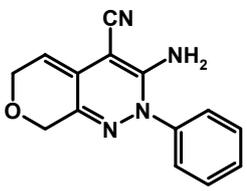
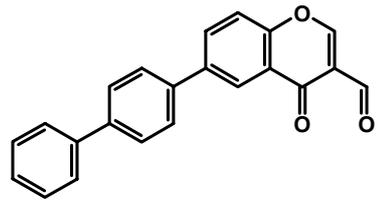
Verbindung	Literatur	Aktivität gegen PTP1B
	[267]	$K_i = 850 \text{ nM}$
	[269]	$IC_{50} = 320 \text{ nM}$
	[264]	$IC_{50} = 2,7 \mu M$
	[268]	$IC_{50} = 0,35 \mu M$
	[271]	$IC_{50} = 4,3 \mu M$

Tabelle 5-13: Nicht peptidische Inhibitoren ohne Phosphonatgruppe

Daneben existiert mittlerweile eine Vielzahl weiterer Substanzen aus den unterschiedlichsten Strukturklassen (Tabelle 5-13). Suramin zum Beispiel, ein antitrypanosomaler Arzneistoff, hat eine hohe *in vitro*-Aktivität ($K_i = 4 \mu\text{M}$), eignet sich aber aufgrund des sehr hohen Molekulargewichtes höchstens als pharmakologische Testsubstanz [262][263]. Verschiedene α -Ketocarbonsäuren besitzen eine geringe Aktivität und strukturelle Ähnlichkeit mit den DFMP-basierten Inhibitoren [264][265]. Als Oxidationsmittel fungierende phenylsubstituierten Periodate können durch Oxidation des aktiven Cysteins ebenso als Inhibitoren fungieren [266], wie einige Inhibitoren für Dual-Spezifische Phosphatasen [267]. Mit den neutralen Pyridazin-Derivaten [268], 1,2-Naphthochinonderivate [269] [270] und halogensubstituierte Formylchromonderivate [271], konnte man zeigen, dass eine Säurefunktion nicht unbedingt erforderlich ist.

Mit Hilfe der kombinatorischen Chemie wurden 5805 Imidazo[1,2-a]pyridine synthetisiert, deren aktivster Vertreter für PTP1B einen IC_{50} -Wert von $0,7 \mu\text{M}$ besitzt, während die selektivste Verbindung gegenüber TCPTP eine Selektivität von 69,7 [272].

Mit dem Verfahren der so genannten kovalenten Anbindung (covalent tethering) wurde das in der Nähe des aktiven Zentrums aber außerhalb der Bindungstasche befindliche Arg47 zu einem Cystein mutiert, an das Inhibitoren, die gut in das aktive Zentrum passen und eine Disulfid-Brücke besitzen, kovalent binden können. Darüber konnte ein pTyr-Mimetikum identifiziert werden, das zwar nur eine relativ schlechte Hemmung besitzt ($K_i = 4100 \text{ nM}$), aber einer neuen Substanzklasse angehört (PDB-Einträge 1NWE und 1NWL) [185].

Seit kurzem sind auch einige selektive allosterische Inhibitoren bekannt, deren Bindungsstelle sich etwa 20 \AA vom aktiven Zentrum entfernt befindet (PDB-Einträge 1T48, 1T49 und 1T4J) [186].

Neben den genannten Inhibitoren gibt es laufend Weiter- und Neuentwicklungen, die aktuell in Übersichtsartikeln zusammengefasst werden, so zuletzt im *Journal Expert Opinion on Investigational Drugs* [273].

5.2.2 Selektivität von PTP1B Inhibitoren

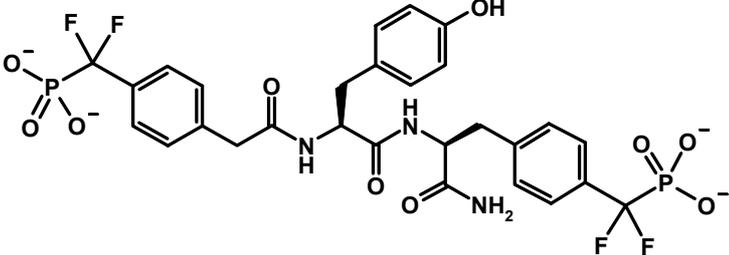
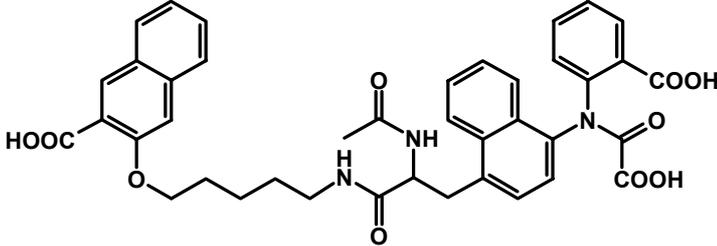
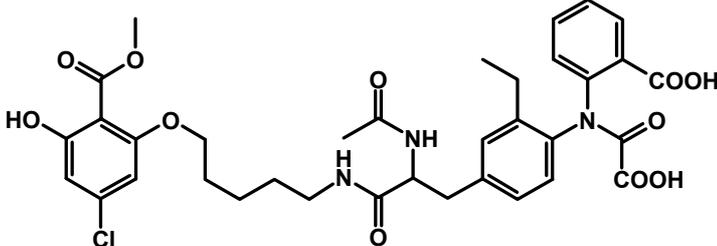
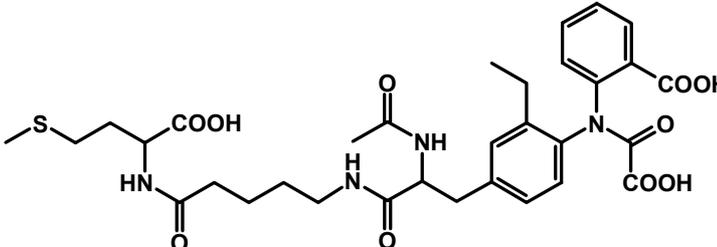
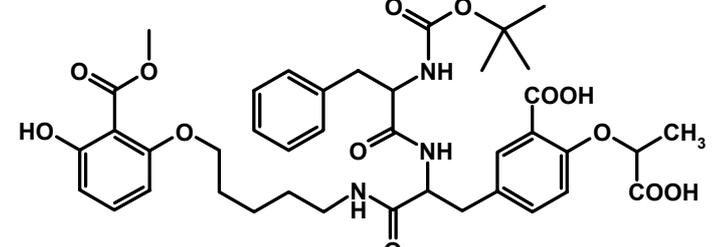
Verbindung	Literatur	K _i [μM]	
		PTP1B	TCPTP
	[184]	0,0049	0,02
	[176]	0,022	0,049
	[178]	0,04	0,23
	[179]	0,076	0,38
	[181]	0,16	1,97

Tabelle 5-14: K_i-Werte für TCPTP und PTP1B im Vergleich

Schon relativ lange versucht man, Inhibitoren zu entwickeln, die nur die PTP1B selektiv inhibieren. Da sich die aktiven Zentren der meisten PTPs von dem der PTP1B relativ stark unterscheiden, besitzen viele Inhibitoren eine ausreichende

Selektivität. Einzige Ausnahme bildet wie schon erwähnt TCPTP, die der PTP1B sehr ähnlich ist.

Seit etwa zwei Jahren versucht man deshalb auch Substanzen zu entwerfen, die auch gegenüber dieser PTP eine größere Selektivität besitzen. In Tabelle 5-14 sind die K_i -Werte für die aktivsten Verbindungen dargestellt.

Vergleicht man publizierte K_i -Werte für PTP1B und TCPTP, so fällt auf, dass diese sehr gut miteinander korrelieren. Leider sind die Verbindungen mit der größten Selektivität nur relativ schwache Inhibitoren ($K_i = 10$ bis $1 \mu\text{M}$), und die potentesten Inhibitoren besitzen nur eine Selektivität von etwa 2 bis 5 (siehe Abbildung 5-13)

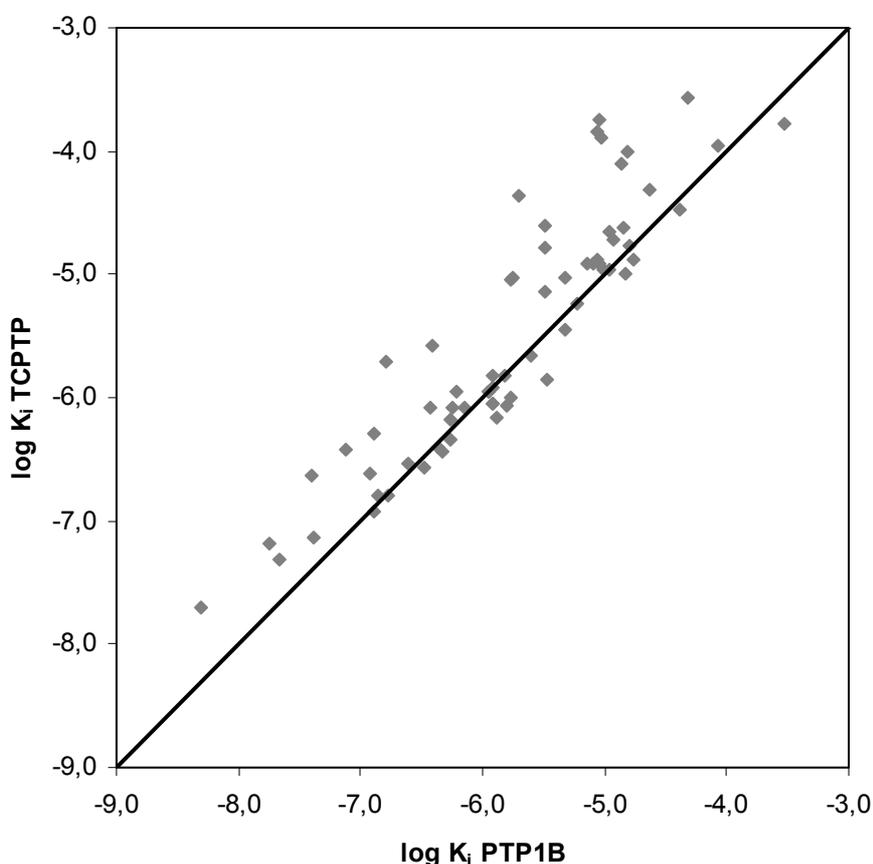
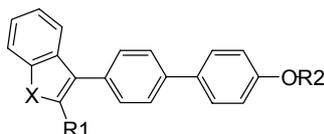


Abbildung 5-13: K_i -Werte für TCPTP und PTP1B im Vergleich

5.2.3 Liganden der 3D-QSAR-Modelle

Begonnen wurde die vorliegende Arbeit mit dem Ziel, für einen Datensatz von aus der Literatur bekannten Inhibitoren der PTP1B [258] eine 3D-QSAR Analyse durchzuführen. Der ausgewählte Datensatz besteht aus insgesamt 137 synthetisierten Verbindungen, für die die Hemmung der PTP1B als IC_{50} -Werte bestimmt wurde (Tabelle 5-15).



Compound	X	R1	R2	$IC_{50}^{(exp.)}$ (μM)
52	O	butyl	H	0.740
56	O	benzyl	H	0.920
57	O	benzoyl	H	0.740
58	S	butyl	H	0.700
59	S	benzyl	H	4.630
60	S	4-OH-benzyl	H	1.080
61	S	2,4-di-OH-benzyl	H	0.580
62	O	butyl	CH ₂ CO ₂ H	2.190
67	O	benzyl	CH(CH ₂ Ph)CO ₂ H (R)	0.350
68	O	benzyl	CH(CH ₂ Ph)CO ₂ H (S)	0.320
69	O	4-F-benzyl	CH(CH ₂ Ph)CO ₂ H (S)	0.330
70	O	benzyl	CH(CH ₂ Ph)CO ₂ CH ₃ (S)	3.470
71	O	benzyl	CH(CH ₂ CH ₂ Ph)CO ₂ H (S)	0.220
72	O	benzyl	CH(CH ₂ CH ₂ -N-phthalinimide)CO ₂ H (S)	0.340
74	O	benzyl	CH(Ph)CO ₂ H (R)	0.400
75	O	benzyl	CH(CH ₃)CO ₂ H (R)	1.320
76	O	benzoyl	CH(CH ₂ Ph)CO ₂ H (R)	0.680
79	S	butyl	CH(CH ₂ Ph)CO ₂ H (R)	0.170
80	S	benzyl	CH(CH ₂ Ph)CO ₂ H (S)	0.095
81	S	butyl	CH(Ph)CO ₂ H (R)	0.110
82	S	4-F-benzyl	CH(CH ₂ Ph)CO ₂ H (R)	0.120
83	S	4-OCH ₃ -benzyl	CH(CH ₂ Ph)CO ₂ H (R)	0.077
84	S	3,4-di-OCH ₃ -benzyl	CH(CH ₂ Ph)CO ₂ H (R)	0.120
85	S	2,4-di-OCH ₃ -benzyl	CH(CH ₂ Ph)CO ₂ H (R)	0.085
86	S	2,4-di-OH-benzyl	CH(CH ₂ Ph)CO ₂ H (R)	0.120
87	S		CH(CH ₂ Ph)CO ₂ H (R)	0.085
88	S		CH(CH ₂ Ph)CO ₂ H (R)	1.160
89	S		CH(CH ₂ Ph)CO ₂ H (R)	1.550

Tabelle 5-15: 3D-QSAR-Datensatz

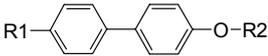
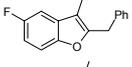
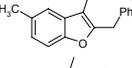
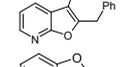
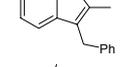
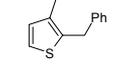
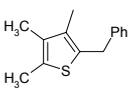
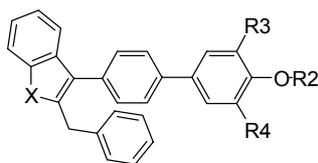
			
Compound	R1	R2	IC ₅₀ (exp.) (μM)
90		CH(CH ₂ Ph)CO ₂ H (S)	0.130
91		CH(CH ₂ Ph)CO ₂ H (R)	0.410
92		CH(CH ₂ Ph)CO ₂ H (S)	0.590
95		CH(CH ₂ Ph)CO ₂ H (R)	0.350
96		CH(CH ₂ Ph)CO ₂ H (R)	0.970
97		CH(CH ₂ Ph)CO ₂ H (R)	0.510

Tabelle 5-15: 3D-QSAR-Daten (Fortsetzung)

Alle Substanzen des Datensatzes sind jeweils aus drei Teilkomponenten zusammengesetzt, einem heteroaromatischen Bicyklus, einem aromatischen Linker und einer sauren Gruppe (Tabelle 5-15).

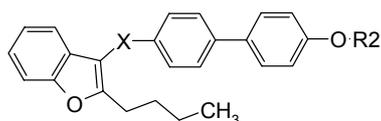
Der Bicyklus ist zumeist ein Benzofuran (77 Verb.), ein Benzothiophen (42 Verb.) oder eines von verschiedenen, kondensierten, stickstoffhaltigen Ringsystemen (6 Verb.). Daneben gibt es noch 4-Trifluormethylphenyl substituierte Oxazole (12 Verb.). Der Linker besteht aus einem Biphenyl- oder Naphthalin-Grundkörper, der mit der 3-Position des Benzothiophen und Benzofuran oder der 5-Position des Oxazol verknüpft ist.

Auf der anderen Seite des Linkers befindet sich entweder eine schwach saure, phenolische Hydroxylgruppe (28 Verb.) oder eine Carbonsäure, die über eine Etherbindung mit dem Linker verknüpft ist. Als Säure wurden Glykolsäure (23 Verb.) und deren in α-Position substituierte Derivate, in erster Linie die chirale Phenylmilchsäure (60 Verb.), verwendet. Daneben enthält der Datensatz einige Verbindungen mit Tetrazolen als isosterem Ersatz für die Carboxylgruppe (4 Verb.), sowie einzelne andere Carbonsäuren (7 Verb.). Ein Teil der Substanzen ist anstelle des Phenothers über eine Sulfongruppe mit Benzoe- oder Salicylsäure verknüpft (15 Verb.). Für alle drei Teilkomponenten der Inhibitoren liegen im Datensatz unterschiedliche Substituenten vor.

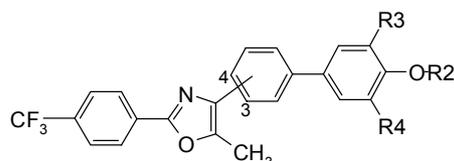


Compound	X	R3	R4	R2	IC _{50(exp.)} (μM)
98	S	Br	H	H	1.070
99	S	Br	Br	H	0.450
100	S	I	I	H	0.520
103	S	Br	H	CH(CH ₂ Ph)CO ₂ H (R)	0.058
104	S	Br	Br	CH(CH ₂ Ph)CO ₂ H (R)	0.025
105	S	4-OCH ₃ -Ph	H	CH(CH ₂ Ph)CO ₂ H (R)	0.053
106	S	4-Cl-Ph	H	CH(CH ₂ Ph)CO ₂ H (R)	0.052
107	S	Br	Br	CH(CH ₂ CH ₂ Ph)CO ₂ H (S)	0.290
111	S	Br	H	CH ₂ CO ₂ H	0.360
112	S	Br	Br	CH ₂ CO ₂ H	0.100
113	S	Ph	H	CH ₂ CO ₂ H	0.100
114	S	4-OCH ₃ -Ph	H	CH ₂ CO ₂ H	0.080
115	S	4-OC ₂ H ₅ -Ph	H	CH ₂ CO ₂ H	0.052
116	S	2,3-di-OCH ₃ -Ph	H	CH ₂ CO ₂ H	0.071
117	S	3,4,5-tri-OCH ₃ -Ph	H	CH ₂ CO ₂ H	0.100
118	S	4-OCH ₃ -Ph	Br	CH ₂ CO ₂ H	0.029
119	S	3-OCH ₃ -Ph	Br	CH ₂ CO ₂ H	0.028
120	S	2,4-di-OCH ₃ -Ph	Br	CH ₂ CO ₂ H	0.047
121	S	4-OCH ₃ -Ph	4-OCH ₃ -Ph	CH ₂ CO ₂ H	0.025
122	S	3-OCH ₃ -Ph	3-OCH ₃ -Ph	CH ₂ CO ₂ H	0.025
123	S	Br	H	CH ₂ CH ₂ CH ₂ CO ₂ H	0.170
124	O	Br	H	CH(CH ₂ Ph)CO ₂ H (S)	0.056
125	O	Br	Br	CH(CH ₂ Ph)CO ₂ H (S)	0.038
126	O	4-OCH ₃ -Ph	H	CH(CH ₂ Ph)CO ₂ H (S)	0.043
127	O	NO ₂	H	CH(CH ₂ Ph)CO ₂ H (R)	0.230
128	O	Br	Br	CH(CH ₂ CH ₃)CO ₂ H (S)	0.130
129	O	Br	Br	CH[CH ₂ CH(CH ₃) ₂]CO ₂ H (R)	0.054
132	O	CH ₃	CH ₃	CH(CH ₂ Ph)CO ₂ H (R)	0.074
133	O	cyclopentyl	H	CH(CH ₂ Ph)CO ₂ H (S)	0.055
134	O	cyclopentyl	H	CH ₂ CO ₂ H	0.170
136	O	NHCH ₂ CO ₂ H	H	CH ₂ CO ₂ H	0.082
137	O	NHCH ₂ CH ₂ CO ₂ H	H	CH ₂ CO ₂ H	0.140
138	O	NHCOCH ₂ CH ₂ CO ₂ H	H	H	0.920
139	O	NHCOCH=CHCO ₂ H	H	H	0.460
140	O	NHCO-C ₆ H ₄ -2-CO ₂ H	H	H	0.160

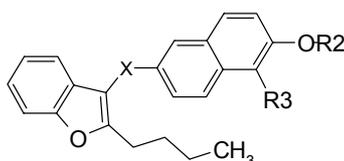
Tabelle 5-15: 3D-QSAR-Daten (Fortsetzung)



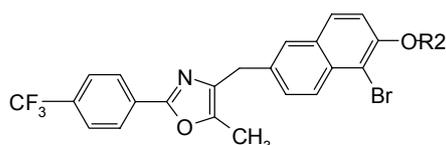
Compound	X	R2	IC _{50(exp.)} (μM)
143	CH ₂	H	1.190
147	CH ₂	CH ₂ CO ₂ H	1.150
149	CH ₂	CH ₂ -tetrazole	0.510



Compound	P.O.A.	R3	R4	R2	IC _{50(exp.)} (μM)
151	4	H	H	CH ₂ CO ₂ H	0.800
153	4	H	H	CH ₂ -tetrazole	0.900
157	4	Br	Br	H	0.650
158	4	Br	Br	CH ₂ CO ₂ H	0.470



Compound	X	R3	R2	IC _{50(exp.)} (μM)
160	CH ₂	H	H	1.300
164	CH ₂	Br	H	0.330
165	CH ₂	I	H	0.380
166	CH ₂	Br	CH ₂ CO ₂ H	1.400
170	CH ₂	Br	CH ₂ -tetrazole	0.700
171	CO	Br	CH ₂ -tetrazole	1.100



Compound	R2	IC _{50(exp.)} (μM)
172	CH ₂ CO ₂ H	1.300

Tabelle 5-15: 3D-QSAR-Daten (Fortsetzung)

Von zwei Verbindungen (80 und 181) aus dem Datensatz wurden die Röntgenkristallstrukturen im Komplex mit der PTP1B aufgeklärt, aber nicht öffentlich zugänglich gemacht, so dass nur eine Beschreibung der Wechselwirkungen und eine Abbildung der Struktur im *Journal of Medicinal Chemistry* verfügbar ist [258].

Die Carboxylatgruppe der Verbindung 80 (Abbildung 5-14) interagiert nicht wie vermutet wurde direkt mit Arg221 und den Amid-NH-Gruppen des P-Loops, sondern bildet stattdessen Wasserstoffbrücken zu zwei Wassermolekülen, die die Carboxylatgruppe mit dem Guanidinrest von Arg221 verbinden. Der erste Phenylring befindet sich in einer Art Sandwichanordnung zwischen Tyr46 und Phe182, während der zweite Phenylring und der Benzylrest über Kation- π -Wechselwirkungen an die beiden basischen Aminosäuren Lys116 und Lys120 gebunden werden. Die lipophilen Bereiche der Seitenketten von Lys116 und Lys120 interagieren zusätzlich mit dem Liganden.

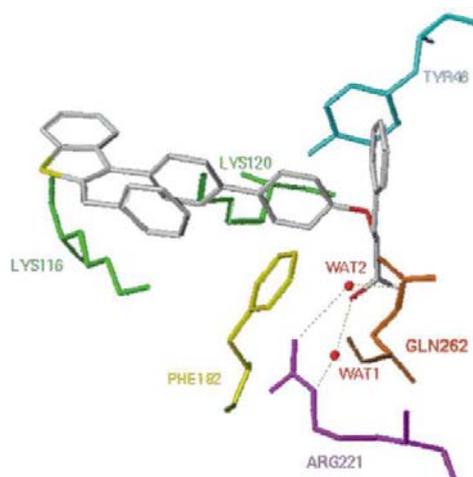


Abbildung 5-14: Verbindung 80 in der PTP1B (aus [258])

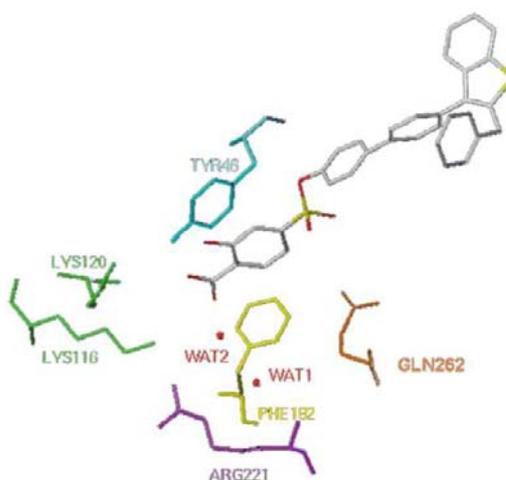


Abbildung 5-15: Verbindung 181 in der PTP1B (aus [258])

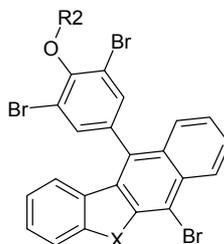
Die zweite Substanz (Verbindung 181, siehe Abbildung 5-15) wurde in einem völlig anderen Bindungsmodus kristallisiert. Sie liegt um nahezu 180° gedreht im aktiven Zentrum, so dass die Carboxylatgruppe direkte Wasserstoffbrücken zu Lys120 und der Hydroxylgruppe von Tyr46 bilden kann. Über ein Wassermolekül wird eine zusätzliche Wechselwirkung mit Arg221 hergestellt. Auch hier bindet der erste Phenylring zwischen Tyr46 und Phe182, aber der lipophile Rest des Moleküls geht unspezifische Wechselwirkungen mit Bereichen des Proteins ein, in denen der andere Ligand nicht bindet [258].

Bei den durchgeführten theoretischen Untersuchungen konnten 19 Verbindungen nicht berücksichtigt werden, da für diese keine IC_{50} -Werte bekannt sind. Sechs Verbindungen wurden ausgeschlossen, da sie in einer Seitenkette eine nicht aufgeklärte Chiralität besitzen. Unter den übrigen Substanzen befinden sich 72 achirale und 59 chirale Verbindungen, die ihr einziges chirales Zentrum in der α -Position der Phenylmilchsäure besitzen. Darunter sind 42 reine Enantiomere (davon 15 S-konfiguriert) und 17 Racemate. Auf die chiralen Verbindungen wird bei der Besprechung der ligandbasierten 3D-QSAR-Modelle genauer eingegangen. Bei der Konformationsanalyse und Überlagerung der Strukturen stellte sich heraus, dass die Sulfonobiphenyle und die Verbindung 78 des Datensatzes sich nicht mit den anderen Strukturen überlagern lassen, was durch die im zugrunde liegenden Artikel veröffentlichten, unterschiedlichen Bindungsmodi unterstützt wird [258].

5.2.4 Liganden zur Datenbanksuche

Als Basis für die Datenbanksuchen, die im zweiten Teil dieser Arbeit durchgeführt wurden, sollten weitere bekannte Inhibitoren zu den Modellen hinzugefügt werden. Aus der großen Vielzahl von veröffentlichten Strukturen wurden diejenigen ausgewählt, die sowohl in einem vergleichbaren *in vitro*-Testsystem untersucht wurden als auch eine ähnliche Größe besitzen. Durch diese Vorgaben wurden alle peptidischen Hemmstoffe sowie alle relativ kleinen Liganden, wie zum Beispiel die Oxylaminobenzoessäurederivate, von den Untersuchungen ausgegrenzt. Nach Ausschluss der in völlig unterschiedlichen Testsystemen getesteten Verbindungen blieben einige Strukturen übrig, die in der gleichen Firma getestet wurden [255][256][257], sowie ein Datensatz, der aus experimentell verifizierten Hits aus einem virtuellen Screening besteht [261].

Die beiden Artikel von Wrobel et al. [255][256] enthalten jeweils nur eine kleine Anzahl von Benzo[b]naphtho[2,3-d]furanen, Benzo[b]naphtho[2,3-d]thiophenen oder 1-oxa-9-thiacyclopenta[b]fluorenen, die in ihren weiteren Substitutionen den Substanzen des ersten Datensatzes entsprechen (Tabelle 5-16 und Tabelle 5-17).



Compound	X	R2	IC _{50(exp.)} (μM)
3	S	H	0.384
4	S	CH ₂ CO ₂ H	0.386
6	S	CH(CH ₂ Ph)CO ₂ H (R)	0.061
7	O	CH(CH ₂ Ph)CO ₂ H (R)	0.083
8	S	CH(CH ₂ CH ₂ -N-phthalimide)CO ₂ H (R)	0.011

Tabelle 5-16: Inhibitoren zur Datenbanksuche [255]

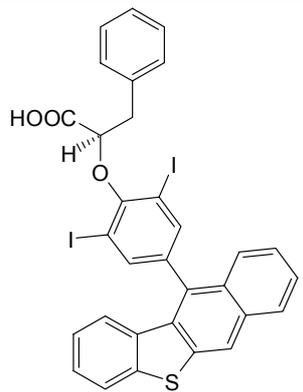
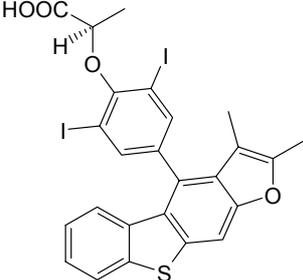
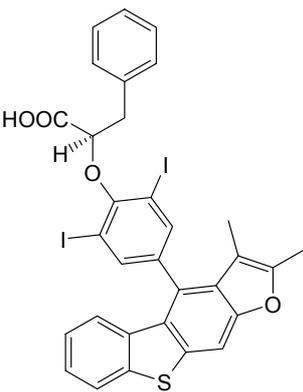
Compound		IC ₅₀ (exp.) (μM)
2		0.179
3		0.284
4		0.074

Tabelle 5-17: Inhibitoren zur Datenbanksuche [256]

Der Artikel von Malamas et al. [257] enthält eine große Anzahl von Azolidin-dionen und davon abgeleitete Verbindungen, von denen aber die meisten ausgeschlossen wurden, da für die *in vitro*-Bestimmungen nicht das humane Enzym sondern das homologe Enzym aus der Ratte verwendet wurde (Tabelle 5-18, Seite 111).

Der Datensatz aus dem Screening-Experiment [261] enthält elf grundsätzlich unterschiedliche Verbindungen mit IC₅₀-Werten in einem geringfügig niedrigeren Bereich als die anderen verwendeten Daten. Eine Verbindung wurde wegen

zu geringer Aktivität aus den Untersuchungen ausgeschlossen (Tabelle 5-19, Seite 112).

Compound	$IC_{50(\text{exp.})}$ (μM)
65	0,3
84	1,9
86	1.4
88	0.3
117	0,37
118	0,85
119	0,16
120	0,12
121	1,3

Tabelle 5-18: Inhibitoren zur Datenbanksuche [257]

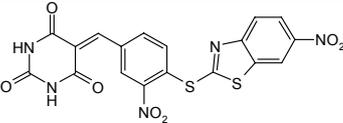
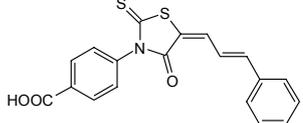
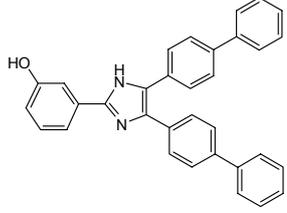
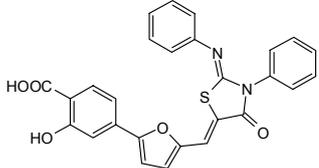
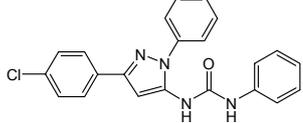
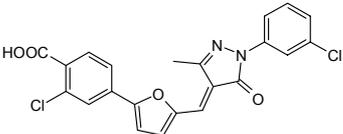
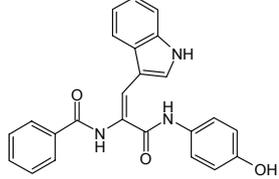
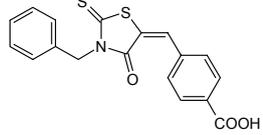
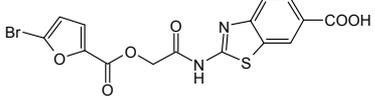
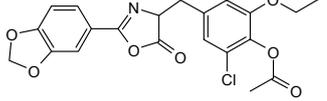
Compound		IC ₅₀ (μ M)
1		4.1
2		4.4
3		8.6
4		9.0
5		12.0
6		19.3
7		21.5
8		21.6
9		34.4
10		35.0

Tabelle 5-19: Inhibitoren zur Datenbanksuche [261]

5.3 Ligandbasierte 3D-QSAR-Modelle

Trotz der Verfügbarkeit von mehreren Kristallstrukturen der PTP1B wurden zu Beginn der Arbeit 3D-QSAR-Modelle erstellt, bei denen die Struktur des Enzyms unberücksichtigt blieb. Diese Methode wird bei der Arzneistoffentwicklung in erster Linie dann angewendet, wenn die dreidimensionale Struktur des Zielenzym oder Rezeptors nicht bekannt ist, aber eine ausreichende Anzahl von Bindungs- oder Inhibitionsdaten für Liganden zur Verfügung steht. Diese Modelle sollten als Test für die Etablierung der Methode und als Grundlage für den Vergleich mit den später erstellten, strukturbasierten 3D-QSAR-Modelle dienen. Dabei ist es wichtig, für die weiteren Berechnungen zuerst eine Konformation der Liganden zu finden, die der bioaktiven Konformation entspricht, auf die dann eine Überlagerung der weiteren Verbindungen aufgebaut werden kann. Anschließend erfolgt die Berechnung der Interaktionsfelder, die dann mit der biologischen Aktivität korreliert wird. Abbildung 5-16 zeigt die generelle Vorgehensweise bei den ligandbasierten Untersuchungen.

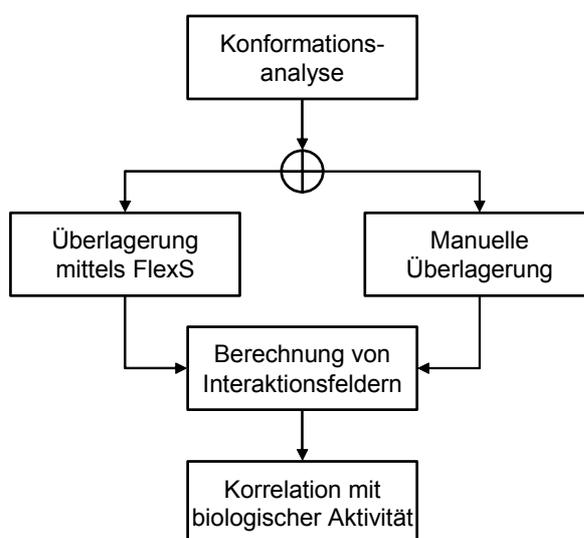


Abbildung 5-16: Ligandbasierte 3D-QSAR-Analyse, Vorgehensweise

Anhand der ligandbasierten 3D-QSAR-Modelle wurde auch der Einfluss der Konformation, der Überlagerungsmethode, des Ladungszustandes und der GRID-Sonden untersucht. Die Ergebnisse dienten als Vergleich mit später erstellten rezeptorbasierten Modellen.

5.3.1 Konformationsanalyse

Als Verfahren zur Konformationsanalyse wurde eine systematische Analyse (siehe 4.2.3) gewählt, da die Inhibitoren keine flexiblen Ringsysteme besitzen, und die Liganden mit etwa zehn rotierbaren Einfachbindungen für dieses Vorgehen nicht zu komplex sind. Verbindung 80, die in Abbildung 5-17 dargestellt ist, wurde als Startstruktur ausgewählt, auf die alle anderen Strukturen überlagert werden sollten. Diese Substanz ist zwar selbst nicht die aktivste Verbindung, sie besitzt aber die gleiche Grundstruktur wie die beiden aktivsten Verbindungen 104 und 142 sowie viele weitere Substanzen des Datensatzes (siehe Abschnitt 5.2.3). Außerdem ist ihre bioaktive Konformation aus einer unveröffentlichten Kristallstruktur bekannt, und der Datensatz enthält keine besonders rigiden Strukturen mit weniger Freiheitsgraden als Verbindung 80. Aufgrund der Biphenylgruppe in der Mitte lassen sich die Substanzen in zwei Teile zerschneiden, die sich in ihrer Konformation gegenseitig kaum beeinflussen können, so dass man sie getrennt voneinander betrachten kann. Zuerst wurde der Benzofuranbiphenyl-Anteil einer systematischen Konformationsanalyse in Sybyl unterzogen, indem die vier Einfachbindungen zwischen den Ringsystemen in Winkelschritten von jeweils 15° gedreht wurden (Magentafarbene Bindungen in Abbildung 5-17). Die Phenylreste wurden aus Symmetriegründen nur im Bereich von 0 bis 180° untersucht, die Bindung zwischen dem Benzofuran und der Benzylgruppe von 0 bis 360° , so dass 41.472 Strukturen resultierten. Die Reduktion der großen Anzahl Konformationen mit dem Programm Ixgros [274] und die anschließende Analyse der Energiewerte lieferte mehrere relevante Strukturen, von denen nach einer Geometrieoptimierung nur noch die besten zwei übrig blieben, die weiter betrachtet wurden.

Anschließend wurden die Winkel der Einfachbindungen im Phenylmilchsäure-Teil der Verbindung in Schritten von 30° gedreht (Orangefarbene Bindungen in Abbildung 5-17). Aufgrund der Symmetrie wurden die Torsionswinkel zum Phenylring und zur Carboxylatgruppe wiederum nur im Bereich von 0 bis 180° untersucht. Auch in diesem Fall konnte die Anzahl der Konformationen durch Ixgros, die Begrenzung der Energie und eine darauf folgende Energieminimierung soweit reduziert werden, dass von den ursprünglichen 124.416 Konforma-

tionen nur noch zwei energetisch günstige Strukturen je Konformer des ersten Teils übrig blieben. Aus diesen beiden wurde durch einen Vergleich mit der Abbildung der unveröffentlichten Kristallstruktur diejenige ausgewählt, die der bioaktiven Konformation der Verbindung 80 entspricht (Abbildung 5-18).

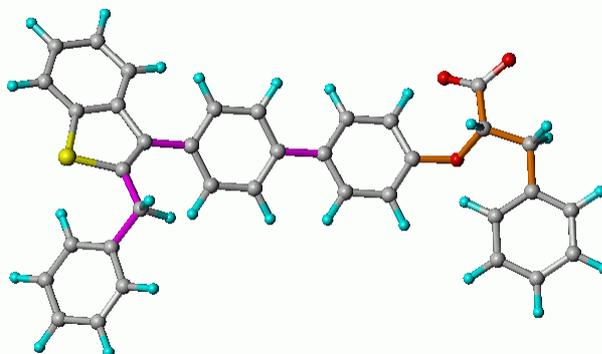


Abbildung 5-17: Konformationsanalyse für Verbindung 80 (drehbare Bindungen in magenta und orange)

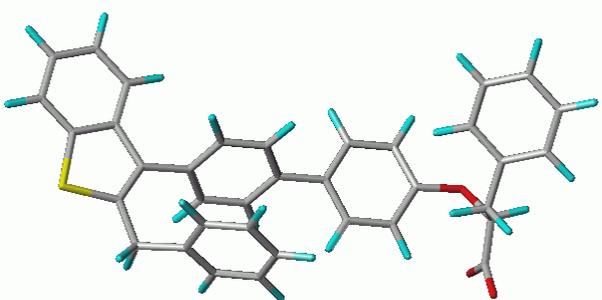


Abbildung 5-18: Angenommene Bioaktive Konformation für Verbindung 80

Um die Eignung der gewählten Konformation für die Überlagerung mit dem Programm FlexS zu überprüfen, wurde die Verbindung 80 auf sich selbst überlagert. Da das Programm FlexS nur bestimmte Torsionswinkel verwendet, die durch eine Analyse der entsprechenden Winkel in den Kristallstrukturen der *Cambridge Structural Database* gewonnen wurden, entsprachen die mit FlexS gefundenen Winkel nicht genau denen der Konformationsanalyse. Die große Ähnlichkeit zwischen der Struktur aus der systematischen Konformationsanalyse und der Überlagerung mit FlexS belegt aber, dass die gefundenen Torsionswinkel nicht nur einem Energieminimum entsprechen, sondern auch häufig in Kristallstrukturen anzutreffen sind. Weil die Abweichungen nur sehr gering waren, und die allgemeine Konformation in beiden Fällen identisch war, wurde die mittels FlexS gewonnene Konformation für weitere Untersuchungen aus-

gewählt, damit die Überlagerung der weiteren Liganden mit vergleichbaren Voraussetzungen stattfinden kann.

