

**Molecular Modelling Studien
am humanen Cytochrom P450 1A2**

Inaugural-Dissertation

Zur Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Andrea Hombrecher

aus Wermelskirchen

Düsseldorf 2005

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

Referent: Univ.-Prof. Dr. Dr. h.c. Hans-Dieter Höltje

Korreferent: Univ.-Prof. Dr. Wolfgang Sippl

Tag der mündlichen Prüfung: 18. Juli 2005

Diese Arbeit wurde als elektronische Dissertation veröffentlicht.

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

Die vorliegende Arbeit wurde von April 2001 bis Juni 2005 am Institut für Pharmazeutische Chemie der Heinrich-Heine-Universität Düsseldorf unter Anleitung von Herrn Prof. Dr. Dr. h.c. H.-D. Höltje angefertigt.

Mein besonderer Dank gilt Herrn Prof. Dr. Dr. h.c. H.-D. Höltje für die Überlassung des interessanten Themas, die hilfreichen Diskussionen und Anregungen und die hervorragenden Arbeitsbedingungen am Institut. Außerdem danke ich ihm für die Unterstützung meines Forschungsaufenthaltes an der Universität Kuopio in Finnland.

Herrn Prof. Dr. Antti Poso danke ich für die Bereitstellung der experimentellen Daten, für viele wertvolle Anregungen und für die Möglichkeit, in seiner Arbeitsgruppe in Kuopio mitarbeiten zu dürfen.

Herrn Prof. Dr. W. Sippl danke ich herzlich für die Anfertigung des Zweitgutachtens.

Weiterhin möchte ich allen Mitgliedern der Arbeitsgruppe für die angenehme Arbeitsatmosphäre, die konstruktive Zusammenarbeit und Hilfsbereitschaft danken.

INHALTSVERZEICHNIS

1	EINLEITUNG	7
1.1	Cytochrome P450	10
1.1.1	Nomenklatur	11
1.1.2	Evolutionäre Entwicklung	11
1.1.3	Verbreitung	12
1.1.4	Funktionen	13
1.2	Das humane Cytochrom P450 1A2	16
1.2.1	Substrate und katalysierte Reaktionen	17
1.2.2	Regulation der Enzymaktivität	28
1.3	Katalytischer Mechanismus der Cytochrome P450	32
1.3.1	Allgemeiner Reaktionszyklus	32
1.3.2	Mechanismus CYP1A2-katalysierter Reaktionen	35
1.4	Proteinstruktur der Cytochrome P450	38
1.4.1	Kristallstrukturen von Säugetiercytochromen	38
1.4.2	Allgemeine Tertiärstruktur	39
2	PROBLEMSTELLUNG UND ZIELSETZUNG	49
3	METHODEN	53
3.1	Analyse der Proteinsequenz	55
3.1.1	Automatisches Sequenzalignment	55
3.1.2	Sekundärstrukturvorhersagen	57
3.1.3	Erkennung von Sequenz-Struktur-Homologien	58
3.2	Homologie-Modelling	58
3.2.1	Behandlung der strukturkonservierten Bereiche	58
3.2.2	<i>Loop Search</i> Routine	59
3.3	Quantenmechanische Berechnung von Atomeigenschaften	60
3.4	Molekülmechanik	61
3.4.1	Das GROMOS96-Kraftfeld	62
3.4.2	Geometrieoptimierung	66
3.4.3	Moleküldynamiksimulationen	67
3.5	Analyse des Proteinmodells	69
3.5.1	Strukturvalidierung mit PROCHECK	69
3.5.2	SURFNET-Kavitäten	70
3.5.3	Clusteranalyse mit NMRCLUST	70
3.6	3D-QSAR Untersuchungen	71
3.6.1	Protein-Ligand-Docking	71
3.6.2	GRID- Wechselwirkungsfelder	74
3.6.3	Statistische Auswertung mit GOLPE	75

4	DURCHFÜHRUNG UND ERGEBNISSE	79
4.1	Konstruktion des Homologiemodells	81
4.1.1	Auswahl der Templatstruktur	81
4.1.2	Sekundärstrukturvorhersagen	82
4.1.3	Sequenzalignment	85
4.1.4	Festlegung der strukturkonservierten Bereiche	89
4.1.5	Modellierung der strukturvariablen Bereiche	90
4.1.6	Energieminimierung	93
4.1.7	Einfügen des Häm in das Modell	94
4.2	Strukturoptimierung und Validierung des Modells	95
4.2.1	Durchführung und Simulationsbedingungen	95
4.2.2	Equilibrierung der Kristallstruktur	98
4.2.3	Auswertung der MDS mit der Kristallstruktur	101
4.2.4	Equilibrierung des Proteinmodells	108
4.2.5	Auswertung der MDS mit dem Modell	111
4.3	Untersuchung von Enzym-Substrat-Komplexen	119
4.3.1	Auswahl von Substraten und Positionierung in das Modell	119
4.3.2	Durchführung der MDS mit Enzym-Substrat-Komplexen	125
4.3.3	MDS mit der Kristallstruktur 1N6B	126
4.3.4	MDS mit der Kristallstruktur 1NR6	128
4.3.5	MDS des Modells mit MelQ	135
4.3.6	MDS des Modells mit 7-Ethoxyresorufin	142
4.3.7	MDS des Modells mit Coffein zur Umsetzung zum Paraxanthin	149
4.3.8	MDS des Modells mit Coffein zur Umsetzung zum Theobromin	154
4.3.9	Vergleich der beiden MDS mit Coffein	159
4.3.10	Vergleich der drei untersuchten Substratkomplexe	160
4.4	3D-QSAR-Studie	162
4.4.1	Validierung des Dockings	162
4.4.2	Analyse des Inhibitor Datensatzes	166
4.4.3	Docking des Datensatzes	170
4.4.4	3D-QSAR-Modell	172
5	DISKUSSION	177
5.1	Konstruktion des Homologiemodells	179
5.2	Validierung des Proteinmodells mittels MDS	180
5.3	Berechnung von Substratkomplexen	182
5.4	Substratkomplexe mit MelQ und 7-Ethoxyresorufin	185
5.5	Substratkomplexe mit Coffein	187
5.6	3D-QSAR-Studie	189
6	ZUSAMMENFASSUNG	191
7	LITERATUR	195
8	ANHANG	209

1 EINLEITUNG

Arzneimittelwechselwirkungen sind in der pharmazeutischen Praxis ein ernst zu nehmendes Problem. Nach Angaben des Niedersächsischen Ärzteblatts wurden in einer kürzlich veröffentlichten englischen Studie 6,5 % aller Krankenhausaufnahmen auf Arzneimittelwechselwirkungen zurückgeführt [1]. Besonders bei älteren multimorbiden Patienten, die oft eine Vielzahl unterschiedlicher Arzneimittel gleichzeitig einnehmen müssen, führen Interaktionen zwischen miteinander unverträglichen Arzneimitteln nicht nur zu einer Verlängerung des Krankenhausaufenthaltes und damit zu einer Erhöhung der Behandlungskosten, sondern sie steigern letztendlich auch das Mortalitätsrisiko der Patienten.

Eine zentrale Rolle im Bereich der pharmakokinetischen Arzneimittelwechselwirkungen spielt die Superfamilie der Cytochrom-P450-Enzyme. In der Phase I der Biotransformation katalysieren die Cytochrome P450 die oxidative Hydroxylierung von Arzneistoffen und anderen lipophilen Xenobiotika, damit diese dann in der Phase II des Metabolismus durch Konjugation mit einer körpereigenen Substanz in hydrophilere, leichter ausscheidbare Verbindungen überführt werden können. Sie sind dabei in der Lage, ein ungewöhnlich breites Spektrum an strukturell höchst unterschiedlichen Substraten umzusetzen. Ebenso groß ist dementsprechend auch die Zahl der Cytochrom-P450-Inhibitoren, weshalb es unter Umständen zu ernsthaften Komplikationen kommen kann, wenn mehrere Arzneistoffe gleichzeitig als Substrate oder Inhibitoren um die Bindungsstelle im selben Enzym konkurrieren. Beispielsweise werden über 50 % aller Arzneistoffe über das Isoenzym Cytochrom P450 3A4 metabolisiert, und es sind darüber hinaus zahlreiche Inhibitoren wie Azol-Antimykotika oder Makrolid-Antibiotika, aber auch Inhaltsstoffe des Grapefruitsaftes bekannt, die toxische Interaktionen verursachen können. Eine nicht ganz unbedeutende Rolle spielt auch das Isoenzym Cytochrom P450 1A2. Zwar metabolisiert dieses Isoenzym nur etwa 15 % der heute verwendeten Pharmaka, es befinden sich darunter jedoch auch solche mit einer engen therapeutischen Breite wie Clozapin oder Theophyllin. Kombiniert man zum Beispiel eines dieser Arzneimittel mit einem Fluorochinolon-Antibiotikum, das die Aktivität des Cytochroms hemmt, kann es durch erhöhte Blutspiegel der langsamer abgebauten Arzneistoffe zu schwerwiegenden Intoxikationen kommen.

Da viele Interaktionen bereits zugelassener Arzneistoffe bekannt sind, können toxische Nebenwirkungen in der Regel durch umsichtige Verschreibung des behandelnden Arztes vermieden werden. Es ist jedoch auch von Interesse, das Interaktionspotential eines Arzneimittels bereits in seiner Entwicklungsphase zu erkennen. Denn in einer Studie des „Centre for Medicines Research“ über die Gründe für das Scheitern von neuen Arzneistoffentwicklungen in der frühklinischen Phase stellte sich heraus, dass bei 16 % aller abgebrochenen Entwicklungsprojekte das Auftreten von Nebenwirkungen, zu denen ja auch die Arzneimittelwechsel-

wirkungen zählen, für das Fehlschlagen des Entwicklungsprojektes verantwortlich waren [2], [3].

Ein möglicher Weg zur Vorhersage von Biotransformationsreaktionen gerade auch in den frühen Stadien der Arzneistoffentwicklung sind theoretische Untersuchungen mit Methoden des Molecular Modelling. Struktur-Wirkungsbeziehungen lassen sich nicht nur für Targets modellieren, an denen der Arzneistoff wirken soll, sondern auch für Anti-Targets, die zu Arzneistoffinteraktionen oder anderen Nebenwirkungen führen könnten. Von einigen Cytochrom-P450-Isoenzymen existieren bereits Kristallstrukturen, die zur Vorhersage von Substrat-Präferenzen herangezogen werden können. Für Isoenzyme, deren dreidimensionale Struktur noch nicht aufgeklärt worden ist, wie beispielsweise das Cytochrom P450 1A2, bilden die Kristallstrukturen eine gute Grundlage für ein Homologiemodell, mit dessen Hilfe dann Vorhersagen durchgeführt werden können.

1.1 Cytochrome P450

Die Cytochrome P450 (EC 1.14.14.1)¹ stellen eine ubiquitär vorkommende Superfamilie von Monooxygenasen dar. Sie besitzen die Fähigkeit, molekularen Sauerstoff in ihrem aktiven Zentrum zu binden und zu aktivieren. Dabei katalysieren sie dessen Spaltung, bei der ein Sauerstoffatom auf das Substrat übertragen und das andere unter Protonenaufnahme zu Wasser reduziert wird. Alle Isoenzyme dieser Superfamilie sind Hämoproteine mit einem Häm (Eisen-Protoporphyrin IX), das als prosthetische Gruppe das katalytische Zentrum der Cytochrome bildet. Das Zentralatom Eisen ist über den Thiolschwefel eines Cysteins kovalent an das Apoenzym gebunden, das aus etwa 400-500 Aminosäuren besteht. Durch die Eisen-Thiolat-Bindung werden die Redox-Eigenschaften von Eisen und Schwefel in einzigartiger Weise genutzt, um molekularen Sauerstoff für chemische Reaktionen zu aktivieren [4]. Ihren Namen erhielten die Cytochrome P450 durch das charakteristische Absorptionsmaximum bei 450 nm, das durch die Bindung von Kohlenmonoxid (CO) an das Häm zustande kommt. „Cytochrom“ kann man mit „Zellfarbstoff“ übersetzen, „P“ steht für „Pigment“ und „450“ für die Absorptionswellenlänge des CO-Komplexes. Garfinkel und Klingenberg beobachteten 1958 unabhängig voneinander dieses Phänomen und gelten damit als die Entdecker der Cytochrome P450 [5], [6].

¹ E.C.: Enzymklassifikationssystem (*Enzyme Commission numbers*); kategorisiert wird durch dieses System die Reaktion, die das Enzym katalysiert

1.1.1 Nomenklatur

Während in den Anfängen der Cytochrom-Forschung neu entdeckte Isoenzyme in der Regel mit einem individuellen Namen bezeichnet wurden, der häufig auf Substratspezifitäten oder physikochemische Eigenschaften der Cytochrome verwies, so hat sich in den letzten Jahren eine systematische Einteilung der Cytochrome in Familien, Unterfamilien und individuelle Isoenzyme entsprechend ihrer Sequenzidentitäten durchgesetzt [7]. Auf die generelle Abkürzung CYP (Cytochrom P450) folgt zunächst die Angabe der Familie als arabische Ziffer (CYP1, 2, 3, ...). Zu einer Familie gehören Isoenzyme mit einer Sequenzidentität > 40 %. Isoenzyme mit einer Sequenzidentität > 55 % bilden eine gemeinsame Unterfamilie, die mit einem lateinischen Buchstaben bezeichnet wird (CYP1A, 1B, 1C, ...). Individuelle Isoenzyme sollten sich in ihrer Sequenzidentität um mehr als 3 % voneinander unterscheiden. Sie werden wiederum durch arabische Ziffern gekennzeichnet (CYP1A1, 1A2, ...).

1.1.2 Evolutionäre Entwicklung

Man nimmt an, dass es in Archaeobakterien bereits seit 3,5 Milliarden Jahren Eisen-Schwefel-Enzyme gibt, die Redoxreaktionen katalysieren können [8]. Die Entwicklung der Superfamilie der Cytochrome P450 aus einem gemeinsamen Vorläufergen begann vermutlich vor etwa 1,4 Milliarden Jahren eng zusammenhängend mit dem Anstieg des Sauerstoffgehalts in der Atmosphäre und verlief dann weitgehend parallel zur Evolution der Lebensformen auf der Erde, was sich heute in der ubiquitären Verbreitung dieser Enzymfamilie widerspiegelt (Abbildung 1.1) [9]. Wahrscheinlich entstanden die Cytochrome P450 entweder aus der Notwendigkeit heraus, den verfügbaren Sauerstoff für die Erschließung von Kohlenstoffquellen mittels Oxidationsreaktionen zu nutzen, oder um den für einige Organismen toxischen molekularen Sauerstoff durch Reduktion zu Wasser zu entgiften. Im weiteren Verlauf der Entwicklung spielten die P450-Enzyme zunächst in erster Linie bei der Biosynthese und dem Metabolismus von Steroiden und anderen endogenen Substanzen eine wichtige Rolle. Erst mit der zunehmenden Besiedelung des Landes durch die Tiere erfolgte eine Abspaltung von überwiegend Xenobiotika metabolisierenden Cytochromfamilien, vermutlich zur Entgiftung von Pflanzentoxinen, die von zahlreichen Pflanzen als Fraßschutz synthetisiert wurden [10]. Es fand eine immer weitere Differenzierung in Unterfamilien mit unterschiedlichen Substratspezifitäten statt, sodass eine breite Spanne von Substanzklassen abgedeckt werden konnte. Auch die Familien CYP1 und CYP2, deren Aufspaltung in spezialisierte Unterarten in diesen Zeitraum fällt, haben vermutlich eine ganze Reihe verschiedener Pflanzentoxine entgiftet. Die Tatsache, dass CYP1A1 in vielen Fischarten gefunden wird, CYP1A2 aber nicht, ist ein Indiz dafür, dass eine Auftrennung der Familie 1 in ihre Unterfamilien erst deutlich nach der Besiedelung

des Landes stattgefunden hat. Dabei ist die Entstehung von CYP1A2 vermutlich auf die Notwendigkeit zurückzuführen heterozyklische Amine zu metabolisieren. Der Anstieg des CYP1A2-Spiegels im Verlauf der Evolution des Menschen ließe sich dann mit dem Verzehr von gekochtem Fleisch erklären, bei dessen Zubereitung heterozyklische Amine entstehen [11].