5.3.2 Superpositionierung

Die Überlagerung der unterschiedlichen Verbindungen des Datensatzes wurde mit zwei verschiedenen Verfahrensweisen durchgeführt. Die Superpositionierung der ähnlichen Grundgerüste der Verbindungen 67, 68, 79 und 82, bei denen sich nur die Konfiguration am chiralen Kohlenstoffatom ändert, oder Benzothiophen durch Benzofuran ersetzt wird, erfolgte für beide Vorgehensweisen auf die gleiche Weise mit FlexS (Abbildung 5-19). Auch die unterschiedlichen Grundgerüste der Verbindungen 147, 152, 167 und 172 wurden immer mit FlexS auf Verbindung 80 überlagert (Abbildung 5-20).

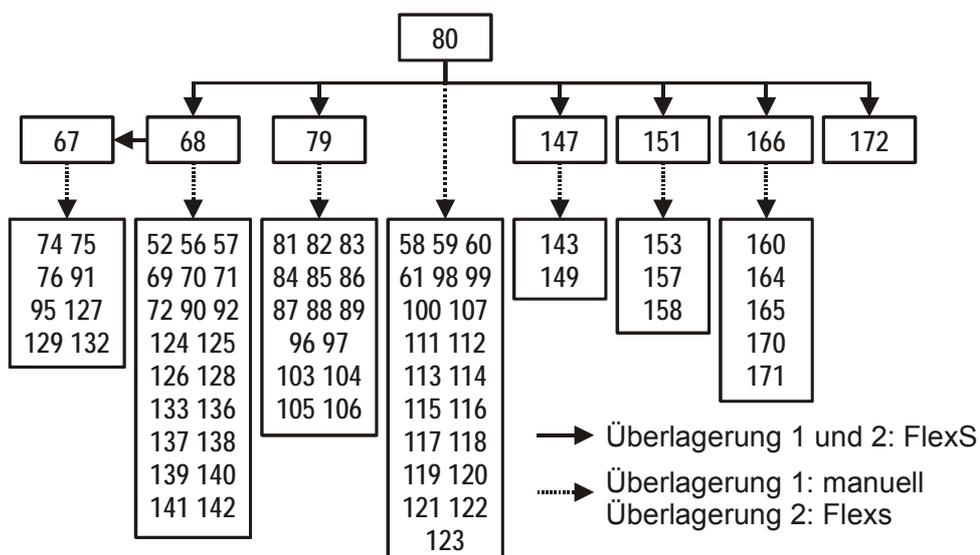


Abbildung 5-19: Vorgehen bei den ligandbasierten Überlagerungen 1 und 2

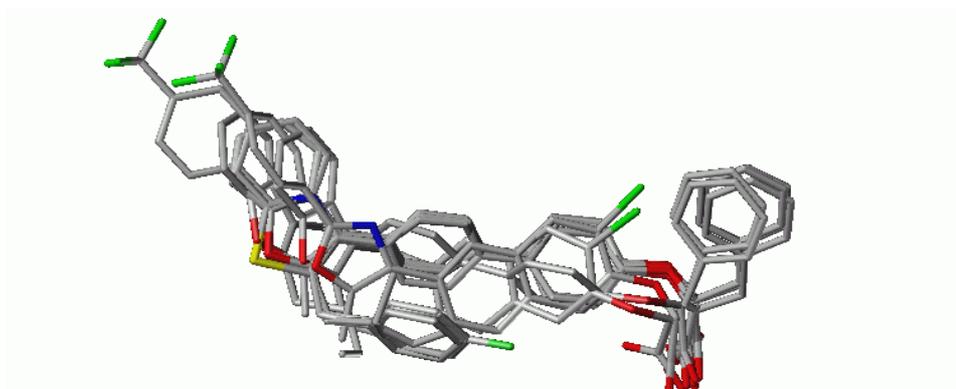


Abbildung 5-20: Ligandbasierte Überlagerungen der Grundgerüste; Verbindungen 67, 68, 79, 80, 82, 147, 151, 166, und 172

Von dieser Überlagerung der Grundgerüste ausgehend wurden dann die Strukturen der einzelnen Substanzen erzeugt (Abbildung 5-19), indem entweder die unterschiedlichen Reste ausgetauscht wurden, ohne die Konformation zu verändern (Überlagerung 1), oder die Strukturen wurden wiederum mit FlexS auf das ähnlichste Templat flexibel superpositioniert (Überlagerung 2). Beide Überlagerungen sind sich ähnlich, allerdings sind die Substituenten in der Überlagerung 1 eher Atom auf Atom überlagert (Abbildung 5-21), während bei der Überlagerung 2 (Abbildung 5-22) mehr das Volumen der Substituenten berücksichtigt ist. Beide Vorgehensweisen liefern die für die jeweilige Methode erwarteten Überlagerungen und wurden bezüglich ihrer Eignung für 3D-QSAR-Analysen untersucht.

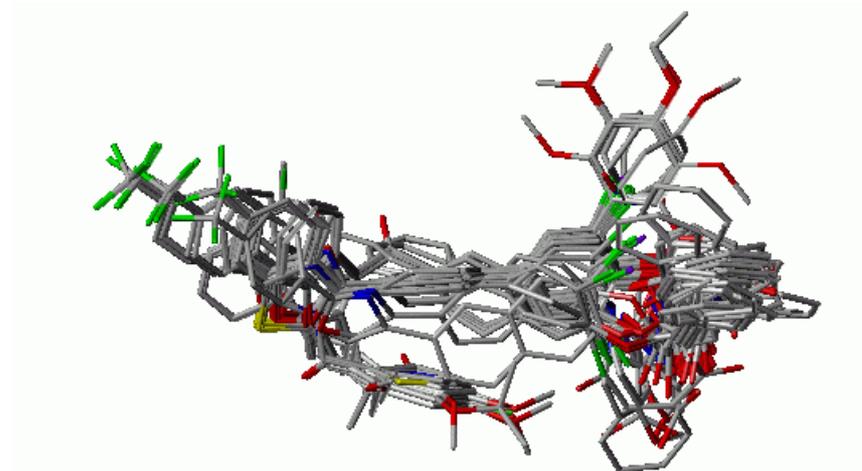


Abbildung 5-21: Überlagerung 1; alle Verbindungen

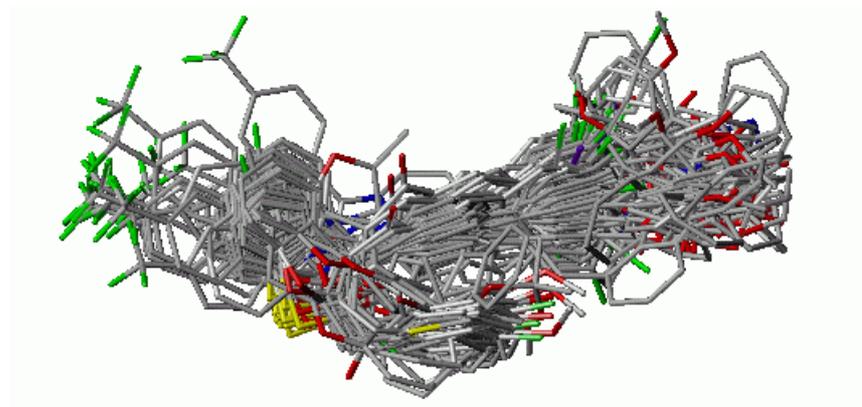


Abbildung 5-22: Überlagerung 2; alle Verbindungen

5.3.3 Auswahl der GRID-Sonden

Um eine geeignete Sonde, die gut zwischen den unterschiedlichen Strukturen entscheiden kann, für das 3D-QSAR auszuwählen, wurden die Interaktionsfelder für eine Vielzahl von Sonden berechnet und mit Hilfe einer PCA in GOLPE analysiert (Abbildung 5-23). Bei allen Auswertungen stieg die Aussagekraft der Modelle mit der Anzahl der Hauptkomponenten stetig an, ab fünf Hauptkomponenten war dieser Anstieg allerdings nicht mehr signifikant.

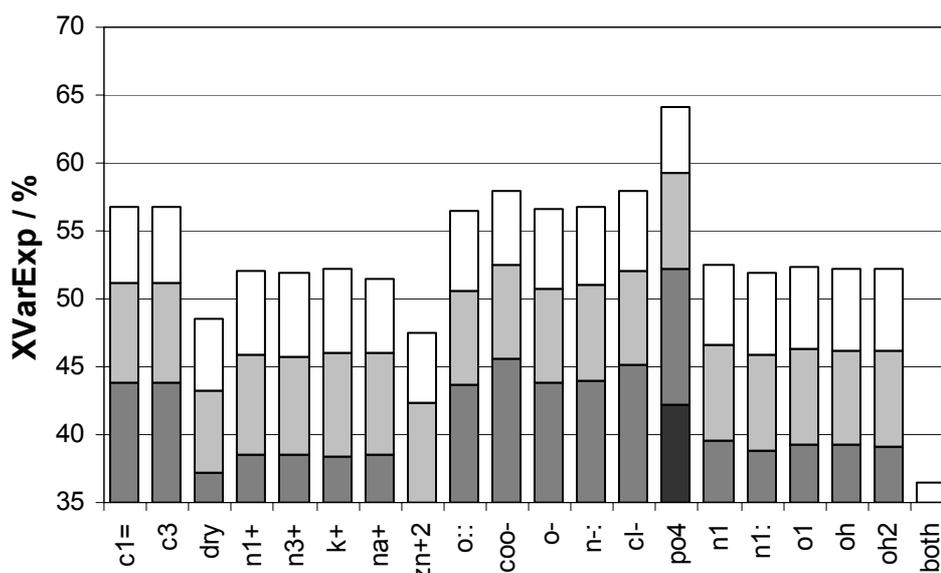


Abbildung 5-23: Vergleichende PCA verschiedener GRID-Sonden, jeweils für drei (dunkelgrau), vier (hellgrau) und fünf (weiß) Hauptkomponenten

Dabei lieferten die lipophilen Sonden (aromatische CH- oder Methylgruppe) relativ gute Ergebnisse und können mit fünf Hauptkomponenten beispielsweise etwa 56,7% der Varianz in den X-Variablen (Interaktionsfelder) erklären.

Eine vergleichbar gute Beschreibung der Variabilität der Interaktionsfelder (57,2%) ist mit negativ geladenen Sonden (Carboxylat-Sauerstoff, Carboxylat-Gruppe, Phenolat-Sauerstoff, Chlorid-Ion oder Tetrazol-Stickstoff) möglich, während die zweifach negativ geladene organische Phosphatgruppe mit 64,1% die beste Erklärung der X-Variablen liefert. Schlechter schneiden positiv geladene Sonden (Natrium-, Kalium- und Ammonium-Ionen) sowie hydrophile neutrale Sonden (Hydroxylgruppe, Wasser, Amine) ab, mit jeweils nur etwa 52,1%. Besonders ungeeignet waren die Dry-Sonde (48,5%) und die Zink-Sonde (47,5%) sowie die Both-Sonde, die sowohl lipophile wie hydrophile

Interaktionen beschreiben soll (36,4%). Bei diesen Untersuchungen, die nur der Auswahl der geeigneten GRID-Sonden diene, wurde die Y-Variable (biologische Aktivität) nicht berücksichtigt.

Die Ergebnisse zeigen, dass die Unterschiede zwischen ähnlichen Sonden einer Gruppe nur sehr gering sind, aber der Auswahl der Sonde eine wichtige Bedeutung zukommt. Sie zeigen weiterhin, dass die Auswahl der Sonde nach rationalen Kriterien nicht zu den gewünschten Ergebnissen beim 3D-QSAR führt, denn bei den betrachteten Liganden spielt die Interaktion mit einer Phosphatgruppe keine Rolle, während Interaktionen mit aromatischen und basischen Gruppen für die Bindung an das Enzym entscheidend sind. Auf den ersten Blick könnte man auf einen Zusammenhang zwischen der guten Aussagekraft der Phosphatsonde und der Hydrolyse von Phosphotyrosin schließen, da aber die Interaktionen der Sonden mit den Liganden untersucht werden, während das Phosphotyrosin mit dem Protein interagiert, gibt es hier keinen Zusammenhang. Die Ursache für die hohe Aussagekraft der Phosphatsonde ist eher darin zu suchen, dass sie mit der Carboxylatgruppe besonders schlechte Wechselwirkungen zeigt, aber mit den Phenolgruppen besser interagiert. Außerdem sind durch die doppelte negative Ladung die Wechselwirkungen mit den unterschiedlich substituierten Aromaten differenzierter als bei den neutralen oder einfach geladenen Sonden. Für die Beurteilung der Unterschiede innerhalb des Datensatzes ist es nicht von Bedeutung, ob die Sonde eher positive oder negative Interaktionen mit den Liganden besitzt.

	OH2	N2=	PO4	CH3
	Wasser	Guanidin	Phosphat	Methyl
VdW-Radius (Å)	1,7	1,7	2,3	1,95
Elektronen	7	8	24	8
Polarisierbarkeit ($\alpha \cdot \text{Å}^3$)	1,2	2,1	6,2	2,17
Ladung	0	0,66	-2	0
H-Brücken Energie (Kcal/M)	-4,5	-3	-2,5	0
H-Brücken Radius (Å)	1,4	1,6	2,3	0
H-Brücken Donoren	2	2	0	0
H-Brücken Akzeptoren	2	0	6	0

Tabelle 5-20: Parameter der verwendeten GRID-Sonden

Da die mit einer nach rationalen Kriterien ausgewählten Sonde erhaltenen Ergebnisse der 3D-QSAR-Modelle bezüglich der Bindung an das Enzym leichter zu interpretieren sind, wurden für weitere Analysen zusätzliche Sonden ausgewählt, die die Eigenschaften der Bindungstasche des Proteins wiedergeben können. Neben der Phosphat-Sonde (PO₄) wurden die Berechnungen deshalb auch für die Methyl-Sonde (CH₃), die Wasser-Sonde (OH₂) und die planare kationische sp²-Stickstoff-Sonde (N₂=), die einem Stickstoff der Aminosäure Arginin entspricht, durchgeführt (Tabelle 5-20).

5.3.4 3D-QSAR-Modelle

Da verschiedene Parameter wie die Überlagerung der Liganden, die Ladung protonierbarer Gruppen oder die Wahl des Datensatzes einen großen Einfluss auf die Güte der 3D-QSAR-Modelle besitzen, wurden deren Auswirkungen auf die Analyse genauer untersucht.

5.3.4.1 Einfluss der Überlagerung

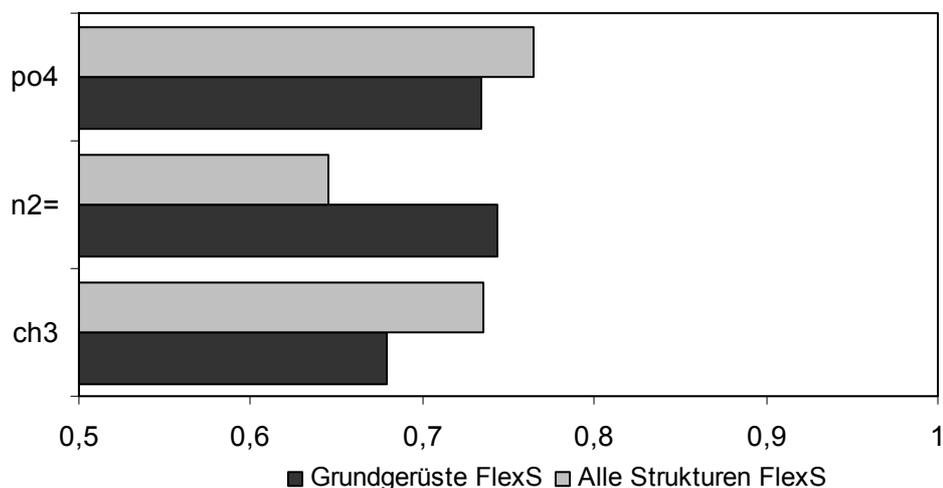


Abbildung 5-24: PLS-Analyse verschiedener ligandbasierter Überlagerungen ($q^2_{L20\%O}$ -Werte)

Für die beiden beschriebenen Überlagerungen und verschiedene GRID-Sonden wurden 3D-QSAR-Modelle erstellt. Diese sind in Abbildung 5-24 miteinander verglichen.

Die Carbonsäuren und Tetrazole waren dabei zunächst ihrem üblichen pK_S -Wert entsprechend deprotoniert, während die Phenole als neutrale Verbindungen vorlagen. Eine eingehende Betrachtung des Ladungszustandes der einzelnen Substanzklassen wurde im Anschluss an die ersten Überlagerungen durchgeführt und wird im Abschnitt 5.3.4.2 beschrieben.

Für die meisten GRID-Sonden lieferte die rein auf FlexS basierende Überlagerung die besseren Ergebnisse für die Vorhersage der internen Testverbindungen. Als Verfahren zur Kreuzvalidierung wurde die *Random-Group*-Methode herangezogen, wobei jeweils nicht nur eine Verbindung aus dem Modell herausgelassen wird sondern eine bestimmte Anzahl, zum Beispiel zwanzig Prozent aller Verbindungen (L20%O). Die beste Korrelation zwischen Interaktionsfeldern und biologischer Aktivität wurde mit der PO4-Sonde und der FlexS-Überlagerung (Überlagerung 2) erzielt, aber die Ergebnisse unterscheiden sich nicht signifikant von den anderen Modellen. Außerdem war bei der N2=-Sonde die andere Vorgehensweise (Überlagerung 1) besser geeignet, so dass letztlich nicht entschieden werden kann, welche von beiden Überlagerungen vorzuziehen ist oder der pharmakophoren Überlagerung entspricht. Die Ergebnisse machen aber nochmals deutlich, dass der Auswahl der GRID-Sonde eine große Bedeutung zukommt.

5.3.4.2 Einfluss des Ladungszustandes

Im nächsten Schritt wurde der Einfluss der Ladungen der Liganden auf die QSAR-Analyse untersucht, da die Verbindungen Säuren und Phenole darstellen und somit protoniert und neutral oder deprotoniert und negativ geladen vorliegen können. Zu diesem Zweck wurden für die Überlagerung 1 drei unterschiedliche Protonierungsstufen analysiert. Die erste Stufe betrachtet alle Verbindungen als protoniert (neutral), die zweite die Carbonsäuren und Tetrazole als deprotoniert, und die dritte Stufe betrachtet zusätzlich die Phenole als negativ geladene Phenolate.

Vergleicht man, wie in Abbildung 5-25, die SDEP- und q^2 -Werte der 3D-QSAR-Modelle, so stellt man fest, dass alle Modelle ähnlich gute Ergebnisse liefern und sich die q^2 -Werte nicht signifikant unterscheiden.

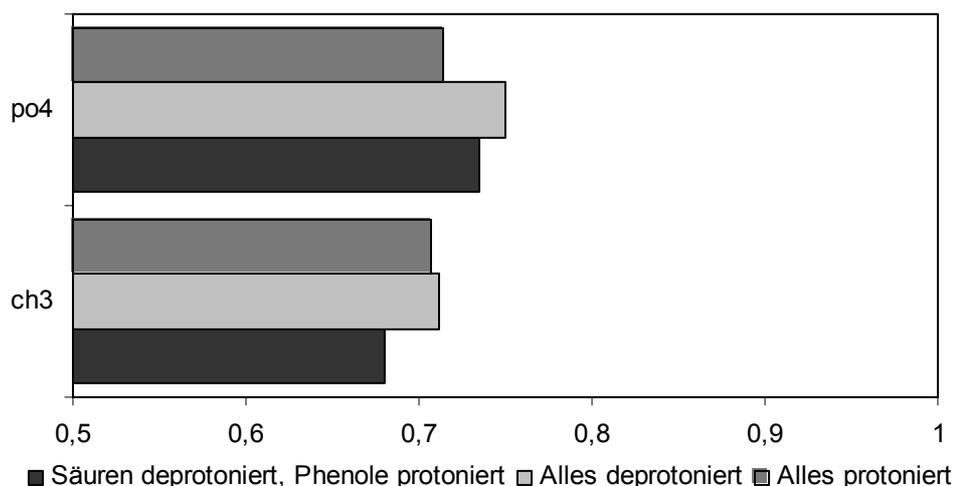


Abbildung 5-25: PLS-Analyse verschiedener Protonierungsstufen der Liganden ($q^2_{L20\%O}$ -Werte)

In einer weiteren Untersuchung wurden nur die Carbonsäuren und Tetrazole betrachtet und die Phenole aus dem Datensatz ausgeschlossen. Die resultierenden Modelle liefern verbesserte q^2 -Werte aber schlechtere SDEP-Werte für die noch enthaltenen Verbindungen. Durch diese Vorgehensweise würde außerdem der Datensatz zu stark eingeschränkt und eine Vorhersage von Phenolen wäre sehr unpräzise. Da die Phenole durchschnittlich etwa eine Zehnerpotenz weniger aktiv sind als die entsprechenden Carbonsäuren, ist die negative Ladung des Inhibitors als wichtiges Strukturmerkmal anzusehen.

Im Folgenden wurde folglich der Datensatz einschließlich deprotonierter Phenole verwendet. Gestützt wird die Annahme, dass die Phenole trotz ihres pK_S -Werts von ungefähr 10 bei der Bindung an das Enzym ebenso wie das Phosphat des Substrates und Inhibitoren mit einer Carboxylatgruppe deprotoniert vorliegen, durch den in der Proteinumgebung der PTP1B stark erniedrigten pK_S -Werte des aktiven Cysteins (siehe Abschnitt 5.1.3). Stabilisiert werden die negativen Ladungen durch das Dipolmoment der zentralen α -Helix, die Amidgruppen des P-Loop und die basische Seitenkette von Arg221.

5.3.4.3 Untersuchung der Racemate

Da der Datensatz 17 Racemate (12,5% aller Verbindungen) enthält, wurde überlegt, auch diese Verbindungen in die 3D-QSAR-Analysen mit einzubeziehen. Diese Verbindungen besitzen ihr einziges Chiralitätszentrum in der

Phenylmilchsäure-Gruppe. Neben den racemischen Verbindungen enthält der Datensatz aber auch 42 Substanzen, die als reine R- oder S-Enantiomere vorliegen. Darunter ist nur eine Verbindung für die sowohl als Racemat als auch in enantiomerenreiner Form die biologische Aktivität bestimmt wurde (Tabelle 5-21). Wie die Strukturen in Abbildung 5-26 zeigen, lassen sich die R-Konfiguration und die S-Konfiguration dieses Moleküls in der angenommenen bioaktiven Konformation sehr gut überlagern.

Substanz	Nummer	IC ₅₀
Racemat	66	0,27 µM
R-Enantiomer	67	0,35 µM
S-Enantiomer	68	0,32 µM

Tabelle 5-21: IC₅₀ der Verbindungen 66, 67 und 68

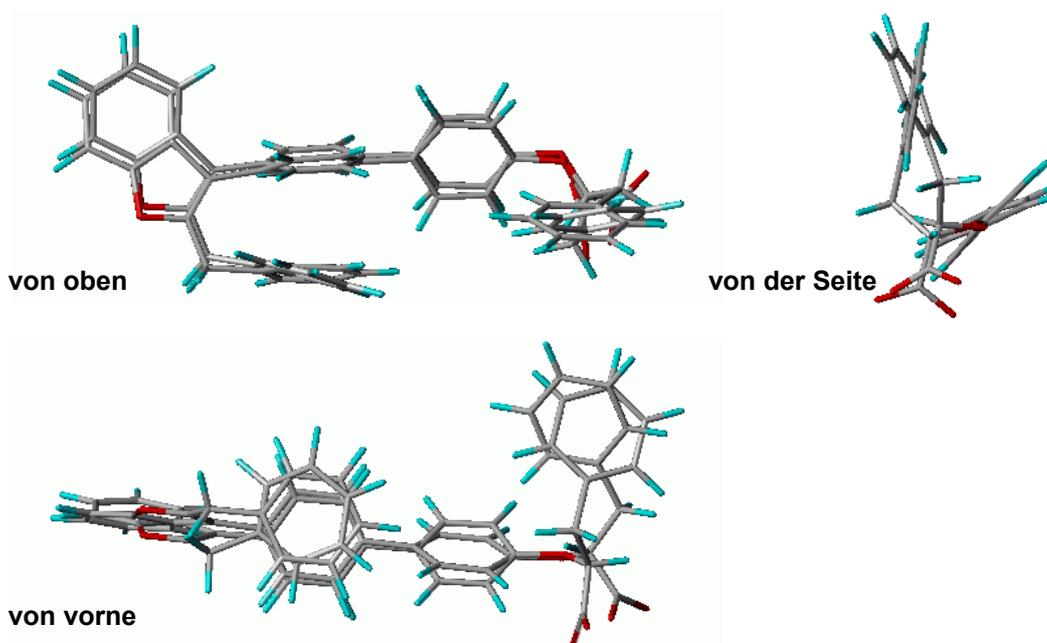


Abbildung 5-26: Überlagerung von R- und S-Enantiomer

Um den Einfluss der Racemate zu untersuchen, wurde der Datensatz ohne diese Verbindungen berechnet und für die chiralen Verbindungen wurde die Aktivität für beide Enantiomeren vorhergesagt. Dazu wurden die Verbindungen im deprotonierten Zustand mit den verschiedenen Sonden analysiert. In Abbildung 5-27 ist die vorhergesagte Aktivität der R- und S-Enantiomeren für die Phosphatsonde gegeneinander aufgetragen, Abbildung 5-28 zeigt die Abweichung der vorhergesagte von der tatsächlichen Aktivität für diese

Verbindungen. Mit einer Ausnahme werden die $\log IC_{50}$ -Werte aller Enantiomerenpaare mit einem Unterschied von kleiner als 0,15 vorhergesagt. Die aktiveren Verbindungen werden aber - für beide Enantiomere gleich - mit geringfügig zu niedriger Aktivität vorhergesagt. Mit den drei anderen Sonden wurden vergleichbare Ergebnisse erzielt.

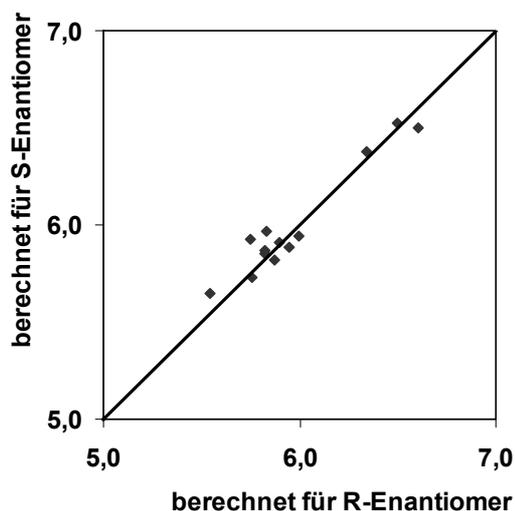


Abbildung 5-27: Berechnete $\log IC_{50}$ -Werte für die R- und S-Enantiomere des Testsatzes

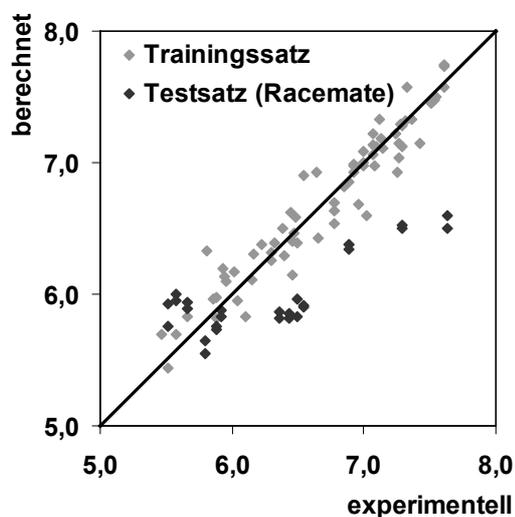


Abbildung 5-28: Vorhergesagte $\log IC_{50}$ -Werte für die Racemate des Testsatzes

5.4 Dockingbasierte 3D-QSAR-Modelle

Neben den ligandbasierten Überlagerungen sollten protein- oder dockingbasierte Überlagerungen erzeugt und die daraus resultierenden 3D-QSAR-Modelle mit den ligandbasierten Verfahren verglichen werden. Das besondere an diesem Vorgehen ist, dass auf diese Weise zwei grundsätzlich verschiedene und meist nicht gemeinsam angewendete Methoden kombiniert werden [11][12][13][14]. Die beschriebenen Verfahren der Konformationsanalyse, Superpositionierung und 3D-QSAR-Analyse werden üblicherweise dann angewendet, wenn die dreidimensionale Struktur des Proteins nicht bekannt ist. Wenn die Proteinstruktur aber durch Röntgendiffraktometrie oder NMR-Spektroskopie aufgeklärt worden ist, dann werden normalerweise proteinbasierte Methoden wie Dockingverfahren, Moleküldynamiksimulation und Berechnung der Bindungsenergie zur Entwicklung neuer Liganden verwendet.

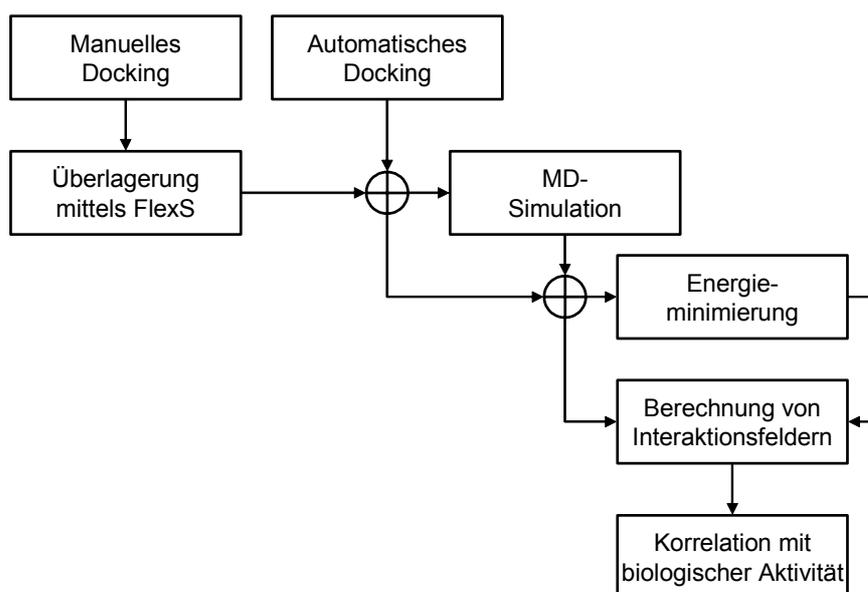


Abbildung 5-29: Proteinbasierte 3D-QSAR-Analyse, Vorgehensweise

Die in Abbildung 5-29 dargestellte Kombination proteinbasierter Dockingverfahren (automatisches und manuelles Docking, Energieminimierung, MD-Simulation) mit der ligandbasierten 3D-QSAR-Analyse (Berechnung von Interaktionsfeldern und Korrelation mit der biologischen Aktivität) soll die Vorteile beider Verfahren nutzen.

5.4.1 Docking in die Kristallstruktur

Die beiden Dockingprogramme FlexX und Autodock, die mit grundsätzlich verschiedenen Verfahren arbeiten (siehe Kapitel 4.4), wurden neben dem manuellen Docking verwendet, um die aktivsten Verbindungen des Datensatzes (104 und 142) in die unterschiedlichen Kristallstrukturen der PTP1B einzupassen. Dazu wurde einerseits die Struktur 1BZH mit offenem WPD-Loop ausgewählt, die ein relativ großes, zyklisches Peptid als Inhibitor enthält, andererseits wurde die Struktur 1ECV verwendet, die im aktiven Zentrum mit geschlossenem WPD-Loop einen relativ kleinen Inhibitor enthält. Mit Autodock und FlexX ließen sich für diese beiden Liganden und weitere Substanzen des Datensatzes zunächst keine Positionierungen des Liganden im aktiven Zentrum finden, bei denen die Säurefunktion mit der pTyr-Bindungsstelle interagiert. Dies war auch mit anderen Kristallstrukturen der PTP1B nicht möglich. Die Ursache für die Probleme, die Liganden zu docken, liegen in erster Linie in den großen strukturellen Unterschieden zwischen den Liganden, mit denen die Kristallstrukturen gewonnen wurden, und den Inhibitoren des untersuchten Datensatzes, die zum Docking verwendet wurden. In den verfügbaren Kristallstrukturen sind entweder sehr kleine Liganden vom OBA-Typ oder große peptidische Liganden mitkristallisiert. Beide Gruppen weisen keinerlei Ähnlichkeit zu den relativ großen, lipophilen, zu dockenden Liganden auf. Man kann davon ausgehen, dass vor allem die Unterschiede in der Größe dazu führen, dass sich das aktive Zentrum an die Inhibitoren anpasst, und die zu Verfügung stehenden Kristallstrukturen für die Liganden nicht ausreichend Platz im aktiven Zentrum besitzen. Aufgrund dieser Unterschiede war es auch nicht möglich, in einem manuellen Docking mit der Überlagerung der Liganden auf die in den Kristallstrukturen vorhandenen Inhibitoren oder Substrate zu beginnen.

Um diese Probleme zu lösen, wurde versucht, für das aktive Zentrum des Enzyms weitere Konformationen zu generieren, die für das Docking besser geeignet sind (siehe Abschnitt 5.4.2). Danach war es möglich mit Autodock plausible Dockingergebnisse zu produzieren (siehe 5.4.4).

5.4.2 Generierung von Enzymkonformationen

Nachdem erste Versuche erfolglos waren, die Liganden in das aktive Zentrum der PTP1B zu docken, sollten mit Hilfe einer Moleküldynamiksimulation weitere Konformationen für das Enzym gewonnen werden. In Abbildung 5-30 ist die Vorgehensweise dargestellt. Aus der Kristallstruktur 1ECV, die einen geschlossenen WPD-Loop besitzt, wurde der enthaltene Inhibitor vom OBA-Typ entfernt und stattdessen der peptidische Inhibitor aus der Struktur 1BZH (WPD-Loop offen) eingefügt.

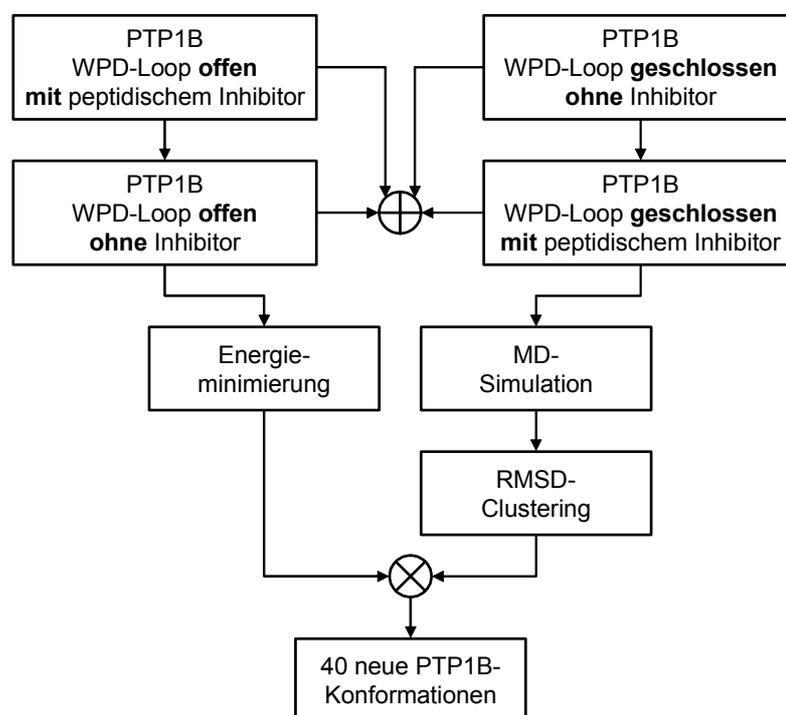


Abbildung 5-30: Erzeugung von neuen Proteinkonformationen, Vorgehensweise

Beide Strukturen wurden jeweils ohne Inhibitor und mit dem peptidischen Inhibitor durch ein Conjugate-Gradient-Verfahren bis zu einem Konvergenzkriterium von 1,0 kcal/Å in ihrer Geometrie optimiert. Anschließend wurden die Strukturen einer Moleküldynamiksimulation bei vergleichbaren Bedingungen unterzogen, in der über 40 ps in jeder Picosekunde eine Struktur herausgeschrieben wurde. Die Simulation wurde mit dem Programm Discover bei einer Temperatur von 350 K durchgeführt. Dadurch sollten weitere Konformationen des aktiven Zentrums generiert werden. Ins Besondere sollte dadurch das aktive Zentrum der geschlossenen Struktur aufgeweitet werden, um den größeren Inhibitoren mehr

Platz zur Verfügung zu stellen. Mit Ausnahme der Loops des aktiven Zentrums (Asn44 bis Ser50, Leu110 bis Cys121, Tyr176 bis Ser187, His214 bis Ser222 und Leu260 bis Thr263) wurde das Protein dabei festgehalten, um die benötigte Simulationszeit zu verkürzen. Die Strukturen aus den Simulationen sind in Abbildung 5-31 exemplarisch für die Struktur 1ECV ohne Inhibitor dargestellt.