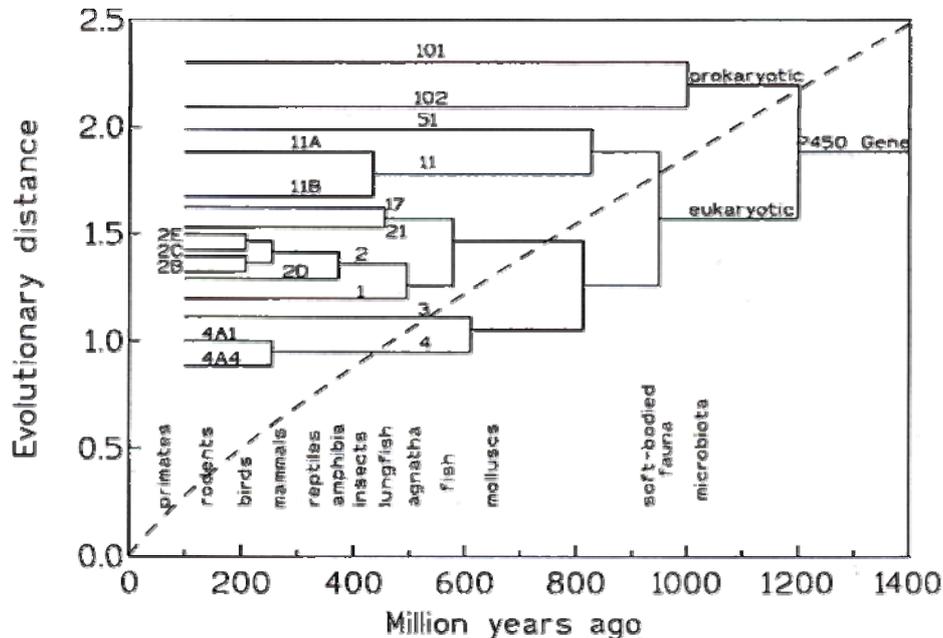


Abbildung 1.1: Vereinfachte Darstellung des phylogenetischen Baumes der Cytochrome P450, der ihre Entwicklung wiedergibt [11]

1.1.3 Verbreitung

Bereits bei der Darstellung der evolutionären Entwicklung der P450-Superfamilie (Abschnitt 1.1.2 und Abbildung 1.1) wurde ihre ubiquitäre Verbreitung angesprochen. Tatsächlich findet man P450-Isoenzyme in Organen und Geweben nahezu jeder Klasse von Lebewesen, angefangen bei den Archaeobakterien über Pilze, Pflanzen, Insekten, Amphibien, Reptilien, Fische und Vögel bis hin zu den Säugetieren und dem Menschen. Im Jahr 2004 waren insgesamt 3043 verschiedene CYP-Gene (inklusive Pseudogene²) bekannt, die sich 368 Familien, bzw. 814 Unterfamilien zuordnen ließen. Davon stammten 1277 Gensequenzen aus dem Tierreich und 207 von den eukaryotischen³ Pilzen und Einzellern. Des Weiteren gehörten 1098 pflanzliche und 461 bakterielle Sequenzen dazu. Im Vergleich dazu sind vom

² Genkopien, die kein funktionelles Protein in voller Länge codieren. Sie spielen eine Rolle bei der Regulierung der Genaktivität.

³ Lebewesen, die einen echten Zellkern und ein Cytoskelett besitzen

Menschen mittlerweile 60 unterschiedliche Cytochrome bekannt. Dass die Forschung auf diesem Gebiet noch nicht abgeschlossen ist, erkennt man daran, dass alleine in der Reispflanze (*Oryza sativa*) kürzlich 450 neue verschiedene CYP-Gene entdeckt wurden. Eine Übersicht über alle bekannten Sequenzen gibt die ständig aktualisierte „Cytochrome P450 Homepage“ von D.R. Nelson [12].

Die Cytochrome aller eukaryotischen Lebewesen sind membranständige Enzyme. Sie sind über einen Membrananker an die Membran des endoplasmatischen Retikulums gebunden und ragen als kugelförmiges Protein ins Cytosol. Einige P450-Isoenzyme, unter anderem das CYP1A1, werden auch in den Mitochondrien gefunden, wo sie ebenfalls membrangebunden vorliegen.

Bei den Säugetieren enthält die Leber den größten Anteil an P450-Isoenzymen. Geringere Mengen einzelner Cytochrome P450 findet man darüber hinaus aber auch in fast jedem anderen Organ oder Gewebe wie den Nieren, der Lunge, den Geschlechtsorganen, der Nebennierenrinde, dem Gehirn, der Nasenschleimhaut, der Plazenta, dem Pankreas, der Milz, dem Gastrointestinaltrakt oder der Haut [13] (vgl. auch Tabelle 1.1).

1.1.4 Funktionen

Häufig hängt der Wirkort der einzelnen CYP-Isoenzyme eng mit ihrer jeweiligen Funktion zusammen. Wie bereits in Abschnitt 1.1.2 erwähnt, sind eine ganze Reihe von Säugetiercytochromen in die Biosynthese von Steroidhormonen eingebunden. Sie kommen daher überwiegend in der Nebennierenrinde vor, wo die Steroidbiosynthese stattfindet. Schon der erste Biosyntheseschritt, die oxidative Abspaltung der Seitenkette des Cholesterols, wird durch ein Cytochrom (CYP11A1) katalysiert. Weitere an der Synthese von Corticosteroiden, Androgenen, Estrogenen und Gestagenen beteiligte Isoenzyme sind CYP11B1, CYP17A1, CYP19A1 und CYP29A1, bei denen man teilweise von ihrem Namen auf die von ihnen katalysierte Reaktion schließen kann (CYP11B1 z.B. katalysiert die 11 β -Hydroxylierung des Progesterons). Weitere Beispiele für spezifische cytochromkatalysierte Reaktionen körpereigener Substanzen sind der Abbau von Cholesterol zu Gallensäure (CYP7, CYP27) sowie die Hydroxylierung von Fettsäuren, Prostaglandinen, Leukotrienen (CYP4) und Vitamin D₃ (CYP27).

Auch die Säugetiercytochrome der Familien 1-3 sind in der Lage, Steroidhormone an verschiedenen Positionen zu hydroxylieren. Diese Reaktionen haben jedoch nichts mit den oben aufgeführten Biosynthesereaktionen zu tun und sind in der Regel von untergeordneter Bedeutung. Die Hauptfunktion dieser Cytochrome besteht in der Umsetzung von exogenen Substanzen. Dabei sind sie wie bereits eingangs erwähnt vor allem an der Phase I der Biotransformation von über 90 % aller Arzneistoffe beteiligt, wo sie über 60 verschiedene Reaktionen katalysieren [11]. Im Gegensatz zu

eng eingegrenzten regio- und stereospezifischen Biosynthesereaktionen katalysieren die Xenobiotika metabolisierenden Cytochrome eine große Anzahl weitaus unspezifischerer Reaktionen und weisen dabei eine deutlich höhere Substratvariabilität auf. Diese relative Vielseitigkeit einzelner Cytochrome einerseits und das Zusammenspiel verschiedener Isoenzyme mit unterschiedlichen Substratspezifitäten andererseits bewirken, dass das P450-Enzymsystem einen hocheffizienten Mechanismus für den Metabolismus körperfremder Stoffe darstellt. Einen Überblick über die humanen Cytochrome der Familien 1-3, die für den Metabolismus von Xenobiotika von Bedeutung sind, ihre Wirkorte und Substratspezifitäten gibt Tabelle 1.1.

Tabelle 1.1: Humane Cytochrome der Familien 1-3, die für den Metabolismus von Xenobiotika maßgeblich sind (nach [14]). Abkürzungen: GIT – Gastrointestinaltrakt, PAH – polyzyklische aromatische Kohlenwasserstoffe (*polycyclic aromatic hydrocarbons*).

CYP	Vorkommen	Typische Substrate
1A1	Lunge, Leber, Gehirn, GIT, Lymphozyten, Herz	PAH, z.B. Benzo[a]pyren
1A2	Leber	aromatische Amine, PAH, Coffein
1B1	steroidbildende Gewebe, Tumorgewebe, Haut, Gehirn, Herz, Lunge, Leber, Niere, GIT, Milz	17 β -Estradiol, PAH
2A6	Leber	Cumarine, Steroide
2B6	Leber, Herz	Nicotin
2C8	Leber, Nieren	Retinoide, Paclitaxel
2C9/10	Leber	Tolbutamid, Diclofenac
2C19	Leber, Herz	(S)-Mephenytoin, Omeprazol, Diazepam
2D6	Leber, Gehirn, Herz	β -Blocker, trizyklische Antidepressiva
2E1	Leber, Lunge, Gehirn, Endothel, Herz, Knochenmark	Ethanol, Nitrosamine, Paracetamol
2F1	Lunge	Cumarine
3A4/5	Leber, GIT, Nieren, Lunge, Gehirn, Endothel, Plazenta, Lymphozyten	Ca-Kanalblocker, Ciclosporin, Paracetamol, Paclitaxel, Steroide
3A7	fötale Leber, Plazenta	s. CYP3A4/5

Viele der hier aufgeführten Isoenzyme kommen wiederum hauptsächlich in der Leber vor und sind dort in unterschiedlichem Ausmaß an Metabolisierungsreaktionen beteiligt. Die größte Bedeutung hat CYP3A4, das etwa 30 % der Lebercytochrome ausmacht und am Metabolismus von über 50 % aller Arzneistoffe beteiligt ist, die von Cytochromen umgesetzt werden. Ebenfalls stark vertreten ist die Unterfamilie CYP2C, der 20 % der Cytochrome in der Leber angehören. Eine wichtige Rolle für den Arzneistoffmetabolismus spielt auch CYP2D6, das etwa 30 % der von Cytochromen umgesetzten Arzneistoffe metabolisiert [14]. Hier ist wiederum der Zusammenhang zwischen Wirkort und Funktion interessant, da dieses Isoenzym neben der Leber auch im Herzen gefunden wird und die Substanzklasse der β -Blocker zu seinen Substraten gehört. Schließlich zählt auch CYP1A2 zu den

wichtigen arzneistoffmetabolisierenden Isoenzymen, wobei es aber darüber hinaus auch eine bedeutende Rolle bei der Metabolisierung anderer körperfremder Chemikalien spielt.

1.2 Das humane Cytochrom P450 1A2

Auch die Familie 1 der Cytochrome ist weit verbreitet. Das CYP1A1-Isoenzym beispielsweise kommt nicht nur beim Menschen, bei einigen Säugetierarten und Hühnern vor, sondern wird vor allem auch in vielen Fischarten gefunden [15] (vgl. auch Abschnitt 1.1.2). Die Verbreitung von CYP1A2 dagegen beschränkt sich auf den Menschen, den Hund, einige Nagetiere und Hühner. Während die CYP1A-Isoenzyme bereits seit langem bekannt und in der Literatur ausführlich dokumentiert sind [16], wurde die Unterfamilie CYP1B mit ihrem Vertreter CYP1B1 erst vor wenigen Jahren entdeckt und kommt ausschließlich beim Menschen, bei der Ratte und bei der Maus vor [17]. Eine Übersicht über alle relevanten Enzyme der Familie CYP1, für die ein Eintrag in der SwissProt-Sequenzdatenbank [18] existiert, und ihr Vorkommen gibt Tabelle 1.2.

Tabelle 1.2: Übersicht über alle relevanten Enzyme der Familie CYP1 [15]

Cytochrom	Organismus
CYP1A1	Mensch , versch. Makakenarten, Schaf, Ratte, Maus, Kaninchen, Meerschweinchen, Goldhamster, Wolfsbarsch, versch. Meeräschenarten, Austernfisch, Tomcod, Flunder, versch. Schollenarten, versch. Meerbrassenarten, Regenbogenforelle, Vieraugen-Falterfisch
CYP1A2	Mensch , Hund, Ratte, Maus, Kaninchen, Meerschweinchen, Goldhamster, Huhn
CYP1A3	Regenbogenforelle
CYP1A4	Huhn
CYP1A5	Huhn
CYP1B1	Mensch , Ratte, Maus

Bezüglich ihres Vorkommens im Körper sind die Cytochrome der Familie 1 ein eher untypisches Beispiel, da zwei der drei humanen Isoenzyme überwiegend extrahepatisch exprimiert werden (vgl. Tabelle 1.1). CYP1A1 wird vor allem in der Lunge gefunden [19], und CYP1B1 ist in den steroidbildenden Geweben wie der Nebennierenrinde, dem Uterus, den Eierstöcken, der Brust, dem Hoden und der Prostata konstitutiv verbreitet [20]. Eine signifikante Expression von CYP1B1 wurde außerdem in zahlreichen Tumorgeweben beobachtet [21]. CYP1A2 kommt als einziges humanes Isoenzym der Familie 1 fast ausschließlich konstitutiv in der Leber vor. Das macht bereits deutlich, dass auch nur dieses Cytochrom eine nennenswerte Rolle im Arzneistoffmetabolismus spielen kann.

Das humane CYP1A2 besteht aus 515 Aminosäuren, hat ein Molekulargewicht von 58,294 kDa und weist mit 72 % eine vergleichsweise geringe Sequenzidentität zum CYP1A1 auf, dem zweiten humanen Isoenzym aus derselben Unterfamilie. Die Sequenzidentität zum CYP1B1 ist erwartungsgemäß mit 41 % nochmals deutlich geringer. Um die Übertragbarkeit von experimentellen Daten abschätzen zu können,

die am selben Isoenzym verschiedener Säugetierspezies gewonnen wurden, sollte erwähnt werden, dass die Sequenzidentität zu CYP1A2-Isoenzymen von Nagetierarten ebenfalls nur bei 70-80 % liegt. Für das entsprechende Rattenenzym beträgt sie beispielsweise 75 %, für das CYP1A2 der Maus 74 % und für das CYP1A2 des Kaninchens 79 %. Beim Vergleich der Aminosäuresequenz mit CYP-Sequenzen anderer Familien fand man heraus, dass eine enge Verwandtschaft zur Familie CYP2 und zu CYP17 besteht, was sich anhand von Abbildung 1.1 gut nachvollziehen lässt. Die Sequenzidentität beträgt hier jeweils etwa 30 % [17].

Bisher wurden mehrere genetische Polymorphismen des CYP1A2-Gens beobachtet, von denen einige auch die codierenden Sequenzabschnitte betreffen [22], [23]. Die Auswirkungen der Mutationen auf die katalytische Aktivität des Proteins konnten bislang allerdings nicht eindeutig geklärt werden.

Im Hinblick auf die Beschaffenheit der Proteinstruktur von CYP1A2 weiß man aus UV-spektroskopischen Studien, dass das aktive Zentrum im Vergleich zu anderen Cytochromen (z.B. CYP102, CYP3A4) extrem widerstandsfähig und rigide ist. Es erweist sich als unempfindlich gegenüber einer Denaturierung durch Druck und zeigt außerdem nur eine geringe Komprimierbarkeit. Darüber hinaus kann man anhand der Vinylbande des Häm ablesen, ob die Vinylsubstituenten in beiden möglichen Konformationen (*in-plane* und *out-of-plane* im Verhältnis zur Hämebene) vorkommen oder nicht. Wenn wie beim CYP1A2 ausschließlich die *in-plane* Konformation existiert, lässt das wiederum auf eine eingeschränkte Beweglichkeit der Vinylsubstituenten und damit auf ein rigides aktives Zentrum schließen [24].

1.2.1 Substrate und katalysierte Reaktionen

In guter Übereinstimmung mit den oben beschriebenen Eigenschaften des aktiven Zentrums von CYP1A2 stehen die von D.F.W. Lewis im Jahr 2000 postulierten charakteristischen Merkmale von Substraten dieses Isoenzym. Wichtigstes Charakteristikum für CYP1A2-Substrate ist demnach ihre Planarität. Typische Substrate weisen darüber hinaus eine im Vergleich zu den Substraten anderer Isoenzyme mittlere Größe auf, die nur in engen Grenzen variiert. Sie sind außerdem neutral bis schwach basisch [25].

Grundsätzlich erfüllt das CYP-Isoenzym 1A2 aufgrund seiner Substratpräferenzen und der Reaktionen, die es katalysiert, zwei Aufgaben. Zum einen ist CYP1A2, wie bereits mehrfach erwähnt, das einzige Isoenzym aus der Familie 1, das eine wichtige Funktion im Phase-I-Metabolismus von Arzneistoffen besitzt. Das ist zwar vor allem dadurch begründet, dass es als einziges Isoenzym konstitutiv in der Leber vorkommt, doch es zeigen sich auch deutlich unterschiedliche Substratspezifitäten der drei Isoenzyme gegenüber Arzneistoffen [26]. Darüber hinaus kommt der gesamten Familie CYP 1 bei der metabolischen Aktivierung von Mutagenen und Prokarzino-

genen eine große Bedeutung zu. Während bei den bisher erwähnten Reaktionen der xenobiotikametabolisierenden Cytochrome davon ausgegangen wurde, dass Biotransformationsreaktionen ausschließlich Abbaureaktionen zur Entgiftung von Arzneistoffen oder anderen Chemikalien darstellen, können P450-Isoenzyme allerdings auch genauso gut an einer oxidativen Aktivierung ihrer Substrate zu toxischen Metaboliten beteiligt sein. In diesem Zusammenhang zählen die polyaromatischen Kohlenwasserstoffe (*polycyclic aromatic hydrocarbons*, PAH), die polyaromatischen und heterozyklischen Amine und Amide sowie das Aflatoxin B₁ zu den Substraten der CYP1, denen gegenüber die einzelnen Isoenzyme unterschiedliche aber überlappende Substratspezifitäten aufweisen.

1.2.1.1 Metabolische Aktivierung von Prokarzinogenen

Die erste in Abschnitt 1.2.1 erwähnte Substanzklasse der Prokarzinogene, die PAH, sind ubiquitär vorkommende Luftschadstoffe, die bei der unvollständigen Verbrennung von organischem Material entstehen und vor allem im Zigarettenrauch, in Autoabgasen, bei der Kohleverbrennung und bei zahlreichen anderen industriellen Prozessen, wie z.B. in Stahlwerken oder Gießereien [27] freigesetzt werden. Auch beim Erhitzen von Fleisch bei hohen Temperaturen, z.B. durch Grillen, kommt es zur Bildung von PAH. Abbildung 1.2 zeigt einige wichtige Verbindungen, an deren Aktivierung CYP1 beteiligt ist [20].

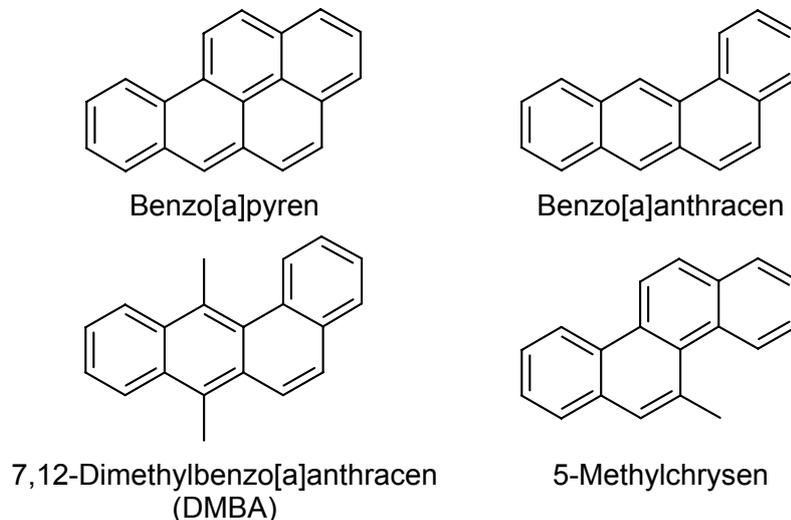


Abbildung 1.2: PAH, die durch die Familie CYP1 aktiviert werden

Die metabolische Aktivierung dieser Substanzen soll am Beispiel von Benzo[a]pyren kurz erklärt werden (Abbildung 1.3). Benzo[a]pyren kann an verschiedenen Ringpositionen von Cytochromen hydroxyliert werden. Während das 3-Hydroxybenzo[a]pyren den Hauptmetaboliten darstellt, ist die kritische Initialisierungsreaktion für die Karzinogenese die stereoselektive Bildung des (+)-7,8-Epoxids, das dann von der Epoxidhydrolase zunächst weiter zum 7,8-Diol hydrolysiert wird. Die isolierte Doppelbindung ist reaktiv genug, um eine zweite cytochromkatalysierte Epoxidierungsreaktion hervorzurufen. Das daraus resultierende (+)-7,8-Diol-9,10-Epoxid

ist ein hochreaktives Elektrophil, das kovalent an die Stickstoffe der Purinbasen der DNA binden kann [29].

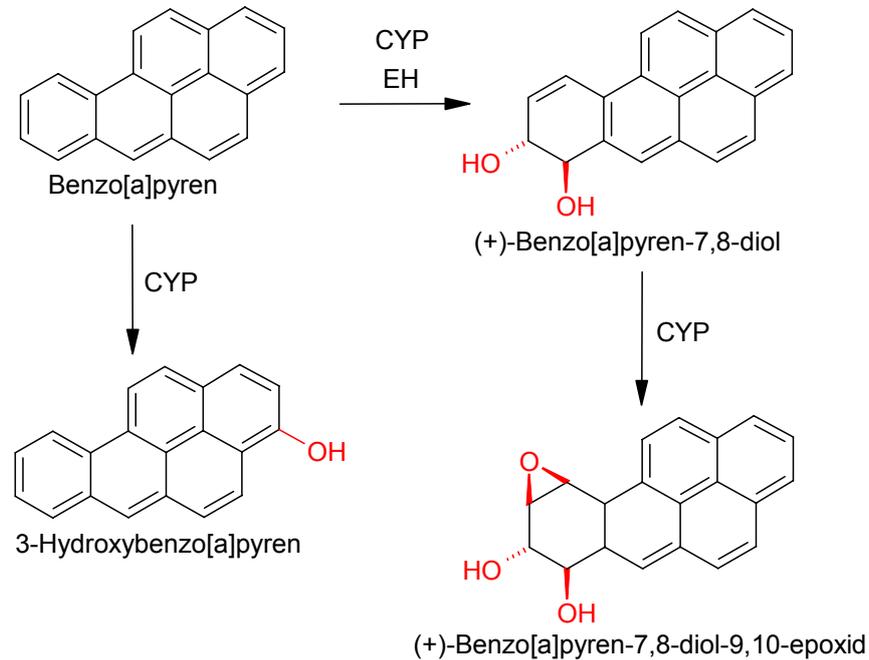


Abbildung 1.3: Metabolismus von Benzo[a]pyren (nach [28]). CYP: Cytochrome, EH: Epoxid-hydrolase

CYP1A2 ist am Metabolismus der neutralen PAH nur in relativ geringem Ausmaß beteiligt. Sie werden bevorzugt von CYP1A1 umgesetzt, weshalb aufgrund der Lokalisation von CYP1A1 in der Lunge hier häufig ein Zusammenhang zwischen einer Inhalation von PAH beispielsweise durch Rauchen und dem Auftreten von Lungenkrebs gesehen wird. Die Substratpräferenz des CYP1A2 liegt dagegen eher bei den stickstoffhaltigen aromatischen Aminen und Amiden [30], auf die im Folgenden näher eingegangen wird.

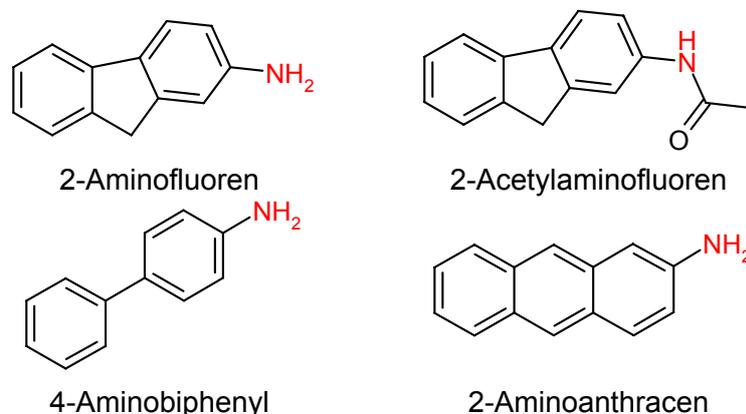


Abbildung 1.4: Aromatische Amine und Amide, deren Umsetzung durch CYP1A2 katalysiert wird

Die in Abbildung 1.4 aufgeführten polyaromatischen Amine und Amide, deren oxidative Umsetzung vor allem durch CYP1A2 katalysiert wird [31], sind synthetische Verbindungen, die bei zahlreichen industriellen Prozessen eingesetzt werden und

darüber hinaus auch in Pflanzenschutzmitteln zu finden sind und bei der Ölgewinnung vorkommen. Einige von ihnen (z.B. Aminobiphenyl und Aminofluoren) sind ebenfalls Bestandteile des Zigarettenrauchs [32].

Heterozyklische aromatische Amine (*heterocyclic aromatic amines*, HAA) entstehen beim starken Erhitzen von eiweißhaltigen Nahrungsmitteln, beispielsweise beim Kochen und Grillen von Fleisch, durch die Pyrolyse von Aminosäuren und werden somit ständig in geringen Konzentrationen mit der Nahrung aufgenommen. Sie besitzen ein deutlich höheres mutagenes Potential als beispielsweise Benzo[a]pyren [33]. An ihrer cytochromkatalysierten Aktivierung ist fast ausschließlich das CYP1A2 beteiligt [31], das diese Verbindungen mit einer hohen Umsatzrate metabolisiert.

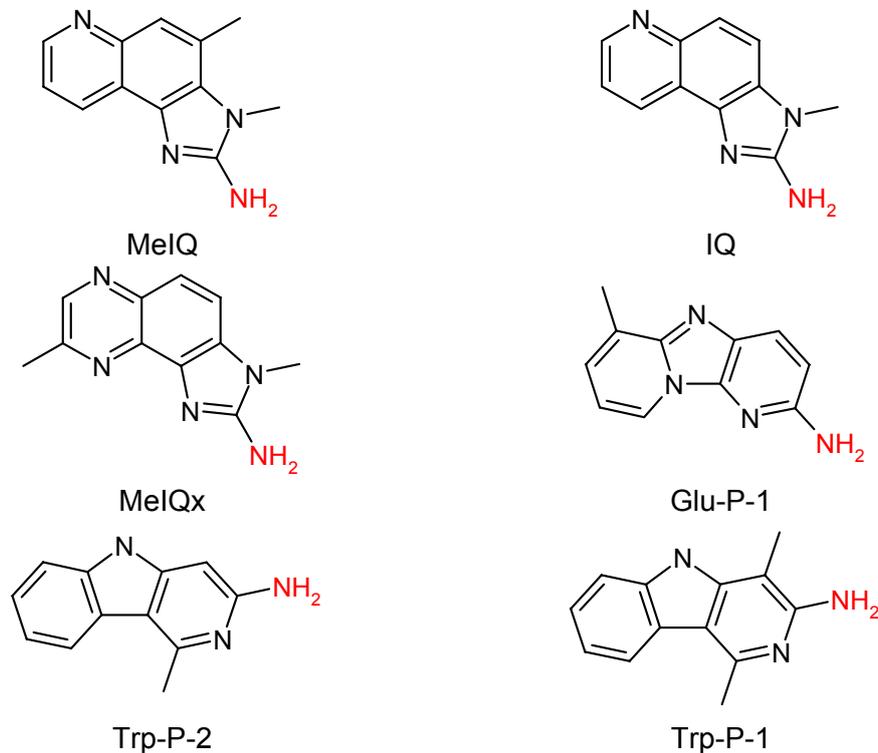


Abbildung 1.5: Mutagene heterozyklische Amine, die spezifisch durch CYP1A2 aktiviert werden. Gängige Abkürzungen für die IUPAC-Namen: MeIQ: (2-Amino-3,4-dimethylimidazo[4,5-f]chinolin), IQ: (2-Amino-3-methylimidazo[4,5-f]chinolin), MeIQx: (2-Amino-3,8-dimethylimidazo[4,5-f]chinoxalin), Glu-P-1: (2-Amino-6-methyl-5H-pyrido[1,2-a:3',2'-d]imidazol), Trp-P-2: (3-Amino-1-methyl-5H-pyrido[4,3-b]indol), Trp-P-1: (3-Amino-1,4-dimethyl-5H-pyrido[4,3-b]indol)

Von den in Abbildung 1.5 dargestellten HAA weist CYP1A2 gegenüber den Chinolin- und Chinoxalinderivaten die höchste Aktivität auf. Unter ihnen wiederum ist MeIQ die aktivste Verbindung. Aufgrund seiner hohen Aktivität und Spezifität gegenüber CYP1A2 wurde MeIQ gerade in letzter Zeit häufig in Mutationsstudien mit diesem Isoenzym eingesetzt [34], [35] und wurde deshalb auch in der vorliegenden Arbeit als repräsentatives Substrat für die Berechnung von Enzym-Substrat-Komplexen verwendet. Ein Datensatz von 12 HAA, darunter die in Abbildung 1.5 abgebildeten Verbindungen, wurde darüber hinaus bereits theoretisch in einer 3D-QSAR-Studie untersucht [36].

Die metabolische Aktivierung aller polyaromatischen und heterozyklischen Amine und Amide verläuft über eine von CYP1A2 katalysierte Oxidation des exozyklischen Amin- bzw. Amidstickstoffs zum Hydroxylamin bzw. zum Hydroxamsäurederivat als Initialisierungsreaktion. Der zweite Reaktionsschritt ist die O-Acetylierung oder O-Sulfatierung des Hydroxylamins mithilfe enzymatischer Esterasen. Bei der Spaltung dieser labilen Ester entstehen hochreaktive elektrophile Nitrenium-Ionen, die wiederum DNA-Addukte bilden können. Häufig findet die enzymatische Acetylierung oder Sulfatierung und damit die endgültige Aktivierung der Verbindungen erst in der Blase oder im Darm statt, weil die entsprechenden Enzyme dort in hohen Konzentrationen vorkommen. Das ist der Grund, weshalb diese Prokarzinogene vielfach mit der Entstehung von Blasen- und Dickdarmkrebs in Verbindung gebracht werden [29].

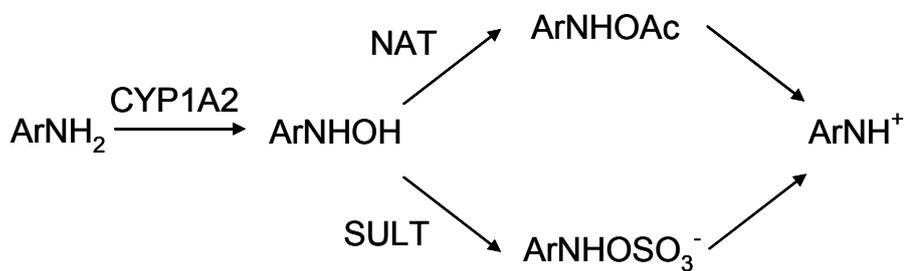


Abbildung 1.6: Schematischer Ablauf der Aktivierung polyaromatischer und heterozyklischer Amine (nach [29]). NAT: N-Acetyltransferase, SULT: Sulfotransferase.

Eines der stärksten bekannten Karzinogene, das Mycotoxin Aflatoxin B_1 (Abbildung 1.7) aus dem Schimmelpilz *Aspergillus flavus* zählt ebenfalls zu den Substraten von CYP1A2. In diesem Fall wird jedoch nicht die metabolische Aktivierung, sondern eine Entgiftungsreaktion, die in der 9a-Hydroxylierung des Toxins zum Aflatoxin M_1 besteht, von CYP1A2 katalysiert [37]. An der metabolischen Aktivierung des Aflatoxin B_1 in Form einer Epoxidierung der 8,9-Doppelbindung sind stattdessen die Isoenzyme CYP2C11 und CYP3A3 beteiligt.

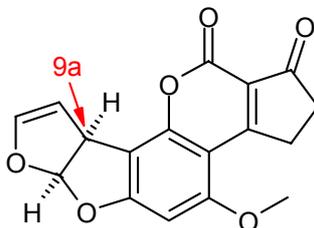


Abbildung 1.7: Strukturformel von Aflatoxin B_1 . Die CYP1A2-Hydroxylierungsstelle ist durch einen Pfeil gekennzeichnet.