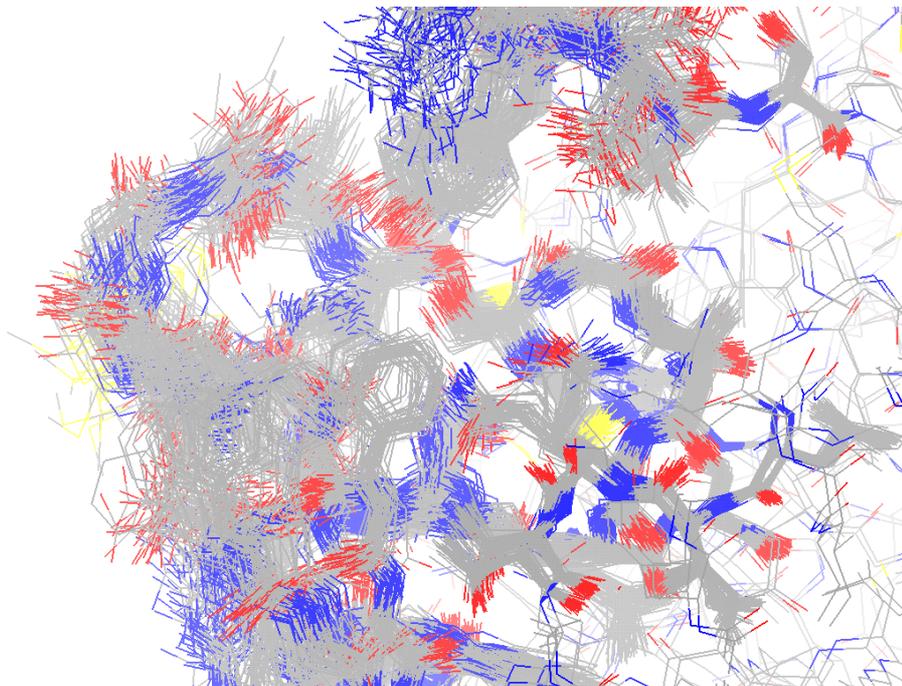


Abbildung 5-31: Proteinstrukturen aus der Dynamik

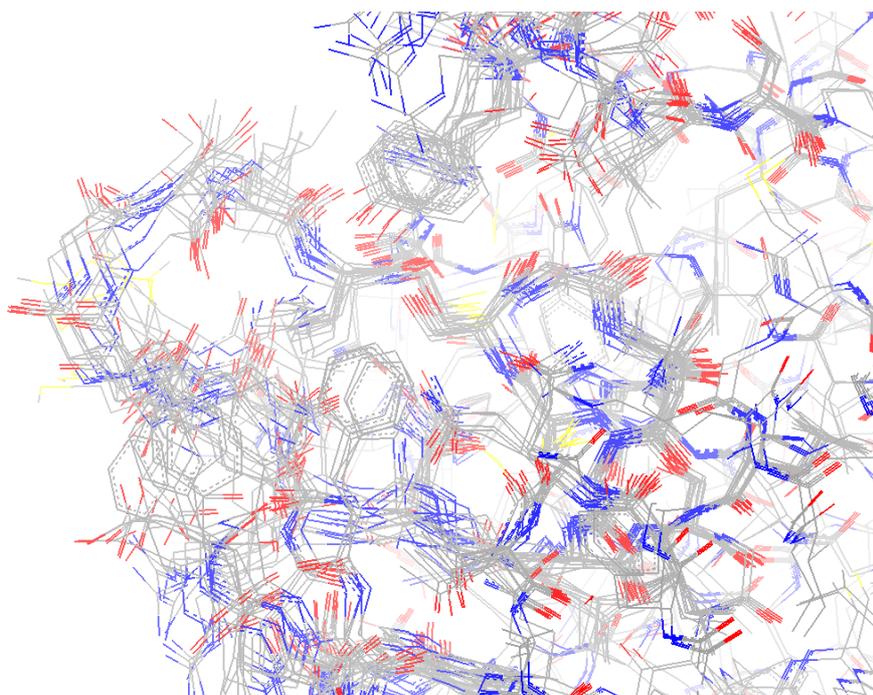


Abbildung 5-32: Neue Proteinstrukturen für das Docking nach dem Clustering

Die erhaltenen Strukturen wurden mit dem Programm NMRClust in Gruppen ähnlicher Strukturen zusammengefasst, so dass aus den ursprünglich 160 Strukturen aus der MD-Simulation 20 unterschiedliche Enzymkonformationen hervorgingen (siehe Abbildung 5-32). Diese wurden gemeinsam mit den vier Strukturen aus den Energieminimierungen als neue Zielstrukturen für das Docking verwendet.

Beim Docking mit dem Programm Autodock wurde hier neben der Verbindung 80, für die der Bindungsmodus bekannt ist, auch die zweifach bromierte Verbindung 104 verwendet, die eine größere Aktivität besitzt und aufgrund der beiden Bromatome mehr Platz im aktiven Zentrum benötigt. Unter Verwendung einer Konformation aus der MD-Simulation der Kristallstruktur 1ECV wurde so der in Abbildung 5-33 dargestellte Komplex mit der Verbindung 104 gefunden, der die in der Literatur beschriebenen Interaktionen [258] aufweist.

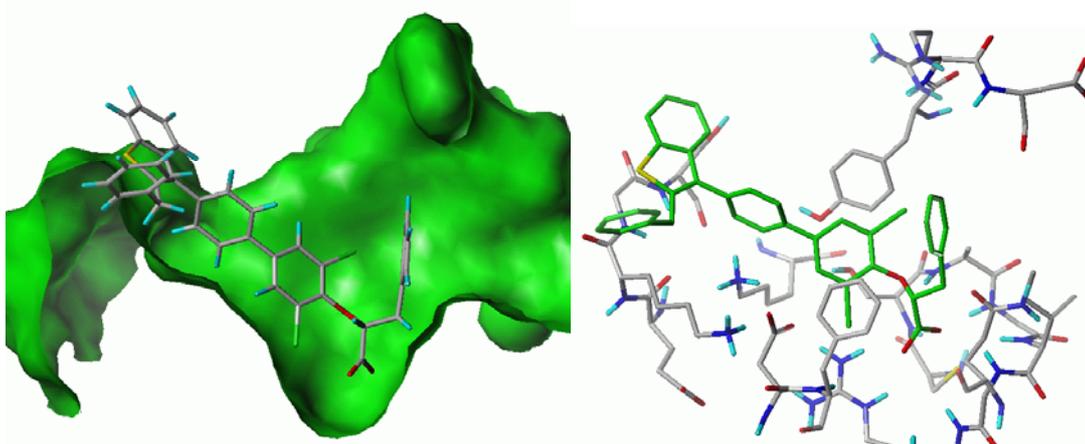


Abbildung 5-33: Docking der Verbindung 104

Der Ligand bindet mit der Säurefunktion im aktiven Zentrum, und der Abstand zum positiv geladenen Arg221 ist in etwa so groß, dass die Wasserstoffbrücken über Wassermoleküle vermittelt werden müssen. Das Biphenyl-Ringsystem wird zwischen Tyr46 und Phe182 in einer Art Sandwich-Anordnung festgehalten, wobei aber der Phenylring von Phe182 zum Phenylring des Inhibitors einen Winkel von 45° bildet. Zusätzlich sind hydrophobe Wechselwirkungen und Kation- π -Interaktionen zwischen dem Benzothiophen und den Seitenketten von Lys120 möglich. Der Komplex wurde einer weiteren Energieminimierung mit

obigen Bedingungen unterzogen und dann für die Erzeugung der proteinbasierten Überlagerung verwendet.

Mit dem Programm FlexX war es auch dann noch nicht möglich, die Inhibitoren zu docken. Die Ursache dafür liegt in der Art und Weise, wie FlexX die Liganden dockt. Ausgehend von einem Startfragment, das zuerst positioniert wird, baut das Programm den kompletten Liganden in die Bindungstasche, wobei zusätzlich mögliche Wassermoleküle platziert und berücksichtigt werden können. Als Startfragment wird sinnvoller Weise die Carbonsäurefunktion ausgewählt. Diese wird aber so tief im aktiven Zentrum gedockt, dass es anschließend nicht mehr möglich ist, die beide sperrigen aromatischen Ringsysteme (Biphenyl und Phenyl der Phenylmilchsäure) in der Bindungstasche unterzubringen. Auch wenn zur Positionierung des Startfragmentes Wassermoleküle berücksichtigt werden sollen, wird die Säurefunktion als Startfragment genau so in der Bindungstasche platziert, als ob keine Wassermoleküle Berücksichtigung finden. Das liegt daran, dass das Programm in diesem Fall die über das Wasser vermittelten ionischen Interaktionen zwischen der Carbonsäure und Arg221 wesentlich schlechter bewertet als die direkte Interaktion und diese deshalb nicht berücksichtigt.

Die von FlexX berechneten Wechselwirkungspositionen für Wassermoleküle befinden sich in drei eng umgrenzten Bereichen im aktiven Zentrum und stimmen mit Berechnungen von GRID-Feldern für Wasser überein. An allen drei Positionen wurde jeweils ein Wassermolekül eingefügt, und anschließend wurde in einer Energieminimierung die Lage der Wasserstoffatome optimiert. Für das weitere Docking mit FlexX wurde dann das Protein ohne Wasser, mit einem, zwei oder drei Wassermolekülen untersucht, so dass also insgesamt acht unterschiedliche Formen betrachtet wurden. Damit war es möglich, für die untersuchten Liganden auch mit FlexX eine korrekte Positionierung zu finden. Dazu war es aber notwendig, bei unterschiedlichen Liganden jeweils eine andere Anordnung oder Anzahl von Wassermolekülen in der Bindungstasche des Proteins zu verwenden. Da aber die am besten bewerteten Dockings nicht dem erwarteten Bindungsmodus entsprechen, und eine automatische Auswahl der richtigen Konformation nicht möglich ist, eignet sich dieses Vorgehen nicht

für das Docking vieler Liganden und ist somit für eine Datenbanksuche völlig ungeeignet.

Vergleicht man die Wechselwirkungsfelder für die Proteinstruktur, in die ein Docking möglich war, mit den Kristallstrukturen in einer PCA-Analyse (Abbildung 5-34), so stellt man fest, dass sich die Struktur durch die Moleküldynamiksimulation relativ stark verändert hat.

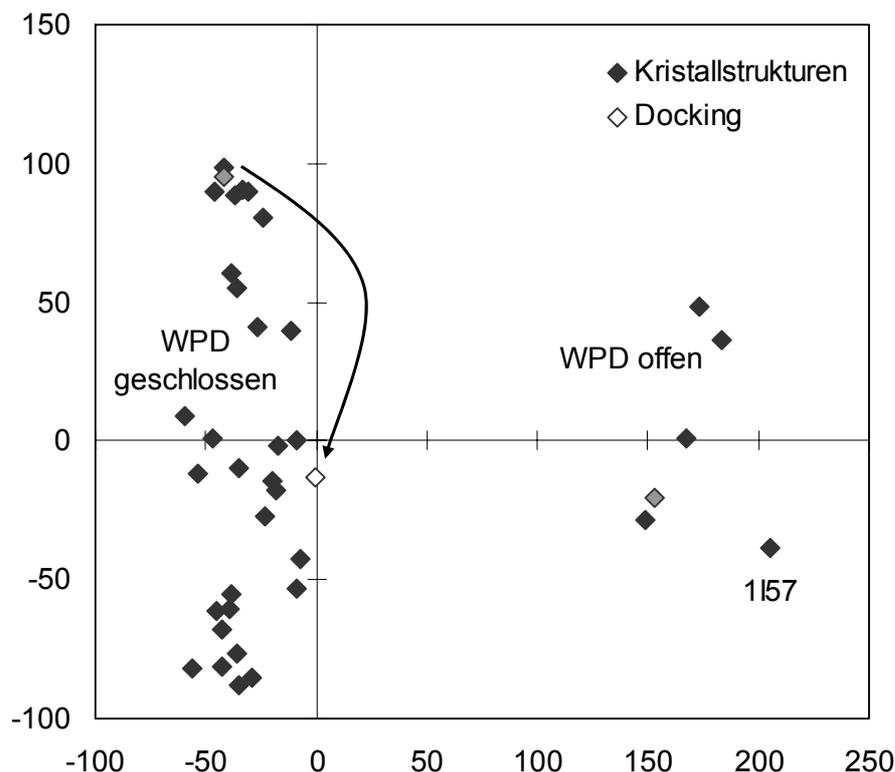


Abbildung 5-34: PCA-Vergleich der Struktur für das Docking mit Kristallstrukturen

Der WPD-Loop ist zwar in der geschlossenen Konformation geblieben, aber die Aminosäuren selbst haben sich zum Teil erheblich bewegt (Abbildung 5-35). Die Lage der Struktur in der PCA-Analyse lässt vermuten, dass in die in der Nähe liegenden Kristallstrukturen das Docking möglich sein könnte. Deshalb wurden diese Strukturen, die den PDB-Einträgen 1PTU, 1KAK, 1KAV, 1G7G oder 1JF7 entsprechen, ebenfalls für das Docking herangezogen. Es war jedoch nicht möglich, die Liganden erfolgreich in diese Proteinstrukturen zu docken.

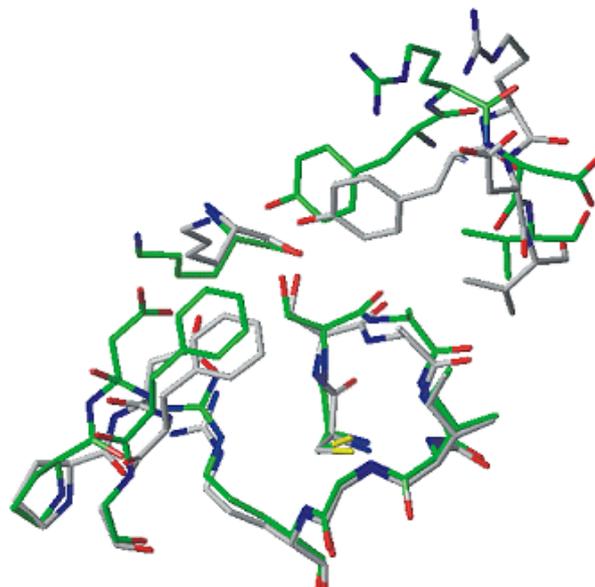


Abbildung 5-35: Vergleich der Struktur für das Docking (grün) mit der Kristallstruktur 1ECV (weiß)

Der Vergleich des Volumens des aktiven Zentrums der einzelnen Strukturen zeigt, dass in den Kristallstrukturen weniger Platz ist als in der Struktur nach der Dynamiksimulation, in die das Docking möglich war. In Abbildung 5-36 erkennt man nicht nur, dass sich einzelne Aminosäuren an der Oberfläche stark bewegen können (rechts oben und ganz links), sondern auch, dass das aktive Zentrum in den Kristallstrukturen wesentlich enger ist als in der Struktur nach der MD-Simulation (Abbildung 5-36 Mitte).

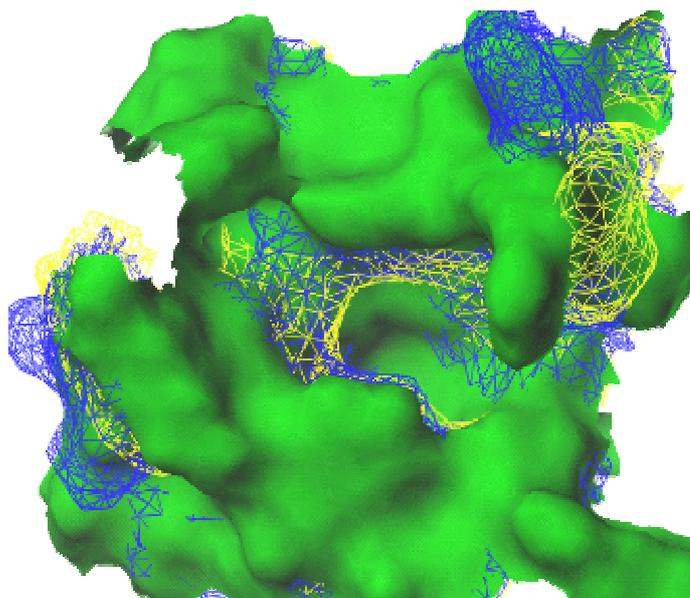


Abbildung 5-36: Vergleich des Platzes im aktiven Zentrum der Struktur für das Docking (grün) und zweier Kristallstrukturen (gelb und blau)

Der geringere Platz in der Bindungsstelle ist zwar ausreichend für viele der kleinen Inhibitoren wie die FOMT-Gruppe von peptidischen Inhibitoren und den pTyr-Rest der Substrate, aber um gleichzeitig die Biphenylgruppe und den Phenylring der Phenylmilchsäure aufzunehmen, ist nicht genügend Platz vorhanden, so dass eine Anpassung notwendig ist.

5.4.3 Optimierung des Programms Autodock

Das Programm Autodock wurde entwickelt, um für eine kleine Anzahl von Protein-Ligand-Paaren möglichst gute Komplexgeometrien zu erzeugen. Es verwendet eine Vielzahl einzelner Unterprogramme und benötigt während der Ausführung manuelle Eingaben durch den Anwender, so dass es für ein automatisches Docking größerer Datensätze oder für die Anwendung bei Datenbanksuchen in dieser Form nur bedingt geeignet ist. Daher wurde ein Programm erstellt, das ohne manuelle Benutzereingaben eine rationelle und automatisierte Nutzung von Autodock auch bei großen Datensätzen erlaubt. Ein Ablaufschema der einzelnen Teilprogramme die zusammengefasst werden sollten ist in Abbildung 5-37 dargestellt. Als Ausgangspunkt für die Berechnungen dienen Dateien im Sybyl-mol2-Format für den Liganden und das Enzym. Die mol2-Datei des Enzyms wird zuerst in eine besondere bdbqs-Datei umgewandelt, die zusätzlich zum pdb-Format, von dem sie abgeleitet ist, noch Partiaalladungen und Solvationsparameter enthält. Der Ligand, der ebenfalls im mol2-Format vorliegt, muss zuerst optimiert werden, da das Programm *autotors*, das den Liganden in das pdbq-Format umwandelt und die drehbaren Bindungen festlegt, spezielle Anforderungen an das mol2-Format stellt.

Sowohl das Protein als auch der Ligand werden im Anschluss mit dem Programm *check-qs* auf korrekte Ladungen geprüft, und das Ergebnis wird in einer err-Datei ausgegeben, so dass dieses Ergebnisprotokoll zur Fehlerüberwachung herangezogen werden kann.

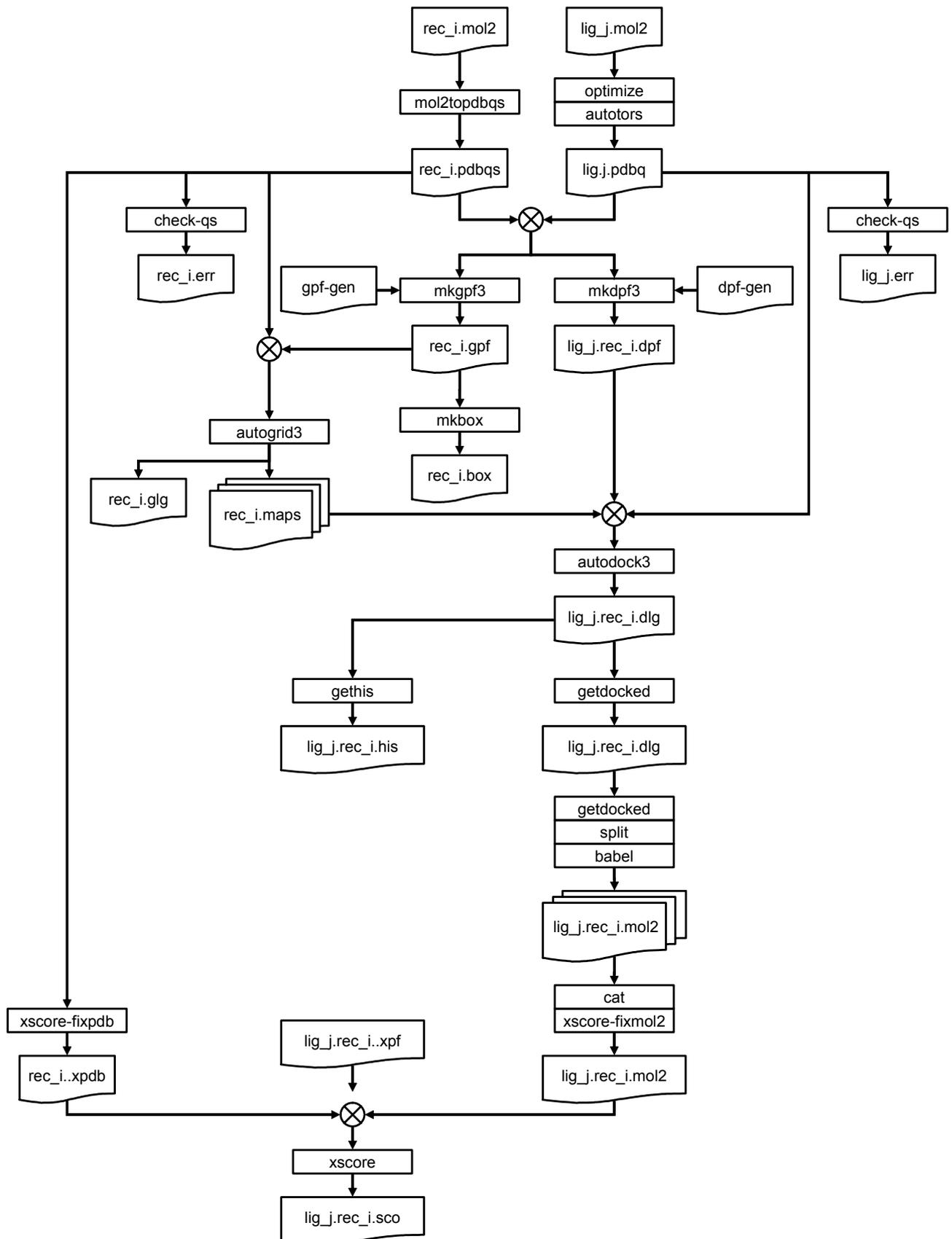


Abbildung 5-37: Ablauf des Programms Autodock

Aus der *pdbq*-Datei des Liganden und der *pdbqs*-Datei des Proteins wird dann mit Hilfe des Programms *mkgpf3* eine *gpf*-Datei erstellt, die die Parameter für die Berechnung der GRID-Felder enthält. Die eigentlichen Parameter werden vom Programm in der Datei *gpf-gen* geliefert und müssen in der erstellten *gpf*-Datei angepasst werden. Dieser Schritt muss ebenso wie die anschließende Berechnung der GRID-Felder mit *autogrid3* nicht für jeden einzelnen Liganden durchgeführt werden, sondern nur für jede Struktur oder Konformation des Proteins. Neben den *map*-Dateien für die einzelnen Atomtypen im Liganden, die das Protein in den Dockingberechnungen in Form von GRID-Feldern repräsentieren, wird eine *glg*-Datei herausgeschrieben, die den Verlauf der Berechnung dokumentiert.

Durch das Programm *mkdopf3* wird ebenfalls unter Verwendung der *pdbq*-Datei des Liganden und der *pdbqs*-Datei des Proteins eine *dpf*-Datei mit den Parametern für das Docking erstellt. Auch bei diesem Schritt, der für jede Protein-Ligand-Kombination durchgeführt werden muss, liefert das Programm die benötigten Parameter (*dpf-gen*-Datei). In der nicht automatisierten Form muss der Benutzer die *dpf*-Datei für das jeweilige Docking anpassen.

In dem ganzen Prozess benötigt das Docking mit *autodock3* selbst die meiste Rechenzeit. Das Programm verwendet neben der *dpf*-Datei die *map*-Dateien und die *pdbq*-Datei des Liganden. Die aus dem Docking resultierende *dlg*-Datei enthält den Verlauf der Berechnung mit *autodock3* und die Ergebnisse in Form von Koordinaten des Liganden in den gedockten Konformationen und Scores für die einzelnen Komplexe. Um diese Ergebnisse aus den *dlg*-Dateien zu extrahieren, können die Programme *getdocked* und *gethis* verwendet werden. Dabei werden die Koordinaten für jeden Komplex in eine separate Datei im *pdb*-Format gespeichert, die aber keine Bindungen zwischen den Atomen enthält. Die *pdb*-Dateien müssen deshalb vor der Verwendung in weiteren Berechnungen mit dem separaten Programm *babel* in *mol2*-Dateien, die auch Bindungen enthalten, umgewandelt werden.

An das Docking anschließend können die Komplexe mit dem Programm X-Score bewertet werden. Dazu ist es aber notwendig die *pdbqs*-Datei des Proteins mit *xscore-fixpdb* und die *mol2*-Dateien der Liganden mit *xscore-*

fixmol2 zu überarbeiten. Auch die Parameter für das Scoring müssen in einer zusätzlichen Eingabedatei bereitgestellt werden.

Ziel der Optimierung war es, möglichst alle der benötigten Programme aus Abbildung 5-37 automatisiert ablaufen zu lassen und dabei Benutzereingaben während des Programmablaufs zu vermeiden. Dazu wurde in der Programmiersprache PERL ein Skript erstellt (Abbildung 5-38), das alle benötigten Programme nacheinander aufruft und den Ablauf in einer Protokolldatei dokumentiert.

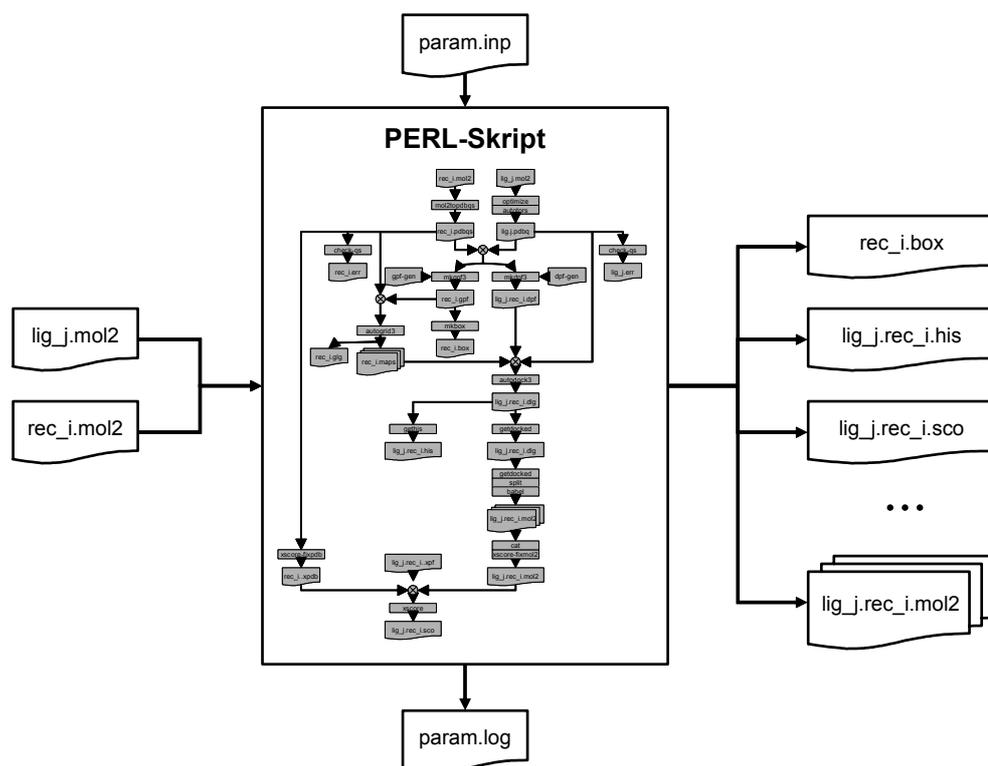


Abbildung 5-38: Automatisiertes Docking

Der Quellcode des Skriptes ist im Anhang 9.2.1 angefügt. Der gesamte Ablauf wurde soweit automatisiert, dass aus einem großen Datensatz alle Liganden automatisch nacheinander gedockt werden können. Ebenso ist es möglich, die Berechnung der GRID-Felder für mehrere Strukturen oder Konformationen des Proteins durchzuführen. Des Weiteren können auch mehrere Liganden in mehrere Proteine gedockt werden. Auch das Extrahieren der Ergebnisse, die Umwandlung der Koordinaten der Liganden und das Scoring mit X-Score wurden in das Skript mit aufgenommen. Damit Eingaben während des Programmablaufs nicht notwendig sind, wurden eine Datei erstellt, die die entscheidenden Parameter für die Steuerung des Ablaufs, die Dateinamen für

die Strukturen der einzelnen Proteine und Liganden sowie die Parameter für die Berechnung der GRID-Felder, das Docking und das Scoring enthält. Ein Beispiel für diese Parameterdatei befindet sich im Anhang 9.2.2.

Die Parameter für das Docking mittels Monte Carlo Simulated Annealing in *autodock3*, wie die Anzahl der Läufe, Zyklen und Schritte, wurden für das virtuelle Screening soweit erniedrigt, dass die benötigte Zeit zum Docking eines Liganden nur noch wenige Minuten dauert. Obwohl sich dadurch die Geometrie der Komplexe etwas verschlechtert, sind die Ergebnisse für eine Datenbanksuche von ausreichend guter Qualität.

5.4.4 Proteinbasierte Überlagerungen

Für die proteinbasierten Überlagerungen wurden zum einen die Verbindungen des Datensatzes, wie bei der ligandbasierten Überlagerung beschrieben, auf die gedockte Konformation von 104 überlagert (Abbildung 5-39).

Zum anderen wurde mit dem Programm Autodock ein Docking für alle Liganden durchgeführt. Die auf Verbindung 104 aus dem Komplex mit der PTP1B-Struktur mit dem Programm FlexS überlagerten anderen Verbindungen des Datensatzes wurden einer weiteren Optimierung mit Discover unterzogen.

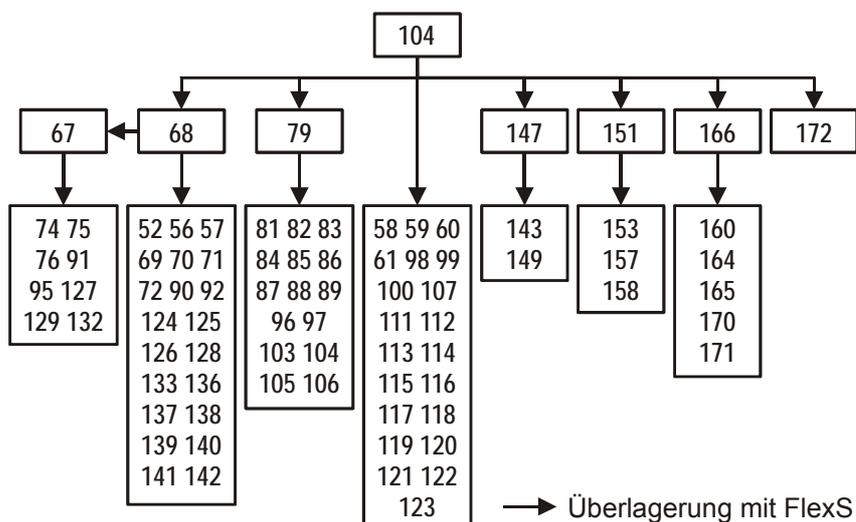


Abbildung 5-39: Vorgehen bei der FlexS-Überlagerung auf die gedockte Verbindung 104

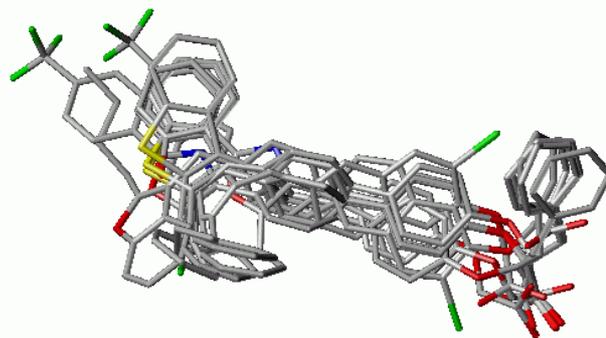


Abbildung 5-40: Automatisches Docking von 104 und Überlagerung der anderen Verbindungen mit FlexS; dargestellt sind nur die Verbindungen 67, 68, 79, 80, 82, 147, 151, 166, und 172

Vergleicht man diese Überlagerung mit den ligandbasierten Überlagerungen, dann stellt man erwartungsgemäß fest, dass zwischen diesen keine nennenswerten Unterschiede bestehen, da die Konformation der Template in beiden Fällen sehr ähnlich ist und die Überlagerung auf den Grundgerüsten basiert. Die Seitenketten und die Säurefunktion haben keinen nennenswerten Einfluss (siehe Abbildung 5-40). Daneben wurde auch eine MD-Simulation der Komplexe mit anschließender Geometrieoptimierung untersucht, die schlechtere Ergebnisse bezüglich der Superpositionierung und der 3D-QSAR-Modelle erzeugte. Auch die Minimierungen mit flexiblem und fixiertem aktiven Zentrum brachten keine Verbesserung der Überlagerung, die Modelle waren mit der nicht optimierten Überlagerung vergleichbar.

Die große Ähnlichkeit der erstellten Modelle äußert sich auch in der 3D-QSAR-Analyse, die für alle diese Überlagerungen vergleichbare Ergebnisse liefert. Eine Ausnahme bildet nur die Überlagerung, in der die Liganden einer Dynamiksimulation in der Bindungstasche unterzogen wurden. Diese Überlagerung liefert wesentlich schlechtere Ergebnisse bezüglich des r^2 -Wertes, des q^2 -Wertes und des SDEP-Wertes wie Abbildung 5-41 und Abbildung 5-42 zeigen.

Basierend auf dem Komplex aus dem Inhibitor 104 und der aus der MD-Simulation gewonnenen PTP1B-Struktur (siehe 5.4.2) wurden die anderen Verbindungen des Datensatzes mit der optimierten Version des Programms Autodock gedockt.

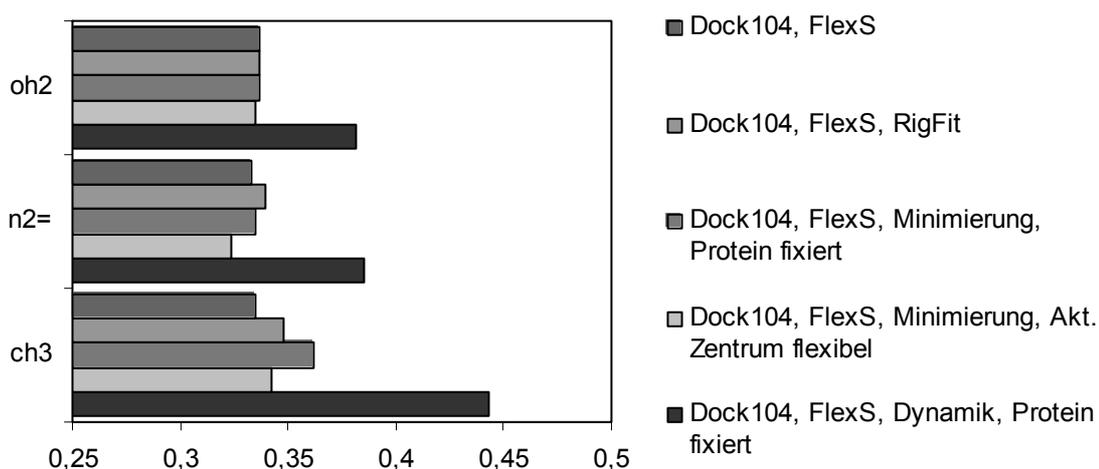


Abbildung 5-41: Vergleich der verschiedenen Minimierungsmethoden (SDEP-Werte)

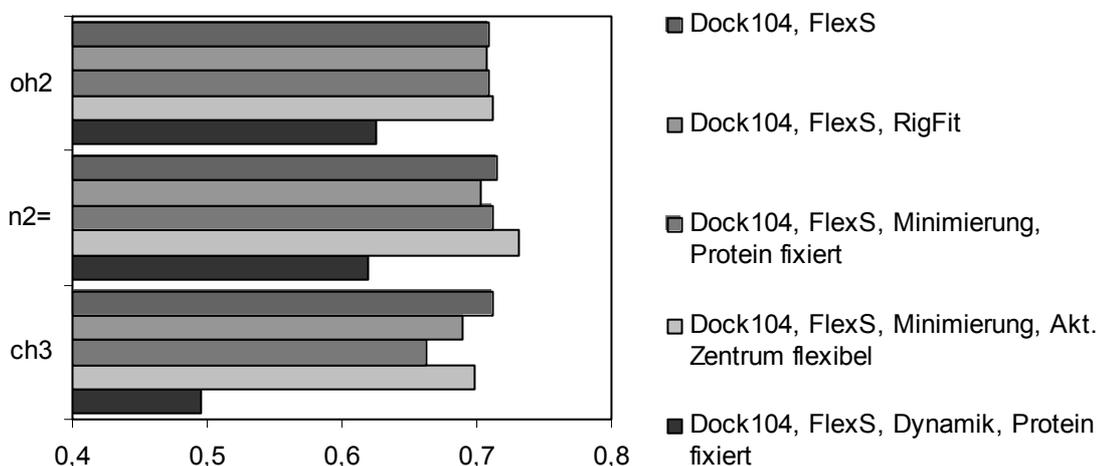


Abbildung 5-42: Vergleich der Minimierungsmethoden (q²-Werte, L20%O)

Aufgrund der Größe von zusätzlichen sperrigen Seitenketten war dies aber für einige Liganden nicht möglich, deshalb wurden diese mit Hilfe von FlexS auf die jeweils ähnlichste gedockte Verbindung überlagert. Alle gedockten und überlagerten Verbindungen wurden anschließend als Komplex mit der fixierten PTP1B minimiert.

Wie in Abbildung 5-43 zu sehen ist, zeigt diese Überlagerung ein ganz anderes Bild als bei der ligandbasierten Analyse. Hier sind jetzt die einzelnen Atome der Grundgerüste nicht mehr optimal überlagert, dafür werden aber die Säurefunktionen und auch die Seitenketten der einzelnen Liganden besser überlagert. Vor dem Hintergrund der Proteinumgebung scheint dieses Ergebnis sinnvoller zu sein als die ligandbasierte Überlagerung.

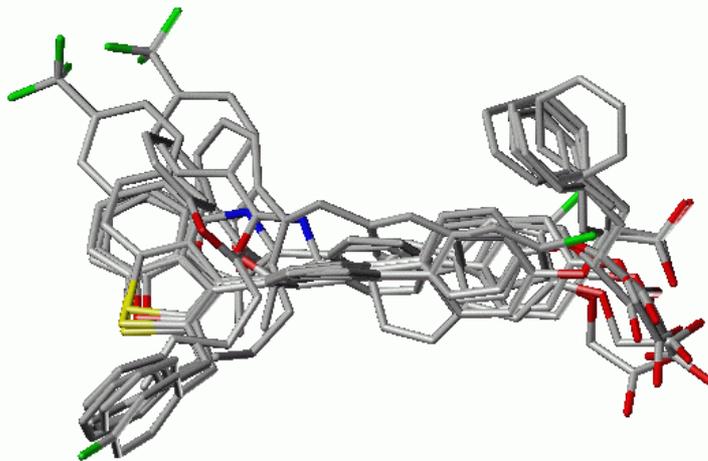


Abbildung 5-43: Automatisches Docking aller Verbindungen; dargestellt sind die Verbindungen 67, 68, 79, 80, 82, 147, 151, 166, und 172

Als Beispiel sind in Abbildung 5-44 zwei Inhibitoren mit gleichem Grundgerüst aber unterschiedlichem Säure-Rest dargestellt und mit einer Oberfläche umgeben, die das elektrostatische Potential visualisiert. Man kann erkennen, dass in der abgebildeten proteinbasierten Überlagerung das negative elektrostatische Potential, das durch die Säurefunktion des Phenylmilchsäure-Restes verursacht wird (Verbindung 68, Gitter), und das Potential des Phenols (Verbindung 52, Fläche) im gleichen Raumsegment liegen.

Bei den ligandbasierten Überlagerungen liegt das Sauerstoffatom des Phenols aber an der Position des Ethersauerstoffs von Verbindung 68, an der sich kein entsprechendes Potential befindet.