1.2.1.2 Metabolisierung von 17 β -Estradiol

Ein weiteres gemeinsames Substrat aller drei CYP1-Isoenzyme ist neben den bereits angesprochenen Prokarzinogenen das 17 β -Estradiol, wobei die hier auftretende Metabolisierungsreaktion nicht als Bestandteil der Steroidhormonbiosynthese zu

verstehen ist. Es handelt sich vielmehr um eine Abbaureaktion, bzw. wiederum um eine metabolische Aktivierung, bei der die Cytochrome der beiden Unterfamilien 1A und 1B eine unterschiedliche Regiospezifität aufweisen (Abbildung 1.8). Während CYP1B1 das 17 β -Estradiol überwiegend zum 4-Hydroxyestradiol hydroxyliert, resultiert aus der Umsetzung mit CYP1A1 und CYP1A2 überwiegend das 2-Hydroxyestradiol [38], [39]. Der größte Teil beider Catecholestrogene wird durch die Catechol-O-Methyltransferase inaktiviert. Es kann jedoch auch unter erneuter Beteiligung von P450-Enzymen oder Peroxidasen zu einer weiteren Oxidation zu den entsprechenden Semichinonen und Chinonen kommen. Beide Catecholestrogen-Chinone sind in der Lage, DNA-Addukte zu bilden, sofern sie nicht durch eine Konjugation mit Glutathion inaktiviert werden.

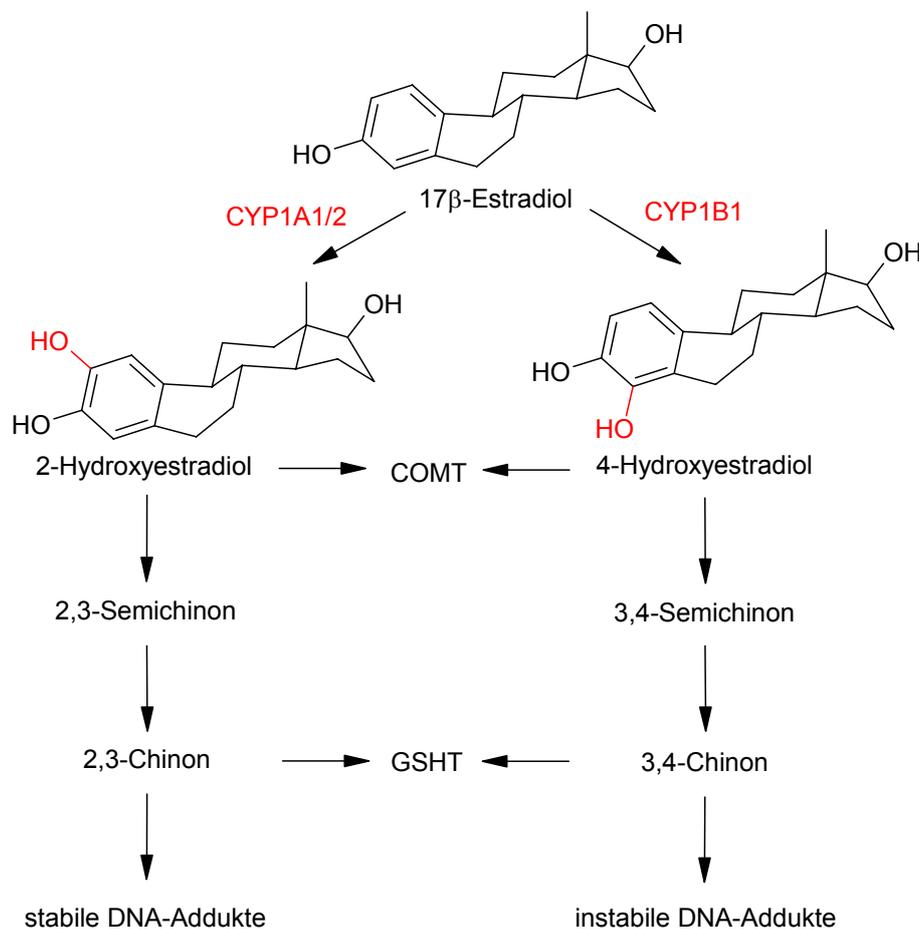


Abbildung 1.8: Metabolismus von 17 β -Estradiol durch CYP1A1/2 und CYP1B1 (nach [41]).
Abkürzungen: COMT: Catechol-O-methyltransferase; GSHT: Glutathion-S-transferase.

Der wesentliche Unterschied zwischen den beiden Reaktionswegen besteht darin, dass das 3,4-Chinon zu instabilen DNA-Addukten weiterreagiert, bei denen das Purin durch Spaltung der Glycosidbindung abgetrennt wird und eine Fehlstelle zurückbleibt [40], [41], während aus dem 2,3-Chinon stabile DNA-Addukte resultieren, die bis zu ihrer Reparatur in der DNA bestehen bleiben. Vermutlich sind die instabilen DNA-Addukte für das karzinogene Potential des 17 β -Estradiols verantwortlich, denn man weiß aus zahlreichen Experimenten, dass das 4-Hydroxyestradiol ein Prokarzinogen

ist, 2-Hydroxyestradiol dagegen nicht. Somit führt im Fall des 17β -Estradiols lediglich die Umsetzung mit CYP1B1 zu einer metabolischen Aktivierung, die Umsetzung mit CYP1A2 und CYP1A1 nicht.

1.2.1.3 Metabolismus von Arzneistoffen und Markersubstraten

Drei Arzneistoffe, die stets im Zusammenhang mit CYP1A2 genannt werden, sind das Paracetamol, das Tacrin und das Clozapin.

Das Analgetikum Paracetamol kann in einer cytochromkatalysierten Reaktion, deren Mechanismus nicht genau bekannt ist, zum N-Acetylbenzochinonimin oxidiert werden [42] (Abbildung 1.9). Während dieser reaktive Metabolit in therapeutischen Dosen in der Regel durch Kopplung mit Glutathion abgefangen wird, ist er bei einer Überdosierung für die Lebertoxizität des Arzneistoffes verantwortlich. Die Oxidationsreaktion wird neben CYP3A4 und CYP2E1 vor allem auch von CYP1A2 katalysiert.

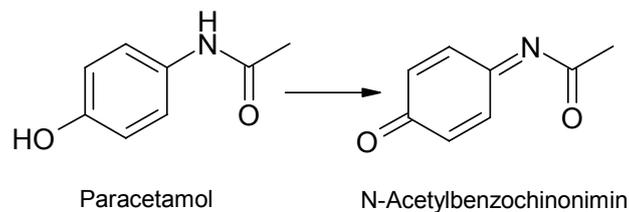


Abbildung 1.9: CYP-katalysierte Oxidation von Paracetamol

Tacrin ist ein zentral wirkender Cholinesteraseinhibitor, der zur Behandlung der Alzheimer-Erkrankung eingesetzt wird. Es wird von CYP1A2 entweder zu 1-, 2- oder 4-Hydroxytacrin abgebaut, oder zum 7-Hydroxytacrin umgesetzt, aus dem sich in einer weiteren CYP1A2-katalysierten Oxidation ebenfalls ein reaktives Chinonimin bilden kann (Abbildung 1.10). Tacrin unterliegt somit wiederum einer metabolischen Aktivierung durch CYP1A2, die ebenso für seine Lebertoxizität verantwortlich ist [43].

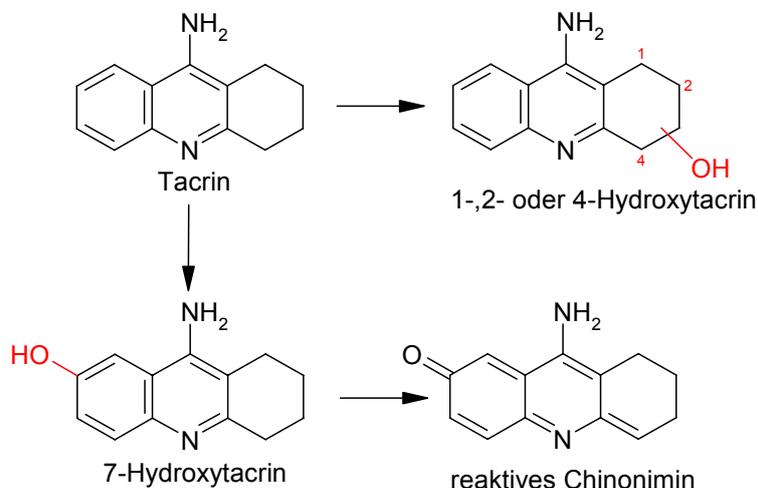


Abbildung 1.10: CYP1A2-katalysierter Metabolismus von Tacrin [43]. Die Hydroxylierungspositionen sind rot markiert.

Das atypische Neuroleptikum Clozapin (Abbildung 1.11) wird ebenfalls von CYP1A2 metabolisiert. Der Hauptmetabolit ist dabei das N-Desmethylclozapin. Die N-Oxidation, die bei der Umsetzung mit CYP3A4 die Hauptreaktion ist, stellt beim CYP1A2-katalysierten Abbau nur ein Nebenprodukt dar [44]. Wie bereits eingangs angedeutet (vgl. Seite 7), besteht besonders beim Clozapin die Gefahr von Arzneimittelinteraktionen. Bereits mehrfach wurden Intoxikationen bei gleichzeitiger Einnahme anderer CYP1A2-Substrate oder -inhibitoren beobachtet [45].

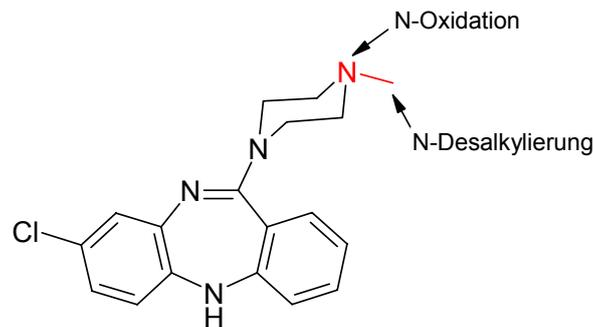


Abbildung 1.11: Metabolismus von Clozapin

Das Auftreten von Agranulozytosen, der gravierendsten Nebenwirkung von Clozapin, ist vermutlich konzentrationsunabhängig und wird folglich nicht von einem Anstieg der Clozapinkonzentration durch Hemmung von CYP1A2 beeinflusst. Das Risiko zentralnervöser Nebenwirkungen nimmt jedoch mit steigender Clozapinkonzentration offensichtlich zu [46].

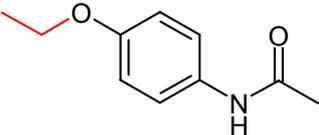
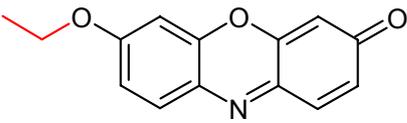
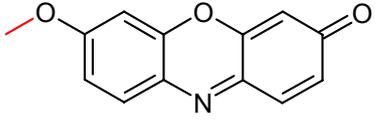
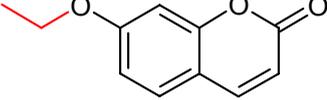
Im Folgenden werden weitere typische Substrate von CYP1A2 beschrieben, die eine Bedeutung als Arzneistoffe oder als Markersubstrate haben und in der vorliegenden Arbeit zusammen mit den beiden bereits in den Abschnitten 1.2.1.1 und 1.2.1.2 vorgestellten Verbindungen Aflatoxin B₁ und 17 β -Estradiol für die Validierung des Dockings verwendet wurden. Von einigen dieser Verbindungen wurden bereits in einer theoretischen Studie Enzym-Substrat-Komplexe generiert und spezifische Interaktionen mit CYP1A2 abgeleitet [47]. CYP1A2 katalysiert bei diesen Verbindungen hauptsächlich drei verschiedene Typen von Metabolisierungsreaktionen, und zwar O-Desalkylierungen, N-Desalkylierungen und aromatische Hydroxylierungen, nach denen sich die folgende Einteilung richtet.

Für das anfangs auch als Phenacetin-O-Deethylase bezeichnete CYP1A2 galt das Phenacetin (Tabelle 1.3) lange Zeit als Markersubstrat. Es wird durch O-Desalkylierung zum Paracetamol oxidiert. An seine Stelle sind mittlerweile die kurzkettigen 7-Alkoxyresorufine (7-Ethoxyresorufin und 7-Methoxyresorufin) getreten, die im EROD- bzw. MROD-Essay⁴ heute als spezifische und sensitive Marker-

⁴ EROD: *Ethoxy-Resorufin-O-Deethylase*; MROD: *Methoxy-Resorufin-O-Deethylase*

substrate für alle drei CYP1-Isoenzyme eingesetzt werden. Dabei weist CYP1A1 eine höhere Aktivität gegenüber 7-Ethoxyresorufin auf, während CYP1A2 das 7-Methoxyresorufin effektiver umsetzt [48].

Tabelle 1.3: Substrate für CYP1A2-katalysierte O-Desalkylierungsreaktionen. Die Alkylreste sind jeweils rot markiert.

Bezeichnung	Strukturformel	Referenz
Phenacetin		[32]
7-Ethoxyresorufin		[48]
7-Methoxyresorufin		[48]
7-Ethoxycumarin		[49]

Auch über ihre Verwendung als Markersubstrate hinaus können die Alkoxyresorufine als die am besten untersuchten und am umfangreichsten in der Literatur dokumentierten Substrate von CYP1A2 gelten. Es wurden zahlreiche Mutationsstudien mit Ethoxy- und Methoxyresorufin als Substrat veröffentlicht [34], [50], so dass das Ethoxyresorufin im Rahmen dieser Arbeit nicht nur zur Validierung des Dockings sondern auch als repräsentatives Substrat für die Berechnung von Enzym-Substrat-Komplexen ausgewählt wurde. Das 7-Ethoxycumarin diene vor allem für Mutationsstudien an Ratten-CYP1A2 (P-450_d) lange Zeit als Markersubstrat [51].

Bei den N-Desalkylierungsreaktionen ist Coffein (Abbildung 1.12) als grundlegendes Beispiel zu nennen. Coffein ist ein weiteres wichtiges Markersubstrat für Affinitätsmessungen am humanen CYP1A2 *in vitro* und auch *in vivo*, weil die N3-Demethylierung zum Paraxanthin, die beim Menschen den Hauptabbauweg des Coffeins darstellt, nahezu ausschließlich durch dieses Isoenzym katalysiert wird. Für die anderen möglichen Abbaureaktionen spielen eventuell auch andere Isoenzyme eine geringfügige Rolle [52], [53]. Drei der vier möglichen CYP1A2-katalysierten Abbaureaktionen de Coffeins sind N-Desalkylierungsreaktionen. Die vierte Reaktion, eine C-Hydroxylierung, spielt nur eine sehr untergeordnete Rolle. HPLC-Messungen von Coffeinmetaboliten nach Umsetzung von Coffein mit humanem CYP1A2 haben gezeigt, dass der Hauptmetabolit Paraxanthin zu 84 % gebildet wird. N1- und N7-Demethylierungen zum Theobromin bzw. zum Theophyllin treten mit 12 % bzw. 4 % deutlich seltener auf. Die Hydroxylierung am C8 liegt mit 1 % nur knapp oberhalb der

Erfassungsgrenze [52]. Ein vergleichbares Verhältnis der vier verschiedenen Coffeinmetabolite wurde auch in anderen Studien gefunden [53], [54]. Interessanterweise lässt es sich nicht auf alle Säugetierarten übertragen. Beispielsweise kommen bei der Ratte N1-Demethylierung und N3-Demethylierung etwa gleich häufig vor [52]. Am Tiermodell gewonnene experimentelle Daten sind also in diesem Fall keineswegs auf den Menschen anwendbar.

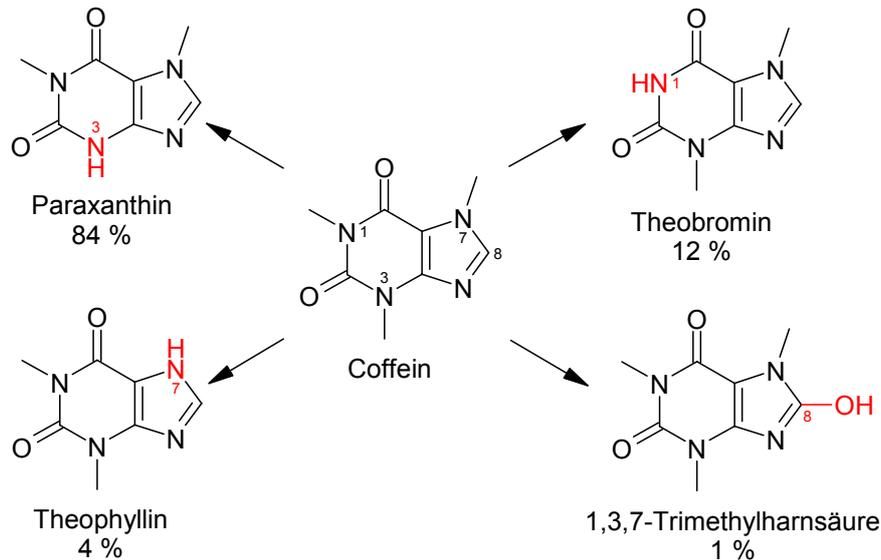


Abbildung 1.12: CYP1A2-katalysierter Metabolismus von Coffein [52]. Die Oxidationsstellen sind durch ihre Atomnummern gekennzeichnet und bei dem jeweiligen Metaboliten rot markiert.

Theophyllin, ein weiteres Xanthin, das in der pharmazeutischen Praxis bei der Behandlung von Patienten mit schwerem Asthma eine Rolle spielt, ist zwar eng mit dem Coffein verwandt, dürfte aber eigentlich nicht im Zusammenhang mit N-Desalkylierungsreaktionen genannt werden, da sie bei dieser Verbindung praktisch kaum vorkommen. Trotz der strukturellen Ähnlichkeit zum Coffein wird bei der Umsetzung von Theophyllin mit CYP1A2 fast ausschließlich 1,3-Dimethylharnsäure gebildet, das Produkt der C8-Hydroxylierung (Abbildung 1.13). Da das Theophyllin ein Arzneistoff mit enger therapeutischer Breite ist, d.h. dass bei erhöhten Plasmaspiegeln schwere Nebenwirkungen auftreten können, gehört es wie das Clozapin zu den CYP1A2-Substraten, die häufiger in klinisch bedeutsame Arzneimittelwechselwirkungen involviert sind [55].

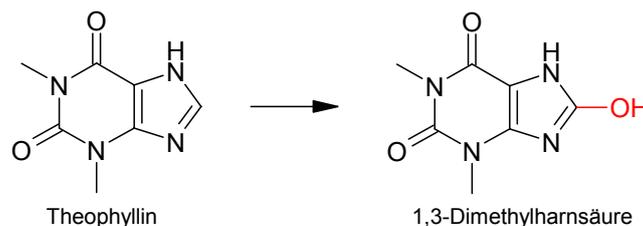
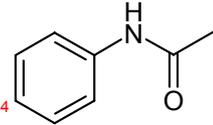
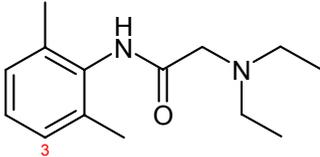
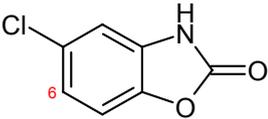
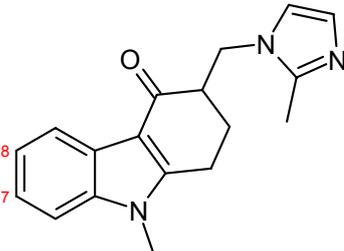


Abbildung 1.13: CYP1A2-katalysierter Metabolismus von Theophyllin [52]

Eine Reihe weiterer Arzneistoffe, die in Tabelle 1.4 dargestellt sind, wird in einer aromatischen Hydroxylierungsreaktion von CYP1A2 umgesetzt. Das Analgetikum Acetanilid beispielsweise wird vermutlich sogar ausschließlich durch eine CYP1A2-katalysierte 4-Hydroxylierung abgebaut. Am Metabolismus des Lokalanästhetikums Lidocain ist neben CYP1A2 auch das CYP3A4 beteiligt, allerdings mit einer unterschiedlichen Regiospezifität. Während bei der Umsetzung mit CYP1A2 das 3'-Hydroxylidocain entsteht, katalysiert CYP3A4 die N-Deethylierungsreaktion. Auch für den Metabolismus des zentral wirksamen Muskelrelaxans Chlorzoxazon ist mehr als ein Isoenzym von Bedeutung. Hier katalysieren CYP1A2 und CYP2E1 etwa gleichwertig die aromatische Hydroxylierung an C6. Der 5-HT₃-Rezeptorantagonist Ondansetron, der bei zytostatikainduziertem Erbrechen eingesetzt wird, wird durch eine Hydroxylierung des Indolrings an den Positionen 7 und 8 metabolisiert. Diese Reaktion wird von CYP1A2, von CYP3A4 und in geringerem Ausmaß auch von CYP2D6 katalysiert.

Tabelle 1.4: Substrate für CYP1A2-katalysierte aromatische Hydroxylierungen. Die Hydroxylierungsstellen sind durch die rote Nummerierung hervorgehoben.

Bezeichnung	Strukturformel	Referenz
Acetanilid		[56]
Lidocain		[57]
Chlorzoxazon		[58]
Ondansetron		[59]

Wie bereits bei den letzten Beispielen deutlich wurde, ist das Isoenzym CYP1A2 am Metabolismus vieler weiterer Arzneistoffe beteiligt, für die es aber nicht den Hauptabbauweg darstellt. Im Einzelnen können sich daraus trotzdem klinisch relevante Arzneimittelinteraktionen ergeben. Als weitere Beispiele seien die N-Desalkylierung von Propafenon, Verapamil und trizyklischen Antidepressiva, die O-Desalkylierung von Naproxen und die aromatische Hydroxylierung von Terbinafin und Warfarin genannt. Im Zusammenhang mit Warfarin sollte die Bedeutung der Stereochemie für den cytochromkatalysierten Abbau erwähnt werden. Das

therapeutisch wirksame (*S*)-Warfarin wird überwiegend von CYP2C9 metabolisiert und das inaktive Enantiomer (*R*)-Warfarin ausschließlich von CYP1A2 [60].

1.2.2 Regulation der Enzymaktivität

Im Fall des CYP1A2 unterliegen sowohl die metabolische Aktivierung verschiedener Karzinogene als auch die durch Antikörperbindung nachweisbare Enzymmenge einer erheblichen interindividuellen Variabilität. Dies ist zum einen auf die Regulation der *CYP1A2*⁵-Genexpression durch unterschiedliche Faktoren zurückzuführen, was unter Umständen einen Einfluss auf die individuell unterschiedliche Anfälligkeit für arylamininduzierte Krebserkrankungen haben kann. Ein signifikanter Zusammenhang besteht beispielsweise zwischen einem Anstieg der CYP1A2-Aktivität und Rauchen [32]. Zum anderen wirkt sich natürlich auch die Bindung spezifischer oder unspezifischer Inhibitoren, die aufgrund verschiedener irreversibler und reversibler Mechanismen wirken können, auf die katalytische Aktivität des Isoenzym aus.

1.2.2.1 Regulation der Genexpression

Induzierbar sind alle CYP-Isoenzyme der Familie 1 über eine Steigerung der Genexpression, die durch den Ah-Rezeptor (*aryl hydrocarbon receptor*) vermittelt wird [61], [62]. Die Bindung eines Liganden an diesen Rezeptor initiiert die Translokation des Ah-Rezeptor-Ligand-Komplexes in den Zellkern, der dort wiederum durch Bindung an regulatorische Genregionen die Transkription von *CYP1*-Genen aktiviert und dadurch schließlich den CYP1-Spiegel in der Mikrosomenmembran steigert [63].

Zu den Liganden des Ah-Rezeptors zählen die im Zigarettenrauch enthaltenen PAH, was den oben erwähnten Zusammenhang erklärt, sowie halogenierte aromatische Kohlenwasserstoffe (*halogenated aromatic hydrocarbons*, HAH). 2,3,7,8-Tetrachloro-dibenzo-p-dioxin (TCDD, Abbildung 1.14), ein Vertreter der so genannten „Dioxine“, die bei der Herstellung von Chlorverbindungen entstehen, und wie die PAH an Staub- und Bodenpartikel gebunden in die Umwelt geraten [64], ist der Ligand mit der höchsten Affinität zum Ah-Rezeptor. Er ist selbst nicht genotoxisch, gilt aber als hoch wirksames Karzinogen, da er z.B. die stärkste Induktion von CYP1A1 hervorruft, was dann wiederum eine metabolische Aktivierung anderer Prokarzinogene zur Folge hat. Dagegen sind einige Ah-Rezeptorliganden wie z.B. das Benzo[a]pyren gleichzeitig Substrate von CYP1 und induzieren damit praktisch ihre eigene metabolische Aktivierung.

⁵ Zur Unterscheidung von der Proteinbezeichnung wird die entsprechende Genbezeichnung üblicherweise kursiv gedruckt

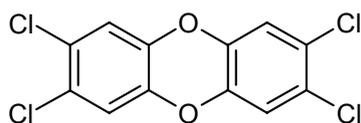


Abbildung 1.14: Strukturformel von TCDD

1.2.2.2 Inhibitoren

Furafyllin (Abbildung 1.15), ein Methylxanthinderivat, das wie Theophyllin in der Asthmatherapie eingesetzt wird und eine enge strukturelle Verwandtschaft zu den beiden bereits erwähnten Xanthinderivaten Coffein und Theophyllin aufweist, ist ein potenter und hochselektiver Inhibitor des humanen CYP1A2. Humanes CYP1A1, das ja aus derselben Unterfamilie stammt, wird durch Furafyllin nicht inhibiert [65].

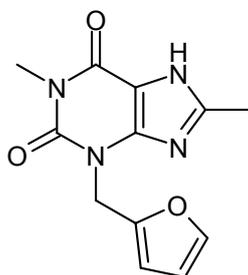
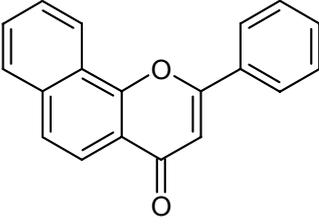
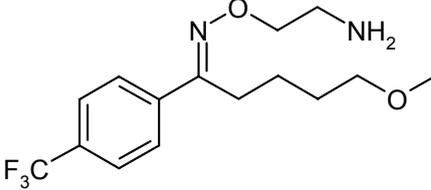


Abbildung 1.15: Strukturformel von Furafyllin

Die Substanz gehört zu den Suizidinhibitoren, das heißt, sie wird von CYP1A2 zu einem reaktiven Metaboliten umgesetzt, der so fest und spezifisch an das Enzym bindet, dass es irreversibel seine Funktionsfähigkeit verliert. Beim Furafyllin stellt die Oxidation der Methylgruppe am C8 die Aktivierungsreaktion dar [66].

Zwei weitere wichtige reversible Inhibitoren des CYP1A2-Isoenzym sind in Tabelle 1.5 dargestellt. Während man beim α -Naphthoflavon vermutet, dass auch andere reversible Inhibierungsmechanismen eine Rolle spielen, ist der selektive Serotonin-Reuptake-Inhibitor (SSRI) Fluvoxamin ein rein kompetitiver Inhibitor. Bezüglich des α -Naphthoflavon geht man davon aus, dass es in einer hydrophoben Tasche nahe dem aktiven Zentrum bindet und dort die Umsetzung durch unterschiedliche vom jeweiligen Substrat abhängige Inhibierungsmechanismen hemmen kann [67]. Neuerdings gibt es Bestrebungen, die inhibitorische Potenz des Fluvoxamins in der antipsychotischen Therapie auszunutzen. So stellte sich in einer kürzlich veröffentlichten *in vivo* Studie heraus, dass man bei Rauchern, die über erhöhte CYP1A2-Plasmaspiegel verfügen (vgl. Abschnitt 1.2.2.1), die Dosis von Olanzapin (vgl. Abschnitt 1.2.1.3) um bis zu 26 % senken kann, wenn man gleichzeitig niedrige subtherapeutische Dosen von Fluvoxamin verabreicht [68].

Tabelle 1.5: Wichtige reversible Inhibitoren von CYP1A2

Bezeichnung	Strukturformel	Referenz
α -Naphthoflavon		[69]
Fluvoxamin		[70]

Die Chinoloncarbonsäuren stellen eine Substanzklasse dar, die als potente kompetitive Inhibitoren des CYP-Isoenzym 1A2 wirken. Bereits früh wurde *in vivo* bei gleichzeitiger Gabe von Theophyllin ein Zusammenhang zwischen der Einnahme von Gyrasehemmern und erhöhten Theophyllin-Plasmaspiegeln festgestellt [71]. Von der in Abbildung 1.16 dargestellten Auswahl an Verbindungen besitzen die Pipemidsäure, das Ciprofloxacin und das Enoxacin die stärkste Inhibitorwirkung [72].

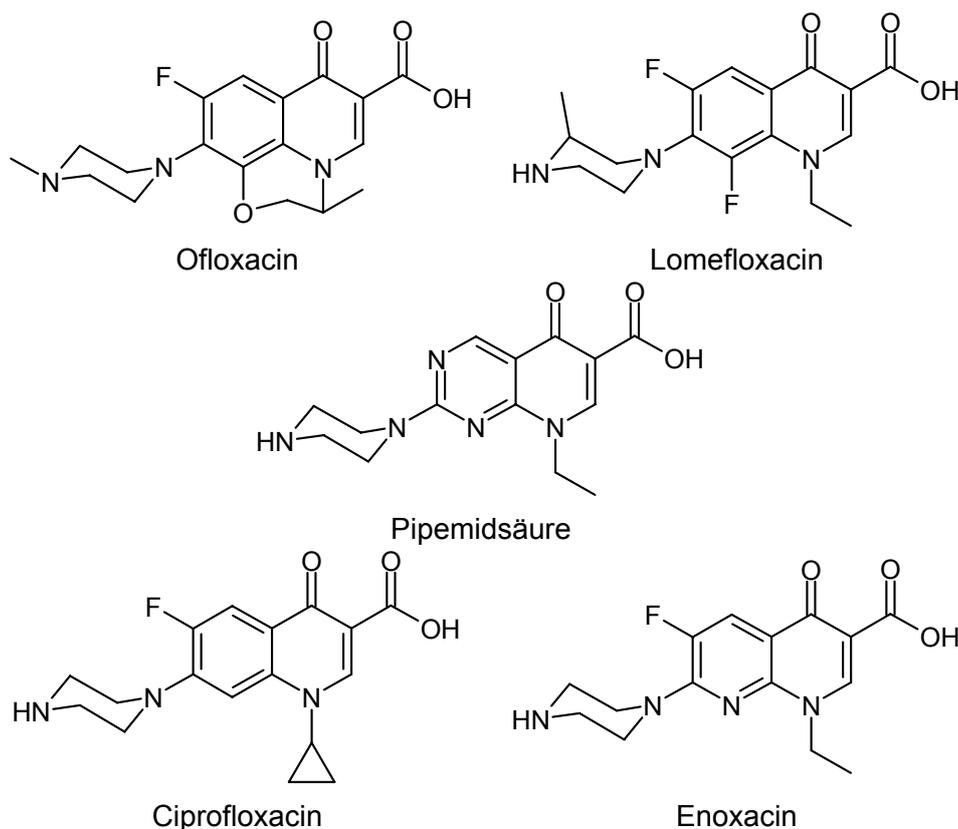


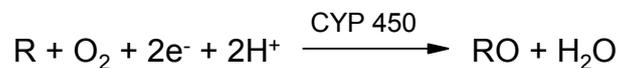
Abbildung 1.16: Chinoloncarbonsäuren als kompetitive CYP1A2-Inhibitoren [72]

Auf der Grundlage von insgesamt 44 Derivaten der 4-Oxochinolin-3-Carbonsäure leiteten Fuhr et al. 1993 Strukturwirkungsbeziehungen für die Inhibierung von CYP1A2 durch diese Verbindungen ab. Demnach setzen Alkylsubstituenten am

Piperazinring die Inhibitorwirkung herab und ein Naphthyridingrundgerüst anstatt eines Chinolins führt zu einer verstärkten inhibitorischen Aktivität. Der Piperazinring ist allerdings nicht essentiell für eine inhibitorische Wirkung. Fluorsubstituenten am Ringgerüst wirken sich kaum auf die Inhibitoraktivität aus. Darüber hinaus stellten Fuhr et al. anhand eines Vergleichs der *Molekularen Elektrostatischen Potentiale* (MEP) ein Pharmakophormodell auf. Bei einer Überlagerung der Chinolonderivate mit Coffein ergaben sich gemeinsame Potentialfelder im Bereich der Carbonylgruppe, der Carboxylatgruppe und des Chinolinstickstoffs, die für die Inhibitorbindung von Bedeutung waren. [73].