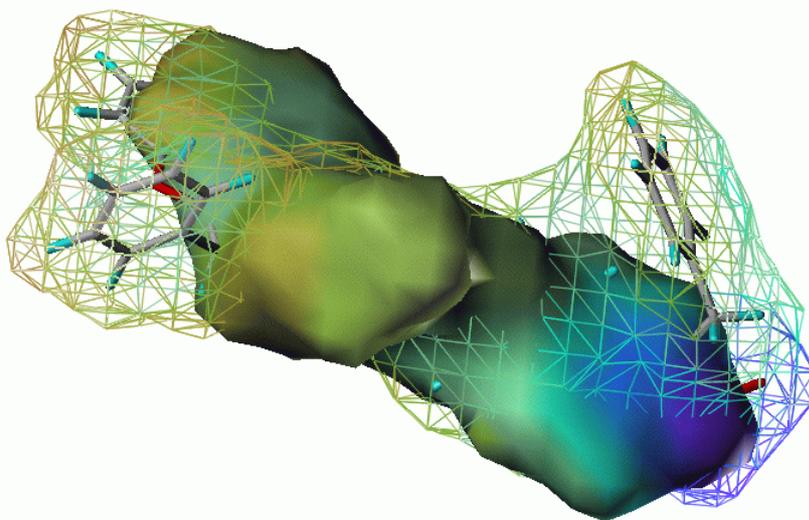


Abbildung 5-44: Elektrostatisches Potential für die rezeptorbasierte Überlagerung der Verbindungen 52 (Oberfläche) und 68 (Gitter)

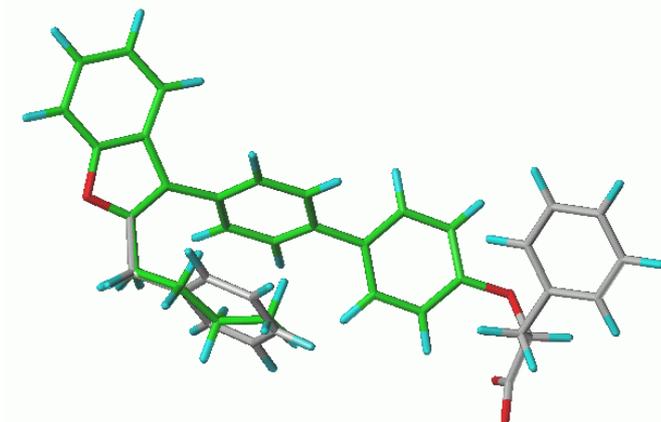


Abbildung 5-45: Ligandbasierte Überlagerungen der Verbindungen 52 (grün) und 68 (weiß)

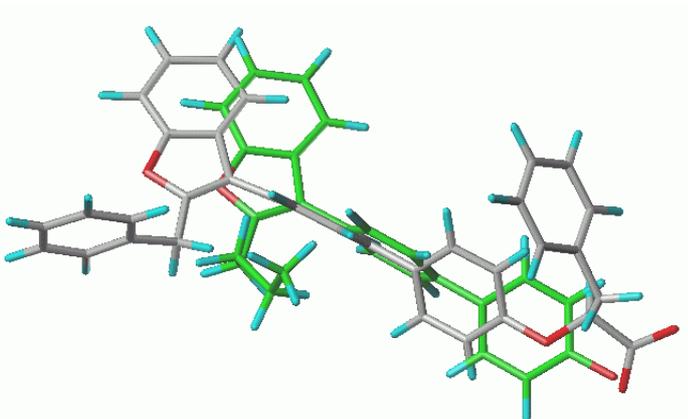


Abbildung 5-46: Proteinbasierte Überlagerungen der Verbindungen 52 (grün) und 68 (weiß)

Abbildung 5-45 zeigt die dazugehörige ligandbasierte und Abbildung 5-46 die proteinbasierte Überlagerung. Beim Vergleich der beiden Abbildungen, kann man die Unterschiede sehr gut erkennen. In den folgenden beiden Abbildungen, in denen die verschiedenen Überlagerungen für den kompletten Datensatz gegenübergestellt sind, erkennt man sehr gut, dass die dockingbasierte Überlagerung (Abbildung 5-47) im Bereich der lipophilen Ringsysteme eine viel größere Variabilität zulässt als die ligandbasierten Modelle (Abbildung 5-48). Da diese Ringsysteme am äußeren Rand der Bindungstasche und an der Oberfläche des Proteins liegen, wo genügend Platz ist und keine gerichteten Wechselwirkungen notwendig sind, sondern nur unspezifische Van-der-Waals-Interaktionen ausgebildet werden, passt dieses Modell sehr gut zu der Proteinumgebung.

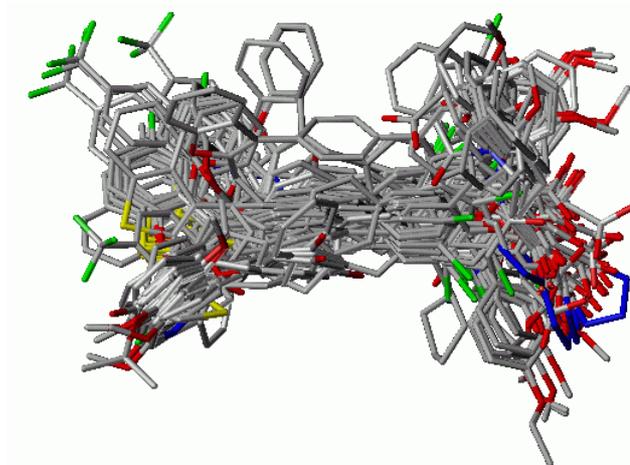


Abbildung 5-47: Automatisches Docking; dargestellt sind alle Verbindungen des Datensatzes

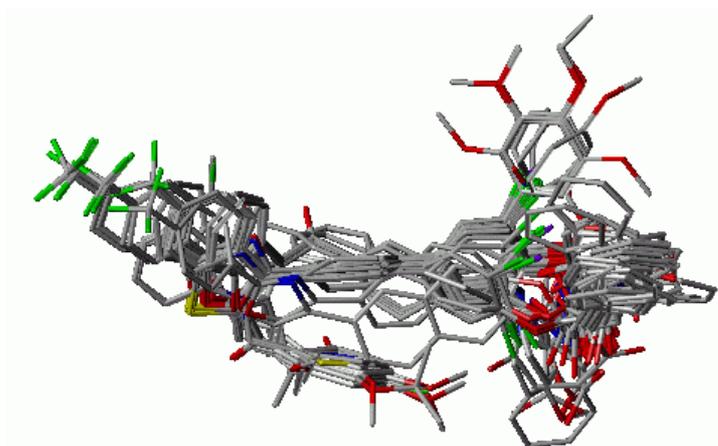


Abbildung 5-48: Ligandbasierte Überlagerung; dargestellt sind alle Verbindungen des Datensatzes

Die bei den ligandbasierten Modellen berechneten Überlagerungen der Ringsysteme sind bei der Betrachtung der Proteinumgebung nicht notwendig, denn die zuvor dargestellte Überlagerung der Säurefunktionen ist wesentlich wichtiger für die weiteren Untersuchungen, da deren Wechselwirkungspunkte mit dem Protein räumlich sehr beschränkt sind. Ebenso wie die ligandbasierten Modelle wurden auch für die proteinbasierten Überlagerungen 3D-QSAR-Modelle berechnet. Dabei wurden alle Liganden deprotoniert und die Racemate aus der Untersuchung ausgeschlossen, da es nach den unter 5.3.4.3 angestellten Überlegungen nicht sinnvoll wäre, diese Verbindung in die Analyse mit einzuschließen.

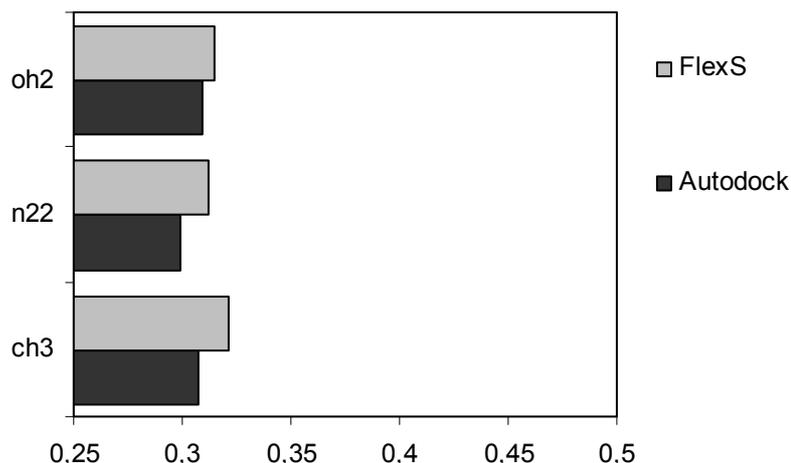


Abbildung 5-49: Vergleich von ligandbasierten (FlexS) und rezeptorbasierten (Autodock) Methoden anhand der SDEP-Werte

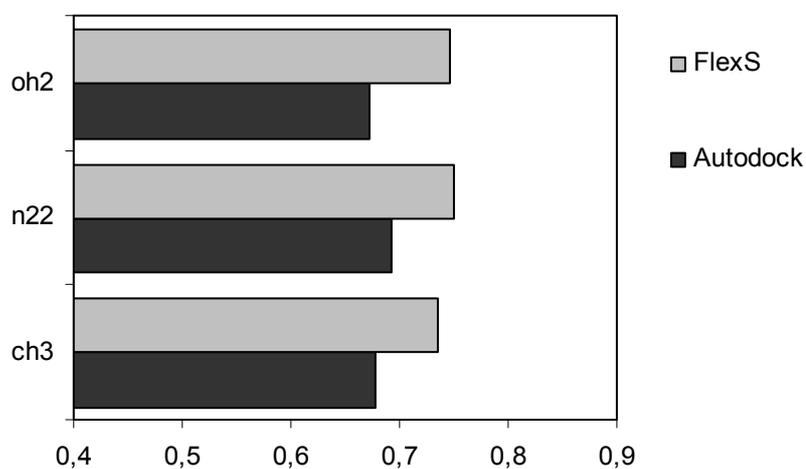


Abbildung 5-50: Vergleich von ligandbasierten und rezeptorbasierten Methoden anhand der q^2 -Werte (L20%O)

Wie in Abbildung 5-49 und Abbildung 5-50 dargestellt, führen die 3D-QSAR-Analysen zu ähnlich guten statistischen Ergebnissen (q^2 , r^2 , SDEP), obwohl die Überlagerungen sich voneinander unterscheiden.

Die Resultate der 3D-QSAR-Analysen weisen darauf hin, dass für den betrachteten Datensatz ligandbasierte und proteinbasierte Methoden gleich gut geeignet sind, da sich die q^2 -Werte nicht signifikant unterscheiden. Die gute Überlagerung der sauren Gruppen, die aus dem Docking hervorgeht, legt jedoch die weitere Verwendung des proteinbasierten Modells bei den Datenbanksuchen nahe.

5.4.5 Interpretation der PLS-Felder der 3D-QSAR-Modelle

Aus den Koeffizienten der PLS-Analyse können Informationen über die Aktivität eines Liganden gewonnen werden. Die Größe der Koeffizienten und deren Vorzeichen bestimmen den Beitrag der Energierterme zur Aktivität für jeden einzelnen Gitterpunkt des GRID-Feldes. Aufgrund der einfacheren und eindeutigeren Interpretierbarkeit wurde für diese Auswertung eine CoMSIA (Comparative Molecular Similarity Indices Analysis) in Sybyl durchgeführt. Diese liefert PLS-Koeffizienten für sterische, elektrostatische und hydrophobe Interaktionen, sowie für Akzeptoren und Donoren von Wasserstoffbrücken.

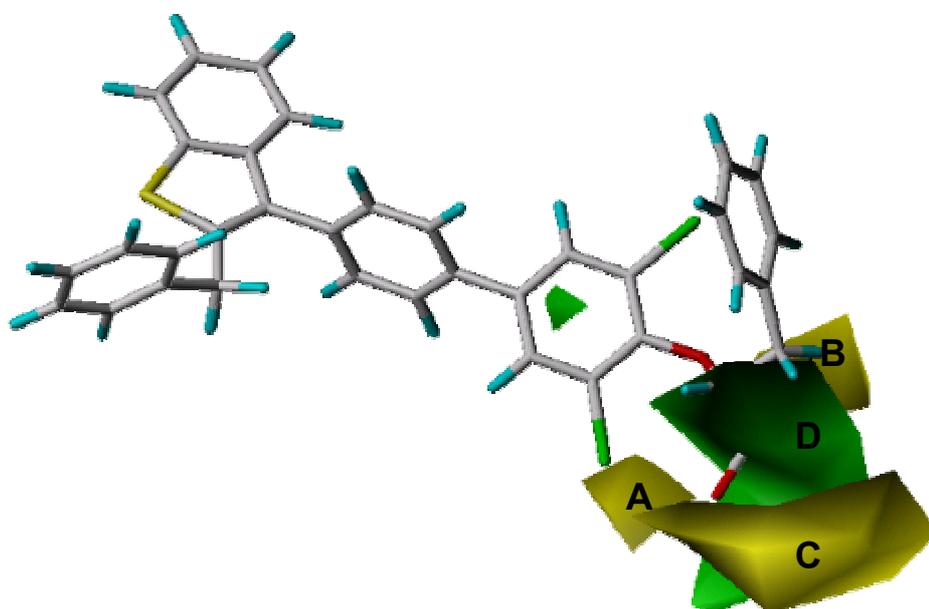


Abbildung 5-51: PLS-Felder der positiven (grün) und negativen (gelb) sterischen Interaktionen; die aktivste Verbindung ist als Vergleich dargestellt.

Die Analyse der sterischen PLS-Felder in Abbildung 5-51 zeigt, dass eine Erweiterung des Moleküls in der Position des Brom-Atoms (Feld A) ungünstig ist. Ebenso ungünstig ist eine Substitution am aliphatischen CH_2 -Kohlenstoff der Phenylmilchsäure (Feld B). Der Bereich, in dem sich die Säurefunktion befindet sollte in jedem Fall von den Verbindungen besetzt werden. Dieser Bereich kann zusätzlich noch nach unten und hinten (Feld D) erweitert werden. Für die Biphenylgruppe und das Benzothiofophen oder Benzofuran gibt es keine sterischen Einschränkungen. Diese Ergebnisse sind in Übereinstimmung mit der räumlichen Umgebung der Liganden im aktiven Zentrum.

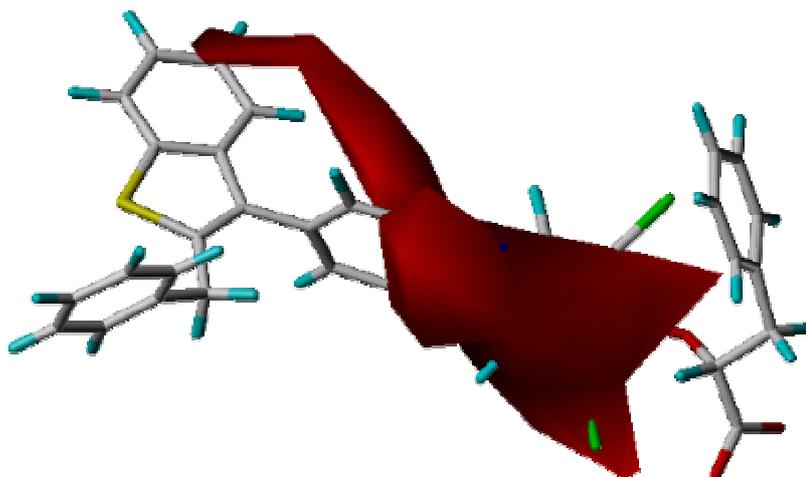


Abbildung 5-52: PLS-Felder der ungünstigen elektrostatischen Interaktionen; die aktivste Verbindung ist als Vergleich dargestellt

Abbildung 5-52 zeigt die PLS-Felder für elektrostatische Wechselwirkungen. Hier gibt es nur ungünstige Interaktionsfelder im Bereich des gesamten Moleküls. Dies ist nicht verwunderlich, da die aktiveren Verbindungen ausschließlich lipophilen Charakter besitzen und außer dem Sauerstoff des Benzofuran oder dem Schwefel des Benzothiophens und dem Ethersauerstoff keine Heteroatome enthalten. Nur einige der weniger aktiven Liganden enthalten weitere Heteroatome, die aber für die biologische Wirkung keinen Beitrag zu leisten scheinen. Ähnliche Ergebnisse werden auch mit Feldern für die Donoren von Wasserstoffbrücken erhalten, die ebenfalls über das ganze Molekül verteilt ungünstige Interaktionen aufzeigen. Für die Akzeptoren gibt es überhaupt keine PLS-Felder und somit sind keine Aussagen zu den Unterschieden zwischen den einzelnen Liganden möglich. Dies kann dadurch erklärt werden, dass die Liganden keine Donoren für Wasserstoffbrücken enthalten, die mit den Feldern der Akzeptoren interagieren könnten.

Für die hydrophoben Interaktionen gibt es vor allem in der Umgebung der Säurefunktion und des Ethersauerstoffs Felder für ungünstige Interaktionen, die darauf zurückzuführen sind, dass alle besonders aktiven Liganden hier entsprechende Funktionen aufweisen und keine hydrophoben Interaktionen zulassen.

5.4.6 Weitere getestete QSAR-Verfahren

Neben der Analyse der Interaktionsfelder wurden weitere proteinbasierte Verfahren zur 3D-QSAR-Analyse untersucht, bei denen in erster Linie eine Auswertung von Wechselwirkungsenergien zwischen Protein und Ligand stattfindet. Analysiert wurden die Scores aus dem Docking mit Autodock, die Wechselwirkungsenergien aus den Minimierungen mit Discover und PrGen, die Wechselwirkungen mit einzelnen Aminosäuren des Proteins (COMBINE-Methode 4.5.2) und die mit dem Programm X-Score berechneten Bindungskonstanten, die ebenfalls in erster Linie auf Wechselwirkungsenergien beruhen. Weder das Scoring mit Autodock, noch die Interaktionsenergien aus Discover oder PrGen führten zu einer Korrelation mit den gemessenen Aktivitäten der Verbindungen, so dass hierüber keine Vorhersage der Bindungsstärke oder eine Einteilung in aktive und inaktive Verbindungen möglich ist. Für die so genannte COMBINE-Analyse wurden die Wechselwirkungen zwischen den Liganden und den Aminosäuren, die die Liganden in der Bindungstasche umgeben, berechnet und in GOLPE analysiert. Neben den Energien für Coulomb-Wechselwirkungen und Van-der-Waals-Interaktionen werden zusätzlich auch Wasserstoffbrücken, die für die hier betrachteten Inhibitoren nicht relevant sind, und die Gesamt-Interaktionsenergien mit einzelnen Aminosäuren analysiert. Die PLS-Analyse in GOLPE zeigt, dass die Van-der-Waals-Wechselwirkungen mit den Aminosäuren Arg47, Asp48, Lys116 und Phe182, die Coulomb-Wechselwirkungen mit Lys116, Asp181 und Arg221, sowie die Gesamt-Interaktionen mit Arg24, Arg47, Asp48, Asp181 und Ser216 für die Einteilung in aktive und weniger aktive Verbindungen entscheidend sind. Die für die Van-der-Waals-Interaktionen gefundenen Aminosäuren befinden sich erwartungsgemäß in der Bindungsstelle und gehen auch alle Wechselwirkungen mit den Liganden ein. Weitere Aminosäuren, wie etwa Tyr46, treten zwar ebenfalls mit den Liganden in Interaktion, doch scheint diese bei allen Liganden vergleichbar zu sein und somit keinen Einfluss auf die Aktivität zu besitzen. Ähnliches gilt für die Coulomb-Wechselwirkungen mit den für die Bindung wichtigen Aminosäuren des P-Loops, die den Untersuchungen zufolge auf die Klassifizierung der aktiven und weniger aktiven Verbindungen keinen

Einfluss haben. Die Interaktion mit der Aminosäure Arg221 unterscheidet sich aber bei unterschiedlich aktiven Liganden, was darauf zurückzuführen sein könnte, dass diese Interaktion je nach Aufbau des Inhibitors mehr oder weniger gut erfolgen kann. Die Korrelation ($r^2 = 0,48$; $q^2_{L20\%O} = 0,36$; SDEP = 0,50) zwischen den Energiewerten und den biologischen Daten ist aber nur mäßig. Dies führt dazu, dass das Modell für Vorhersagen nicht geeignet ist.

Auch die Bewertung mit dem Programm X-Score lieferte weder mit den drei einzelnen Scoringfunktionen ($q^2_{LOO} = 0,19$ bis $0,39$; SDEP = $0,49$ bis $0,43$), noch für den Gesamtscore ($q^2_{LOO} = 0,36$; SDEP = $0,44$) eine gute Korrelation. Hingegen fiel bei diesen Berechnungen auf, dass sowohl der logP-Wert ($q^2_{LOO} = 0,55$; SDEP = $0,37$), als auch das Molekulargewicht ($q^2_{LOO} = 0,57$; SDEP = $0,36$) relativ gut mit der Aktivität korrelieren. Bei Verwendung der sonst benutzten, robusteren Methoden $q^2_{L20\%O}$ waren die Ergebnisse noch schlechter.

5.5 Datenbanksuchen

Auf der Suche nach neuen Leitstrukturen ist das High-Throughput-Screening (HTS) seit Beginn der 90er Jahre des 20. Jahrhunderts eine weit verbreitete Methode. Allerdings wird das HTS durch die stetig anwachsende Größe der in der pharmazeutischen Industrie zur Verfügung stehenden Substanzbibliotheken, die oft mit Methoden der kombinatorischen Chemie erzeugt werden, immer zeit- und kostenintensiver. Eine Möglichkeit, diese Zeit- und Kostenproblematik zu umgehen, besteht darin, nicht die ganze Datenbank experimentell zu analysieren, sondern stattdessen zuerst über ein virtuelles Screening (VS) einen Teil der Datenbank zu extrahieren, in der die aktiven Verbindungen angereichert sind. Das ist besonders dann erforderlich, wenn ein HTS nicht möglich ist, und nur wenige Verbindungen untersucht werden können, wie zum Beispiel beim Screening mittels NMR-Spektroskopie (Kernresonanz-Spektroskopie). Beim VS werden verschiedene ligandbasierte oder proteinbasierte Filtermethoden verwendet, die entweder nach Ähnlichkeiten zu bekannten Liganden suchen oder bei bekannter Zielstruktur die im HTS-Experiment untersuchte Enzymbindung simulieren. Für ein effektives VS ist entscheidend, dass die notwendigen Berechnungen innerhalb weniger Tage durchgeführt werden können. Nachdem es möglich war, aussagekräftige, ligandbasierte und proteinbasierte 3D-QSAR-Modelle für den betrachteten Datensatz bekannter PTP1B-Inhibitoren zu erstellen (siehe Abschnitte 5.3 und 5.4), sollte untersucht werden, ob diese Modelle geeignet sind, in größeren Datenbanken neue PTP1B-Inhibitoren zu finden. Dabei wurden die Methoden der proteinbasierten Überlagerung und die 3D-QSAR-Modelle, die mit dem erweiterten Datensatz erstellt wurden, zugrunde gelegt. Als Datenbanken für das VS standen der World Drug Index (WDI) der Firma Derwent [275] die Maybridge Datenbank (MaybridgeDB) der Maybridge Chemical Company [276] und eine firmeninterne Datenbank der Firma Evotec OAI (EvotecDB) zu Verfügung.

Der WDI enthält etwa 57.700 Strukturen von Verbindungen, die sich in verschiedenen Phasen der Arzneistoffentwicklung befinden. Neben bereits in

der Therapie verwendeten Arzneistoffen sind darin Verbindungen aufgenommen, für die aus wissenschaftlichen Publikationen Bindungsdaten bekannt sind. Die MaybridgeDB enthält etwa 56.600 Strukturen von Verbindungen, die mit Hilfe der kombinatorischen Chemie synthetisiert wurden und von der Maybridge Chemical Company erworben werden können.

Die Strukturen, die im SD-Format vorliegen, enthalten keine Wasserstoffatome und befinden sich energetisch in einer sehr ungünstigen Konformation, da es sich eigentlich um zweidimensionale Strukturen handelt, bei denen die Z-Koordinate immer den Wert 0 besitzt. Um die Strukturen für die Datenbanksuchen verwenden zu können, wurden sie mit dem Programm Corina in dreidimensionale, energetisch günstige Konformationen mit Wasserstoffatomen und das mol2-Dateiformat umgewandelt.

Die EvotecDB enthält circa 44.600 Verbindungen aus verschiedenen Datensätzen, die mit mehreren Verfahren des HTS untersucht wurden. Die Hits aus dem HTS wurden verifiziert und bei den bestätigten Verbindungen wurden die IC_{50} -Werte bestimmt. Zu Beginn der Untersuchungen der Verbindungen lagen diese Aktivitätsdaten allerdings nicht vor, sondern es standen nur die Strukturen der einzelnen Substanzen im mol2-Format und SD-Format zu Verfügung.

5.5.1 Vorgehensweise

Mit Hilfe der bereits zur Erstellung der 3D-QSAR-Modelle verwendeten Methoden sollten die verschiedenen Strukturdatenbanken nach neuen PTP1B-Inhibitoren durchsucht werden. Eine Datenbank mit über 50.000 Verbindungen lässt sich jedoch mit den im Rahmen dieser Arbeit zu Verfügung stehenden Methoden und Mitteln nicht vollständig und in vertretbarer Zeit in eine Proteinbindungstasche docken, sodass vorher eine Auswahl getroffen werden musste, welche Verbindungen gedockt werden sollten. Deshalb wurden zuerst relativ grobe Verfahren angewandt, die solche Verbindungen herausfiltern, die mit großer Wahrscheinlichkeit inaktiv sind. Anschließend wurden genauere Methoden benutzt, die weiter nach aktiven Verbindungen suchen, bis letztendlich mit Hilfe eines 3D-QSAR-Modells die Aktivität einzelner Verbindungen vorhergesagt wurde.

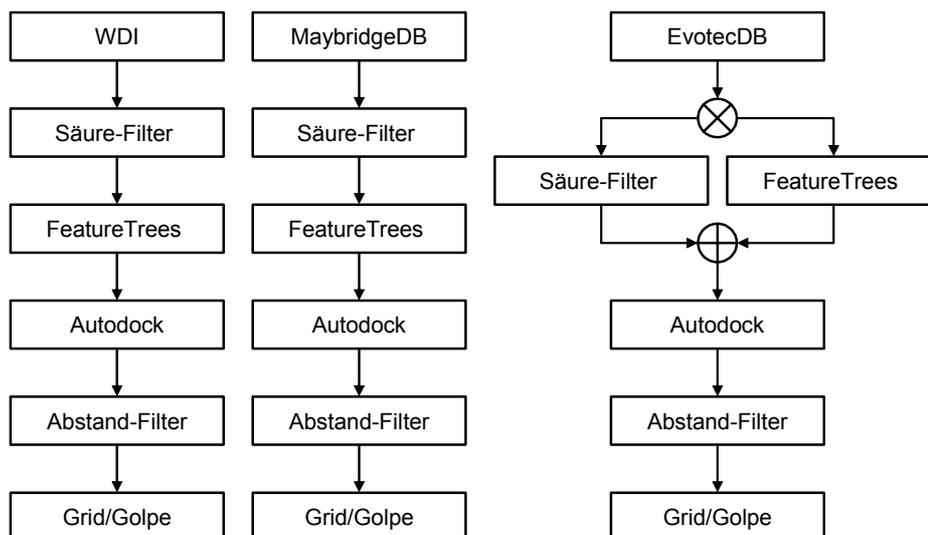


Abbildung 5-53: Vorgehensweise beim virtuellen Screening

Die gesamte Vorgehensweise ist in der Abbildung 5-53 für das Screening der drei in dieser Arbeit analysierten Datenbanken dargestellt.

Basierend auf der Annahme, dass die Säurefunktion für die Bindung an die PTP1B von entscheidender Bedeutung ist, wurden aus dem WDI und der MaybridgeDB zunächst alle Carbonsäuren herausgesucht. Diese Gruppe von Verbindungen wurde dann mit dem Programm FeatureTrees analysiert und es wurden die Verbindungen selektiert, die den bereits bekannten Inhibitoren am ähnlichsten sind, wobei als Ausgangspunkt für diese Ähnlichkeitssuche die Verbindungen 80, 142, 164 und 172 aus [258] und die Verbindungen 1 bis 10 aus [261] verwendet wurden. Jede Struktur, die in Bezug auf mindestens eine der Suchstrukturen mit einem Ähnlichkeitsindex über 0,83 bewertet wurde oder sich unter den besten 500 Substanzen befand, wurde für weitere Untersuchungen ausgewählt. Es wurden hier 14 verschiedene Suchstrukturen und die beiden genannten Grenzkriterien verwendet, um die Anzahl an Strukturen, die weiter analysiert werden müssen, beträchtlich einzugrenzen aber trotzdem nur wenige der möglichen Kandidaten auszuschließen. Der Ähnlichkeitsindex 0,83 schließt mit den oben genannten Suchverbindungen alle Strukturen von bekannten Inhibitoren des Datensatzes mit ein. Da aber bei den Ähnlichkeitsanalysen der Ähnlichkeitsindex auch von der verwendeten Suchstruktur abhängig ist, und somit bei gegebenen Ähnlichkeitsindex von 0,83 je nach verwendeter Suchstruktur einige wenige bis über 1.000 ähnliche Verbindungen gefunden werden,

wurden zusätzlich für jede Suchverbindung die 500 ähnlichsten für die weitergehende Analyse ausgewählt, um für jede Suchstruktur eine Mindestanzahl von ähnlichen Kandidaten zu berücksichtigen.

Für das virtuelle Screening der EvotecDB wurden nicht nur die Säuren sondern alle Verbindungen des Datensatzes nach dem obigen Schema mittels Feature-Trees analysiert, weil der Datensatz nur wenige Säuren enthält. Die hierbei untersuchten Säuren schnitten bei der FeatureTrees Analyse in der Bewertung nicht gut ab, und es war anzunehmen, dass auch unter den Verbindungen, die keine Carbonsäure darstellen, aktive Kandidaten enthalten sind. Deshalb wurde der gesamte Datensatz mit FeatureTrees analysiert und es wurden zusätzlich alle Säuren des Datensatzes für das Docking ausgewählt.

Anschließend wurden für alle drei Datensätze die ausgewählten Verbindungen aus den Multi-Mol2-Dateien herausgeschrieben und in einzelne Mol2-Dateien umgewandelt. Für die einzelnen Verbindungen wurden Ladungen berechnet, wobei für den Protonierungszustand pH 7 angenommen wurde, und die Strukturen für das Docking, das mit dem Programm Autodock durchgeführt werden sollte, vorbereitet.

Aus dem Docking resultierten pro Ligand bis zu 25 Konformationen, unter denen die besten ausgewählt werden sollten. Da sich der Score aus dem Programm Autodock nicht gut mit der biologischen Aktivität korreliert und sich deshalb nicht eignet, um unterschiedliche Verbindungen nach ihrer Aktivität zu klassifizieren, wurde als erstes grobes Filterkriterium nach dem Docking der Abstand der gedockten Konformation zur Bindungstasche gemessen. Bei WDI und MaybridgeDB, bei denen alle gedockten Strukturen wegen der Vorselektion eine Carbonsäure enthalten, sollte der Abstand eines der Sauerstoffatome nicht weiter als 7,5 Å vom aktiven Zentrum entfernt sein. Da die gedockten Verbindungen der EvotecDB jedoch nicht nur Carbonsäuren enthalten, wurden aus diesem Datensatz die Konformationen verworfen, bei denen kein Atom näher als 7,5 Å an das aktive Zentrum heran reicht. Für die Verbindungen, die den Auswahlkriterien genügten, wurden mit dem Programm GRID Wechselwirkungsfelder berechnet, auf deren Grundlage die biologischen Aktivitäten der

einzelnen Verbindungen mit dem 3D-QSAR-Modell aus den proteinbasierten Untersuchungen vorhergesagt wurde.

5.5.2 World Drug Index und Maybridge Datenbank

Das Screening wurde wie oben beschrieben durchgeführt. Von den etwa 57.700 Verbindungen des WDI enthielten etwa 9.300 (16,1%) eine oder mehrere Carbonsäurefunktionen und ein Molekulargewicht zwischen 200 und 1.000. Daraus wurden mit FeatureTrees 1.700 Verbindungen (18,4%) herausgefiltert, was etwa 3,1% des gesamten WDI entspricht. Die MaybridgeDB enthält weniger Carbonsäuren, weniger Verbindungen mit Arzneistoffcharakter, und daher viele Strukturen, die nur eine sehr geringe Ähnlichkeit zu den bekannten Inhibitoren besitzen. Nach der Selektion der Carbonsäuren und der Ähnlichkeitssuche mit FeatureTrees blieben deshalb nur etwa 800 Verbindungen (1,45%) übrig, die gedockt wurden. Von diesen 2.500 Verbindungen aus den beiden Datensätzen zusammen wurden mit Autodock unter Berücksichtigung des Abstandes der Säurefunktion zum aktiven Zentrum etwa 1.200 erfolgreich in die Bindungstasche der PTP1B gedockt von denen etwa 75% aus dem WDI stammen. Da für einige Liganden mehrere Konformationen den Kriterien entsprachen, lieferte das Docking insgesamt 1.930 Konformationen für die 3D-QSAR-Analyse. Für diese Konformationen wurden Interaktionsfelder mit der N2=-Sonde berechnet und die biologische Aktivität vorhergesagt, wobei Verbindungen mit einem IC_{50} -Wert $> 1 \mu\text{M}$ als aktiv angesehen werden. Diese aktiven Verbindungen, von denen 324 aus dem WDI und 103 aus der MaybridgeDB stammen, wurden miteinander verglichen und der Bindungsmodus analysiert (Abbildung 5-54). In den beiden links abgebildeten Komplexen binden die Liganden eher auf der linken Seite, auf der auch die Benzofuran- oder Benzothiophen- und die Biphenylgruppe der bekannten Liganden binden. Die beiden Liganden in den rechten Abbildungen sind in der rechten Seite des aktiven Zentrums platziert, in der bei den größeren der bekannten Inhibitoren der zusätzliche Phenylrest der Phenylmilchsäure-Gruppe bindet.

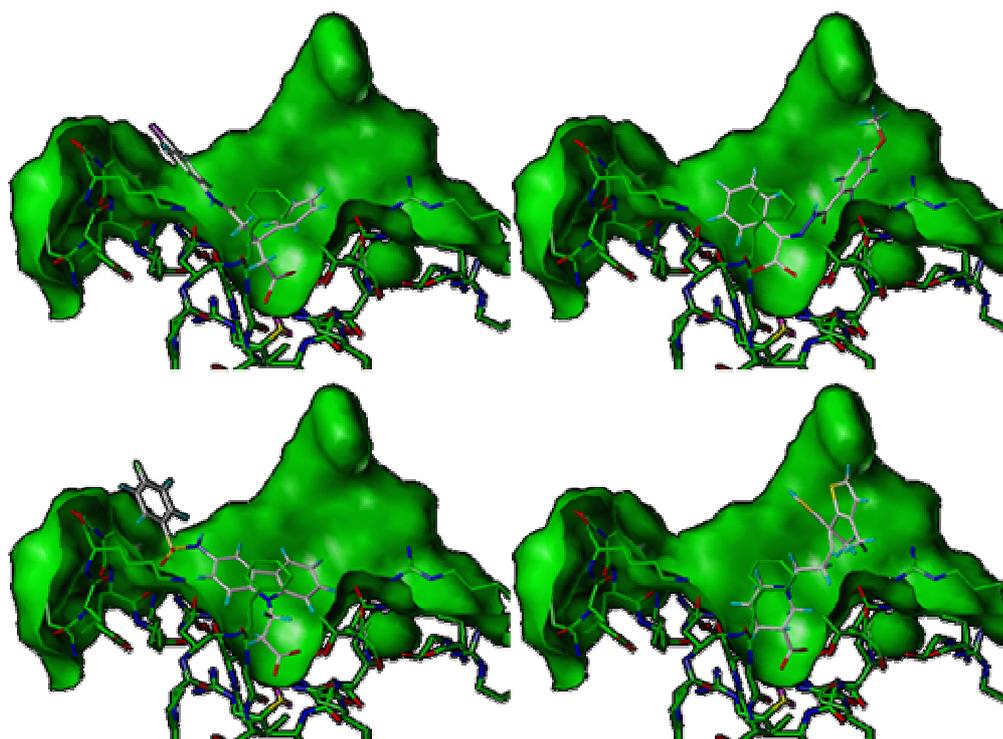


Abbildung 5-54: Gedockte Verbindungen aus dem WDI

Fügt man die bekannten Inhibitoren, die für die 3D-QSAR-Modelle verwendet wurden in den Datensatz aus WDI und MaybridgeDB ein, so beträgt der Anteil an bekannten aktiven Verbindungen etwa 0,07 %. Eine Gesamtübersicht über die im folgenden beschriebene Anreicherung in den einzelnen Schritten der Datenbanksuche, die im Folgenden genauer beschrieben ist, ist in Abbildung 5-55 und Abbildung 5-56 dargestellt.

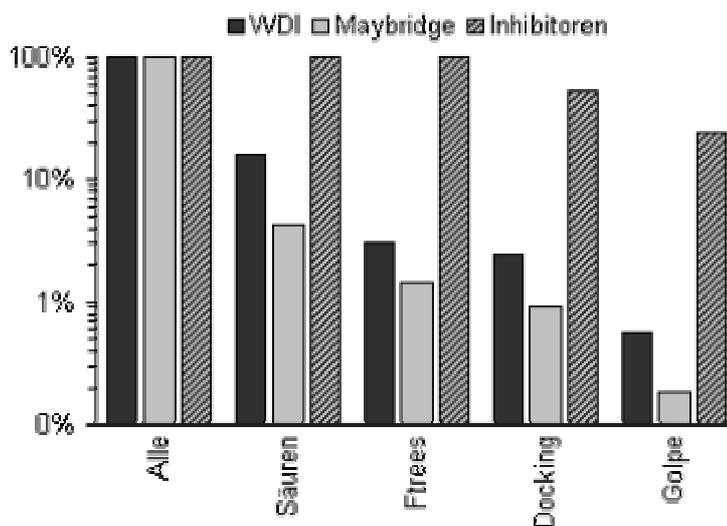


Abbildung 5-55: Filterung von WDI, Maybridge und bekannten Inhibitoren

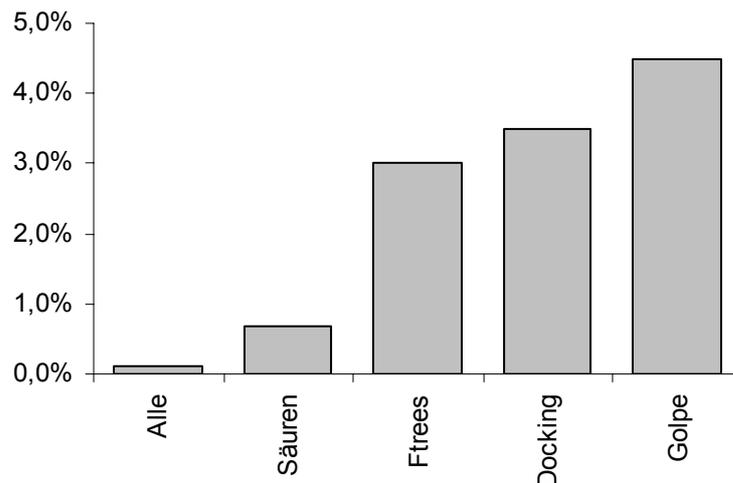


Abbildung 5-56: Anreicherung der Inhibitoren in der Datenbank

Durch die Filterung nach Carbonsäuren, die etwa 90% des Datensatzes ausschließt, steigt der Anteil auf etwa 0,7 %. Die Ähnlichkeitsanalyse mit dem Programm FeatureTrees, die von den noch vorhandenen Strukturen nochmals ungefähr 75% verwirft, reduziert den Datensatz zusammen mit dem vorherigen Schritt auf nur noch 2.640 Verbindungen (2,2% vom gesamten Datensatz), so dass der Anteil an aktiven Verbindungen auf 3,0 % steigt. Die Ähnlichkeitsuche stellt somit den effektivsten Filterungsschritt im gewählten Szenario dar. Das Docking mit Autodock bringt nur eine geringe Anreicherung an aktiven Substanzen, die in erster Linie durch die Bedingung, dass eine Säurefunktion im aktiven Zentrum binden muss, zustande kommt und nicht durch die Scoringfunktion des Programms selbst. Von den 2.640 Verbindungen wurden 1.200 im aktiven Zentrum gedockt, die mit dem Programm PrGen nochmals optimiert wurden. Da mit der gewählten Dockingstrategie nicht alle bekannten Inhibitoren gedockt werden können, steigt die Anreicherung nur von 3,0% auf 3,5%. Die Filterung mit GOLPE ist abhängig vom Grenzwert für die vorhergesagte Aktivität ($\log IC_{50}$ -Wert). Die bekannten Inhibitoren besitzen Aktivitäten im Bereich von 4,28 bis 7,96. Da ein QSAR-Modell nur Werte innerhalb des Bereiches des Trainingsatzes vorhersagen kann, liegen natürlich auch alle vorhergesagten Verbindungen der beiden Datenbanken in diesem Intervall. Von den 1.200 Verbindungen die erfolgreich gedockt wurden, wurde die beste mit einem $\log IC_{50}$ von 7,18 vorhergesagt, die schlechteste mit 4,89. Da der Grenzwert innerhalb des Aktivitätsintervalls liegen muss, werden

natürlich auch die bekannten Inhibitoren mit der niedrigsten Aktivität ausgeschlossen. In der Tabelle 5-22 ist dargestellt, wie viele der Verbindungen der Datenbank und der bekannten Inhibitoren in den jeweiligen Aktivitätsbereichen der GOLPE-Analyse liegen.