1.3 Katalytischer Mechanismus der Cytochrome P450

Die allgemeine Reaktionsgleichung für die von Cytochrom P450 katalysierten Oxidationsreaktionen lässt sich folgendermaßen formulieren:



Ein Substrat R wird durch Einführung eines Sauerstoffatoms aus molekularem Sauerstoff zu RO oxidiert, während das zweite Sauerstoffatom zu Wasser reduziert wird. Auf diese Weise katalysieren die Cytochrome P450 wie in Abschnitt 1.2.1 dargestellt Hydroxylierungen von CH-Bindungen, Epoxidierungen von Doppelbindungen, Oxidationen und Desalkylierungen von Heteroatomen und viele weitere Reaktionen. Die beiden für die Reaktion benötigten Elektronen liefern NADH bzw. NADPH über verschiedene Elektronentransfersysteme, nach denen man die Cytochrome in unterschiedliche Klassen einteilen kann. Säugetiercytochrome der Mitochondrienmembran und Bakterien-cytochrome gehören zur Klasse I, bei der ein einzelnes Elektron über eine FAD-abhängige Reduktase auf ein Eisen-Schwefel-Redoxin und von dort weiter auf das Cytochrom übertragen wird. Bei der Klasse II, zu der die mikrosomalen Säugetiercytochrome (und das bakterielle CYP102) gehören, verläuft der Elektronentransport über ein FAD und FMN enthaltendes Flavoprotein, die NADPH-Cytochrom-P450-Reduktase, bei der die FAD-Untereinheit das Elektron von NADPH übernimmt und die FMN-Untereinheit das Elektron an das Cytochrom weitergibt [74]. Als Quelle für das zweite Elektron kann bei den Säugetiercytochromen auch Cytochrom b_5 dienen.

1.3.1 Allgemeiner Reaktionszyklus

Der genaue Ablauf des Reaktionszyklus cytochromkatalysierter Reaktionen ist bis heute ungeklärt, da einige Zwischenstufen schnell zerfallen oder weiterreagieren und experimentell äußerst schwer zu erfassen sind. Wertvolle Erkenntnisse wurden aus spektroskopischen und kristallographischen Untersuchungen bei sehr tiefen Temperaturen (77-100 K) gewonnen, die anhand der Hydroxylierungsreaktion von Campher durch CYP101 ($P450_{cam}$), einem Cytochrom aus dem Bakterium *Pseudomonas putida*, durchgeführt wurden [75], [76]. CYP101 ist das am besten untersuchte Isoenzym der P450-Superfamilie, dessen dreidimensionale Struktur auch als erstes aufgeklärt wurde. Abbildung 1.17 gibt den Reaktionszyklus nach dem derzeitigen Kenntnisstand wieder.

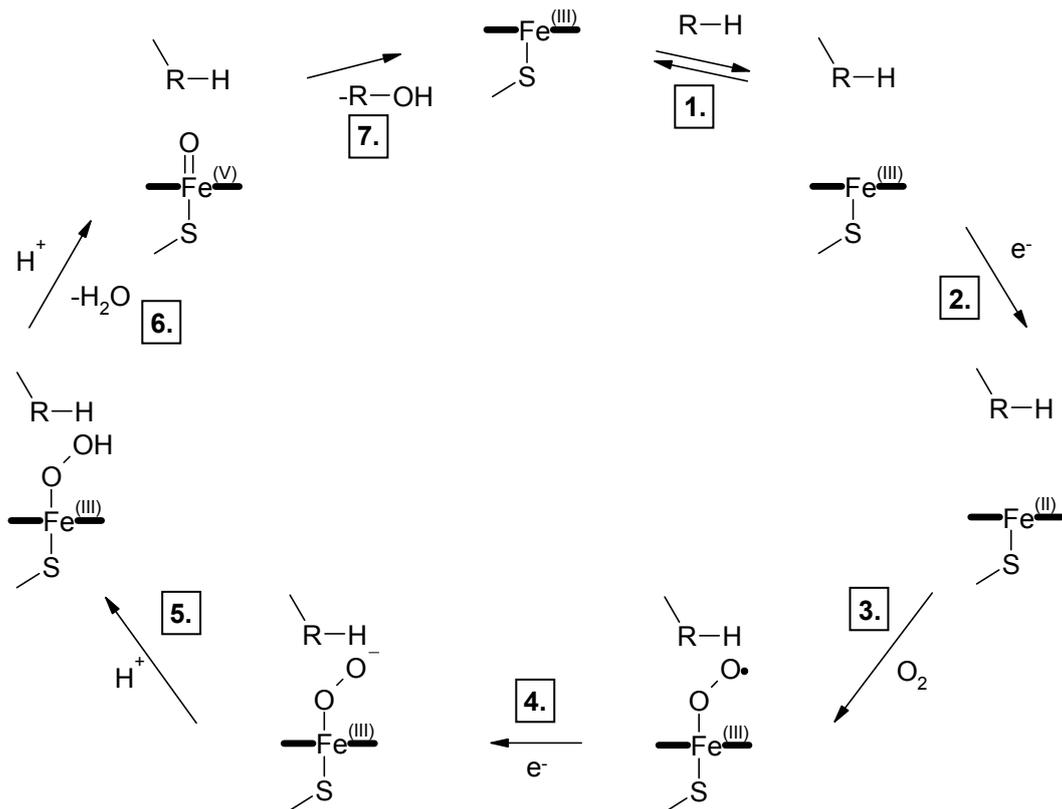


Abbildung 1.17: Schematische Darstellung des Reaktionszyklus cytochromkatalysierter Oxidationen (nach [79]). Der Porphyrinring ist durch einen schwarzen Balken angedeutet. – R – H symbolisiert das Substrat und – R – OH das Reaktionsprodukt.

1. Der erste Schritt im Reaktionszyklus ist die Bindung des Substrats, die mit einer Verdrängung von Wassermolekülen aus dem aktiven Zentrum einhergeht, und dadurch einen Entropiegewinn für das System darstellt. Die Substratbindung ist somit ein entropiegesteuerter Vorgang.

Bei vielen CYP-Isoenzymen liegt das Hämeisen im substratfreien Zustand in der *low-spin*-Konfiguration vor, wobei ein Wassermolekül als sechster Ligand an das Eisen gebunden ist. Auch dieses Wassermolekül wird in der Regel während der Substratbindung verdrängt und das Hämeisen geht in die *high-spin*-Konfiguration über, bei der aus sterischen Gründen nur fünf Koordinationsstellen besetzt sind. Diese Spinänderung lässt sich anhand spektroskopischer Messungen verfolgen, man erhält ein charakteristisches Typ-I-Differenzspektrum, bei dem sich die Absorptionsbande von etwa 418 nm (*low-spin*-Eisen) nach etwa 390 nm (*high-spin*-Eisen) verschiebt.

Das Isoenzym CYP1A2 bildet hier eine Ausnahme. Man beobachtete in spektroskopischen Studien, dass das Hämeisen des humanen CYP1A2 bereits im substratfreien Zustand fast ausschließlich in der *high-spin*-Konfiguration vorliegt [77]. Im Vergleich dazu liegt das Eisen im CYP1A2 der Ratte zu zwei Drittel im *high-spin*-Zustand vor, während das Eisen im CYP1A2 des Kaninchens vollständig die *low-spin*-Konfiguration aufweist. Die Autoren vermuteten daraufhin,

dass das aktive Zentrum des humanen CYP1A2 zu eng und zu lipophil ist, um die Bindung von Wasser und damit den *low-spin*-Zustand zu ermöglichen. Kombinierte quanten- und molekülmechanische Berechnungen mit der QM/MM-Methode haben in der Tat bestätigt, dass der Spinzustand eines substratfreien Isoenzym durch das jeweilige elektrische Feld des Proteins bestimmt wird [78].

Andererseits kann die Bindung eines Substrates an ein CYP-Isoenzym auch ein reverses Typ-I-Spektrum erzeugen, das heißt, eine verstärkte Absorption bei etwa 418 nm, was bedeuten würde, dass das Eisen trotz Substratbindung überwiegend in der *low-spin*-Konfiguration vorliegt. Eine mögliche Erklärung dafür ist entweder, dass das Substrat das an das Hämeisen gebundene Wasser nicht verdrängt, oder dass ein Sauerstoff- oder Stickstoffatom des Substrates mit dem Hämeisen interagiert [11].

2. Im zweiten Schritt wird ein Elektron von der Reduktase auf das Hämeisen übertragen, wodurch das Eisen(III) zu Eisen(II) reduziert wird. Da der Ionenradius des *high-spin*-Eisens größer ist als der des *low-spin*-Eisens, wird das Eisen leicht aus der Hämebene gehoben, was wiederum einen Einfluss auf das Redoxpotential hat und die Elektronenübertragung erleichtert [79].
3. Der dritte Schritt ist die Bindung eines Sauerstoffmoleküls als sechster Ligand des Hämeisens, wodurch das Eisen in den *low-spin*-Zustand übergeht. Da das Eisen in der Hämumgebung aufgrund seiner Oxidationszahl (II) und der *high-spin*-Konfiguration eine hohe Affinität zu molekularem Sauerstoff aufweist, läuft dieser Reaktionsschritt sehr schnell ab. Vermutlich bildet sich hier ein Gleichgewicht zwischen Fe(II)O_2 und dem reaktiveren $\text{Fe(III)O}_2^\bullet$. Der zuletzt genannte Eisen-Sauerstoff-Komplex kann in einer so genannten Entkopplungsreaktion unter Freisetzung von Wasserstoffperoxid zerfallen und den Reaktionszyklus unterbrechen.
4. Die Übertragung des zweiten Elektrons verläuft sehr langsam und stellt für viele cytochromkatalysierte Reaktionen den geschwindigkeitsbestimmenden Schritt dar. Auch hier existiert möglicherweise ein Gleichgewicht zwischen Fe(II)O_2^\bullet und der reaktiveren negativ geladenen Verbindung Fe(III)O_2^- . Mit Hilfe dieser reaktiven Eisen-Peroxo-Verbindungen lassen sich einige cytochromkatalysierte Reaktionen erklären, bei denen ein nukleophiles Agens benötigt wird, wie z.B. der letzte Schritt der Aromatasereaktion im Rahmen der Biosynthese des Estrons durch CYP17A1 [80].
5. Die Protonierung der Peroxo-Verbindung zu einer Hydroperoxo-Verbindung folgt sehr schnell, wobei auch die Hydroperoxo-Zwischenstufe noch eine gewisse Nukleophilie aufweist und bei nukleophilen Additionen eine Rolle spielen könnte (s. 4.).

6. Aus einer zweiten sich ebenfalls schnell anschließenden Protonierungsreaktion resultiert unter Abspaltung von Wasser ein weiteres reaktives Zwischenprodukt, das das oxidierende Agens der meisten cytochromkatalysierten Reaktionen darstellt. Aufgrund der hohen Geschwindigkeit der Protonierungsreaktionen und der starken Reaktivität des entstehenden Zwischenproduktes konnte dieses bisher noch nicht eindeutig charakterisiert werden und seine genaue Struktur ist unbekannt. Während man früher vielfach ein Eisen(III)-Oxo als elektrophile Eisen-Sauerstoff-Verbindung vermutete, wird heute überwiegend in Analogie zur Häm-Peroxidase, bei der die Struktur der entsprechenden reaktiven Zwischenstufe (*Compound I*) bereits aufgeklärt ist, eine elektrophile Eisen(V)-Oxo-Verbindung postuliert. Die Bezeichnung „Eisen(V)“ ist dabei eine Vereinfachung, da das Eisen formal eine Ladung von +4 besitzt und eine weitere positive Ladung über den Porphyrinring verteilt ist, sodass der Porphyrinring ein Radikalkation darstellt. Eisen und Sauerstoff sind über eine Doppelbindung verbunden, die der im molekularen Sauerstoff ähnelt [79].
7. Im letzten Reaktionsschritt wird das Sauerstoffatom aus dem reaktiven Eisen-Oxo-Komplex in das Substrat eingeführt und das Reaktionsprodukt dissoziiert vom Enzym. Dabei gibt es für die verschiedenen von Cytochromen P450 katalysierten Reaktionen unterschiedliche Vorstellungen zum genauen Mechanismus des Reaktionsablaufes. Die Reaktionen, die für CYP1A2 eine Rolle spielen (vgl. Abschnitt 1.2.1), werden im Folgenden kurz dargestellt.

1.3.2 Mechanismus CYP1A2-katalysierter Reaktionen

1.3.2.1 Aliphatische Hydroxylierung

Vermutlich läuft die aliphatische Hydroxylierung über den so genannten „*oxygen rebound*“ Mechanismus ab. Das bedeutet, dass keine konzertierte Reaktion stattfindet, sondern dass zuerst ein Wasserstoffatom vom Substrat auf die Eisen-Oxo-Verbindung übertragen wird, und dass sich dann in einem zweiten Schritt eine Bindung zwischen dem Substrat, das nun als Radikalkation vorliegt, und dem an das Eisen gebundenen Sauerstoff neu ausbildet (Abbildung 1.18). Wichtig ist hierbei, dass man mit diesem Reaktionsablauf erklären kann, warum aliphatische Hydroxylierungen nicht immer stereospezifisch verlaufen. In der radikalischen Zwischenstufe kann leicht eine Inversion der Konfiguration des Radikalkations stattfinden.

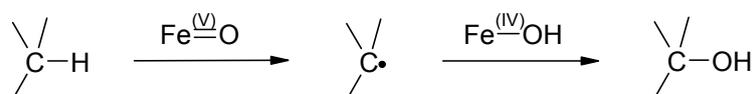


Abbildung 1.18: Mechanismus der aliphatischen Hydroxylierung (nach [79])

1.3.2.2 Aromatische Oxidation

Die Vorstellung vom Ablauf aromatischer Oxidationen basiert auf Experimenten mit deuteriertem Benzen und Warfarin, bei denen eine 1,2-Migration des Deuteriums festgestellt wurde. Demnach findet hier keine Wasserstoffübertragung statt, sondern die Eisen-Oxo-Verbindung bindet an den Aromaten. Durch die Übertragung eines Elektrons vom radikalischen Zwischenprodukt der aromatischen Verbindung auf das Hämeisen entsteht ein Kation, das die Migration des Deuteriums auslöst (Abbildung 1.19).

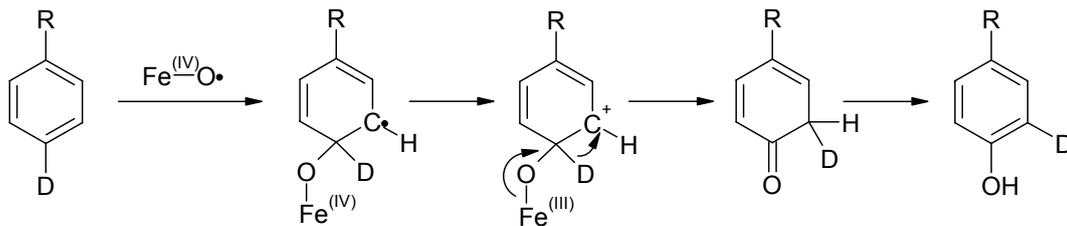


Abbildung 1.19: Mechanismus der aromatischen Oxidation (nach [79])

Für die Oxidation der PAH wird ein sehr ähnlicher Mechanismus vermutet, der ebenfalls über ein kationisches Zwischenprodukt verläuft. Da bei der Epoxidierung von PAH allerdings in der Regel die Stereochemie erhalten bleibt, könnte es andererseits auch sein, dass diese Reaktion eher konzertiert abläuft.

1.3.2.3 N-Oxidation

Für die N-Oxidation beispielsweise von aromatischen Aminen wird eine schrittweise Elektronenübertragung postuliert, bei der zuerst aus der Stickstoffverbindung ein Radikal entsteht und anschließend die Eisen-Sauerstoff-Bindung homolytisch gespalten wird.



Abbildung 1.20: Mechanismus der N-Oxidation (nach [79])

1.3.2.4 N- und O-Desalkylierungen

Zum Ablauf von N-Desalkylierungsreaktionen sind grundsätzlich zwei Mechanismen denkbar, die unter Umständen vom jeweiligen Substrat oder Isoenzym abhängen. Beide sind in Abbildung 1.21 dargestellt. Die erste Möglichkeit verläuft in Analogie zur aliphatischen Hydroxylierung über die Übertragung eines Wasserstoffatoms vom Alkylrest des Stickstoffs auf den Eisen-Oxo-Komplex. Daraufhin findet wiederum ein „oxygen rebound“ statt, das heißt, dass durch die Bindung des aktiven Sauerstoffs an den Alkylrest des Stickstoffs ein Carbinolamin entsteht, das dann weiter zu Amin und Aldehyd zerfällt. Denkbar ist hier aber auch als zweite Möglichkeit, dass wie bei der N-Oxidation zunächst ein Elektron vom Stickstoff auf die Eisen-Sauerstoff-

Verbindung übertragen wird und sich durch Übertragung eines weiteren Elektrons und eines Protons ein Iminiumion als Zwischenprodukt bildet. Erst daraus würde dann durch Bindung des reaktiven Sauerstoffs an den Alkylrest das Carbinolamin entstehen. Das Substrat müsste hierbei folglich so im aktiven Zentrum plaziert sein, dass eine Elektronenübertragung auf den Stickstoff als erster Reaktionsschritt stattfinden könnte.

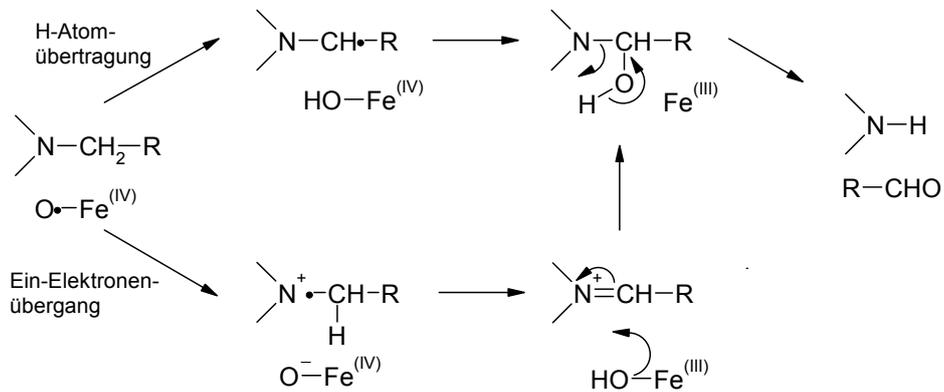


Abbildung 1.21: Zwei mögliche Mechanismen der N-Desalkylierung (nach [79])

Weitgehend einig ist man sich darüber, dass die O-Desalkylierung von Etherverbindungen über den oberen Reaktionsweg in Abbildung 1.21 stattfindet, also über die Übertragung eines Wasserstoffatoms auf den Eisen-Oxo-Komplex und anschließenden „oxygen rebound“.

1.4 Proteinstruktur der Cytochrome P450

1.4.1 Kristallstrukturen von Säugetiercytochromen

Bis vor wenigen Jahren standen ausschließlich Kristallstrukturen von nicht membrangebundenen, löslichen Bakteriencytochromen als Template für Homologie-Modelle von menschlichen Cytochromen zur Verfügung. Die Kristallisation eines normalerweise membrangebundenen Säugetiercytochroms gelang erstmalig im Jahr 2000, als die Struktur des Kaninchencytochroms CYP2C5 veröffentlicht wurde [81] (PDB-Code 1DT6). Um die Kristallisation zu ermöglichen, wurde der N-terminale Membrananker entfernt und es wurde eine vier Aminosäuren lange Histidinsequenz an den C-Terminus angefügt, damit das Protein per Affinitätschromatographie aufgereinigt werden konnte. Man vermutete, dass der Bereich zwischen den Helices F und G auch an der Membranbindung beteiligt ist. Deshalb wurden weiterhin einige Aminosäuren in diesem Bereich gegen die entsprechenden Aminosäuren aus CYP2C3 ausgetauscht, was die Löslichkeit des Enzyms verbesserte und die Aggregation von Monomeren verhinderte [82]. Man sollte hierbei aber stets berücksichtigen, dass sich derartige Mutationen, auch wenn sie außerhalb des aktiven Zentrums liegen, durchaus auf die allgemeine Konformation des Proteins auswirken können. Die Aufklärung dieser Struktur brachte einerseits neue wesentliche Erkenntnisse über die strukturellen Merkmale der mikrosomalen Cytochrome P450 und andererseits einen entscheidenden Fortschritt für die Modellierung von menschlichen Cytochromen, da die Sequenzidentität der Säugetiercytochrome untereinander deutlich höher ist als die Sequenzidentität zu den Bakteriencytochromen und damit deutlich bessere Modelle erstellt werden können. Der Nachteil dieser Kristallstruktur bestand jedoch darin, dass Helix B'' aufgrund der schlechten Auflösung der Struktur im Bereich zwischen den Helices B und C nicht als Helix zu erkennen war, und dass der Loop zwischen den Helices F und G ganz fehlte, der, wie man heute weiß, ebenfalls zwei kurze Helices enthält, die das aktive Zentrum wie ein Deckel abschließen. Dadurch eignete sich diese Struktur zwar für Docking-Studien und andere Untersuchungen an einer starren Proteinstruktur, erwies sich aber für Moleküldynamiksimulationen (MDS, s. Abschnitt 3.4.3) als relativ ungünstig.

Erst seit 2003 sind weitere Kristallstrukturen von Säugetiercytochromen veröffentlicht worden. Dies sind unter anderem Strukturen mit höherer Auflösung, die ersten menschlichen Cytochromstrukturen, sowie eine Struktur in der offenen Konformation. Wichtig für die Durchführung von MDS von Protein-Substrat-Interaktionen ist jedoch, dass nun auch – abgesehen vom Membrananker – Strukturen vollständiger Enzyme im substratgebundenen Zustand vorliegen. Tabelle 1.6 gibt einen Überblick über alle bisher veröffentlichten Kristallstrukturen von Säugetiercytochromen.

Tabelle 1.6: Bisher veröffentlichte Kristallstrukturen von Säugetiercytochromen

Organismus	Name	PDB	Jahr	Auflösung (Å)	Ligand	Einordnung
Kaninchen	CYP2B4	1PO5	2003	1.60	-	Offene Konformation
		1SUO	2004	1.90	Chlorophenyl- imidazol	Koordinative Bindung an das Hämeisen
	CYP2C5	1DT6	2000	3.00	-	Helix B'' und Loop FG fehlen
		1N6B	2003	2.30	DMZ	Aliphatische Hydroxylierung
		1NR6	2003	2.10	Diclofenac	Aromatische Hydroxylierung
Mensch	CYP2C8	1PQ2	2003	2.70	Palmitinsäure	Substratbindung an den Loop FG
		CYP2C9	1OG2	2003	2.60	-
	1OG5		2003	2.55	Warfarin	„Allosterische Bindungsstelle“
	1R9O		2004	2.00	Flurbiprofen	Aromatische Hydroxylierung
	CYP3A4	1W0E	2004	2.80	-	Geschlossene Konformation
		1W0F	2004	2.65	Progesteron	Bindung an den Loop FG
		1W0G	2004	2.74	Metyrapon	Koordinative Bindung an das Hämeisen
		1TQN	2004	2.05	-	Wasser als sechster Hämiligand

1.4.2 Allgemeine Tertiärstruktur

Seit die Anzahl an Kristallstrukturen von P450-Isoenzymen in letzter Zeit stark angestiegen ist, weiß man aus Vergleichen der unterschiedlichen Strukturen von Säugetieren und Bakterien, dass die allgemeine Tertiärstruktur über die gesamte Superfamilie konserviert ist, auch wenn die Sequenzidentität der einzelnen Isoenzyme untereinander sehr gering ist. Fast alle Sekundärstrukturelemente sind konserviert, unterscheiden sich in ihrer Länge und Ausrichtung aber mitunter deutlich voneinander. Grundsätzlich kann man feststellen, dass die N-terminale Hälfte der Isoenzyme bis einschließlich der Helix H wesentlich geringer konserviert ist als die C-terminale Hälfte [83]. Wie am Beispiel von CYP2C5 [84] (PDB-Code 1N6B) in Abbildung 1.22 zu erkennen ist, setzen sich alle P450-Isoenzyme aus einer α -Helixdomäne auf der einen Seite und einer β -Faltblattdomäne auf der anderen Seite zusammen.

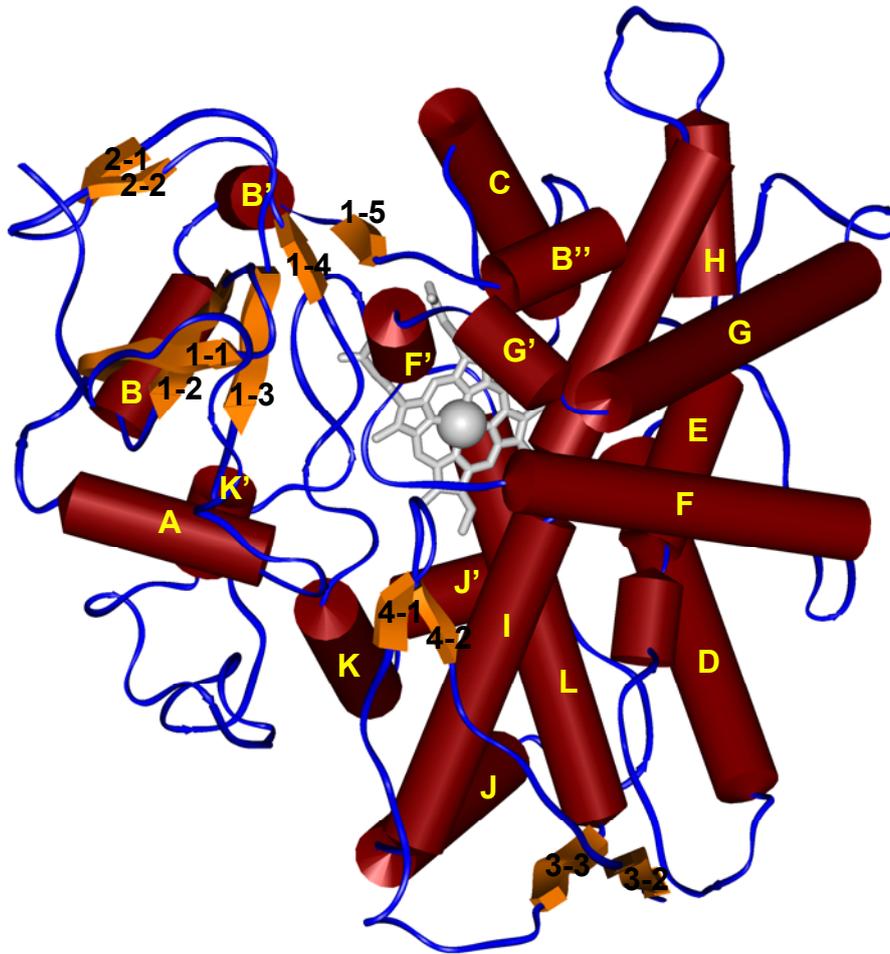


Abbildung 1.22: Tertiärstruktur der Cytochrome P450 am Beispiel von CYP2C5 (1N6B). Das Proteinrückgrat ist als blaues Band dargestellt und das Häm mit dem zentralen Eisenatom ist grau eingefärbt. Die Helices sind als dunkelrote Zylinder dargestellt und mit lateinischen Großbuchstaben gekennzeichnet. Die Faltblätter sind als orange gefärbte Blockpfeile dargestellt und mit arabischen Ziffern durchnummeriert.

Üblicherweise werden die Helices mit lateinischen Großbuchstaben von A bis L durchnummeriert und die Faltblätter mit arabischen Ziffern von 1 bis 4, wobei die Einzelstränge wiederum durch arabische Ziffern (1-1 bis 4-2) gekennzeichnet werden. Die Helices D, E, K, L und das C-terminale Ende von Helix I sind um das Häm herum angeordnet und bilden den am höchsten konservierten Kern der Proteinstruktur. Die Helices D, E, I und L stellen gleichzeitig das Zentrum der α -Helixdomäne dar. Dabei liegt das Häm in einer sandwichartigen Form zwischen den Helices I und L. Die β -Faltblattdomäne besteht im Wesentlichen aus den Faltblättern β 1 und β 2 [85]. Im Gegensatz zu den Bakterien-cytochromen besitzen die Säugetierstrukturen drei zusätzliche Nebenhelices, A', J' und K', und die Helix J ist deutlich verlängert. Darüber hinaus ist auch der Loop zwischen den Helices F und G bei den mikrosomalen Cytochromen stark erweitert und enthält zwei weitere Nebenhelices F' und G'.

1.4.2.1 Hämbindung

Die am strengsten konservierten Aminosäuren der P450-Isoenzyme sind die an der Hämbindung beteiligten Reste. Dazu gehören die so genannte Cystein-Tasche und das N-terminale Ende von Helix L, sowie Faltblatt β_{1-4} und das N-terminale Ende von Helix C. Abbildung 1.23 zeigt wiederum anhand der Kristallstruktur 1N6B wichtige Interaktionen zwischen dem Häm und dem Apoenzym. Das zentrale Eisenatom des Häm ist über seine fünfte Koordinationsstelle kovalent an den Thiolschwefel eines invarianten Cysteins (hier Cys 432) in der Cystein-Tasche des Proteins gebunden. Man weiß aus Mutationsstudien, dass die Eisen-Cystein-Koordination eine wichtige Rolle für die Bildung der aktiven Eisen-Oxo-Verbindung spielt. Im Vergleich zum Cys-Wildtyp zeigt beispielsweise eine His-Mutante einen deutlich langsameren Elektronentransfer, eine erschwerte Spaltung der O-O-Bindung und eine geringere Substrataffinität beim Ablauf des katalytischen Mechanismus [86].

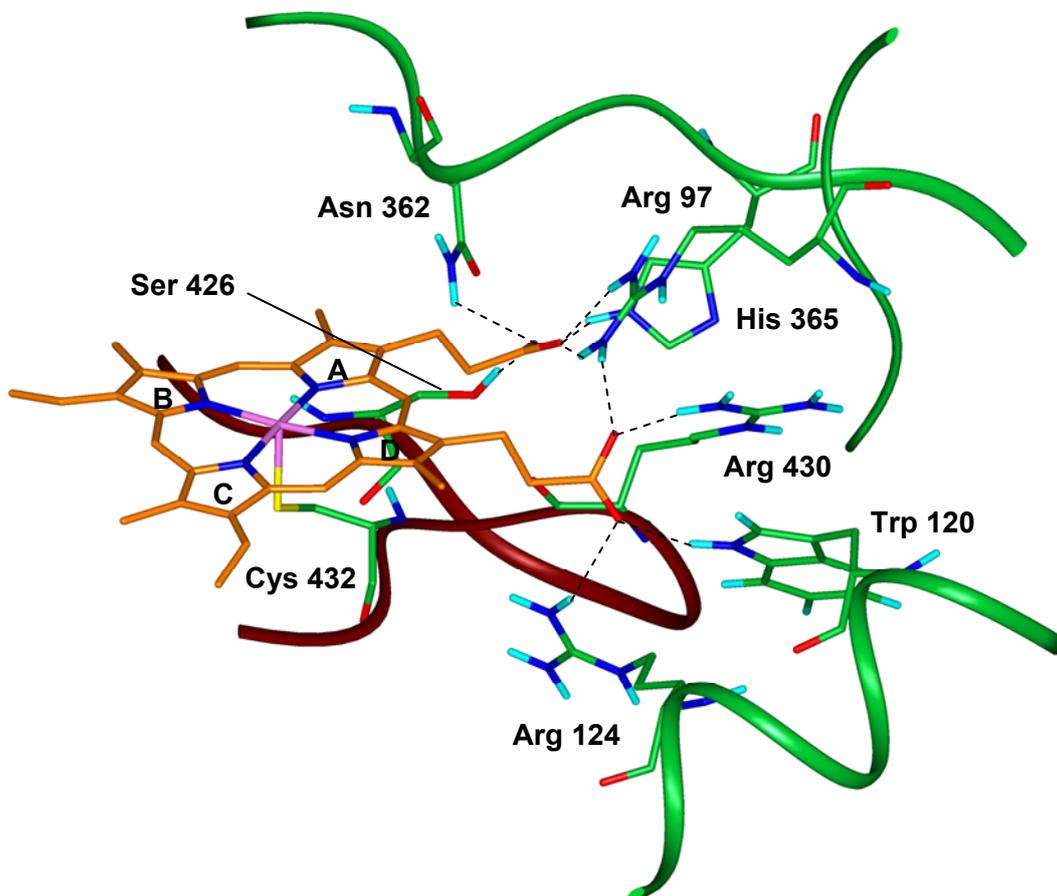


Abbildung 1.23: Konservierte Protein-Häm-Interaktionen der Cytochrome P450 am Beispiel von CYP2C5 (1N6B). Das Kohlenstoffgerüst des Häm ist in orange dargestellt und das Proteingerüst in grün. Die rotbraune Schleife markiert die Cystein-Tasche. Alle Heteroatome sind atomkodierte eingefärbt. Die schwarzen gestrichelten Linien markieren Salzbrücken zwischen den Propionatseitenketten des Häm und Aminosäureseitenketten des Proteins.