Vorhergesagte Aktivität	gedockte Verbindungen	bekannte Inhibitoren
> 4,0	100%	100%
> 5,0	98%	93%
> 5,5	37%	91%
> 6,0	21%	77%
> 6,3	7,7%	61%
> 6,5	3,3%	48%
> 7,0	0,3%	29%

Tabelle 5-22: Mit GOLPE berechnete logIC₅₀-Werte

Die gefundenen Verbindungen stellen mögliche neue Grundkörper als PTP1B-Inhibitoren dar, die nach der experimentellen Bestätigung ihrer PTP1B-Aktivität zu neuen Leitstrukturen optimiert werden könnten.

5.5.3 Evotec OAI Datenbank

Neben der Maybridge Datenbank und dem World Drug Index wurde eine Datenbank der Firma Evotec OAI untersucht. Auch bei dieser Datenbank wurden zuerst mit dem Programm FeatureTrees den bekannten Strukturen ähnliche Verbindungen herausgesucht. Alle Verbindungen, die sich entweder unter den 500 Besten für jede Suchstruktur befanden oder mit einem Ähnlichkeitsindex über 0,83 vorhergesagt wurden, wurden ausgewählt. Zu diesen 6.559 Verbindungen wurden noch die 2.595 Carbonsäuren der Datenbank hinzugefügt, so dass 8.775 Strukturen für weitere Untersuchungen berücksichtigt wurden. Davon konnten 8.642 Verbindungen für das Docking mit Autodock verwendet werden und für 8.495 wurde wenigstens eine gedockte Konformation gefunden. In 10.382 gedockten Konformationen von 5.326 Verbindungen befand sich mindestens ein Atom des Liganden in einem Abstand von maximal 5,0 Å zum aktiven Zentrum der PTP1B. 9.737 dieser Konformationen wurden einer Analyse mit GOLPE unterzogen, nach dem Interaktionsfelder für die N2=-Sonde in GRID berechnet wurden.

Bei diesen Untersuchungen stellte sich heraus, dass die im Datensatz enthaltenen, aktiven Verbindungen mit der gewählten Vorgehensweise nicht wieder gefunden werden konnten. Eine genauere Analyse ergab, dass schon bei der ersten Einschränkung mit dem Programm FeatureTrees der größte Teil der aktiven Verbindungen herausgefiltert wird. Vergleicht man die Strukturen, die als Vorgabe dienten, mit den Verbindungen des Evotec Datensatzes, so stellt man fest, dass sich die beiden Gruppen extrem voneinander unterscheiden. Dies lässt sich auch zeigen, wenn man die Verbindungen wie in Abbildung 5-57 im gemeinsamen chemischen Raum anordnet, wo sie völlig unterschiedliche Bereiche einnehmen. Außerdem trifft die Annahme, dass die aktiven Verbindungen eine negative Ladung besitzen, auf diesen Datensatz nicht zu.

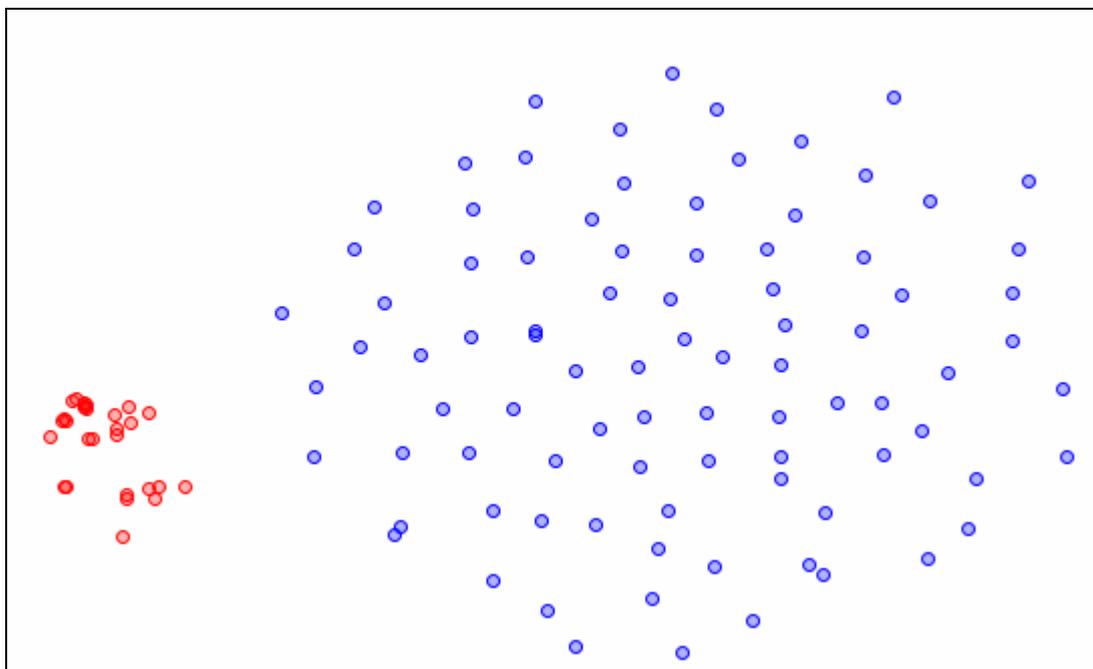


Abbildung 5-57: Anordnung der aktivsten Verbindungen der 3D-QSAR-Analysen (rot) und der aktiven Substanzen der EvotecDB (blau) im chemischen Raum (basierend auf Unity-Fingerprints und einer PCA)

5.5.4 Evotec OAI Subset

Um die Eignung weiterer Verfahren, insbesondere von Docking-Methoden, zur Eingrenzung des Datensatzes zu analysieren, wurde eine Gruppe von 996 Verbindungen aus der EvotecDB herausgegriffen. Diese Gruppe repräsentiert den gesamten Datensatz hinsichtlich der Anzahl an aktiven Verbindungen und der chemischen Diversität. Unter den Verbindungen befinden sich 88 Substan-

zen, die von Evotec im Screening-Experiment als Hits gefunden wurden. Es sind also 8,8 % der Verbindungen als aktiv anzusehen.

Die Verbindungen wurden alle mit den Programmen Autodock, FlexX und Gold gedockt, und mit den unterschiedlichen in den Programmen enthaltenen Scoringfunktionen, sowie mit der separaten Software X-Score bewertet. Dabei ließ sich für mehrere Kombinationen aus Docking und Scoring eine Anreicherung von aktiven Verbindungen finden.

Zuerst wurde das Programm Autodock verwendet, mit dem sich 98 Verbindungen, darunter 9 aktive Substanzen, nicht docken lassen. Die Rate der erfolglosen Dockings ist unter den aktiven und den inaktiven Verbindungen also annähernd gleich. Sortiert man die gedockten Konformation in einer Prioritätsliste nach dem besten Scoringwert, dann kann man unter den Top-10-% der Verbindungen 14 aktive finden, was einer Hitrate von 14 % entspricht. Im Vergleich zu den 8,8 % aktiver Verbindungen im untersuchten Datensatz entspricht das einer Anreicherung um den Faktor 1,5, jedoch werden lediglich 16 % aller aktiven Verbindungen gefunden.

Die Geometrieoptimierung der Komplexe mit dem Programm PrGen und die Auswertung der damit berechneten Wechselwirkungsenergien führten zu keiner Verbesserung der Vermehrung an aktiven Verbindungen.

Alternativ zu dieser Optimierung wurden die Komplexe mit dem Programm X-Score erneut bewertet. Von den drei Scoringfunktionen, die mit X-Score berechnet werden, schnitt hier der HM-Score am besten ab, bei dem unter den Top-10-% 19 und unter den Top-1-% immer noch 7 aktive Verbindungen gefunden werden, was einer Hitrate von 70 % und einer Anreicherung um den Faktor 7,9 entspricht. Daneben wurde auch mit dem Consensus-Score (Mittelwert der drei Funktionen) eine Anreicherung um den Faktor 5,7 (Top-1-%) erzielt (Siehe Abbildung 5-58).

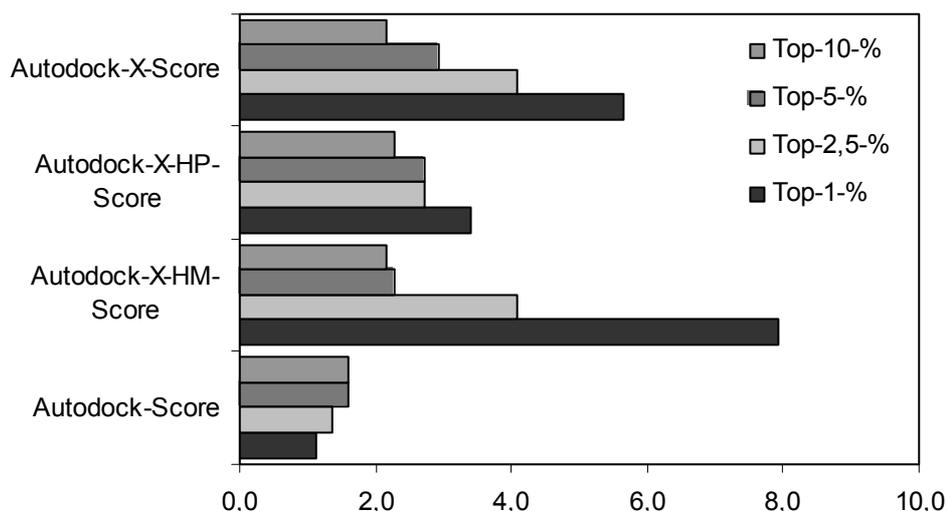


Abbildung 5-58: Anreicherungs-faktoren des Screening mit Autodock

Um zu testen, ob mit anderen Dockingprogrammen und Scoringfunktionen bessere Ergebnisse erzielt werden können, wurden anschließend das Programm Gold, das die beiden Scoringfunktionen Gold-Score und ChemScore enthält, und das Programm FlexX mit dem eingebauten Scoring sowie den Scoringfunktionen aus C-Score (D-Score, PMF-Score, G-Score und Chem-Score) angewendet.

In Gold wurde bereits das Docking mit beiden Scoringfunktionen durchgeführt und die Komplexe wurden, wie schon bei den Autodock-Untersuchungen, jeweils mit X-Score erneut bewertet.

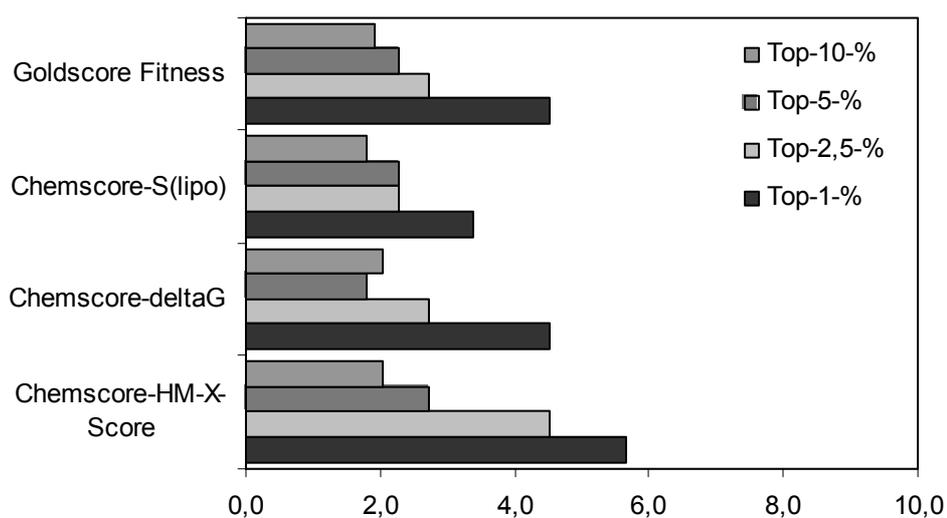


Abbildung 5-59: Anreicherungs-raten des Screening mit Gold

Wie aus Abbildung 5-59 ersichtlich ist, konnte mit der Gold-Scoringfunktion ebenso wie mit der ChemScore-Funktion für die Top-1-% der Verbindungen ein Anreicherungsfaktor von 4,5 erzielt werden. Auch mit der Teilfunktion, die die Lipophilie berechnet, konnte eine Anreicherung um den Faktor 3,4 erreicht werden. Für die mit der ChemScore-Funktion erhaltenen Komplexe konnte mit X-Score eine Verbesserung des Ranking erzielt werden, so dass eine Anreicherung um den Faktor 5,7 resultiert.

Die Darstellung der Anreicherungsfaktoren in Abbildung 5-60 zeigt, dass auch beim Docking mit FlexX das Ranking besser wird, wenn man ein erneutes Scoring mit weiteren Scoringfunktionen durchführt. Mit der Scoringfunktion von FlexX wird nur eine Anreicherung um den Faktor 3,4 erzielt, mit den beiden Scoringfunktionen D-Score und ChemScore beträgt der Faktor jedoch 4,6.

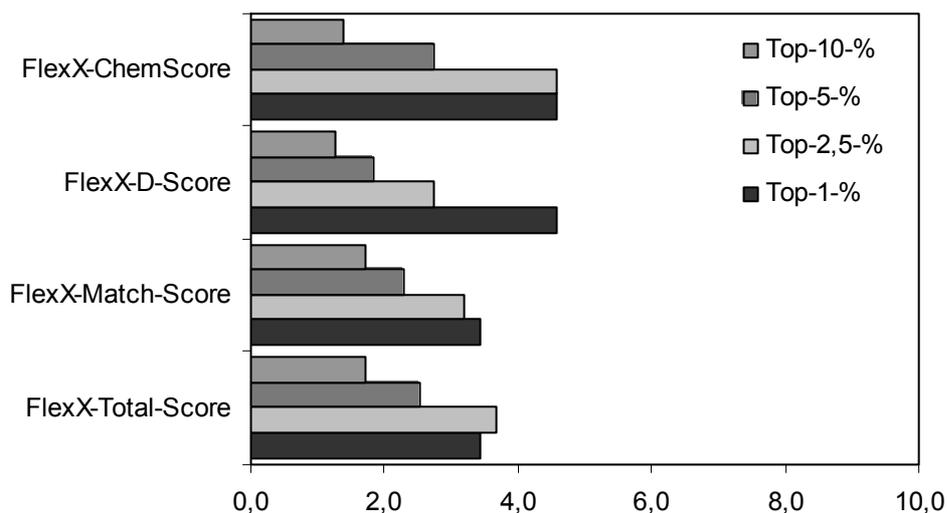


Abbildung 5-60: Anreicherungsdaten des Screening mit FlexX

Zusammenfassend kann man über alle drei Dockingprogramme und die verschiedenen Scoringfunktionen feststellen, dass bei der Betrachtung einer kleineren Anzahl von topgerankten Verbindungen die Anreicherung steigt, also der prozentuale Anteil an Hits größer wird.

Als Beispiel ist in Abbildung 5-61 der prozentuale Anteil von verifizierten Hits unter den topgerankten Komplexen aufgetragen, und man erkennt, dass bei der Betrachtung der Top 20 (2 %) ein Anteil von über 50 % an Hits erzielt wird.

Dieser Wert nimmt dann rapide ab, erreicht bei der Betrachtung der Top 50 (5 %) eine Hitrate von etwa 20 % und fällt dann relativ stetig bis zum Ende ab.

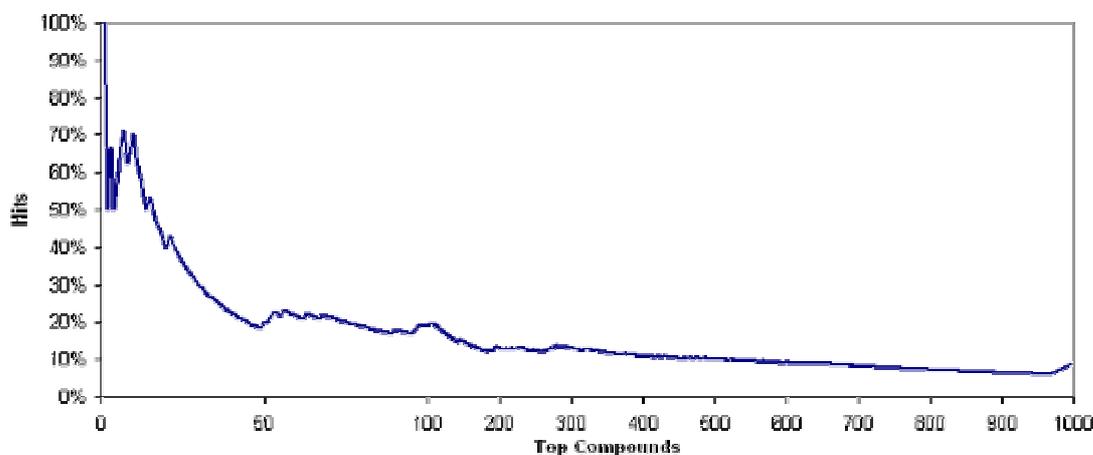


Abbildung 5-61: Prozentualer Anteil an Hits in Abhängigkeit von der Anzahl der betrachteten topgerankten Verbindungen des Screening mit Autodock und X-Score

In den folgenden Tabellen sind die Ergebnisse exemplarisch für das Screening mit Gold (Tabelle 5-23), Autodock und X-Score (Tabelle 5-24) sowie FlexX und ChemScore (Tabelle 5-25) dargestellt. Anhand der Daten erkennt man, dass eine gute Anreicherung erzielt werden kann, wobei trotzdem nur ein kleiner Anteil der Hits gefunden wird. Es ist also immer ein Kompromiss zwischen Anreicherungsfaktor und prozentualen Anteil der gefundenen Hits zu treffen. Werden diese Methoden als virtuelles Screening einem experimentellen Screening vorgeschaltet, so wäre es möglich, nur etwa 2,5 % der Verbindungen *in vitro* zu untersuchen, um etwa 10 % der Hits zu finden (Tabelle 5-24). Auch das Testen von einem Zehntel der Verbindungen würde die Anzahl der Treffer pro getestete Verbindung noch um mehr als den Faktor zwei erhöhen.

Top X Verbindungen	Hits / Verbindungen	gefundene Hits / Hits im Datensatz	Anreicherung
0,5 %	60%	3%	6,8
1 %	40%	5%	4,5
2 %	30%	7%	3,4
2,5 %	24%	7%	2,7
5 %	20%	11%	2,3
10 %	17%	19%	1,9

Tabelle 5-23: Ergebnisse des Screening mit Gold und Goldscore

Top X Verbindungen	Hits / Verbindungen	gefundene Hits / Hits im Datensatz	Anreicherung
0,5 %	60%	3%	6,8
1 %	70%	8%	7,9
2 %	40%	9%	4,5
2,5 %	36%	10%	4,1
5 %	20%	11%	2,3
10 %	19%	22%	2,2

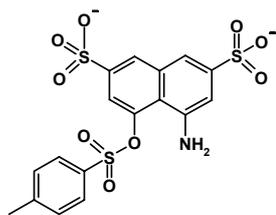
Tabelle 5-24: Ergebnisse des Screening mit Autodock und X-Score

Top X Verbindungen	Hits / Verbindungen	gefundene Hits / Hits im Datensatz	Anreicherung
0,5 %	40%	2%	4,6
1 %	40%	5%	4,6
2 %	45%	10%	5,2
2,5 %	40%	11%	4,6
5 %	24%	14%	2,7
10 %	12%	14%	1,4

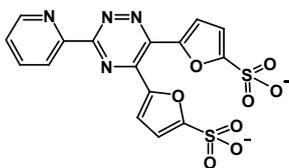
Tabelle 5-25: Ergebnisse des Screening mit FlexX und ChemScore

Da die Strukturen der Verbindungen des Screening der EvotecDB aus Gründen des Patentschutzes nicht abgebildet werden dürfen, wurden die gleichen Dockingmethoden auf einen Datensatz mit etwa 2.000 Strukturen des NCI (National Cancer Institute) [277] angewendet. Aus dem ursprünglichen Datensatz wurden zuvor etwa 300 Verbindungen herausgefiltert, die Metallatome oder mehrere Enantiomere enthalten oder mehr als 15 drehbare Bindungen besitzen. Die anderen Verbindungen wurden mit den Programmen Autodock, FlexX und Gold gedockt und die Resultate mit X-Score beziehungsweise der ChemScore-Funktion aus C-Score erneut bewertet. Die Verbindungen mit den besten Scores in den einzelnen Scoringfunktionen sind in Tabelle 5-26 zusammengefasst.

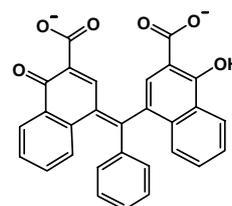
Auch hier zeigt sich wie bereits bei der Untersuchung des Evotec-Datensatzes, dass die Programme jeweils unterschiedliche Hits finden. Alle Hits, die abgebildet sind, wurden entweder in einer der Scoringfunktionen unter den besten zehn gefunden, oder sie befanden sich in zwei Scoringfunktionen unter den 50 am besten bewerteten Verbindungen.



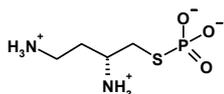
Gold: Top 1
FlexX: Top 4



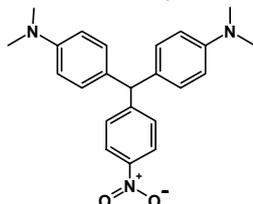
FlexX: Top 1
Gold: Top 2



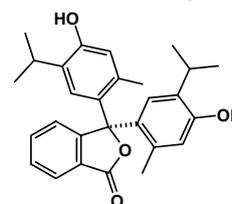
Autodock: Top 1



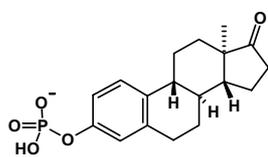
Gold: Top 10



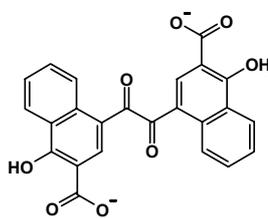
FlexX: Top 10



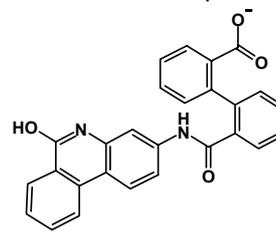
Autodock: Top 10



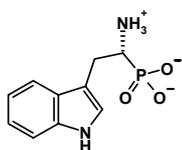
Gold: Top 10



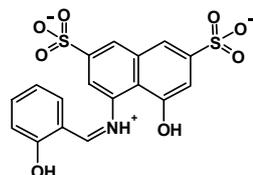
FlexX: Top 10



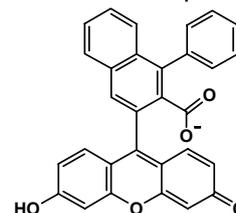
Autodock: Top 10



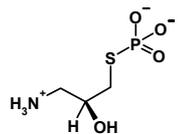
Gold: Top 10



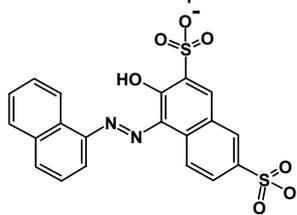
FlexX: Top 10



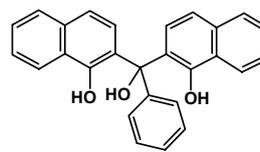
Autodock: Top 10



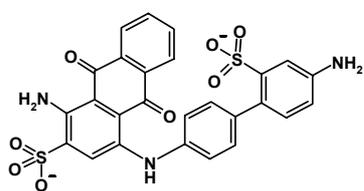
Gold: Top 10



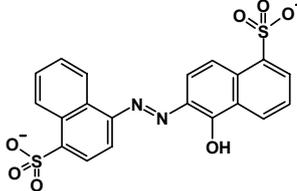
FlexX: Top 10



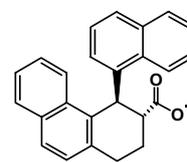
Autodock: Top 10



Gold: Top 10



FlexX: Top 10



Autodock: Top 10

Tabelle 5-26: Strukturen des NCI-Datensatzes mit dem besten Scores im virtuellen Screening

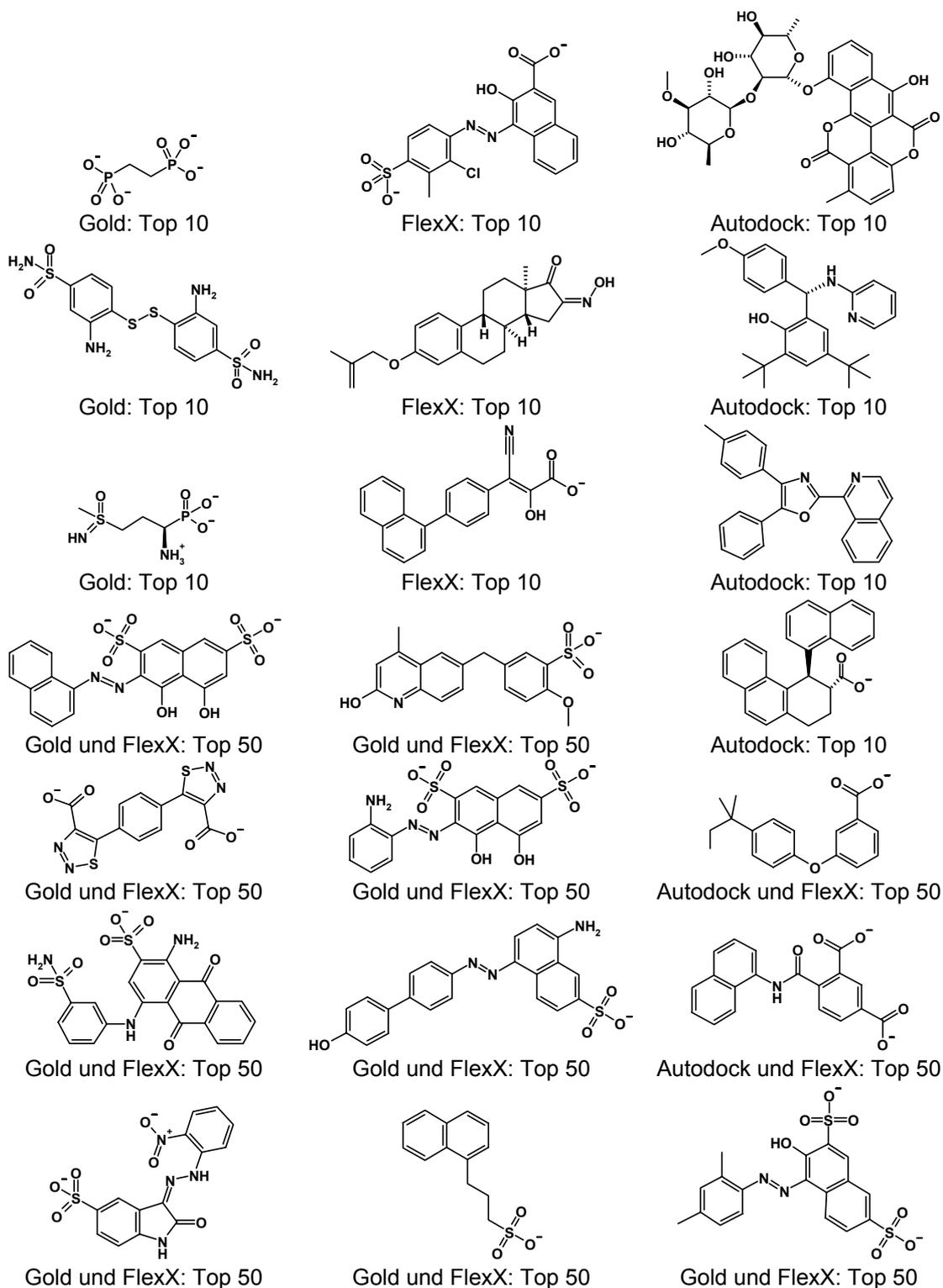


Tabelle 5-26: Strukturen des NCI-Datensatzes mit dem besten Scores im virtuellen Screening (Fortsetzung)

Von diesen enthalten mehr als 80 % eine Säurefunktion in Form einer Carboxylat-, Phosphat-, oder Sulfatgruppe. Darunter sind neben neuen Strukturen auch einige Verbindungen, die in dieser oder ähnlicher Form schon als Inhibitoren der PTP1B bekannt sind, wie zum Beispiel die mit einer Sulfonsäure substituierten Naphthalinderivate oder das substituierte Phenolphthalein, wodurch nochmals die Anwendbarkeit der Methode für die Suche nach neuen PTP1B Inhibitoren gezeigt werden konnte.

6 Diskussion

In der vorliegenden Arbeit wurde für die aus der Literatur bekannten Röntgenkristallstrukturen der PTP1B mit Hilfe einer Hauptkomponentenanalyse (PCA) ein Vergleich der dreidimensionalen Struktur des Enzyms auf der Basis von Interaktionsfeldern durchgeführt. Durch diesen Vergleich ließen sich die Strukturen nicht nur in zwei Gruppen mit offenem oder geschlossenem WPD-Loop einteilen, sondern alle im Jahre 2003 neu in der PDB veröffentlichten Strukturen konnten mit diesem Modell den beiden Gruppen widerspruchsfrei zugeordnet werden. Zusätzlich war eine Aufteilung in Untergruppen mit ähnlichen Konformationen möglich, bei der sich gezeigt hat, dass Kristallstrukturen, die mit ähnlichen Liganden im aktiven Zentrum gewonnen wurden, in eine Untergruppe der PCA zusammenfallen. Die Unterschiede in der PCA konnten exemplarisch an der dreidimensionalen Struktur der jeweiligen Protein-Ligand-Komplexe erklärt werden. Anhand dieser Einordnung verschiedener Kristallstrukturen konnten besonders geeignete, also möglichst unterschiedliche Strukturen für weitere Untersuchungen wie Moleküldynamiksimulationen und Docking-Analysen ausgewählt werden. Da ein erfolgreiches Docking in die ausgewählten Strukturen nicht auf Anhieb möglich war, wurde eine MD-Simulation durchgeführt, aus der weitere Konformationen der Bindungstasche resultierten. Die repräsentativen Strukturen aus dieser Simulation wurden für das Docking eingesetzt und getestet. Eine der Proteinkonformationen die für das Docking geeignet war, wurde ebenfalls in die PCA-Analyse eingefügt und bildet dort eine separate Gruppe, die sich von den verfügbaren Kristallstrukturen erheblich unterscheidet. Diese besondere Konformation wurde für die Dockinguntersuchungen bei den späteren proteinbasierten 3D-QSAR-Modellen und den Datenbanksuchen nach verwendet, weil nur mit dieser Struktur die Liganden sinnvoll gedockt werden konnten.

In einer ähnlichen vergleichenden Untersuchung wurden zusätzlich die Kristallstrukturen anderer PTPs betrachtet. Die große Sequenzähnlichkeit zwischen PTP1B und TCPTP konnte auch in den dreidimensionalen Strukturen und der PCA-Analyse bestätigt werden. Die anderen Phosphatasen, bei denen viele Inhibitoren der PTP1B nicht aktiv sind, lassen sich hingegen eindeutig von der PTP1B differenzieren. Gäbe es auch von den anderen PTPs insbesondere der

TCPTP mehrere Kristallstrukturen, ließe sich das Verfahren auch zur Auswahl von Proteinstrukturen zum Design selektiver Inhibitoren nutzen.

Die PCA der Interaktionsfelder hat gegenüber den üblichen Analysemethoden, die nur auf einem Clustering der RMS-Werte der überlagerten Proteinstrukturen basieren, die Vorteile, dass Ergebnisse leichter zu interpretieren sind, da die Strukturen nicht nur in Cluster eingeordnet, sondern die einzelnen Strukturen auch zueinander angeordnet werden. Außerdem können Strukturen mit unterschiedlichen Sequenzen verglichen werden und man kann die Analyse auf die Umgebung des aktiven Zentrums beschränken.

Die negative Ladung des Schwefelatoms von Cys215 konnte durch die UHBD-Berechnungen bestätigt werden. Die negative Ladung wird hauptsächlich durch die Amidgruppen des P-Loops und das Dipolmoment der zentralen Helix stabilisiert. Für die anderen Aminosäuren, deren berechneter pK_S -Wert vom theoretischen Standardwert abweicht, konnte ebenfalls eine Erklärung für diese Abweichung gefunden werden.

Aus der Vielzahl der in der Fachliteratur vorhandenen Inhibitoren der PTP1B mussten für die weiteren Arbeiten geeignete Verbindungen ausgewählt werden. Die Inhibitoren aus verschiedenen Klassen lassen sich aufgrund starker struktureller Unterschiede nur bedingt miteinander vergleichen und in einem gemeinsamen Modell betrachten. Mit einem relativ großen Datensatz aus der Literatur [258] war jedoch eine homogene Gruppe von geeigneten Substanzen vorhanden, die für die Untersuchung von Struktur-Wirkungs-Beziehungen genutzt werden konnte. Da für einige Strukturen dieses Datensatzes die Konfiguration von chiralen Zentren nicht bekannt ist, wurden diese Verbindungen in separaten Untersuchungen genauer betrachtet.

Aufgrund von vorhandenen Literaturdaten wurde Verbindung 80 ausgewählt und in zwei Schritten einer systematische Konformationsanalyse unterzogen, die vier verschiedene, energetisch günstige Konformere für diese Verbindung lieferte. Eine dieser Konformationen entsprach der angenommenen bioaktiven Konformation, die aus einer unveröffentlichten Kristallstruktur hervorgeht [258]. Auf die berechnete Konformation der Verbindung 80 wurden die unterschiedlichen Grundgerüste der weiteren Substanzen superpositioniert. Ausgehend

von dieser Überlagerung wurden unterschiedliche Superpositionierungs-Ansätze für die weiteren Liganden, die sich von den Grundstrukturen nur in Seitenketten unterschieden, getestet und miteinander verglichen. Sowohl die Überlagerung mit dem Programm FlexS als auch die Überlagerung mit der „Fit-Atoms-Routine“ in SYBYL führten zu einer konsistenten Superpositionierung und lieferten bei den anschließend durchgeführten 3D-Struktur-Wirkungs-Beziehungen vergleichbar gute Ergebnisse.

Anhand der Überlagerungen wurden zunächst unterschiedliche GRID-Sonden getestet, um die für eine QSAR Analyse aussagekräftigsten Sonden auswählen zu können. Neben der Phosphat-Sonde, die die Unterschiede zwischen den Liganden am besten erklärt, wurden noch drei weitere Sonden (Wasser-, Methyl- und Guanidin-Sonde) ausgewählt, da die Ergebnisse, die mit diesen Sonden gewonnenen wurden, aufgrund des Vorkommens in Aminosäuren des aktiven Zentrum des Enzyms leichter auf die Proteinumgebung zu übertragen sind.

Die anschließende Untersuchung des Einflusses des Ladungszustandes der Liganden auf die 3D-QSAR-Analyse zeigte, dass die Vorhersagen bessere Ergebnisse liefert, wenn nicht nur die Liganden, die eine Carbonsäure enthalten, sondern auch die phenolischen Liganden deprotoniert und somit negativ geladen vorlagen. Dieses Ergebnis deckte sich mit den Untersuchungen zum Ladungszustand von Cys215 im aktiven Zentrum, das ebenfalls deprotoniert vorliegt, obwohl es ebenso wie das Phenol einen theoretischen pK_S -Wert von etwa zehn besitzt. Der pK_S -Wert des Cysteins wird durch den elektrostatischen Einfluss der Enzymumgebung auf etwa vier gesenkt. Da die phenolische Hydroxylgruppe in unmittelbarer Nähe im aktiven Zentrum bindet, trifft diese Veränderung des pK_S -Werts mit hoher Wahrscheinlichkeit auch auf diese Gruppe zu.

Für die racemisch vorliegenden Verbindungen des Datensatzes wurde wegen der Lage des Asymmetriezentrums und der Ähnlichkeit der überlagerten Enantiomere die Vermutung aufgestellt, dass beide Enantiomere im aktiven Zentrum binden können. Durch die 3D-QSAR-Untersuchungen konnte diese

Vermutung bestätigt werden. Zusätzlich wurde gezeigt, dass beide Enantiomere zur Aktivität des Gemisches etwa gleich viel beitragen.

Bei 3D-QSAR-Untersuchungen ergibt sich im Allgemeinen das Problem, dass auch bei intern validierten Modellen (hohe q^2 -Werte) keine genaue Vorhersage eines externen Testsatzes möglich ist. Häufig sind Modelle mit hohen q^2 -Werten übertrainiert, und deshalb nur in der Lage Vorhersagen innerhalb des Modells zu machen. Deshalb ist eine Einteilung des Datensatzes in einen Trainingssatz und einen Testsatz sinnvoll, so wie es bei der Vorhersage der Racemate durchgeführt wurde.

Für die aktivste Verbindung des Datensatzes (Verbindung 104) war es nach anfänglichen Schwierigkeiten möglich, das Docking in die Bindungstasche der Kristallstruktur durchzuführen, Wechselwirkungskomplexe mit der Proteinstruktur der Phosphatase zu erzeugen. Darauf aufbauend wurde für die weiteren Verbindungen des Datensatzes eine dockingbasierte Überlagerung erzeugt.

Dazu war es einerseits notwendig, Röntgenkristallstrukturen der PTP1B einer Moleküldynamiksimulation zu unterziehen, um weitere Konformationen für das aktive Zentrum zu generieren. Bei der Proteinkonformation, mit der das Docking möglich war, ist das Protein-Rückrat leicht verschoben und die Konformation der Seitenketten hat sich geändert, wodurch im aktiven Zentrum mehr Platz für den Liganden zu Verfügung steht. Ein Vergleich dieser Konformation mit anderen Röntgenkristallstrukturen, anhand der PCA-Analyse, legte nahe, dass man für das Docking auch die Strukturen 1PTU, 1KAK, 1KAV, 1G7G oder 1JF7 hätte verwenden können. Dabei ist jedoch zu berücksichtigen, dass es sich bei 1PTU um die Cys215Ser-Mutante des Enzyms handelt, und dass die anderen Strukturen nicht genügend Platz im aktiven Zentrum besitzen.

Des Weiteren war es notwendig, das Programm Autodock so zu optimieren, dass nicht nur einzelne Verbindungen, sondern ein größerer Datensatz automatisch gedockt werden konnte. Mit der Schaffung dieser Voraussetzungen für ein rationelles Docking konnten dockingbasierte Überlagerungen, die sich signifikant von den ligandbasierten Überlagerungen unterscheiden, generiert werden. Daraus wurden anschließend 3D-QSAR-Modelle erstellt, die ebenfalls eine gute Vorhersage der biologischen Aktivität erlaubten. Diese durchgeführte

Kombination von proteinbasierten (Docking) und ligandbasierten (3D-QSAR) Vorgehensweisen, die die Vorteile beider Verfahren nutzen und die grundsätzlichen Probleme der Konformationsanalyse und Überlagerung einerseits und der Energieberechnung andererseits umgehen sollte, kann als protein- oder rezeptorbasiertes 3D-QSAR bezeichnet werden [11][12][13][14]. Die Unterschiede in den Überlagerungen lassen sich sehr gut mit den zugrunde liegenden Verfahren erklären. Das proteinbasierte Vorgehen führt zu einer sehr guten Superpositionierung der sauren Funktion der Inhibitoren, die für die Bindung an das Enzym besonders wichtig ist. Da die großen lipophilen Grundgerüste keine gerichteten Wechselwirkungen mit der Phosphatase eingehen, ist es nachvollziehbar, dass eine exakte Überlagerung der einzelnen Ringsysteme nicht erfolgt. Hingegen erkennt man beim ligandbasierten Ansatz, dass die Überlagerung in erster Linie auf den Ringsystemen der Grundkörper basierte und Seitenketten eine untergeordnete Rolle spielten.

Zwischen Interaktionsfeldern mit verschiedenen GRID-Sonden und den biologischen Aktivitäten ergab sich auch in den proteinbasierten Untersuchungen eine hohe Korrelation. Dadurch besteht die Möglichkeit, mit den erstellten 3D-QSAR-Modellen Vorhersagen der Aktivität nicht getesteter Verbindungen zu treffen. Diese Vorhersagen waren für andere Verfahren, die nur Wechselwirkungsenergien oder Scoringfunktionen in der Korrelationsanalyse berücksichtigen, nicht möglich. Für das schlechte Abschneiden dieser Verfahren gibt es eine Reihe von Ursachen, die in der Literatur beschrieben sind [278][279][280][281]. Einerseits werden Solvation und Desolvation von Protein und Ligand nicht berücksichtigt, andererseits haben geringe Änderungen in der relativen Position und Konformation des Liganden oft starke Auswirkungen in den Kraftfeldberechnungen. Auch die Änderung der Anzahl der Freiheitsgrade der beteiligten Bindungspartner und die Dynamik und Flexibilität des Proteins bleiben bisher ebenfalls unberücksichtigt. Ein weiteres Problem stellt das korrekte Platzieren von Wassermolekülen bei Docking Studien dar.

Für die Datenbanksuchen ist eine breitere Datenbasis notwendig als für die Untersuchung von Struktur-Wirkungsbeziehungen. Bei den QSAR-Analysen sollte der Datensatz möglichst homogen sein, aber bei der Suche nach neuen

potentiellen Inhibitoren sollten auch Verbindungen mit einbezogen werden, die sich für die gemeinsame Betrachtung in einem QSAR-Modell nicht eignen. Daher wurden zu diesem Zweck weitere Verbindungen herangezogen, die sich vom ursprünglichen 3D-QSAR Datensatz hinsichtlich ihrer chemischen Diversität unterscheiden, aber unter den gleichen Bedingungen biologisch getestet wurden.

Für die virtuelle Suche nach neuen Inhibitoren für PTP1B bieten sich mehrere Methoden an, die sich in ligandbasierte Methoden wie zum Beispiel die Ähnlichkeitssuche mit FeatureTrees und die protein- oder strukturbasierten Methoden des molekularen Dockings unterteilen lassen. Bei den durchgeführten Datenbankrecherchen stellte sich heraus, dass ein auf einer Ähnlichkeitssuche mit den bekannten Inhibitoren basierendes Screening nur wenig geeignet ist, um in den verwendeten Moleküldatenbanken nach neuen Hemmstoffen für die PTP1B zu suchen. Aufgrund der besonderen strukturellen Eigenschaften der bekannten Inhibitoren, deren hohe Lipophilie, die fast immer vorhandene Carbonsäure- oder Phenolgruppe und das Fehlen weiterer spezifischer Interaktionspunkte, weisen diese Verbindungen große Unterschiede zu den in den verfügbaren Moleküldatenbanken enthaltenen Verbindungen auf, so dass eine auf Ähnlichkeit basierende Suche nicht zum gewünschten Ziel führen konnte.

Der erfolgreichere Ansatz war die proteinbasierte Vorgehensweise, bei der alle Moleküle oder eine diverse, repräsentative Auswahl aus dem Datensatz gedockt wurden. Diese Methode ist auch dann geeignet, wenn neben der Proteininformation keine weiteren Informationen wie zum Beispiel bekannte Liganden zu Verfügung stehen. Je nach verwendeter Kombination aus Dockingmethode und Scoringfunktion sowie der gewählten Grenzwerte lassen sich mit diesem Verfahren gute Anreicherungen der aktiven Verbindungen erzielen. Als besonders geeignet für das virtuelle Screening nach PTP1B-Inhibitoren haben sich die Kombination von Autodock mit X-Score und die Kombination von FlexX mit ChemScore erwiesen, mit denen ein Anreicherungsfaktor von acht erreicht wurde. Damit ließ sich die Anzahl der Verbindungen, die experimentell unter-

sucht werden müssen, um eine vorgegebene Anzahl von aktiven Substanzen zu erhalten, stark einschränken und die Hitrate erhöhen.

Weitere Ansätze könnten in Zukunft angewandt werden um die Hitrate des virtuellen Screenings zu erhöhen. Eine Möglichkeit der Optimierung besteht in der Verwendung von mehreren Proteinkonformationen, um die Flexibilität des aktiven Zentrums in gewissem Umfang zu berücksichtigen. Diese Vorgehensweise wurde nur für einige wenige bekannte Liganden durchgeführt, um eine geeignete Konformation des aktiven Zentrums zu finden, sie würde sich aber auch für das Docking beim virtuellen Screening eignen, da davon auszugehen ist, dass die unterschiedlichen Verbindungen im Screening auch unterschiedliche Konformationen des aktiven Zentrums benötigen. Diese Berechnungen benötigen jedoch ein Vielfaches der Rechenzeit, so dass sich andere Verfahren eher anbieten wie zum Beispiel das so genannte Soft-Docking, bei dem die repulsiven Terme in einer stark abgeschwächten Form in die Berechnungen eingehen [278]. Eine weitere Möglichkeit, das Docking zu verbessern oder in eine bestimmte Richtung zu lenken, ist die Verwendung von so genannten „pharmakophoren Constraints“. Dabei werden beim Docking nur solche Komplexe zugelassen, die gewisse Nebenbedingungen erfüllen wie beispielsweise das Vorkommen bestimmter Wasserstoffbrücken. Dadurch lässt sich vor allem beim virtuellen Screening die Anreicherung verbessern [279]. Das gleiche Ziel verfolgt die Vorselektion bestimmter Verbindungen für das Docking anhand von einfachen physikochemischen Deskriptoren zu einer für das Target spezifischen fokussierten Datenbank [279]. Durch diese beiden Einschränkungen soll insbesondere eine artifizielle Anreicherung beim Screening verhindert werden. Des Weiteren wird an einer Optimierung oder Neuentwicklung der verwendeten Scoringfunktionen gearbeitet. In der aktuellen Literatur findet sich zum Beispiel ein aktueller Vergleich verschiedener Scoringfunktionen und der Kombination aus mehreren unterschiedlichen Funktionen in [280] und [281]. Neue Scoringfunktionen werden häufig speziell für bestimmte Zielproteine konzipiert, indem viele mögliche Deskriptoren für Protein-Ligand-Interaktionen berechnet werden. Mittels verschiedener geeigneter Methoden wie zum Beispiel die Regressionsanalyse [282] und die

Verwendung artifizierlicher neuronaler Netze [283] werden die Deskriptoren ausgewählt, die mit den biologischen Bindungsdaten am besten korrelieren. Diese vielfältigen Optimierungsansätze des Screenings mit Hilfe des Dockings sollen in Zukunft analysiert werden und können gegebenenfalls in Kombination eingesetzt werden, um die Effektivität dieses Verfahrens weiter zu verbessern.

Seit Beginn dieser Arbeit sind zwei Veröffentlichungen erschienen [284][285], die sich mit Struktur-Wirkungs-Beziehungen des hier behandelten Datensatzes beschäftigen. In beiden Fällen wurden die Racemate und die Verbindungen, für die keine IC_{50} -Werte bestimmt wurden, von den Untersuchungen ausgeschlossen. Die in der vorliegenden Arbeit nicht betrachteten Sulfonylbiphenyle, die einen anderen Bindungsmodus aufweisen, wurden jedoch in die Untersuchungen mit aufgenommen. Für die erste Untersuchung wurden die Verbindungen in einen Trainingssatz mit 92 Substanzen und einen Testsatz mit 26 Substanzen eingeteilt. Die zweite Analyse benutzte diese Einteilung, verwendete aber zusätzlich die Verbindungen ohne IC_{50} -Werte als unbekanntes Testsatz.

In der ersten Studie [284] wurden CoMFA- und CoMSIA-Verfahren eingesetzt. Die Konformationsanalyse und die Überlagerung wurden mittels der in Sybyl enthaltenen Methoden Simulated Annealing, RMS-Fit, Multiple-Fit und QSAR-Rigid-Body-Field-Fit durchgeführt. Die sterischen Interaktionen trugen zu 47 % zur Aktivität der Verbindungen bei, die elektrostatischen zu 53 %. Wurde neben den CoMFA-Feldern der ALogP-Wert berücksichtigt, so war der Beitrag der sterischen, elektrostatischen und ALogP-Terme 35 %, 43 % und 22 %. Die ebenfalls untersuchten HOMO-Deskriptoren leisteten nur einen geringen Beitrag zu den Modellen. Das beste CoMFA-Modell resultierte aus der RMS-Fit-Routine und besitzt einen r^2 -Wert von 0,84 und einen q^2_{LOO} -Wert von 0,62.

Das CoMSIA-Modell erreichte die besten Ergebnisse, berücksichtigte aber nur die sterischen (Beitrag 37 %), elektrostatischen (36 %) und lipophilen (27 %) Interaktionen, und besaß einen r^2 -Wert von 0,91 und einen q^2_{LOO} -Wert von 0,60. Die Interaktionen für Wasserstoffbrücken-Akzeptoren und Donoren hatten nur einen kleinen Beitrag an den Modellen und wurden deshalb nicht verwen-

det. Die resultierenden Interaktionsfelder sind den in dieser Arbeit erzielten Ergebnissen sehr ähnlich.

Die zweite Untersuchung verfolgte einen neuen und völlig anderen Ansatz. An Stelle von Interaktionsfeldern wurden neben berechneten physikochemischen Eigenschaften so genannte chemische Building-Blocks als Deskriptoren verwendet. Insgesamt enthielten die 118 Verbindungen 509 Building-Blocks und 71 so genannte Makrostrukturen. Nach einer einfachen Vorselektion der Deskriptoren aufgrund der diesen assoziierten IC_{50} -Werte, wurden die wichtigsten Deskriptoren mittels einer PLS-Analyse ausgewählt. Das beste Modell ohne die physikochemischen Deskriptoren verwendete noch 50 Building-Blocks oder Makrostrukturen in 12 PLS-Hauptkomponenten und besaß einen r^2 -Wert von 0,83 und einen q^2_{LOO} -Wert von 0,71. Wurden nur die 71 Makrostrukturen und die physikochemischen Eigenschaften in einem Modell mit 5 Hauptkomponenten benutzt, so wurde ein r^2 -Wert von 0,83 und ein q^2_{LOO} -Wert von 0,71 erzielt. Die Auswertung der einflussreichen Building-Blocks und Makrostrukturen ist auf sehr ähnliche Derivate der untersuchten Substanzen beschränkt, ebenso wie die Analyse der CoMFA- oder CoMSIA-Felder.

Mit den in der vorliegenden Arbeit verwendeten Methoden wurden bessere q^2 -Werte erzielt, obwohl die durchgeführte Kreuzvalidierung mit der anspruchsvolleren L20%O-Methode erfolgte und nicht mit der anfälligeren LOO-Methode. Beispielsweise lieferte das ligandbasierte Modell einen $q^2_{L20\%O}$ -Wert von 0,76 im Vergleich zu einem q^2_{LOO} -Wert von 0,71 der beiden anderen Veröffentlichungen. Proteinbasierte Überlagerungen für die QSAR-Analysen und weitergehende Untersuchungen zu neuen Liganden wie das virtuelle Screening wurden in den beiden Arbeiten nicht durchgeführt.

7 Zusammenfassung

Im Rahmen der vorliegenden Arbeit wurden die zurzeit verfügbaren Röntgenkristallstrukturen der Protein-Tyrosin-Phosphatase (PTP) 1B untersucht. Dazu wurde ein Ansatz gewählt, der die Interaktionsfelder in der Nähe der Bindungstasche mit den statistischen Mitteln der Hauptkomponentenanalyse auswertet und es ermöglicht, die Konformationen des aktiven Zentrums auf dieser Grundlage zu klassifizieren. Dieser Ansatz konnte auf die Strukturen anderer Phosphatasen übertragen werden und ermöglicht somit den Vergleich von homologen Proteinen auf der Basis von Wechselwirkungen mit der Bindungsstelle.

Im nächsten Schritt wurden ligandbasierte und rezeptorbasierte 3D-QSAR-Modelle für Inhibitoren der PTP1B aufgestellt, die in der Lage sind, die unterschiedliche inhibitorische Aktivität der einzelnen Liganden quantitativ zu erklären. Dazu wurden die Wechselwirkungsfelder der Inhibitoren und die biologischen Aktivitäten mittels multivariater Datenanalyse korreliert. Die zugrunde liegenden Superpositionierungen entstammen bei den ligandbasierten Modellen einer Überlagerung der Grundgerüste der Inhibitoren. Die proteinbasierten Überlagerungen wurden mit Hilfe des Dockings mit einer zu diesem Zweck optimierten Version des Programms Autodock modelliert. Sie zeigen zu den ligandbasierten Überlagerungen signifikante Unterschiede, die sehr gut durch die Eigenheiten der verwendeten Methoden erklärt werden können. In allen Fällen ergaben sich nach einer Überprüfung durch eine Kreuzvalidierung signifikante und prädiktive Modelle. Vermutungen über den Ladungszustand und die Konfiguration am chiralen Zentrum einiger Liganden konnten anhand der 3D-QSAR-Modelle bestätigt werden.

Für das virtuelle Screening war neben dem World Drug Index und der Maybridge Datenbank eine proprietäre Datenbank der Firma EvotecOAI verfügbar, die eine Reihe aktiver Verbindungen enthält. Mittels Docking und unter Verwendung geeigneter Scoringfunktionen war es möglich, die Inhibitoren im Screening wieder zu finden und es konnte durch eine gute Anreicherung der aktiven Verbindungen gezeigt werden, dass es mit den gewählten Methoden möglich ist, Inhibitoren für die PTP1B aufzuspüren, die sich von den bekannten Liganden strukturell unterscheiden.

8 Literatur

- [1] WHO Fact Sheet 138: "Diabetes mellitus"
<http://www.who.int/mediacentre/factsheets/fs138/en/>, (2002)
- [2] WHO Fact Sheet 236: "Diabetes: the cost of diabetes"
<http://www.who.int/mediacentre/factsheets/fs236/en/>, (2002)
- [3] Neel BG, Tonks NK: "Protein tyrosine phosphatases in signal transduction." *Curr Opin Cell Biol* 9, 193-204 (1997)
- [4] Elchebly M, Payette P, Michaliszyn E, Cromlish W, Collins S, Loy AL, Normandin D, Cheng A, Himms-Hagen J, Chan CC, Ramachandran C, Gresser MJ, Tremblay ML, Kennedy BP: "Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene." *Science* 283, 1544-8 (1999)
- [5] Klamon LD, Boss O, Peroni OD, Kim JK, Martino JL, Zabolotny JM, Moghal N, Lubkin M, Kim YB, Sharpe AH, Stricker-Krongrad A, Shulman GI, Neel BG, Kahn BB: "Increased energy expenditure, decreased adiposity, and tissue-specific insulin sensitivity in protein-tyrosine phosphatase 1B-deficient mice." *Mol Cell Biol* 20, 5479-89 (2000)
- [6] Ahmad F, Azevedo JL, Cortright R, Dohm GL, Goldstein BJ: "Alterations in skeletal muscle protein-tyrosine phosphatase activity and expression in insulin-resistant human obesity and diabetes." *J Clin Invest* 100, 449-58 (1997)
- [7] Di Paola R, Frittitta L, Miscio G, Bozzali M, Baratta R, Centra M, Spampinato D, Santagati MG, Ercolino T, Cisternino C, Soccio T, Mastroianno S, Tassi V, Almgren P, Pizzuti A, Vigneri R, Trischitta V: "A variation in 3' UTR of hPTP1B increases specific gene expression and associates with insulin resistance." *Am J Hum Genet* 70, 806-12 (2002)
- [8] Jia Z, Barford D, Flint AJ, Tonks NK: "Structural basis for phosphotyrosine peptide recognition by protein tyrosine phosphatase 1B." *Science* 268, 1754-8 (1995)
- [9] Zabolotny JM, Haj FG, Kim YB, Kim HJ, Shulman GI, Kim JK, Neel BG, Kahn BB: "Transgenic overexpression of protein-tyrosine phosphatase 1B in muscle causes insulin resistance, but overexpression with leukocyte antigen-related phosphatase does not additively impair insulin action." *J Biol Chem* 279, 24844-51 (2004)
- [10] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: "The Protein Data Bank." *Nucleic Acids Res* 28, 235-42 (2000)
- [11] Sippl W: "Receptor-based 3D QSAR analysis of estrogen receptor ligands--merging the accuracy of receptor-based alignments with the computational efficiency of ligand-based methods." *J Comput Aided Mol Des* 14, 559-72 (2000)
- [12] Sippl W, Contreras JM, Parrot I, Rival YM, Wermuth CG: "Structure-based 3D QSAR and design of novel acetylcholinesterase inhibitors." *J Comput Aided Mol Des* 15, 395-410 (2001)

- [13] Sippl W: "Development of biologically active compounds by combining 3D QSAR and structure-based design methods." *J Comput Aided Mol Des* 16, 825-30 (2002)
- [14] Sippl W: "Binding affinity prediction of novel estrogen receptor ligands using receptor-based 3-D QSAR methods." *Bioorg Med Chem* 10, 3741-55 (2002)
- [15] Krebs EG, Fischer EH: "Phosphorylase activity of skeletal muscle extracts." *J Biol Chem* 216, 113-20 (1955)
- [16] Fischer EH, Krebs EG: "Conversion of phosphorylase b to phosphorylase a in muscle extracts." *J Biol Chem* 216, 121-32 (1955)
- [17] Sutherland EW Jr, Wosilait WD: "Inactivation and activation of liver phosphorylase." *Nature* 175, 169-70 (1955)
- [18] Krebs EG, Beavo JA: "Phosphorylation-dephosphorylation of enzymes." *Annu Rev Biochem* 48, 923-59 (1979)
- [19] Karin M, Hunter T: "Transcriptional control by protein phosphorylation: signal transmission from the cell surface to the nucleus." *Curr Biol* 5, 747-57 (1995)
- [20] Zhang ZY: "Protein tyrosine phosphatases: prospects for therapeutics." *Curr Opin Chem Biol* 5, 416-23 (2001)
- [21] Kahn CR: "Banting Lecture. Insulin action, diabetogenes, and the cause of type II diabetes." *Diabetes* 43, 1066-84 (1994)
- [22] Czech MP, Corvera S: "Signaling mechanisms that regulate glucose transport." *J Biol Chem* 274, 1865-8 (1999)
- [23] White MF: "The IRS-signaling system: a network of docking proteins that mediate insulin and cytokine action." *Recent Prog Horm Res* 53, 119-38 (1998)
- [24] Ruderman NB, Kapeller R, White MF, Cantley LC: "Activation of phosphatidylinositol 3-kinase by insulin." *Proc Natl Acad Sci U S A* 87, 1411-5 (1990)
- [25] Cheatham B, Vlahos CJ, Cheatham L, Wang L, Blenis J, Kahn CR: "Phosphatidylinositol 3-kinase activation is required for insulin stimulation of pp70 S6 kinase, DNA synthesis, and glucose transporter translocation." *Mol Cell Biol* 14, 4902-11 (1994)
- [26] Saltiel AR, Kahn CR: "Insulin signalling and the regulation of glucose and lipid metabolism." *Nature* 414, 799-806 (2001)
- [27] Kruszynska YT, Olefsky JM: "Cellular and molecular mechanisms of non-insulin dependent diabetes mellitus." *J Investig Med* 44, 413-28 (1996)
- [28] Cicirelli MF, Tonks NK, Diltz CD, Weiel JE, Fischer EH, Krebs EG: "Microinjection of a protein-tyrosine-phosphatase inhibits insulin action in *Xenopus* oocytes." *Proc Natl Acad Sci U S A* 87, 5514-8 (1990)

- [29] Tonks NK, Cicirelli MF, Diltz CD, Krebs EG, Fischer EH: "Effect of microinjection of a low-Mr human placenta protein tyrosine phosphatase on induction of meiotic cell division in *Xenopus* oocytes." *Mol Cell Biol* 10, 458-63 (1990)
- [30] Kenner KA, Anyanwu E, Olefsky JM, Kusari J: "Protein-tyrosine phosphatase 1B is a negative regulator of insulin- and insulin-like growth factor-I-stimulated signaling." *J Biol Chem* 271, 19810-6 (1996)
- [31] Seely BL, Staubs PA, Reichart DR, Berhanu P, Milarski KL, Saltiel AR, Kusari J, Olefsky JM: "Protein tyrosine phosphatase 1B interacts with the activated insulin receptor." *Diabetes* 45, 1379-85 (1996)
- [32] Bandyopadhyay D, Kusari A, Kenner KA, Liu F, Chernoff J, Gustafson TA, Kusari J: "Protein-tyrosine phosphatase 1B complexes with the insulin receptor in vivo and is tyrosine-phosphorylated in the presence of insulin." *J Biol Chem* 272, 1639-45 (1997)
- [33] Walchli S, Curchod ML, Gobert RP, Arkinstall S, Hooft van Huijsduijnen R: "Identification of tyrosine phosphatases that dephosphorylate the insulin receptor. A brute force approach based on "substrate-trapping" mutants." *J Biol Chem* 275, 9792-6 (2000)
- [34] Goldstein BJ, Bittner-Kowalczyk A, White MF, Harbeck M: "Tyrosine dephosphorylation and deactivation of insulin receptor substrate-1 by protein-tyrosine phosphatase 1B. Possible facilitation by the formation of a ternary complex with the Grb2 adaptor protein." *J Biol Chem* 275, 4283-9 (2000)
- [35] Calera MR, Vallega G, Pilch PF: "Dynamics of protein-tyrosine phosphatases in rat adipocytes." *J Biol Chem* 275, 6308-12 (2000)
- [36] Arregui CO, Balsamo J, Lilien J: "Impaired integrin-mediated adhesion and signaling in fibroblasts expressing a dominant-negative mutant PTP1B." *J Cell Biol* 143, 861-73 (1998)
- [37] Liu F, Hill DE, Chernoff J: "Direct binding of the proline-rich region of protein tyrosine phosphatase 1B to the Src homology 3 domain of p130(Cas)." *J Biol Chem* 271, 31290-5 (1996)
- [38] Liu F, Sells MA, Chernoff J: "Protein tyrosine phosphatase 1B negatively regulates integrin signaling." *Curr Biol* 8, 173-6 (1998)
- [39] Cheng A, Bal GS, Kennedy BP, Tremblay ML: "Attenuation of adhesion-dependent signaling and cell spreading in transformed fibroblasts lacking protein tyrosine phosphatase-1B." *J Biol Chem* 276, 25848-55 (2001)
- [40] Balsamo J, Arregui C, Leung T, Lilien J: "The nonreceptor protein tyrosine phosphatase PTP1B binds to the cytoplasmic domain of N-cadherin and regulates the cadherin-actin linkage." *J Cell Biol* 143, 523-32 (1998)
- [41] Rhee J, Lilien J, Balsamo J: "Essential tyrosine residues for interaction of the non-receptor protein-tyrosine phosphatase PTP1B with N-cadherin." *J Biol Chem* 276, 6640-4 (2001)

- [42] Flint AJ, Tiganis T, Barford D, Tonks NK: "Development of "substrate-trapping" mutants to identify physiological substrates of protein tyrosine phosphatases." *Proc Natl Acad Sci U S A* 94, 1680-5 (1997)
- [43] Lee SR, Kwon KS, Kim SR, Rhee SG: "Reversible inactivation of protein-tyrosine phosphatase 1B in A431 cells stimulated with epidermal growth factor." *J Biol Chem* 273, 15366-72 (1998)
- [44] Liu F, Chernoff J: "Protein tyrosine phosphatase 1B interacts with and is tyrosine phosphorylated by the epidermal growth factor receptor." *Biochem J* 327, 139-45 (1997)
- [45] Milarski KL, Zhu G, Pearl CG, McNamara DJ, Dobrusin EM, MacLean D, Thieme-Seffler A, Zhang ZY, Sawyer T, Decker SJ, et al.: "Sequence specificity in recognition of the epidermal growth factor receptor by protein tyrosine phosphatase 1B." *J Biol Chem* 268, 23634-9 (1993)
- [46] Roome J, O'Hare T, Pilch PF, Brautigan DL: "Protein phosphotyrosine phosphatase purified from the particulate fraction of human placenta dephosphorylates insulin and growth-factor receptors." *Biochem J* 256, 493-500 (1988)
- [47] Myers MP, Andersen JN, Cheng A, Tremblay ML, Horvath CM, Parisien JP, Salmeen A, Barford D, Tonks NK: "TYK2 and JAK2 are substrates of protein-tyrosine phosphatase 1B." *J Biol Chem* 276, 47771-4 (2001)
- [48] Zabolotny JM, Bence-Hanulec KK, Stricker-Krongrad A, Haj F, Wang Y, Minokoshi Y, Kim YB, Elmquist JK, Tartaglia LA, Kahn BB, Neel BG: "PTP1B regulates leptin signal transduction in vivo." *Dev Cell* 2, 489-95 (2002)
- [49] Cheng A, Uetani N, Simoncic PD, Chaubey VP, Lee-Loy A, McGlade CJ, Kennedy BP, Tremblay ML: "Attenuation of leptin action and regulation of obesity by protein tyrosine phosphatase 1B." *Dev Cell* 2, 497-503 (2002)
- [50] Flint AJ, Gebbink MF, Franza BR Jr, Hill DE, Tonks NK: "Multi-site phosphorylation of the protein tyrosine phosphatase, PTP1B: identification of cell cycle regulated and phorbol ester stimulated sites of phosphorylation." *EMBO J* 12, 1937-46 (1993)
- [51] Schievella AR, Paige LA, Johnson KA, Hill DE, Erikson RL: "Protein tyrosine phosphatase 1B undergoes mitosis-specific phosphorylation on serine." *Cell Growth Differ* 4, 239-46 (1993)
- [52] Shifrin VI, Neel BG: "Growth factor-inducible alternative splicing of nontransmembrane phosphotyrosine phosphatase PTP-1B pre-mRNA." *J Biol Chem* 268, 25376-84 (1993)
- [53] Shifrin VI, Davis RJ, Neel BG: "Phosphorylation of protein-tyrosine phosphatase PTP-1B on identical sites suggests activation of a common signaling pathway during mitosis and stress response in mammalian cells." *J Biol Chem* 272, 2957-62 (1997)

- [54] Dube N, Cheng A, Tremblay ML: "The role of protein tyrosine phosphatase 1B in Ras signaling." *Proc Natl Acad Sci U S A* 101, 1834-9 (2004)
- [55] Dadke S, Kusari J, Chernoff J: "Down-regulation of insulin signaling by protein-tyrosine phosphatase 1B is mediated by an N-terminal binding region." *J Biol Chem* 275, 23642-7 (2000)
- [56] Salmeen A, Andersen JN, Myers MP, Tonks NK, Barford D: "Molecular basis for the dephosphorylation of the activation segment of the insulin receptor by protein tyrosine phosphatase 1B." *Mol Cell* 6, 1401-12 (2000)
- [57] Ahmad F, Li PM, Meyerovitch J, Goldstein BJ: "Osmotic loading of neutralizing antibodies demonstrates a role for protein-tyrosine phosphatase 1B in negative regulation of the insulin action pathway." *J Biol Chem* 270, 20503-8 (1995)
- [58] Chen H, Cong LN, Li Y, Yao ZJ, Wu L, Zhang ZY, Burke TR Jr, Quon MJ: "A phosphotyrosyl mimetic peptide reverses impairment of insulin-stimulated translocation of GLUT4 caused by overexpression of PTP1B in rat adipose cells." *Biochemistry* 38, 384-9 (1999)
- [59] Maegawa H, Ide R, Hasegawa M, Ugi S, Egawa K, Iwanishi M, Kikkawa R, Shigeta Y, Kashiwagi A: "Thiazolidine derivatives ameliorate high glucose-induced insulin resistance via the normalization of protein-tyrosine phosphatase activities." *J Biol Chem* 270, 7724-30 (1995)
- [60] Forsell PA, Boie Y, Montalibet J, Collins S, Kennedy BP: "Genomic characterization of the human and mouse protein tyrosine phosphatase-1B genes." *Gene* 260, 145-53 (2000)
- [61] Lembertas AV, Perusse L, Chagnon YC, Fislser JS, Warden CH, Purcell-Huynh DA, Dionne FT, Gagnon J, Nadeau A, Lusia AJ, Bouchard C: "Identification of an obesity quantitative trait locus on mouse chromosome 2 and evidence of linkage to body fat and insulin on the human homologous region 20q." *J Clin Invest* 100, 1240-7 (1997)
- [62] Echwald SM, Bach H, Vestergaard H, Richelsen B, Kristensen K, Drivsholm T, Borch-Johnsen K, Hansen T, Pedersen O: "A P387L variant in protein tyrosine phosphatase-1B (PTP-1B) is associated with type 2 diabetes and impaired serine phosphorylation of PTP-1B in vitro." *Diabetes* 51, 1-6 (2002)
- [63] Mok A, Cao H, Zinman B, Hanley AJ, Harris SB, Kennedy BP, Hegele RA: "A single nucleotide polymorphism in protein tyrosine phosphatase PTP-1B is associated with protection from diabetes or impaired glucose tolerance in Oji-Cree." *J Clin Endocrinol Metab* 87, 724-7 (2002)
- [64] Santaniemi M, Ukkola O, Kesaniemi YA: "Tyrosine phosphatase 1B and leptin receptor genes and their interaction in type 2 diabetes." *J Intern Med* 256, 48-55 (2004)

- [65] Kipfer-Coudreau S, Eberle D, Sahbatou M, Bonhomme A, Guy-Grand B, Froguel P, Galan P, Basdevant A, Clement K: "Single nucleotide polymorphisms of protein tyrosine phosphatase 1B gene are associated with obesity in morbidly obese French subjects." *Diabetologia* 47, 1278-84 (2004)
- [66] Tobin JF, Tam S: "Recent advances in the development of small molecule inhibitors of PTP1B for the treatment of insulin resistance and type 2 diabetes." *Curr Opin Drug Discov Devel* 5, 500-12 (2002)
- [67] Rotella DP: "Novel "Second-Generation" Approaches for the Control of Type 2 Diabetes" *J Med Chem* 47, 4111-2 (2004)
- [68] Skyler JS: "Diabetes Mellitus: Pathogenesis and treatment Strategies" *J Med Chem* 47, 4113-7 (2004)
- [69] Hooft van Huijsduijnen R, Sauer WHB, Bombrun A, Swinnen D.: "Prospects for Inhibitors of Protein Tyrosine Phosphatase 1B as Antidiabetic Drugs" *J Med Chem* 47, 4142-6 (2004)
- [70] Knudsen LB: "Glucagon-like Peptide-1: The Basis of a New Class of Treatment for Type 2 Diabetes" *J Med Chem* 47, 4128-34 (2004)
- [71] Weber AE: "Dipeptidyl Peptidase IV Inhibitors for the Treatment of Diabetes" *J Med Chem* 47, 4135-41 (2004)
- [72] Henke BR: "Peroxisome Proliferator-Activated Receptor alpha/gamma Dual Agonists for the Treatment of Type 2 Diabetes" *J Med Chem* 47, 4118-27 (2004)
- [73] Friedman JM: "Obesity in the new millennium." *Nature* 404, 632-4 (2000)
- [74] Tartaglia LA, Dembski M, Weng X, Deng N, Culpepper J, Devos R, Richards GJ, Campfield LA, Clark FT, Deeds J, et al.: "Identification and expression cloning of a leptin receptor, OB-R." *Cell* 83, 1263-71 (1995)
- [75] Ahima RS, Prabakaran D, Mantzoros C, Qu D, Lowell B, Maratos-Flier E, Flier JS: "Role of leptin in the neuroendocrine response to fasting." *Nature* 382, 250-2 (1996)
- [76] Bjorbaek C, Uotani S, da Silva B, Flier JS: "Divergent signaling capacities of the long and short isoforms of the leptin receptor." *J Biol Chem* 272, 32686-95 (1997)
- [77] Cook WS, Unger RH: "Protein tyrosine phosphatase 1B: a potential leptin resistance factor of obesity." *Dev Cell* 2, 385-7 (2002)
- [78] Lam NT, Lewis JT, Cheung AT, Luk CT, Tse J, Wang J, Bryer-Ash M, Kolls JK, Kieffer TJ: "Leptin increases hepatic insulin sensitivity and protein tyrosine phosphatase 1B expression." *Mol Endocrinol* 18, 1333-45 (2004)

- [79] Rondinone CM, Trevillyan JM, Clampit J, Gum RJ, Berg C, Kroeger P, Frost L, Zinker BA, Reilly R, Ulrich R, Butler M, Monia BP, Jirousek MR, Waring JF: "Protein tyrosine phosphatase 1B reduction regulates adiposity and expression of genes involved in lipogenesis." *Diabetes* 51, 2405-11 (2002)
- [80] Hoffman BT, Nelson MR, Burdick K, Baxter SM: "Protein tyrosine phosphatases: strategies for distinguishing proteins in a family containing multiple drug targets and anti-targets." *Curr Pharm Des* 10, 1161-81 (2004)
- [81] Griffith JP, Kim JL, Kim EE, Sintchak MD, Thomson JA, Fitzgibbon MJ, Fleming MA, Caron PR, Hsiao K, Navia MA: "X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex." *Cell* 82, 507-22 (1995)
- [82] Kissinger CR, Parge HE, Knighton DR, Lewis CT, Pelletier LA, Tempczyk A, Kalish VJ, Tucker KD, Showalter RE, Moomaw EW, et al.: "Crystal structures of human calcineurin and the human FKBP12-FK506-calcineurin complex." *Nature* 378, 641-4 (1995)
- [83] Goldberg J, Huang HB, Kwon YG, Greengard P, Nairn AC, Kuriyan J: "Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1." *Nature* 376, 745-53 (1995)
- [84] Barford D, Das AK, Egloff MP: "The structure and mechanism of protein phosphatases: insights into catalysis and regulation." *Annu Rev Biophys Biomol Struct* 27, 133-64 (1998)
- [85] Li L, Dixon JE: "Form, function, and regulation of protein tyrosine phosphatases and their involvement in human diseases." *Semin Immunol* 12, 75-84 (2000)
- [86] Denu JM, Lohse DL, Vijayalakshmi J, Saper MA, Dixon JE: "Visualization of intermediate and transition-state structures in protein-tyrosine phosphatase catalysis." *Proc Natl Acad Sci U S A* 93, 2493-8 (1996)
- [87] Hooft van Huijsduijnen R: "Protein tyrosine phosphatases: counting the trees in the forest." *Gene* 225, 1-8 (1998)
- [88] Tonks NK, Diltz CD, Fischer EH: "Purification of the major protein-tyrosine-phosphatases of human placenta." *J Biol Chem* 263, 6722-30 (1988)
- [89] Charbonneau H, Tonks NK, Kumar S, Diltz CD, Harrylock M, Cool DE, Krebs EG, Fischer EH, Walsh KA: "Human placenta protein-tyrosine-phosphatase: amino acid sequence and relationship to a family of receptor-like proteins." *Proc Natl Acad Sci U S A* 86, 5252-6 (1989)
- [90] Tonks NK, Diltz CD, Fischer EH: "Purification of protein-tyrosine phosphatases from human placenta." *Methods Enzymol* 201, 427-42 (1991)

- [91] Chernoff J, Schievella AR, Jost CA, Erikson RL, Neel BG: "Cloning of a cDNA for a major human protein-tyrosine-phosphatase." *Proc Natl Acad Sci U S A* 87, 2735-9 (1990)
- [92] Brown-Shimer S, Johnson KA, Lawrence JB, Johnson C, Bruskin A, Green NR, Hill DE: "Molecular cloning and chromosome mapping of the human gene encoding protein phosphotyrosyl phosphatase 1B." *Proc Natl Acad Sci U S A* 87, 5148-52 (1990)
- [93] Pallen CJ, Lai DS, Chia HP, Boulet I, Tong PH: "Purification and characterization of a higher-molecular-mass form of protein phosphotyrosine phosphatase (PTP 1B) from placental membranes." *Biochem J* 276, 315-23 (1991)
- [94] Frangioni JV, Beahm PH, Shifrin V, Jost CA, Neel BG: "The nontransmembrane tyrosine phosphatase PTP-1B localizes to the endoplasmic reticulum via its 35 amino acid C-terminal sequence." *Cell* 68, 545-60 (1992)
- [95] Dadke S, Chernoff J: "Interaction of protein tyrosine phosphatase (PTP) 1B with its substrates is influenced by two distinct binding domains." *Biochem J* 364, 377-83 (2002)
- [96] Barford D: "The mechanism of protein kinase regulation by protein phosphatases." *Biochem Soc Trans* 29, 385-91 (2001)
- [97] Andersen JN, Mortensen OH, Peters GH, Drake PG, Iversen LF, Olsen OH, Jansen PG, Andersen HS, Tonks NK, Moller NP: "Structural and evolutionary relationships among protein tyrosine phosphatase domains." *Mol Cell Biol* 21, 7117-36 (2001)
- [98] Puius YA, Zhao Y, Sullivan M, Lawrence DS, Almo SC, Zhang ZY: "Identification of a second aryl phosphate-binding site in protein-tyrosine phosphatase 1B: a paradigm for inhibitor design." *Proc Natl Acad Sci U S A* 94, 13420-5 (1997)
- [99] Asante-Appiah E, Patel S, Dufresne C, Roy P, Wang Q, Patel V, Friesen RW, Ramachandran C, Becker JW, Leblanc Y, Kennedy BP, Scapin G: "The structure of PTP-1B in complex with a peptide inhibitor reveals an alternative binding mode for bisphosphonates." *Biochemistry* 41, 9043-51 (2002)
- [100] Scapin G, Patel S, Patel V, Kennedy B, Asante-Appiah E: "The structure of apo protein-tyrosine phosphatase 1B C215S mutant: more than just an S --> O change." *Protein Sci* 10, 1596-605 (2001)
- [101] Xie L, Zhang YL, Zhang ZY: "Design and characterization of an improved protein tyrosine phosphatase substrate-trapping mutant." *Biochemistry* 41, 4032-9 (2002)
- [102] Zhang ZY: "Protein-tyrosine phosphatases: biological function, structural characteristics, and mechanism of catalysis." *Crit Rev Biochem Mol Biol* 33, 1-52 (1998)

- [103] Barford D, Flint AJ, Tonks NK: "Crystal structure of human protein tyrosine phosphatase 1B." *Science* 263, 1397-404 (1994)
- [104] Pannifer AD, Flint AJ, Tonks NK, Barford D: "Visualization of the cysteinyl-phosphate intermediate of a protein-tyrosine phosphatase by x-ray crystallography." *J Biol Chem* 273, 10454-62 (1998)
- [105] Deng H, Callender R, Huang Z, Zhang ZY: "Is the PTPase-vanadate complex a true transition state analogue?" *Biochemistry* 41, 5865-72 (2002)
- [106] Peters GH, Frimurer TM, Olsen OH: "Electrostatic evaluation of the signature motif (H/V)CX5R(S/T) in protein-tyrosine phosphatases." *Biochemistry* 37, 5383-93 (1998)
- [107] Zhang ZY, Wang Y, Dixon JE: "Dissecting the catalytic mechanism of protein-tyrosine phosphatases." *Proc Natl Acad Sci U S A* 91, 1624-7 (1994)
- [108] Zhang YL, Hollfelder F, Gordon SJ, Chen L, Keng YF, Wu L, Herschlag D, Zhang ZY: "Impaired transition state complementarity in the hydrolysis of O-arylphosphorothioates by protein-tyrosine phosphatases." *Biochemistry* 38, 12111-23 (1999)
- [109] Wu L, Zhang ZY: "Probing the function of Asp128 in the lower molecular weight protein-tyrosine phosphatase-catalyzed reaction. A pre-steady-state and steady-state kinetic investigation." *Biochemistry* 35, 5426-34 (1996)
- [110] Hengge AC, Sowa GA, Wu L, Zhang ZY: "Nature of the transition state of the protein-tyrosine phosphatase-catalyzed reaction." *Biochemistry* 34, 13982-7 (1995)
- [111] Zhang ZY, Dixon JE: "Protein tyrosine phosphatases: mechanism of catalysis and substrate specificity." *Adv Enzymol Relat Areas Mol Biol* 68, 1-36 (1994)
- [112] Lohse DL, Denu JM, Santoro N, Dixon JE: "Roles of aspartic acid-181 and serine-222 in intermediate formation and hydrolysis of the mammalian protein-tyrosine-phosphatase PTP1." *Biochemistry* 36, 4568-75 (1997)
- [113] Zhang ZY: "Kinetic and mechanistic characterization of a mammalian protein-tyrosine phosphatase, PTP1." *J Biol Chem* 270, 11199-204 (1995)
- [114] Zhang ZY, Dixon JE: "Active site labeling of the Yersinia protein tyrosine phosphatase: the determination of the pKa of the active site cysteine and the function of the conserved histidine 402." *Biochemistry* 32, 9340-5 (1993)
- [115] Guan KL, Dixon JE: "Evidence for protein-tyrosine-phosphatase catalysis proceeding via a cysteine-phosphate intermediate." *J Biol Chem* 266, 17026-30 (1991)

- [116] Gregoret LM, Rader SD, Fletterick RJ, Cohen FE: "Hydrogen bonds involving sulfur atoms in proteins." *Proteins* 9, 99-107 (1991)
- [117] Stuckey JA, Schubert HL, Fauman EB, Zhang ZY, Dixon JE, Saper MA: "Crystal structure of Yersinia protein tyrosine phosphatase at 2.5 Å and the complex with tungstate." *Nature* 370, 571-5 (1994)
- [118] Tonks NK, Diltz CD, Fischer EH: "Characterization of the major protein-tyrosine-phosphatases of human placenta." *J Biol Chem* 263, 6731-7 (1988)
- [119] Zhang M, Van Etten RL, Stauffacher CV: "Crystal structure of bovine heart phosphotyrosyl phosphatase at 2.2-Å resolution." *Biochemistry* 33, 11097-105 (1994)
- [120] Sarmiento M, Zhao Y, Gordon SJ, Zhang ZY: "Molecular basis for substrate specificity of protein-tyrosine phosphatase 1B." *J Biol Chem* 273, 26368-74 (1998)
- [121] Pedersen AK, Guo XL, Moller KB, Peters GH, Andersen HS, Kastrup JS, Mortensen SB, Iversen LF, Zhang ZY, Moller NP: "Residue 182 influences the second step of protein-tyrosine phosphatase-mediated catalysis." *Biochem J* 378, 421-33 (2004)
- [122] Peters GH, Iversen LF, Branner S, Andersen HS, Mortensen SB, Olsen OH, Moller KB, Moller NP: "Residue 259 is a key determinant of substrate specificity of protein-tyrosine phosphatases 1B and alpha." *J Biol Chem* 275, 18201-9 (2000)
- [123] Zhang YL, Yao ZJ, Sarmiento M, Wu L, Burke TR Jr, Zhang ZY: "Thermodynamic study of ligand binding to protein-tyrosine phosphatase 1B and its substrate-trapping mutants." *J Biol Chem* 275, 34205-12 (2000)
- [124] Vetter SW, Keng YF, Lawrence DS, Zhang ZY: "Assessment of protein-tyrosine phosphatase 1B substrate specificity using "inverse alanine scanning"." *J Biol Chem* 275, 2265-8 (2000)
- [125] Sarmiento M, Puius YA, Vetter SW, Keng YF, Wu L, Zhao Y, Lawrence DS, Almo SC, Zhang ZY: "Structural basis of plasticity in protein tyrosine phosphatase 1B substrate recognition." *Biochemistry* 39, 8171-9 (2000)
- [126] Ibarra-Sanchez MJ, Simoncic PD, Nestel FR, Duplay P, Lapp WS, Tremblay ML: "The T-cell protein tyrosine phosphatase." *Semin Immunol* 12, 379-86 (2000)
- [127] Romsicki Y, Kennedy BP, Asante-Appiah E: "Purification and characterization of T cell protein tyrosine phosphatase reveals significant functional homology to protein tyrosine phosphatase-1B." *Arch Biochem Biophys* 414, 40-50 (2003)
- [128] Iversen LF, Moller KB, Pedersen AK, Peters GH, Petersen AS, Andersen HS, Branner S, Mortensen SB, Moller NP: "Structure determination of T cell protein-tyrosine phosphatase." *J Biol Chem* 277, 19982-90 (2002)

- [129] Mahadev K, Zilbering A, Zhu L, Goldstein BJ: "Insulin-stimulated hydrogen peroxide reversibly inhibits protein-tyrosine phosphatase 1b in vivo and enhances the early insulin action cascade." *J Biol Chem* 276, 21938-42 (2001)
- [130] Wu X, Hoffstedt J, Deeb W, Singh R, Sedkova N, Zilbering A, Zhu L, Park PK, Arner P, Goldstein BJ: "Depot-specific variation in protein-tyrosine phosphatase activities in human omental and subcutaneous adipose tissue: a potential contribution to differential insulin sensitivity." *J Clin Endocrinol Metab* 86, 5973-80 (2001)
- [131] Meng TC, Buckley DA, Galic S, Tiganis T, Tonks NK: "Regulation of insulin signaling through reversible oxidation of the protein tyrosine phosphatases TC45 and PTP1B." *J Biol Chem* 279, 37716-25 (2004)
- [132] Tjernberg A, Hallen D, Schultz J, James S, Benkestock K, Bystrom S, Weigelt J: "Mechanism of action of pyridazine analogues on protein tyrosine phosphatase 1B (PTP1B)." *Bioorg Med Chem Lett* 14, 891-5 (2004)
- [133] Salmeen A, Andersen JN, Myers MP, Meng TC, Hinks JA, Tonks NK, Barford D: "Redox regulation of protein tyrosine phosphatase 1B involves a sulphenyl-amide intermediate." *Nature* 423, 769-73 (2003)
- [134] van Montfort RL, Congreve M, Tisi D, Carr R, Jhoti H: "Oxidation state of the active-site cysteine in protein tyrosine phosphatase 1B." *Nature* 423, 773-7 (2003)
- [135] Li S, Whorton AR: "Regulation of protein tyrosine phosphatase 1B in intact cells by S-nitrosothiols." *Arch Biochem Biophys* 410, 269-79 (2003)
- [136] Kenner KA, Hill DE, Olefsky JM, Kusari J: "Regulation of protein tyrosine phosphatases by insulin and insulin-like growth factor I." *J Biol Chem* 268, 25455-62 (1993)
- [137] Fukada T, Tonks NK: "Identification of YB-1 as a regulator of PTP1B expression: implications for regulation of insulin and cytokine signaling." *EMBO J* 22, 479-93 (2003)
- [138] Clark M, Cramer RD,: "Validation of the General Purpose Tripos 5.2 Force Field" *J Comp Chem* 10, 982-1012 (1989)
- [139] Tripos Inc., St. Louis, USA: "Sybyl"
<http://www.tripos.com/sciTech/inSilicoDisc/emmt.html>
- [140] Hagler AT, Huler E, Lifson S: "Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals." *J Am Chem Soc* 96, 5319-27 (1974)
- [141] Accelrys Inc., San Diego, USA: "Insight" <http://www.accelrys.com/insight/>
- [142] Vedani A, Zbinden P, Snyder JP, Greenidge PA: "Pseudoreceptor Modeling: The Construction of Three-Dimensional Receptor Surrogates" *J Am Chem Soc* 117, 4987-94 (1995)

- [143] Zbinden P, Dobler M, Folkers G, Vedani A: "PrGen: Pseudoreceptor Modeling Using Receptor-mediated Ligand Alignment and Pharmacophore Equilibration" *Quant Struct-Act Relat* 17, 122-30 (1998)
- [144] Kelley LA, Gardner SP, Sutcliffe MJ: "An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies." *Protein Eng* 9, 1063-5 (1996)
- [145] Antosiewicz J, McCammon JA, Gilson MK: "The determinants of pKas in proteins." *Biochemistry* 35, 7819-33 (1996)
- [146] Davis ME, Madura JD, Luty BA, McCammon JA: "Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian dynamics program" *Comp Phys Comm* 62, 187-97 (1991)
- [147] Madura JD, Briggs JM, Wade RC, Davis ME, Luty BA, Ilin A, Antosiewicz A, Gilson MK, Bagheri B, Scott LR, McCammon JA: "Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics program" *Comp Phys Comm* 91, 57-95 (1995)
- [148] Lemmen C, Lengauer T, Klebe G: "FLEXS: a method for fast flexible ligand superposition." *J Med Chem* 41, 4502-20 (1998)
- [149] Allen FH: "The Cambridge Structural Database: a quarter of a million crystal structures and rising." *Acta Crystallogr B* 58, 380-8 (2002)
- [150] Klebe G, Mietzner T: "A fast and efficient method to generate biologically relevant conformations." *J Comput Aided Mol Des* 8, 583-606 (1994)
- [151] Rarey M, Dixon JS: "Feature trees: a new molecular similarity measure based on tree matching." *J Comput Aided Mol Des* 12, 471-90 (1998)
- [152] Rarey M, Kramer B, Lengauer T, Klebe G: "A fast flexible docking method using an incremental construction algorithm." *J Mol Biol* 261, 470-89 (1996)
- [153] Goodsell DS, Morris GM, Olson AJ: "Automated docking of Flexible ligands: applications of Autodock." *J Mol Recognit* 9, 1-5 (1996)
- [154] Jones G, Willett P, Glen RC: "A genetic algorithm for flexible molecular overlay and pharmacophore elucidation." *J Comput Aided Mol Des* 9, 532-49 (1995)
- [155] Jones G, Willett P, Glen RC, Leach AR, Taylor R: "Development and validation of a genetic algorithm for flexible docking." *J Mol Biol* 267, 727-48 (1997)
- [156] Wang R, Lai L, Wang S: "Further development and validation of empirical scoring functions for structure-based binding affinity prediction." *J Comput Aided Mol Des* 16, 11-26 (2002)
- [157] Wang R, Gao Y, Lai L: "Calculating partition coefficient by atom-additive method." *Perspect Drug Discov Des* 19, 47-66 (2000)

- [158] Kuntz ID, Blaney JM, Oatley SJ, LanGRIDge R, Ferrin TE: "A geometric approach to macromolecule-ligand interactions." *J Mol Biol* 161, 269-88 (1982)
- [159] Muegge I, Martin YC: "A general and fast scoring function for protein-ligand interactions: a simplified potential approach." *J Med Chem* 42, 791-804 (1999)
- [160] Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP: "Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes." *J Comput Aided Mol Des* 11, 425-45 (1997)
- [161] Goodford PJ: "A computational procedure for determining energetically favorable binding sites on biologically important macromolecules." *J Med Chem* 28, 849-57 (1985)
- [162] Multivariate Infometric Analysis S.r.l., Perugia, Italy: "GOLPE 4.5." <http://www.miasrl.com/GOLPE.htm>
- [163] Cramer RD 3rd, Patterson DE, Bunce JD: "Recent advances in comparative molecular field analysis (CoMFA)." *Prog Clin Biol Res* 291, 161-5 (1989)
- [164] Ortiz AR, Pisabarro MT, Gago F, Wade RC: "Prediction of drug binding affinities by comparative binding energy analysis." *J Med Chem* 38, 2681-91 (1995)
- [165] Pastor M, Cruciani G, Clementi S: "Smart region definition: a new way to improve the predictive ability and interpretability of three-dimensional quantitative structure-activity relationships." *J Med Chem* 40, 1455-64 (1997)
- [166] Klebe G, Abraham U, Mietzner T: "Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity." *J Med Chem* 37, 4130-46 (1994)
- [167] Romsicki Y, Scapin G, Beaulieu-Audy V, Patel S, Becker JW, Kennedy BP, Asante-Appiah E: "Functional characterization and crystal structure of the C215D mutant of protein-tyrosine phosphatase-1B." *J Biol Chem* 278, 29009-15 (2003)
- [168] Larsen SD, Barf T, Liljebris C, May PD, Ogg D, O'Sullivan TJ, Palazuk BJ, Schostarez HJ, Stevens FC, Bleasdale JE: "Synthesis and biological activity of a novel class of small molecular weight peptidomimetic competitive inhibitors of protein tyrosine phosphatase 1B." *J Med Chem* 45, 598-622 (2002)
- [169] Groves MR, Yao ZJ, Roller PP, Burke TR Jr, Barford D: "Structural basis for inhibition of the protein tyrosine phosphatase 1B by phosphotyrosine peptide mimetics." *Biochemistry* 37, 17773-83 (1998)

- [170] Jia Z, Ye Q, Dinaut AN, Wang Q, Waddleton D, Payette P, Ramachandran C, Kennedy B, Hum G, Taylor SD: "Structure of protein tyrosine phosphatase 1B in complex with inhibitors bearing two phosphotyrosine mimetics." *J Med Chem* 44, 4584-94 (2001)
- [171] Scapin G, Patel SB, Becker JW, Wang Q, Desponts C, Waddleton D, Skorey K, Cromlish W, Bayly C, Therien M, Gauthier JY, Li CS, Lau CK, Ramachandran C, Kennedy BP, Asante-Appiah E: "The structural basis for the selectivity of benzotriazole inhibitors of PTP1B." *Biochemistry* 42, 11451-9 (2003)
- [172] Andersen HS, Iversen LF, Jeppesen CB, Branner S, Norris K, Rasmussen HB, Moller KB, Moller NP: "2-(oxalylamino)-benzoic acid is a general, competitive inhibitor of protein-tyrosine phosphatases." *J Biol Chem* 275, 7101-8 (2000)
- [173] Iversen LF, Andersen HS, Branner S, Mortensen SB, Peters GH, Norris K, Olsen OH, Jeppesen CB, Lundt BF, Ripka W, Moller KB, Moller NP: "Structure-based design of a low molecular weight, nonphosphorus, nonpeptide, and highly selective inhibitor of protein-tyrosine phosphatase 1B." *J Biol Chem* 275, 10300-7 (2000)
- [174] Peters GH, Iversen LF, Andersen HS, Moller NP, Olsen OH: "Residue 259 in Protein-Tyrosine Phosphatase PTP1B and PTPalpha Determines the Flexibility of Glutamine 262." *Biochemistry* 43, 8418-8428 (2004)
- [175] Iversen LF, Andersen HS, Moller KB, Olsen OH, Peters GH, Branner S, Mortensen SB, Hansen TK, Lau J, Ge Y, Holsworth DD, Newman MJ, Hundahl Moller NP: "Steric hindrance as a basis for structure-based design of selective inhibitors of protein-tyrosine phosphatases." *Biochemistry* 40, 14812-20 (2001)
- [176] Szczepankiewicz BG, Liu G, Hajduk PJ, Abad-Zapatero C, Pei Z, Xin Z, Lubben TH, Trevillyan JM, Stashko MA, Ballaron SJ, Liang H, Huang F, Hutchins CW, Fesik SW, Jirousek MR: "Discovery of a potent, selective protein tyrosine phosphatase 1B inhibitor using a linked-fragment strategy." *J Am Chem Soc* 125, 4087-96 (2003)
- [177] Liu G, Szczepankiewicz BG, Pei Z, Janowick DA, Xin Z, Hajduk PJ, Abad-Zapatero C, Liang H, Hutchins CW, Fesik SW, Ballaron SJ, Stashko MA, Lubben T, Mika AK, Zinker BA, Trevillyan JM, Jirousek MR: "Discovery and structure-activity relationship of oxalylarylamino benzoic acids as inhibitors of protein tyrosine phosphatase 1B." *J Med Chem* 46, 2093-103 (2003)
- [178] Liu G, Xin Z, Liang H, Abad-Zapatero C, Hajduk PJ, Janowick DA, Szczepankiewicz BG, Pei Z, Hutchins CW, Ballaron SJ, Stashko MA, Lubben TH, Berg CE, Rondinone CM, Trevillyan JM, Jirousek MR: "Selective protein tyrosine phosphatase 1B inhibitors: targeting the second phosphotyrosine binding site with non-carboxylic acid-containing ligands." *J Med Chem* 46, 3437-40 (2003)

- [179] Xin Z, Oost TK, Abad-Zapatero C, Hajduk PJ, Pei Z, Szczepankiewicz BG, Hutchins CW, Ballaron SJ, Stashko MA, Lubben T, Trevillyan JM, Jirousek MR, Liu G: "Potent, selective inhibitors of protein tyrosine phosphatase 1B." *Bioorg Med Chem Lett* 13, 1887-90 (2003)
- [180] Liu G, Xin Z, Pei Z, Hajduk PJ, Abad-Zapatero C, Hutchins CW, Zhao H, Lubben TH, Ballaron SJ, Haasch DL, Kaszubska W, Rondinone CM, Trevillyan JM, Jirousek MR: "Fragment screening and assembly: a highly efficient approach to a selective and cell active protein tyrosine phosphatase 1B inhibitor." *J Med Chem* 46, 4232-5 (2003)
- [181] Pei Z, Li X, Liu G, Abad-Zapatero C, Lubben T, Zhang T, Ballaron SJ, Hutchins CW, Trevillyan JM, Jirousek MR: "Discovery and SAR of novel, potent and selective protein tyrosine phosphatase 1B inhibitors." *Bioorg Med Chem Lett* 13, 3129-32 (2003)
- [182] Xin Z, Liu G, Abad-Zapatero C, Pei Z, Szczepankiewicz BG, Li X, Zhang T, Hutchins CW, Hajduk PJ, Ballaron SJ, Stashko MA, Lubben TH, Trevillyan JM, Jirousek MR: "Identification of a monoacid-based, cell permeable, selective inhibitor of protein tyrosine phosphatase 1B." *Bioorg Med Chem Lett* 13, 3947-50 (2003)
- [183] Bleasdale JE, Ogg D, Palazuk BJ, Jacob CS, Swanson ML, Wang XY, Thompson DP, Conradi RA, Mathews WR, Laborde AL, Stuchly CW, Heijbel A, Bergdahl K, Bannow CA, Smith CW, Svensson C, Liljebris C, Schostarez HJ, May PD, Stevens FC, Larsen SD: "Small molecule peptidomimetics containing a novel phosphotyrosine bioisostere inhibit protein tyrosine phosphatase 1B and augment insulin action." *Biochemistry* 40, 5642-54 (2001)
- [184] Sun JP, Fedorov AA, Lee SY, Guo XL, Shen K, Lawrence DS, Almo SC, Zhang ZY: "Crystal structure of PTP1B complexed with a potent and selective bidentate inhibitor." *J Biol Chem* 278, 12406-14 (2003)
- [185] Erlanson DA, McDowell RS, He MM, Randal M, Simmons RL, Kung J, Waight A, Hansen SK: "Discovery of a new phosphotyrosine mimetic for PTP1B using breakaway tethering." *J Am Chem Soc* 125, 5602-3 (2003)
- [186] Wiesmann C, Barr KJ, Kung J, Zhu J, Erlanson DA, Shen W, Fahr BJ, Zhong M, Taylor L, Randal M, McDowell RS, Hansen SK: "Allosteric inhibition of protein tyrosine phosphatase 1B." *Nat Struct Mol Biol* 11, 730-7 (2004)
- [187] Yang J, Cheng Z, Niu T, Liang X, Zhao ZJ, Zhou GW: "Structural basis for substrate specificity of protein-tyrosine phosphatase SHP-1" *J Biol Chem* 275, 4066-71 (2000)
- [188] Yang J, Liang X, Niu T, Meng W, Zhao Z, Zhou GW: "Crystal structure of the catalytic domain of protein-tyrosine phosphatase SHP-1." *J Biol Chem* 273, 28199-207 (1998)

- [189] Szedlacsek SE, Aricescu AR, Fulga TA, Renault L, Scheidig AJ: "Crystal structure of PTP-SL/PTPBR7 catalytic domain: implications for MAP kinase regulation." *J Mol Biol* 311, 557-68 (2001)
- [190] Hoffmann KM, Tonks NK, Barford D: "The crystal structure of domain 1 of receptor protein-tyrosine phosphatase mu." *J Biol Chem* 272, 27505-8 (1997)
- [191] Bilwes AM, den Hertog J, Hunter T, Noel JP: "Structural basis for inhibition of receptor protein-tyrosine phosphatase-alpha by dimerization." *Nature* 382, 555-9 (1996)
- [192] Hof P, Pluskey S, Dhe-Paganon S, Eck MJ, Shoelson SE: "Crystal structure of the tyrosine phosphatase SHP-2." *Cell* 92, 441-50 (1998)
- [193] Gordon JA: "Use of vanadate as protein-phosphotyrosine phosphatase inhibitor." *Methods Enzymol* 201, 477-82 (1991)
- [194] Poucheret P, Verma S, Grynepas MD, McNeill JH: "Vanadium and diabetes." *Mol Cell Biochem* 188, 73-80 (1998)
- [195] Huyer G, Liu S, Kelly J, Moffat J, Payette P, Kennedy B, Tsaprailis G, Gresser MJ, Ramachandran C: "Mechanism of inhibition of protein-tyrosine phosphatases by vanadate and pervanadate." *J Biol Chem* 272, 843-51 (1997)
- [196] Cuncic C, Desmarais S, Detich N, Tracey AS, Gresser MJ, Ramachandran C: "Bis(N,N-dimethylhydroxamido)hydroxooxovanadate inhibition of protein tyrosine phosphatase activity in intact cells: comparison with vanadate." *Biochem Pharmacol* 58, 1859-67 (1999)
- [197] Tracey AS: "Hydroxamido vanadates: aqueous chemistry and function in protein tyrosine phosphatases and cell cultures." *J Inorg Biochem* 80, 11-6 (2000)
- [198] Mohammad A, Wang J, McNeill JH: "Bis(maltolato)oxovanadium(IV) inhibits the activity of PTP1B in Zucker rat skeletal muscle in vivo." *Mol Cell Biochem* 229, 125-8 (2002)
- [199] Peters KG, Davis MG, Howard BW, Pokross M, Rastogi V, Diven C, Greis KD, Eby-Wilkens E, Maier M, Evdokimov A, Soper S, Genbauffe F: "Mechanism of insulin sensitization by BMOV (bis maltolato oxo vanadium); unliganded vanadium (VO₄) as the active component." *J Inorg Biochem* 96, 321-30 (2003)
- [200] Kanamori K, Nishida K, Miyata N, Okamoto K, Miyoshi Y, Tamura A, Sakurai H: "Syntheses, structures, stability, and insulin-like activities of peroxovanadium(V) complexes with a heteroligand." *J Inorg Biochem* 86, 649-56 (2001)
- [201] Cebula RW, Blanchard JL, Boisclair MD, Pal K, Bockovich NJ: "Synthesis and phosphatase inhibitory activity of analogs of sulfircin" *Bioorg Med Chem Lett* 15, 2015-20 (1997)

- [202] Hiroki A, Hatakeyama H, Kawakami M, Watanabe T, Takei I, Umezawa K: "Antidiabetic effect of a nitrosamine-free dephostatin analogue, methoxime-3,4-dephostatin, in db/db mice." *Biomed Pharmacother* 56, 179-85 (2002)
- [203] Watanabe T, Suzuki T, Umezawa Y, Takeuchi T, Otsuka M, Umezawa K: "Structure–Activity Relationship and Rational Design of 3,4-Dephostatin Derivatives as Protein Tyrosine Phosphatase Inhibitors" *Tetrahedron* 56, 741-52 (2000)
- [204] Suzuki T, Hiroki A, Watanabe T, Yamashita T, Takei I, Umezawa K: "Potentiation of insulin-related signal transduction by a novel protein-tyrosine phosphatase inhibitor, Et-3,4-dephostatin, on cultured 3T3-L1 adipocytes." *J Biol Chem* 276, 27511-8 (2001)
- [205] Chen RM, Hu LH, An TY, Li J, Shen Q: "Natural PTP1B Inhibitors from *broussonetia papyrifera*." *Bioorg Med Chem Lett* 12, 3387-90 (2002)
- [206] Ripka O: "Epidemiology of high blood pressure" *Med Klin* 63, 89-93 (1968)
- [207] Ham SW, Park J, Lee SJ, Yoo JS: "Selective inactivation of protein tyrosine phosphatase PTP1B by sulfone analogue of naphthoquinone." *Bioorg Med Chem Lett* 9, 185-6 (1999)
- [208] Urbanek RA, Suchard SJ, Steelman GB, Knappenberger KS, Sygowski LA, Veale CA, Chapdelaine MJ: "Potent reversible inhibitors of the protein tyrosine phosphatase CD45." *J Med Chem* 44, 1777-93 (2001)
- [209] Fu H, Park J, Pei D: "Peptidyl aldehydes as reversible covalent inhibitors of protein tyrosine phosphatases." *Biochemistry* 41, 10700-9 (2002)
- [210] Xian M, Wang K, Chen X, Hou Y, McGill A, Zhou B, Zhang ZY, Cheng JP, Wang PG: "Inhibition of protein tyrosine phosphatases by low-molecular-weight S-nitrosothiols and S-nitrosylated human serum albumin." *Biochem Biophys Res Commun* 268, 310-4 (2000)
- [211] Takakura K, Beckman JS, MacMillan-Crow LA, Crow JP: "Rapid and irreversible inactivation of protein tyrosine phosphatases PTP1B, CD45, and LAR by peroxynitrite." *Arch Biochem Biophys* 369, 197-207 (1999)
- [212] Burke TR Jr, Kole HK, Roller PP: "Potent inhibition of insulin receptor dephosphorylation by a hexamer peptide containing the phosphotyrosyl mimetic F2Pmp." *Biochem Biophys Res Commun* 204, 129-34 (1994)
- [213] Chen L, Wu L, Otaka A, Smyth MS, Roller PP, Burke TR Jr, den Hertog J, Zhang ZY: "Why is phosphonodifluoromethyl phenylalanine a more potent inhibitory moiety than phosphonomethyl phenylalanine toward protein-tyrosine phosphatases?" *Biochem Biophys Res Commun* 216, 976-84 (1995)
- [214] Huyer G, Kelly J, Moffat J, Zamboni R, Jia Z, Gresser MJ, Ramachandran C: "Affinity selection from peptide libraries to determine substrate specificity of protein tyrosine phosphatases." *Anal Biochem* 258, 19-30 (1998)

- [215] Xie L, Lee SY, Andersen JN, Waters S, Shen K, Guo XL, Moller NP, Olefsky JM, Lawrence DS, Zhang ZY: "Cellular effects of small molecule PTP1B inhibitors on insulin signaling." *Biochemistry* 42, 12792-804 (2003)
- [216] Kole HK, Smyth MS, Russ PL, Burke TR Jr: "Phosphonate inhibitors of protein-tyrosine and serine/threonine phosphatases." *Biochem J* 311, 1025-31 (1995)
- [217] Burke TR Jr, Ye B, Akamatsu M, Ford H Jr, Yan X, Kole HK, Wolf G, Shoelson SE, Roller PP: "4'-O-[2-(2-fluoromalonyl)]-L-tyrosine: a phosphotyrosyl mimic for the preparation of signal transduction inhibitory peptides." *J Med Chem* 39, 1021-7 (1996)
- [218] Burke TR Jr, Ye B, Yan X, Wang S, Jia Z, Chen L, Zhang ZY, Barford D: "Small molecule interactions with protein-tyrosine phosphatase PTP1B and their use in inhibitor design." *Biochemistry* 35, 15989-96 (1996)
- [219] Taylor SD, Kotoris CC, Dinaut AN, Wang Q, Ramachandran C, Huang Z: "Potent non-peptidyl inhibitors of protein tyrosine phosphatase 1B." *Bioorg Med Chem* 6, 1457-68 (1998)
- [220] Yao ZJ, Ye B, Wu XW, Wang S, Wu L, Zhang ZY, Burke TR Jr: "Structure-based design and synthesis of small molecule protein-tyrosine phosphatase 1B inhibitors." *Bioorg Med Chem* 6, 1799-810 (1998)
- [221] Liu DG, Gao Y, Voigt JH, Lee K, Nicklaus MC, Wu L, Zhang ZY, Burke TR Jr: "Acylsulfonamide-containing PTP1B inhibitors designed to mimic an enzyme-bound water of hydration." *Bioorg Med Chem Lett* 13, 3005-7 (2003)
- [222] Hum G, Grzyb J, Taylor SD: "Synthesis of non-peptidyl alpha,alpha-difluoromethylenephosphonic acids on a soluble polymer support" *J Comb Chem* 2, 234-42 (2000)
- [223] Leung C, Grzyb J, Lee J, Meyer N, Hum G, Jia C, Liu S, Taylor SD: "The difluoromethylenesulfonic acid group as a monoanionic phosphate surrogate for obtaining PTP1B inhibitors." *Bioorg Med Chem* 10, 2309-23 (2002)
- [224] Hum G, Lee J, Taylor SD: "Synthesis of [difluoro-(3-alkenylphenyl)-methyl]-phosphonic acids on non-crosslinked polystyrene and their evaluation as inhibitors of PTP1B." *Bioorg Med Chem Lett* 12, 3471-4 (2002)
- [225] Yokomatsu T, Murano T, Umesue I, Soeda S, Shimeno H, Shibuya S: "Synthesis and biological evaluation of alpha,alpha-difluorobenzylphosphonic acid derivatives as small molecular inhibitors of protein-tyrosine phosphatase 1B." *Bioorg Med Chem Lett* 9, 529-32 (1999)
- [226] Murthy VS, Kulkarni VM: "Molecular modeling of protein tyrosine phosphatase 1B (PTP 1B) inhibitors." *Bioorg Med Chem* 10, 897-906 (2002)

- [227] Lau CK, Bayly CI, Gauthier JY, Li CS, Therien M, Asante-Appiah E, Cromlish W, Boie Y, Forghani F, Desmarais S, Wang Q, Skorey K, Waddleton D, Payette P, Ramachandran C, Kennedy BP, Scapin G: "Structure based design of a series of potent and selective non peptidic PTP-1B inhibitors." *Bioorg Med Chem Lett* 14, 1043-8 (2004)
- [228] Vetter SW, Zhang ZY: "Probing the phosphopeptide specificities of protein tyrosine phosphatases, SH2 and PTB domains with combinatorial library methods." *Curr Protein Pept Sci* 3, 365-97 (2002)
- [229] Montserat J, Chen L, Lawrence DS, Zhang ZY: "Potent low molecular weight substrates for protein-tyrosine phosphatase." *J Biol Chem* 271, 7868-72 (1996)
- [230] Wang Q, Huang Z, Ramachandran C, Dinaut AN, Taylor SD: "Naphthalenebis[alpha,alpha-difluoromethylenephosphonates] as potent inhibitors of protein tyrosine phosphatases." *Bioorg Med Chem Lett* 8, 345-50 (1998)
- [231] Ibrahim OA, Wu L, Zhao K, Zhang ZY: "Synthesis and characterization of a novel class of protein tyrosine phosphatase inhibitors." *Bioorg Med Chem Lett* 10, 457-60 (2000)
- [232] Shen K, Keng YF, Wu L, Guo XL, Lawrence DS, Zhang ZY: "Acquisition of a specific and potent PTP1B inhibitor from a novel combinatorial library and screening procedure." *J Biol Chem* 276, 47311-9 (2001)
- [233] Taing M, Keng YF, Shen K, Wu L, Lawrence DS, Zhang ZY: "Potent and highly selective inhibitors of the protein tyrosine phosphatase 1B." *Biochemistry* 38, 3793-803 (1999)
- [234] Desmarais S, Friesen RW, Zamboni R, Ramachandran C: "[Difluoro(phosphono)methyl]phenylalanine-containing peptide inhibitors of protein tyrosine phosphatases." *Biochem J* 337, 219-23 (1999)
- [235] Asante-Appiah E, Ball K, Bateman K, Skorey K, Friesen R, Desponts C, Payette P, Bayly C, Zamboni R, Scapin G, Ramachandran C, Kennedy BP: "The YRD motif is a major determinant of substrate and inhibitor specificity in T-cell protein-tyrosine phosphatase." *J Biol Chem* 276, 26036-43 (2001)
- [236] Skorey KI, Kennedy BP, Friesen RW, Ramachandran C: "Development of a robust scintillation proximity assay for protein tyrosine phosphatase 1B using the catalytically inactive (C215S) mutant." *Anal Biochem* 291, 269-78 (2001)
- [237] Liu G, Trevillyan JM: "Protein tyrosine phosphatase 1B as a target for the treatment of impaired glucose tolerance and type II diabetes." *Curr Opin Investig Drugs* 3, 1608-16 (2002)
- [238] Guo XL, Shen K, Wang F, Lawrence DS, Zhang ZY: "Probing the molecular basis for potent and selective protein-tyrosine phosphatase 1B inhibition." *J Biol Chem* 277, 41014-22 (2002)

- [239] Kole HK, Akamatsu M, Ye B, Yan X, Barford D, Roller PP, Burke TR Jr: "Protein-tyrosine phosphatase inhibition by a peptide containing the phosphotyrosyl mimetic, L-O-malonyltyrosine." *Biochem Biophys Res Commun* 209, 817-22 (1995)
- [240] Akamatsu M, Roller PP, Chen L, Zhang ZY, Ye B, Burke TR Jr: "Potent inhibition of protein-tyrosine phosphatase by phosphotyrosine-mimic containing cyclic peptides." *Bioorg Med Chem* 5, 157-63 (1997)
- [241] Roller PP, Wu L, Zhang ZY, Burke TR Jr: "Potent inhibition of protein-tyrosine phosphatase-1B using the phosphotyrosyl mimetic fluoro-O-malonyl tyrosine (FOMT)." *Bioorg Med Chem Lett* 8, 2149-50 (1998)
- [242] Liotta AS, Kole HK, Fales HM, Roth J, Bernier M: "A synthetic tris-sulfotyrosyl dodecapeptide analogue of the insulin receptor 1146-kinase domain inhibits tyrosine dephosphorylation of the insulin receptor in situ." *J Biol Chem* 269, 22996-3001 (1994)
- [243] Kole HK, Liotta AS, Garant MJ, Kole S, Bernier M: "Specific inhibition of insulin receptor dephosphorylation by a synthetic dodecapeptide containing sulfotyrosyl residues as phosphotyrosyl mimetic." *Indian J Biochem Biophys* 34, 50-5 (1997)
- [244] Desmarais S, Jia Z, Ramachandran C: "Inhibition of protein tyrosine phosphatases PTP1B and CD45 by sulfotyrosyl peptides." *Arch Biochem Biophys* 354, 225-31 (1998)
- [245] Glover NR, Tracey AS: "Structure, modelling, and molecular dynamics studies of the inhibition of protein tyrosine phosphatase 1B by sulfotyrosine peptides." *Biochem Cell Biol* 77, 469-86 (1999)
- [246] Gao Y, Voigt J, Zhao H, Pais GC, Zhang X, Wu L, Zhang ZY, Burke TR Jr: "Utilization of a peptide lead for the discovery of a novel PTP1B-binding motif." *J Med Chem* 44, 2869-78 (2001)
- [247] Burke TR Jr, Yao ZJ, Ye B, Miyoshi K, Otaka A, Wu L, Zhang ZY: "Phospho-Azatyrosine, a less effective protein-tyrosine phosphatase substrate than phosphotyrosine." *Bioorg Med Chem Lett* 11, 1265-8 (2001)
- [248] Gao Y, Wu L, Luo JH, Guo R, Yang D, Zhang ZY, Burke TR Jr: "Examination of novel non-phosphorus-containing phosphotyrosyl mimetics against protein-tyrosine phosphatase-1B and demonstration of differential affinities toward Grb2 SH2 domains." *Bioorg Med Chem Lett* 10, 923-7 (2000)
- [249] Lee K, Gao Y, Yao ZJ, Phan J, Wu L, Liang J, Waugh DS, Zhang ZY, Burke TR Jr: "Tripeptide inhibitors of Yersinia protein-tyrosine phosphatase." *Bioorg Med Chem Lett* 13, 2577-81 (2003)

- [250] Liljebris C, Larsen SD, Ogg D, Palazuk BJ, Bleasdale JE: "Investigation of potential bioisosteric replacements for the carboxyl groups of peptidomimetic inhibitors of protein tyrosine phosphatase 1B: identification of a tetrazole-containing inhibitor with cellular activity." *J Med Chem* 45, 1785-98 (2002)
- [251] Larsen SD, Stevens FC, Lindberg TJ, Bodnar PM, O'Sullivan TJ, Schostarez HJ, Palazuk BJ, Bleasdale JE: "Modification of the N-terminus of peptidomimetic protein tyrosine phosphatase 1B (PTP1B) inhibitors: identification of analogues with cellular activity." *Bioorg Med Chem Lett* 13, 971-5 (2003)
- [252] Arabaci G, Yi T, Fu H, Porter ME, Beebe KD, Pei D: "alpha-bromoacetophenone derivatives as neutral protein tyrosine phosphatase inhibitors: structure-Activity relationship." *Bioorg Med Chem Lett* 12, 3047-50 (2002)
- [253] Kotoris CC, Chen MJ, Taylor SD: "Novel phosphate mimetics for the design of non-peptidyl inhibitors of protein tyrosine phosphatases." *Bioorg Med Chem Lett* 8, 3275-80 (1998)
- [254] Andersen HS, Olsen OH, Iversen LF, Sorensen AL, Mortensen SB, Christensen MS, Branner S, Hansen TK, Lau JF, Jeppesen L, Moran EJ, Su J, Bakir F, Judge L, Shahbaz M, Collins T, Vo T, Newman MJ, Ripka WC, Moller NP: "Discovery and SAR of a novel selective and orally bioavailable nonpeptide classical competitive inhibitor class of protein-tyrosine phosphatase 1B." *J Med Chem* 45, 4443-59 (2002)
- [255] Wrobel J, Sredy J, Moxham C, Dietrich A, Li Z, Sawicki DR, Seestaller L, Wu L, Katz A, Sullivan D, Tio C, Zhang ZY: "PTP1B inhibition and antihyperglycemic activity in the ob/ob mouse model of novel 11-arylbenzo[b]naphtho[2,3-d]furans and 11-arylbenzo[b]naphtho[2,3-d]thiophenes." *J Med Chem* 42, 3199-202 (1999)
- [256] Wrobel J, Li Z, Sredy J, Sawicki DR, Seestaller L, Sullivan D: "Synthesis and PTP1B inhibition of novel 4-aryl-1-oxa-9-thiacyclopenta[b]fluorenes." *Bioorg Med Chem Lett* 10, 1535-8 (2000)
- [257] Malamas MS, Sredy J, Gunawan I, Mihan B, Sawicki DR, Seestaller L, Sullivan D, Flam BR: "New azolidinediones as inhibitors of protein tyrosine phosphatase 1B with antihyperglycemic properties." *J Med Chem* 43, 995-1010 (2000)
- [258] Malamas MS, Sredy J, Moxham C, Katz A, Xu W, McDevitt R, Adebayo FO, Sawicki DR, Seestaller L, Sullivan D, Taylor JR: "Novel benzofuran and benzothiophene biphenyls as inhibitors of protein tyrosine phosphatase 1B with antihyperglycemic properties." *J Med Chem* 43, 1293-310 (2000)
- [259] Malamas MS, Sredy J, McCaleb M, Gunawan I, Mihan B, Sullivan D: "Antihyperglycemic activity of new 1,2,4-oxadiazolidine-3,5-diones." *Eur J Med Chem* 36, 31-42 (2001)

- [260] Sarmiento M, Wu L, Keng YF, Song L, Luo Z, Huang Z, Wu GZ, Yuan AK, Zhang ZY: "Structure-based discovery of small molecule inhibitors targeted to protein tyrosine phosphatase 1B." *J Med Chem* 43, 146-55 (2000)
- [261] Doman TN, McGovern SL, Witherbee BJ, Kasten TP, Kurumbail R, Stallings WC, Connolly DT, Shoichet BK: "Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B." *J Med Chem* 45, 2213-21 (2002)
- [262] Zhang YL, Keng YF, Zhao Y, Wu L, Zhang ZY: "Suramin is an active site-directed, reversible, and tight-binding inhibitor of protein-tyrosine phosphatases." *J Biol Chem* 273, 12281-7 (1998)
- [263] McCain DF, Wu L, Nickel P, Kassack MU, Kreimeyer A, Gagliardi A, Collins DC, Zhang ZY: "Suramin derivatives as inhibitors and activators of protein-tyrosine phosphatases." *J Biol Chem* 279, 14713-25 (2004)
- [264] Chen YT, Seto CT: "Divalent and trivalent alpha-ketocarboxylic acids as inhibitors of protein tyrosine phosphatases." *J Med Chem* 45, 3946-52 (2002)
- [265] Chen YT, Seto CT: "Parallel synthesis of a library of bidentate protein tyrosine phosphatase inhibitors based on the alpha-ketoacid motif." *Bioorg Med Chem* 12, 3289-98 (2004)
- [266] Leung KW, Posner BI, Just G: "Periodinates: a new class of protein tyrosine phosphatase inhibitors." *Bioorg Med Chem Lett* 9, 353-6 (1999)
- [267] Rice RL, Rusnak JM, Yokokawa F, Yokokawa S, Messner DJ, Boynton AL, Wipf P, Lazo JS: "A targeted library of small-molecule, tyrosine, and dual-specificity phosphatase inhibitors derived from a rational core design and random side chain variation." *Biochemistry* 36, 15965-74 (1997)
- [268] Liljebris C, Martinsson J, Tedenborg L, Williams M, Barker E, Duffy JE, Nygren A, James S: "Synthesis and biological activity of a novel class of pyridazine analogues as non-competitive reversible inhibitors of protein tyrosine phosphatase 1B (PTP1B)." *Bioorg Med Chem* 10, 3197-212 (2002)
- [269] Ahn JH, Cho SY, Ha JD, Chu SY, Jung SH, Jung YS, Baek JY, Choi IK, Shin EY, Kang SK, Kim SS, Cheon HG, Yang SD, Choi JK: "Synthesis and PTP1B inhibition of 1,2-naphthoquinone derivatives as potent anti-diabetic agents." *Bioorg Med Chem Lett* 12, 1941-6 (2002)
- [270] Cheon HG, Kim SM, Yang SD, Ha JD, Choi JK: "Discovery of a novel protein tyrosine phosphatase-1B inhibitor, KR61639: potential development as an antihyperglycemic agent." *Eur J Pharmacol* 485, 333-339 (2004)
- [271] Shim YS, Kim KC, Chi DY, Lee KH, Cho H: "Formylchromone derivatives as a novel class of protein tyrosine phosphatase 1B inhibitors." *Bioorg Med Chem Lett* 13, 2561-3 (2003)

- [272] Behnke D, Taube R, Hess S, Schauer-Vukasinovic V, Thormann M : "Imidazo[1,2-a]pyridines via a three component reaction as inhibitors of protein tyrosine phosphatase 1B" Posterbeitrag QSAR 2002, (2002)
- [273] Taylor SD, Hill B: "Recent advances in protein tyrosine phosphatase 1B inhibitors." *Expert Opin Investig Drugs* 13, 199-214 (2004)
- [274] Sippl W: "Molecular Modelling Untersuchungen zum Bindungsverhalten von Histamin H2- und H3-Rezeptor Liganden" Dissertation, Heinrich-Heine-Universität Düsseldorf, Seite 210 (1997)
- [275] Derwent: "World Drug Index (WDI)"
<http://www.derwent.com/products/lr/wdi/>
- [276] Maybridge, Trevillett, Tintagel, Cornwall PL34 OHW, England:
"Maybridge Screening Collection"
http://www.maybridge.com/html/m_msc_unit.htm
- [277] National Cancer Institute: "NCI Diversity Set"
http://dtp.nci.nih.gov/branches/dscb/diversity_explanation.html
- [278] Ferrari AM, Wei BQ, Costantino L, Shoichet BK: "Soft docking and multiple receptor conformations in virtual screening." *J Med Chem* 47, 5076-84 (2004)
- [279] Verdonk ML, Berdini V, Hartshorn MJ, Mooij WT, Murray CW, Taylor RD, Watson P: "Virtual screening using protein-ligand docking: avoiding artificial enrichment." *J Chem Inf Comput Sci* 44, 793-806 (2004)
- [280] Ferrara P, Gohlke H, Price DJ, Klebe G, Brooks CL 3rd: "Assessing scoring functions for protein-ligand interactions." *J Med Chem* 47, 3032-47 (2004)
- [281] Krovat EM, Langer T: "Impact of Scoring Functions on Enrichment in Docking-Based Virtual Screening: An Application Study on Renin Inhibitors" *J Chem Inf Comput Sci* 44, 1123-9 (2004)
- [282] Giordanetto F, Cotesta S, Catana C, Trosset JY, Vulpetti A, Stouten PF, Kroemer RT: "Novel scoring functions comprising QXP, SASA, and protein side-chain entropy terms." *J Chem Inf Comput Sci* 44, 882-93 (2004)
- [283] Klon AE, Glick M, Thoma M, Acklin P, Davies JW: "Finding more needles in the haystack: A simple and efficient method for improving high-throughput docking results." *J Med Chem* 47, 2743-9 (2004)
- [284] Murthy VS, Kulkarni VM: "3D QSAR CoMFA and CoMSIA on protein tyrosine phosphatase 1B inhibitors." *Bioorg Med Chem* 10, 2267-82 (2002)
- [285] Yang C, Cross K, Myatt GJ, Blower PE, Rathman JF: "Building Predictive Models for Protein Tyrosine Phosphatase 1B Inhibitors Based on Discriminating Structural Features by Reassembling Medicinal Chemistry Building Blocks" *J Med Chem* (in Press), (2004)

9 Anhang

9.1 Abkürzungen

9.1.1 Aminosäuren

Aminosäure	Drei-	Ein-
	Buchstaben-Code	
Alanin	Ala	A
Cystein	Cys	C
Aspartat	Asp	D
Glutamat	Glu	E
Phenylalanin	Phe	F
Glycin	Gly	G
Histidin	His	H
Isoleucin	Ile	I
Lysin	Lys	K
Leucin	Leu	L
Methionin	Met	M
Asparagin	Asn	N
Prolin	Pro	P
Glutamin	Gln	Q
Arginin	Arg	R
Serin	Ser	S
Threonin	Thr	T
Valin	Val	V
Tyrosin	Tyr	Y
Phosphotyrosin	pTyr	pY
Beliebige Aminosäure	Xxx	X

9.1.2 Programmablaufsymbole

Symbol	Beschreibung
	Programm, Prozedur oder Funktion
	Datei
	Mehrere ähnliche Dateien
	Und-Verknüpfung
	Oder-Verknüpfung

9.2 Skripte

9.2.1 Skript zur Automatisierung von Autodock

```

#! /usr/bin/perl

# Functions:
# read_inputfile
# log_parameters
# preprec
# mkgpf
# calcgrd
# preplig
# optimize_mol2
# mkdpf
# dockall
# extract
# xscore
#-----

# main function, opens the summary-files and controls the use of the other
# functions

open (LOG , ">>@ARGV[0].log") or die "@ARGV[0].log could not be written";
print LOG ("\nInputfile: @ARGV[0]\n\n");

read_inputfile (@ARGV[0]);
log_parameters ();
if ($keywords =~ /results/){
    open (AD_ONE , ">>@ARGV[0].ad_list");
}
if ($keywords =~ /xscore/){
    open (XS_ONE , ">>@ARGV[0].xs_list");
    foreach $receptor (@receptors){
        system("xscore -fixpdb $receptor.pdbq $receptor.x.pdb");
        print ("Preparing $receptor for X-Score Calculation\n");
    }
}

if ($keywords =~ /timeout/){timeout ();}
if ($keywords =~ /preprec/){preprec ();}
if ($keywords =~ /calcgrd/){calcgrd ();}
if ($keywords =~ /preplig/){preplig ();}
if (($keywords =~ /makedpf/) or ($keywords =~ /dockall/) or
($keywords =~ /results/) or ($keywords =~ /xscore/)){
    foreach $ligand (@ligands){
        foreach $receptor (@receptors){
            if ($keywords =~ /makedpf/){mkdpf ($ligand, $receptor);}
            if ($keywords =~ /dockall/){dockall ($ligand, $receptor);}
            if ($keywords =~ /results/){extract ($ligand, $receptor);}
            if ($keywords =~ /xscore/){xscore ($ligand, $receptor);}
        }
    }
}
if ($keywords =~ /results/){close (AD_ONE);}
if ($keywords =~ /xscore/){close (XS_ONE);}
print LOG ("\n-----");

```

```
print LOG ("-----\n");
close (LOG);
exit;
# end of main
#-----

sub read_inputfile{
# reads the inputfile and stores the values in certain variables
my ($file) = @_;
open (INFILE , "$file") or die "$file could not be read";
while (<INFILE>){
    $inline = $_; chomp $inline;
    if ($inline =~ /# Configuration/){last}
}
# Read Configuration
while (<INFILE>){
    $inline = $_; chomp $inline;
    if ($inline =~ /# Keywords/){last}
}
# Read Keywords
print ("Reading Keywords\n");
$keywords = "Keywords:";
while (<INFILE>){
    $inline = $_; chomp $inline;
    if ($inline =~ /# Macromolecule/){last}
    if ($inline eq ""){next}
    if (substr($inline, 0, 1) eq "#"){next}
    ($inline) = split (/s+/, $inline);
    $keywords = $keywords." ".$inline
}
print LOG ("$keywords\n\n");
# Read Macromolecules
print ("Reading Macromolecule(s)\n");
$n = 0;
while (<INFILE>){
    $inline = $_; chomp $inline;
    if ($inline =~ /# Ligand Molecule/){last}
    if ($inline eq ""){next}
    if (substr($inline, 0, 1) eq "#"){next}
    ($receptor) = split (/\.\/, $inline);
    if (-e "$receptor.mol2"){
        print LOG ("Macromolecule $n: $receptor\n");
        $receptors[$n] = $receptor; $n++;
    }else{
        print LOG ("$receptor.mol2 does not exist and will be ignored\n");
    }
}
}
$nreceptors = $n;
if ($nreceptors == 0){die "No Macromolecule\n";}
print LOG ("$nreceptors Macromolecule(s) read\n\n");
# Read Ligands
print ("Reading Ligand(s)\n");
$n = 0;
while (<INFILE>){
    $inline = $_; chomp $inline;
    if ($inline =~ /# Timeout Parameters/){last}
    elsif ($inline eq ""){next}
    elsif (substr($inline, 0, 1) eq "#"){next}
    ($ligand) = split (/\.\/, $inline);
```

```

    if (-e "$ligand.mol2"){
        print LOG ("Ligand $n: $ligand\n");
        $ligands[$n] = $ligand; $n++;
    }else{
        print LOG ("$ligand.mol2 does not exist and will be ignored\n");
    }
}
}
$nligands = $n;
if ($nligands == 0){die "No Ligand\n";}
print LOG ("$nligands Ligand(s) read\n\n");
# Read timeout Parameters
print ("Reading timeout Parameters\n");
$n = 0;
while (<INFILE>){
    $inline = $_;
    if ($inline =~ /# Grid Maps Calculation Parameters/){last}
    elsif ($inline =~ /ad_max/){($trash, $admax,) = split /\s+/, $inline;}
    elsif ($inline =~ /xs_max/){($trash, $xsmax,) = split /\s+/, $inline;}
}
($trash, $gridx, $gridy, $gridz) = split /\s+/, $gridcenter;
# Read Grid Maps Parameters
print ("Reading Grid Maps Parameters\n");
$n = 0;
while (<INFILE>){
    $inline = $_;
    if ($inline =~ /# Docking Calculation Parameters/){last}
    elsif ($inline =~ /types/){$types = $inline;}
    elsif ($inline =~ /npts/){$npts = $inline;}
    elsif ($inline =~ /spacing/){$spacing = $inline;}
    elsif ($inline =~ /gridcenter/){$gridcenter = $inline;}
    elsif ($inline =~ /smooth/){$smooth = $inline;}
    $grid_params[$n] = $inline; $n++;
}
($trash, $gridx, $gridy, $gridz) = split /\s+/, $gridcenter;
# Read Docking Parameters
print ("Reading Docking Parameters\n");
$n = 0;
while (<INFILE>){
    $inline = $_;
    if ($inline =~ /# X-Score Calculation Parameters/){last}
    $dock_params[$n] = $inline; $n++;
}
# Read X-Score Parameters
print ("Reading X-Score Parameters\n");
$n = 0;
while (<INFILE>){
    $inline = $_;
    if ($inline =~ /REFERENCE_MOL2_FILE/){$reference = $inline;}
    $xsco_params[$n] = $inline; $n++;
}
close (INFILE);
}
# end of read_inputfile
#-----

sub log_parameters{
# writes the found parameters to the logfile
    if ($keywords =~ /calcgrd/){
        print LOG ("Grid Maps Parameters\n\n");
    }
}

```

```

        foreach $param (@grid_params){
            if (substr($param, 0, 1) eq "#"){next}
            if ($param eq "\n"){next}
            print LOG (" $param");
        }
    print LOG ("\n");
}
if ($keywords =~ /dockall/){
    print LOG ("Docking Parameters\n\n");
    foreach $param (@dock_params){
        if (substr($param, 0, 1) eq "#"){next}
        if ($param eq "\n"){next}
        print LOG (" $param");
    }
    print LOG ("\n");
}
if ($keywords =~ /xscore/){
    print LOG ("X-Score Parameters\n\n");
    foreach $param (@xsco_params){
        if (substr($param, 0, 1) eq "#"){next}
        if ($param eq "\n"){next}
        print LOG (" $param");
    }
    print LOG ("\n");
}
}
# end of log_parameters
#-----

sub timeout{
# starts a separate programm which check for autodock or xscore jobs with
# a long (values given in the inputfile) runtime period
# used to kill hangup jobs
    print ("Starting timeout\n");
    system("timeout $ad_max $xsmax &");
}
# timeout
#-----

sub preprec{
# prepares all macromolecules given in the inputfile:
# converts the mol2-files into pdbqs-files
# generates a temporary (ligand-)pdbq-file for the atomtypes given in
# the inputfile
# prepares gpf-file using function mkgpf
# creates a box.pdb-file to display the size and position of the gridfields
# checks partial charges for errors
    print LOG ("Preparing $nreceptors Macromolecule(s)\n");
    print ("Preparing $nreceptors Macromolecule(s)\n");
    open (OUTFILE , ">tmp.pdbq") or die "tmp.pdbq could not be written";
    $n = 0;
    ($trash, $atoms) = split /\s+/, $types;
    while ($atoms ne ""){
        $atom = chop $atoms; $n++;
        printf OUTFILE
            "ATOM      %2s  %1s   TMP      0          0.000  0.000  0.000\n", $n, $atom;
    }
    $ntypes = $n;
    close (OUTFILE);
}

```

```

    foreach $receptor (@receptors){
        system("mol2topdbqs $receptor.mol2");
        mkgpf ($receptor);
        system("mkbox $receptor.gpf > $receptor.box.pdb");
        system("check-qs $receptor");
    }
    system("rm tmp.pdbq");
}
# end of preprec
#-----

sub mkgpf{
# generates the gpf-file using the temporary ligand-pdbq-file, the pdbqs-file of
# the receptor and the standard gpf-gen file provided with autodock
# replaces some parameters with the parameters given in the inputfile
    my ($receptor) = @_ ;
    system("mkgpf3 tmp.pdbq $receptor.pdbqs");
    open (INFILE , "$receptor.gpf")
        or die "$receptor.gpf could not be read";
    open (OUTFILE , ">$receptor.new.gpf")
        or die "$receptor.new.gpf could not be written";
    while (<INFILE>){
        $inline = $_;
        if ($inline =~ /types/){$inline = $types;}
        elsif ($inline =~ /npts/){$inline = $npts;}
        elsif ($inline =~ /spacing/){$inline = $spacing;}
        elsif ($inline =~ /gridcenter/){$inline = $gridcenter;}
        elsif ($inline =~ /smooth/){$inline = $smooth;}
        print OUTFILE "$inline";
    }
    close (INFILE);
    close (OUTFILE);
    system("mv $receptor.new.gpf $receptor.gpf");
}
# end of mkgpf
#-----

sub calcgrd{
# calculates the gridmaps for all macromolecules given in the inputfile
    print LOG ("Calculating $ntypes Grids for $nreceptors Macromolecule(s)\n");
    print ("Calculating $ntypes Grids for $nreceptors Macromolecule(s)\n");
    print ("please wait ...\n");
    foreach $receptor (@receptors){
        system("autogrid3 -p $receptor.gpf -l $receptor.glg");
    }
}
# end of calcgrd
#-----

sub preplig{
# prepares the ligands given in the inputfile:
# uses function optimize_mol2 to revise the mol2-file
# uses autotors to define the torsions and convert the ligand to pdbq-file
# checks partial charges for errors
    print LOG ("Preparing $nligands Ligand(s) \t(preplig)\n");
    print ("Preparing $nligands Ligand(s) \t(preplig)\n");
    foreach $ligand (@ligands){
        optimize_mol2 ($ligand);
        system("autotors -A +15.0 -r -a -m $ligand.mol2 $ligand.pdbq");
    }
}

```

```

        # -A +15.0    check rings for aromaticity (aromatic atoms -> type A)
        # -r        Root is the non-hydrogen atom closest to the center
        # -a        disallow torsion rotations in amide and peptide bonds
        # -h        can't be used in combination with x-score
        # -h        merge nonpolar hydrogens with the corresponding carbon
        # -m        input in mol2 format
        system("check-qs $ligand");
    }
}
# end of preplig
#-----

sub optimize_mol2{
# revises the mol2-file of the ligand
# centers the ligand in the grid
# sets atomtype of Chlorine to 'cl'
# sets atomtype of Bromine to 'br'
# sets aromatic bondtype to 'ar'
    my ($file) = @_ ;
    open (INFILE , "$file.mol2") or die "$file.mol2 could not be read";

    $inline = <INFILE>;
    until ($inline =~ /@<TRIPOS>ATOM/){$inline = <INFILE>;}
    $inline = <INFILE>;
    until ($inline =~ /@<TRIPOS>BOND/){
        @line = split /\s+/, $inline;
        shift(@line);
        $x += $line[2];
        $y += $line[3];
        $z += $line[4];
        $n ++;
        $inline = <INFILE>;
    }
    $x = $x/$n;          # Actual center coordinates
    $y = $y/$n;
    $z = $z/$n;
    $x = $gridx-$x;     # New center coordinates
    $y = $gridy-$y;
    $z = $gridz-$z;
    close (INFILE);
    open (INFILE , "$file.mol2")
        or die "$file.mol2 could not be read";
    open (OUTFILE , ">$file.new.mol2")
        or die "$file.new.mol2 could not be written";
    $inline = <INFILE>;
    until ($inline =~ /@<TRIPOS>MOLECULE/){$inline = <INFILE>;}
    print OUTFILE "$inline";
    $inline = <INFILE>;    chomp $inline;
    ($molecule) = split /\./, $inline;
    print OUTFILE "$molecule\n";
    until ($inline =~ /@<TRIPOS>ATOM/){
        $inline = <INFILE>;
        print OUTFILE "$inline";
    }
    $inline = <INFILE>;
    until ($inline =~ /@<TRIPOS>BOND/){
        @line = split /\s+/, $inline;
        shift(@line);
        $line[2] = $line[2] + $x;

```

```

$line[3] = $line[3] + $y;
$line[4] = $line[4] + $z;
$line[1] =~ s/(C|c)(L|l)/cl/; $line[5] =~ s/(C|c)(L|l)/cl/;
$line[1] =~ s/(B|b)(R|r)/br/; $line[5] =~ s/(B|b)(R|r)/br/;
printf (OUTFILE "%7d %-8s %9.4f %9.4f %9.4f %-9s %ld %-4s %8.4f\n",
        $line[0], $line[1], $line[2], $line[3], $line[4], $line[5], $line[6],
        $molecule, $line[8]);
$inline = <INFILE>;
}
until ($inline =~ /@<TRIPOS>SUBSTRUCTURE/ or $inline eq ""){
    print OUTFILE "$inline";
    $inline = <INFILE>; $inline =~ s/ AR / ar /;
}
close (INFILE);
close (OUTFILE);
system("mv $file.new.mol2 $file.mol2");
}
# end of optimize_mol2
#-----

sub mkdpf{
# generates the dpf-file using actual ligand-pdbq-file, the actual pdbqs-file
# of the receptor and the standard dpf-gen file provided with autodock
# replaces some parameters with the parameters given in the inputfile
my ($ligand, $receptor) = @_;
system("mkdpf3 $ligand.pdbq $receptor.pdbqs");

open (INFILE , "$ligand.$receptor.dpf")
    or die "$ligand.$receptor.dpf could not be read";
open (OUTFILE , ">$ligand.$receptor.new.dpf")
    or die "$ligand.$receptor.new.dpf could not be written";
$inline = <INFILE>;
until ($inline =~ /Simulated Annealing Parameters/){
    if ($inline =~ /about/){
        $inline = "about ".substr($gridcenter,10);
    }
    print OUTFILE "$inline";
    $inline = <INFILE>;
}
foreach $dock_param (@dock_params){
    if ($dock_param =~ /trjout/){
        if (substr ($dock_param,0,1) eq "#"){print OUTFILE "#";}
        print OUTFILE "trjout\t$ligand.trj\t# trajectory file\n";
    }elseif ($dock_param =~ /watch/){
        if (substr ($dock_param,0,1) eq "#"){print OUTFILE "#";}
        print OUTFILE "watch\t$ligand.watch.pdb\t# real-time monitoring\n";
    }elseif ($dock_param =~ /rmsref/){
        if (substr ($dock_param,0,1) eq "#"){print OUTFILE "#";}
        print OUTFILE
            "rmsref\t$ligand.pdbq\t# reference structure file for RMS calc.\n";
    }else{
        print OUTFILE "$dock_param";
    }
}
close (INFILE);
close (OUTFILE);
system("mv $ligand.$receptor.new.dpf $ligand.$receptor.dpf");
}
# end of mkdpf

```

```
#-----  
sub dockall{  
# performs the docking of the actual ligand to the actual receptor  
  my ($ligand, $receptor) = @_;  
  print LOG ("Docking $ligand to $receptor\t");  
  print ("Docking $ligand to $receptor\n");  
  system("autodock3 -p $ligand.$receptor.dpf -l $ligand.$receptor.dlg");  
  print LOG ("Successful\n");  
}  
# end of dockall  
#-----  
sub extract{  
# extracts the clustering histogram  
# extracts the docked ligands  
# converts ligands to mol2-file using babel  
  my ($ligand, $receptor) = @_;  
  $file = "$ligand.$receptor";  
  open (INFILE , "$file.dlg") or die "$file.dlg could not be read";  
  while (<INFILE>){  
    $inline = $_;  
    if ($inline =~ /Unsuccessful Completion/){return;}  
    if ($inline =~ /CLUSTERING HISTOGRAM/){last;}  
  }  
  $inline = <INFILE>;  
  $inline = <INFILE>;  
  $outfile = $file.".clust";  
  open (OUTFILE , ">$outfile");  
  print (OUTFILE "\t$file\n");  
  
  while (<INFILE>){  
    $inline = $_;  
    if ($inline =~ /LOWEST ENERGY DOCKED/){last;}  
    if ($inline =~ /multi-member/){last;}  
    if ($inline =~ /RMSD TABLE/){last;}  
    if (substr($inline, 0, 5) eq "  1 "){print (AD_ONE "$file\t$inline");}  
    print (OUTFILE $inline);  
  }  
  close (OUTFILE);  
  
  $ligand = 1;  
  system ("mkdir $file.mdb");  
  while (<INFILE>){  
    $inline = $_;  
    if ($inline =~ /MODEL/){  
      if (length $ligand == 1) {$outfile = $file."_0".$ligand};  
      if (length $ligand == 2) {$outfile = $file."_".$ligand};  
      open (OUTFILE , ">$outfile");  
    }  
    if ($inline =~ /ATOM/){  
      if (substr($inline, 13, 1) eq "A"){ substr ($inline, 12, 1, "C");}  
      if (substr($inline, 13, 2) eq "cl"){ substr ($inline, 12, 1, "Cl");}  
      if (substr($inline, 13, 2) eq "br"){ substr ($inline, 12, 1, "Br");}  
      $inline =~ s/\'/ /;  
      print (OUTFILE $inline);  
    }  
    if ($inline =~ /ENDMDL/){  
      close (OUTFILE);  
    }  
  }  
}
```

```
        system ("babel -ipdb $outfile -omol2 $outfile.mol2");
        system ("rm $outfile");
        system ("cat $outfile.mol2 >> $file.mol2");
        system ("mv $outfile.mol2 $file.mdb");
        $ligand++;
    }
}
close (INFILE);
}
# end of extract
#-----

sub xscore{
# prepares xpf-inputfile for xscore using the parameters of the inputfile
# revises ligands using xscore-fixmol2
# performs scoring with xscore
    my ($ligand, $receptor) = @_;
    open (OUTFILE , ">$ligand.$receptor.xpf")
        or die "$ligand.$receptor.xpf could not be written";
    print OUTFILE "FUNCTION SCORE\n";
    print OUTFILE "RECEPTOR_PDB_FILE    ./ $receptor.x.pdb\n";
    print OUTFILE "$reference";
    print OUTFILE "LIGAND_MOL2_FILE      ./ $ligand.$receptor.mol2\n";
    print OUTFILE "OUTPUT_TABLE_FILE      ./ $ligand.$receptor.xsc\n";
    print OUTFILE "OUTPUT_LOG_FILE        ./ $ligand.$receptor.xsl\n";
    foreach $xsco_param (@xsco_params){
        print OUTFILE "$xsco_param";
    }
    close (OUTFILE);
    system ("xscore -fixmol2 $ligand.$receptor.mol2 $ligand.$receptor.tmp");
    system ("mv $ligand.$receptor.tmp $ligand.$receptor.mol2");
    system ("xscore $ligand.$receptor.xpf");
    open (INFILE , "$ligand.$receptor.xsc");
    while (<INFILE>){
        $inline = $_;
        if (substr( $inline, 0, 5) eq "1    "){print (XS_ONE "$inline"); last;}
    }
    close (INFILE);
}
# end of xscore
#-----
```

9.2.2 Eingabedatei für das automatisierte Autodock

```
#####
#
#   inputfile for autodock_ix   #
#
#####

# Limitations
# You should not use this script to dock ligands to different proteins
# as the grids are put in the same place for each protein structure.
# But you can use it perfectly for some alligned structres of one
# protein, like different NMR or X-ray structures.

# Caution!!!
# You can use multiple ligands and macromolecules and perform multiple
# docking runs, but 1000 ligands and 20 macromolecules and 50 runs
# result in 100.000 docking calculations and x-score calculations.

# Needs
# Babel is needed to convert the resulting pdb-files of the docked
# ligands to correct mol2 files.
# X-Score is needed to rescore the docked ligands.
# As could be expected, a full installation of autodock 3.0.5 is
# also needed.

# To do before starting #-----
# Prepare Macromolecules (Sybyl)
# Delete hydrogens
# Add essential hydrogens
# Load Kollman Charges
# Save as mol2-file
# Prepare Ligands (Sybyl)
# Add hydrogens
# Compute Gasteiger-Hückel-Charges
# Minimize
# Save as mol2-file
#-----

# Configuration #-----
# Not used, but these variables should be set the shell configuration file.

#set path =(/home/SOFTWARE/autodock3/bin/i86Linux2 $path)
#set path =(/home/SOFTWARE/autodock3/share $path)
#setenv AUTODOCK_UTI /home/SOFTWARE/autodock3/share
#limit stacksize unlimited

#setenv XTOOL_HOME /home/SOFTWARE/xscore_v1.2/
#setenv XTOOL_PARAMETER $XTOOL_HOME/parameter
#setenv XTOOL_BIN $XTOOL_HOME/bin
#set path = ($path $XTOOL_BIN)

#setenv BABEL_DIR /home/SOFTWARE/babel-1.6/
#set path =(/home/SOFTWARE/babel-1.6/ $path)
#-----
```

```
# Keywords #-----

timeout      # Kill Jobs running for a long time
preprec      # Prepare Macromolecule(s)
calcgrd      # Calculate Macromolecule Grids
preplig      # Prepare Ligands
makedpf      # Make all dpf-files           Ligand(s) x Macromolecule(s)
dockall      # Perform all dockings       Ligand(s) x Macromolecule(s)
#parallel    # parallel execution of jobs (not implemented)
results      # Postprocess results (make mol2-files of the docked ligands)
xscore       # Perform x-score calculation for all dockings
#-----

# Macromolecules #-----

receptor1.mol2
receptor2.mol2
#receptor3.mol2

# Ligand Molecules #-----

ncd-10134.mol2
ncd-10408.mol2
ncd-10424.mol2
#ncd-10458.mol2
#ncd-10459.mol2
#ncd-10460.mol2

#-----

# Timeout Parameters #-----

ad_tmax      30      # Time (minutes) before an autodock job is killed
xs_tmax      3       # Time (minutes) before a x-score job is killed

#-----

# Grid Maps Calculation Parameters #-----

npts         80 50 50      # num.grid points in xyz
spacing      .375        # spacing (Angstroms)
gridcenter   2.30 -2.70 -3.30 # xyz-coordinates or "auto"
types        CANOSPHFcbl  # atom type names
smooth       0.500        # store minimum energy within radius (Angstroms)
#-----

# Docking Calculation Parameters #-----

# Simulated Annealing Parameters

rt0 500.      # SA: initial RT
rtrf 0.95     # SA: RT reduction factor/per cycle
#linear_schedule # SA: do not use geometric cooling
runs 20       # SA: number of runs
cycles 80     # SA: cycles
```

```
accs 800          # SA: steps accepted
rejs 800          # SA: steps rejected
select m          # SA: minimum or last

# Trajectory Parameters (Simulated Annealing Only)

#trjfrq 100       # trajectory frequency
#trjbeg 1         # start trj output at cycle
#trjend 50        # end trj output at cycle
#trjout  gtp.trj  # trajectory file
#trjsel  E        # A=acc only;E=either acc or rej
#watch  gtp.watch.pdb # real-time monitoring file

outlev 0          # diagnostic output level

# Docked Conformation Clustering Parameters for "analysis" command
rmstol 2.0        # cluster tolerance (Angstroms)
rmsref  gtp.pdbq  # reference structure file for RMS calc.
#rmsnosym # do no symmetry checking in RMS calc.
#write_all # write all conformations in a cluster

extnrg 1000.      # external grid energy
e0max 0.          # max. allowable initial energy, max. num. retries

# Genetic Algorithm (GA) and Lamarckian Genetic Algorithm Parameters (LGA)

#ga_pop_size 50   # number of individuals in population
#ga_num_evals 250000 # maximum number of energy evaluations
#ga_num_generations 27000 # maximum number of generations
#ga_elitism 1     # num. of top individuals that automatically survive
#ga_mutation_rate 0.02 # rate of gene mutation
#ga_crossover_rate 0.80 # rate of crossover
#ga_window_size 10 # num. of generations for picking worst individual
#ga_cauchy_alpha 0 # ~mean of Cauchy distribution for gene mutation
#ga_cauchy_beta 1 # ~variance of Cauchy distribution for gene mutation
#set_ga          # set the above parameters for GA or LGA

# Local Search (Solis & Wets) Parameters (for LS alone and for LGA)

#sw_max_its 3000 # number of iterations of Solis & Wets local search
#sw_max_succ 4   # number of consecutive successes before changing rho
#sw_max_fail 4   # number of consecutive failures before changing rho
#sw_rho 1.0      # size of local search space to sample
#sw_lb_rho 0.01  # lower bound on rho
#ls_search_freq 0.06 # probability of performing local search on an indiv.
#set_pswl       # set the above pseudo-Solis & Wets parameters

# Perform Dockings

#do_local_only 50 # do only local search
#do_global_only 10 # do only global search (traditional GA)
simanneal # do as many SA runs as set by the "runs" command above
#ga_run 100 # do this many GA or LGA runs

# Perform Cluster Analysis
analysis # do cluster analysis on results
#-----
```

```
# X-Score Calculation Parameters #-----
#REFERENCE_MOL2_FILE ./atp.mol2 # Optional,
# if empty, the first pose of each
# ligand is used to define the pocket

NUMBER_OF_HITS 0 # to extract for each ligand
HITS_DIRECTORY ./hits.mdb
###
### want to include atomic binding scores in the resulting Mol2 files?
###
SHOW_ATOM_BIND_SCORE YES [YES/NO]

# set up scoring functions
APPLY_HPSCORE YES [YES/NO]
  HPSCORE_CVDW 0.004
  HPSCORE_CHB 0.053
  HPSCORE_CHP 0.011
  HPSCORE_CRT -0.061
  HPSCORE_C0 3.448
APPLY_HMSCORE YES [YES/NO]
  HMSCORE_CVDW 0.004
  HMSCORE_CHB 0.094
  HMSCORE_CHM 0.394
  HMSCORE_CRT -0.099
  HMSCORE_C0 3.585
APPLY_HSSCORE YES [YES/NO]
  HSSCORE_CVDW 0.004
  HSSCORE_CHB 0.069
  HSSCORE_CHS 0.004
  HSSCORE_CRT -0.092
  HSSCORE_C0 3.349

# set up chemical rules for pre-screening ligand molecules
APPLY_CHEMICAL_RULES NO [YES/NO]
  MAXIMAL_MOLECULAR_WEIGHT 600.0
  MINIMAL_MOLECULAR_WEIGHT 200.0
  MAXIMAL_LOGP 6.00
  MINIMAL_LOGP 1.00
  MAXIMAL_HB_ATOM 8
  MINIMAL_HB_ATOM 2
#-----
#-----
```


Lebenslauf

Name, Vorname: Claßen-Houben, Dirk

Geburtstag und -ort: 07. Februar 1975 in Heinsberg

Familienstand: Verheiratet seit 06. Dezember 2002

Staatsangehörigkeit: Deutsch

Seit 03/2004 Wissenschaftlicher Leiter Chemoinformatik bei der Biovertis AG in Wien

01/2001 bis 12/2004 Erarbeitung der vorliegenden Dissertation unter der Anleitung von Prof. Dr. Wolfgang Sippl und Prof. Dr. Hans-Dieter Höltje am Institut für pharmazeutische Chemie der Heinrich Heine Universität Düsseldorf

01/2001 Approbation als Apotheker

12/2000 Drittes Staatsexamen

05/2000 bis 12/2000 Praktische Ausbildung in der Arbeitsgruppe von PD Dr. Klaus-Jürgen Schleifer am Institut für pharmazeutische Chemie der Heinrich-Heine-Universität Düsseldorf

11/1999 bis 04/2000 Praktische Ausbildung in der Schloss-Apotheke, Bernd Barsch in Mönchengladbach/Rheydt

09/1999 Zweites Staatsexamen

08/1997 Erstes Staatsexamen

07/1997 bis 05/1999 Mitarbeit im Arbeitskreis von Prof. Dr. Rudolf Bauer am Institut für pharmazeutische Biologie der Heinrich-Heine-Universität Düsseldorf; Beteiligung an vier Veröffentlichungen (Chinese Drug Monographs and Analysis)

10/1995 bis 07/1999 Studium der Pharmazie an der Heinrich-Heine-Universität Düsseldorf

07/1994 bis 06/1995 Wehrdienst: Sanitätsdienst in Würzburg und Budel/NL

08/1991 bis 06/1994 Cusanusgymnasium Erkelenz (Abitur)

08/1985 bis 07/1991 Städt. Realschule Heinsberg (Fachoberschulreife)

08/1981 bis 07/1985 Kath. Grundschule Braunsrath