Das invariante Cystein ist Bestandteil eines zehn Aminosäuren langen konservierten Motivs, das bei allen Cytochromen P450 aber in keinem anderen Protein gefunden

wird [11]. Dieses Motiv mit der Sequenz FxxGxR/HxCxG bildet die Cystein-Tasche, die an den mit einem x gekennzeichneten variablen Positionen häufig hydrophobe Aminosäuren enthält, sich aufgrund der Glycine wie eine hydrophobe Schleife um das Cystein herum legt und durch das Phenylalanin abgeschlossen wird [83]. In Abbildung 1.23 ist sie als rotbraune Schleife eingezeichnet. Vermutlich hat die Hydrophobizität dieses Motivs auch einen Einfluss auf das Redoxpotential [87].

Zusätzlich zu den hydrophoben Interaktionen sind die Propionatseitenketten der Hämringe A und D über Salz- und Wasserstoffbrücken an Aminosäuren des Proteins gebunden. Innerhalb der gesamten Superfamilie hochkonserviert sind die Interaktionen der Propionatseitenkette des Hämring D mit einem Histidin oder Tryptophan und einem Arginin aus der Helix C (hier Trp120 und Arg 124) sowie die Interaktionen der Propionatseitenkette des Hämring A mit einem Arginin oder Histidin des Faltblattstrangs β_{1-4} (hier His 365) und mit einem Histidin oder Arginin aus der Cystein-Tasche (hier Arg 430). Daneben findet man meist weitere weniger streng konservierte Wechselwirkungen wie im Beispiel des CYP2C5 zwischen der Propionatgruppe des Hämring A und Ser 426 aus der Cystein-Tasche sowie Asn 362 vor dem Faltblattstrang β_{1-4} . Des Weiteren ist hier Arg 97 des Faltblattstrangs β_{1-5} mit beiden Propionatgruppen über Salzbrücken verbunden.

1.4.2.2 Wasserstoffbrückenmuster in Helix I

Ein charakteristisches Strukturmerkmal aller P450-Isoenzyme ist ein Knick in der zentralen Helix I, die oberhalb des Häm quer durch das gesamte Enzym verläuft. Dieser Knick kommt durch eine Unterbrechung des regulären Wasserstoffbrückenmusters in der Helix direkt über dem katalytischen Zentrum zustande. In Abbildung 1.24 ist der betreffende Helixabschnitt der Kristallstruktur 1N6B exemplarisch dargestellt. Ein konserviertes Threonin (hier Thr 298) bildet über die Hydroxylgruppe seiner Seitenkette eine Wasserstoffbrücke zum Carbonylsauerstoff des Proteinrückgrats der Aminosäure n-4. Im Fall des CYP2C5 ist das Ala 294. In der dadurch entstandenen aufgeweiteten Helixwindung befinden sich häufig ein oder zwei Wassermoleküle, die die fehlenden Wasserstoffbrücken zwischen Proteinrückgrat-Atomen ersetzen. Das Threonin bildet gemeinsam mit einer benachbarten sauren Aminosäure (hier Glu 297) ein konserviertes Motiv, das wahrscheinlich am Transfer der für die katalytische Reaktion benötigten Protonen ins aktive Zentrum beteiligt ist. Möglicherweise spielen auch die eingelagerten Wassermoleküle bei diesem Transfermechanismus eine Rolle. Aus Mutationsstudien am CYP101 (P450cam) ist bekannt, dass eine Entfernung der sauren Aminosäure zu einem drastischen Aktivitätsverlust des Enzyms führt, während eine Mutation des Threonins wesentlich geringere Auswirkungen auf die Enzymaktivität hat [88]. Bei vergleichbaren Mutationsstudien am humanen CYP1A2 stellte man fest, dass Asp320Ser- bzw. Asp320Ala-Mutanten kaum exprimiert wurden und nahezu inaktiv waren [34], [35].

Demnach scheint vor allem die hochkonservierte saure Aminosäure für die Katalyse von Bedeutung zu sein.

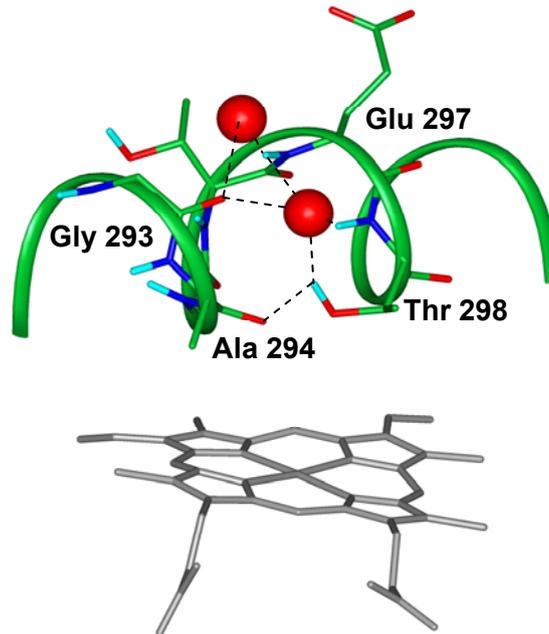


Abbildung 1.24: Charakteristische Aufweitung der Helix I über dem Häm am Beispiel von CYP2C5 (1N6B). Das Proteingerüst ist in grün dargestellt und das Häm in grau. Alle Heteroatome sind atomkodierte eingefärbt. Die roten Kugeln in der aufgeweiteten Helixwindung symbolisieren Wassermoleküle. Potentielle Wasserstoffbrücken sind als gestrichelte Linien eingezeichnet.

1.4.2.3 Zugangskanal zum aktiven Zentrum

Erste Hinweise auf die Lokalisation des Zugangskanals zum aktiven Zentrum erhielt man aus Kristallstrukturen des Bakterien Cytochroms CYP102 (P450_{BM3}), die sowohl im substratfreien als auch im substratgebundenen Zustand gewonnen werden konnten. Bei diesen Strukturen kommt eine offene Konformation durch ausgeprägte strukturelle Veränderungen im Bereich der Helices F und G und dem dazwischen liegenden Loop zustande. Diese Bewegungen wirken sich auch auf die Konformation der nachfolgenden Helices H und I aus [89]. Aus einer Kristallstruktur des bakteriellen Cytochroms CYP51 im geöffneten Zustand weiß man, dass dort neben dem FG-Loop und den Helices H und I auch der BC-Loop an der Öffnung des Zugangskanals beteiligt ist [90].

Ein recht drastisches Beispiel für mögliche Konformationsänderungen in geöffneten Cytochromstrukturen ist die kürzlich erschienene Kristallstruktur des Säugetiercytochroms CYP2B4 [91] (PDB-Code 1PO5). Im weit geöffneten aktiven Zentrum ist ein zweites Proteinmolekül über ein Histidin kovalent an das Hämeisen des ersten Moleküls gebunden, sodass praktisch ein Dimer entsteht. Ob diese Struktur tatsächlich eine physiologische Relevanz hat, ist fraglich, aber sie zeigt deutlich, welche Proteinbereiche besonders flexibel und höchstwahrscheinlich an der Bildung des Zugangskanals beteiligt sind. Abbildung 1.25 zeigt eine Überlagerung dieser

Kristallstruktur mit der geschlossenen Konformation von CYP2C5. Die Helix B'' hat sich in der offenen Konformation vollständig zu einer Spirale entwunden, ebenso die letzte Windung der Helix F. Der Loop, in dem sich die Nebenhelices F' und G' befinden, ist dadurch flexibel genug, um weit nach außen zu schwingen. Das Faltblatt $\beta 4$ weicht ebenfalls zurück und lässt Raum über dem Häm. Wie auch bei CYP51 kann man beobachten, dass das N-terminale Ende der Helix I deutlich nach oben klappt. Im Gegensatz dazu liegt der Loop FG in der geschlossenen Konformation von CYP2C5 wie ein Deckel über dem aktiven Zentrum, und auch die Helix B'' und der β -Turn des Faltblatts $\beta 4$ tragen zur Abdeckung des aktiven Zentrums bei.

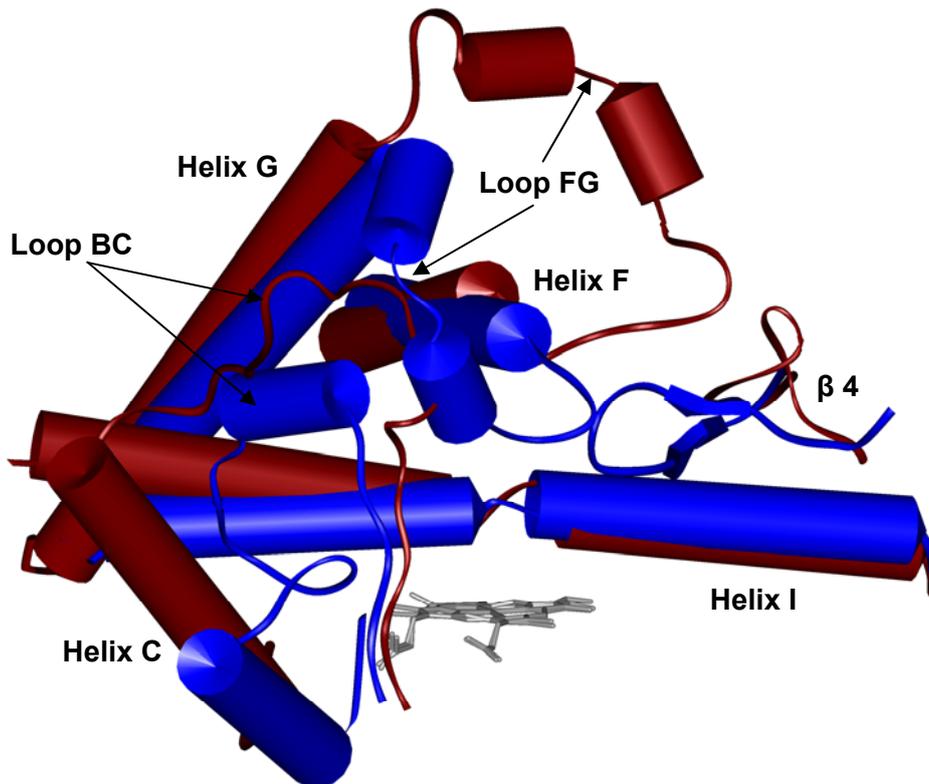


Abbildung 1.25: Überlagerung der Kristallstrukturen des substratgebundenen CYP2C5 (blau, 1N6B) mit dem weit geöffneten CYP2B4 (dunkelrot, 1PO5). Abgebildet sind die Proteinbereiche, die vermutlich den Zugangskanal zum aktiven Zentrum bilden.

1.4.2.4 Bindung von Liganden im aktiven Zentrum

An der Bindung von Substraten oder Inhibitoren im aktiven Zentrum sind sechs verschiedene Proteinbereiche beteiligt, die Substraterkennungsregionen SRS1-6 (*Substrate Recognition Sites 1-6*). Sie wurden von Gotoh auf der Grundlage eines Alignments von Sequenzen der Familie CYP2 mit Bakterienstrukturen identifiziert und lassen sich prinzipiell auf jede Cytochrom P450-Struktur übertragen [92]. Zu diesen Proteinregionen, die in Abbildung 1.26 anhand der Kristallstruktur 1N6B mit dem mitkristallisierten Sulfaphenazol-Derivat DMZ im aktiven Zentrum dargestellt sind, zählen einerseits hochkonservierte Strukturelemente wie der zentrale Abschnitt

der Helix I (SRS4, vgl. auch Abschnitt 1.4.2.2), aber andererseits auch so konformatorisch flexible und innerhalb der Superfamilie stark variable Bereiche wie die Helix B'' (SRS1) und Teile der Helices F und G (SRS2 und 3), gegebenenfalls unter Beteiligung des verbindenden Loops (vgl. auch Abschnitt 1.4.2.3). Die beiden noch verbleibenden Substratbinderegionen bestehen aus dem Abschnitt zwischen Helix K und dem Faltblattstrang β_{1-4} (SRS5), räumlich gesehen gegenüber von Helix I gelegen, und dem β -Turn des Faltblatts β_4 (SRS6).

In Kristallstrukturen, die Enzym-Substrat-Komplexe in einem katalytisch aktiven Zustand darstellen, weisen die Substrate stets mit ihrer Hydroxylierungsstelle auf das Hämeisen, wobei sich das zu oxidierende Atom etwa 4-5 Å vom Eisen entfernt befindet. Beispielsweise wird DMZ hauptsächlich an der 4-Methylgruppe des Phenylrings hydroxyliert und dieser Methylsubstituent ist in der Kristallstruktur im Abstand von 4,30 Å auf das Zentralatom des Häm orientiert.

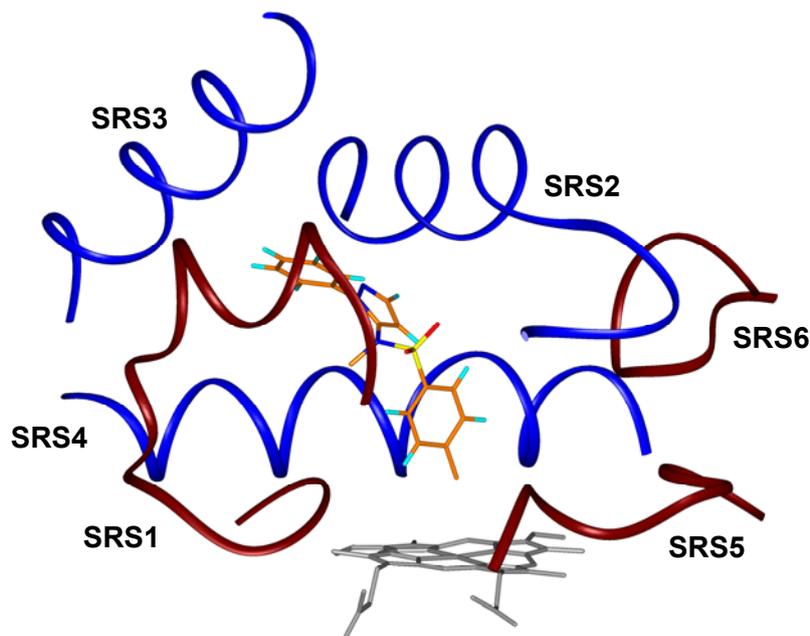


Abbildung 1.26: Das aktive Zentrum von CYP2C5 mit den Substraterkennungsregionen SRS1-6. Aus Gründen der Übersichtlichkeit sind SRS1, 5 und 6 in dunkelrot dargestellt und SRS2-4 in blau. Die Heteroatome des Substrats DMZ (4-Methyl-N-methyl-N-(2-phenyl-2H-pyrazol-3-yl)benzensulfonamid) sind atomkodiert eingefärbt.

Bemerkenswert sind die konformatorischen Veränderungen und der unterschiedliche Solvationsgrad des aktiven Zentrums desselben Isoenzym im Komplex mit verschiedenen Substraten. Zwei Kristallstrukturen von CYP2C5, von denen die eine das Substrat DMZ enthält (s.o.) und die andere mit dem Substrat Diclofenac kristallisiert wurde (PDB-Code 1NR6), ermöglichen hier einen direkten Vergleich [93] (Abbildung 1.27). Da Diclofenac deutlich kleiner als DMZ ist, rücken in diesem Komplex die Helix F, das Faltblatt β_4 und vor allem die Helix B'' tiefer ins aktive Zentrum hinein, um van-der-Waals-Kontakte zum Substrat auszubilden. Eine wichtige Funktion erfüllen bei der Bindung des Diclofenac im aktiven Zentrum von

CYP2C5 auch die in Abbildung 1.27 dargestellten Wassermoleküle, die im DMZ-Komplex fehlen. Sie dienen einerseits dazu, Hohlräume im aktiven Zentrum aufzufüllen, in denen sich sonst Substituenten größerer Substrate befinden wie hier beispielsweise der Phenylring des DMZ. Andererseits sind sie in der Lage, polare Substituenten des Substrates wie hier die Carboxylatgruppe des Diclofenac zu solvatisieren und Wasserstoffbrücken zu Aminosäuren der SRS des Enzyms zu vermitteln. Folglich sind xenobiotikametabolisierende Säugetiercytochrome wie das CYP2C5 in der Lage, die Form ihres aktiven Zentrums an die Strukturen verschiedener Substrate anzupassen und polare Substratbereiche mit Hilfe von Wasserclustern zu solvatisieren.

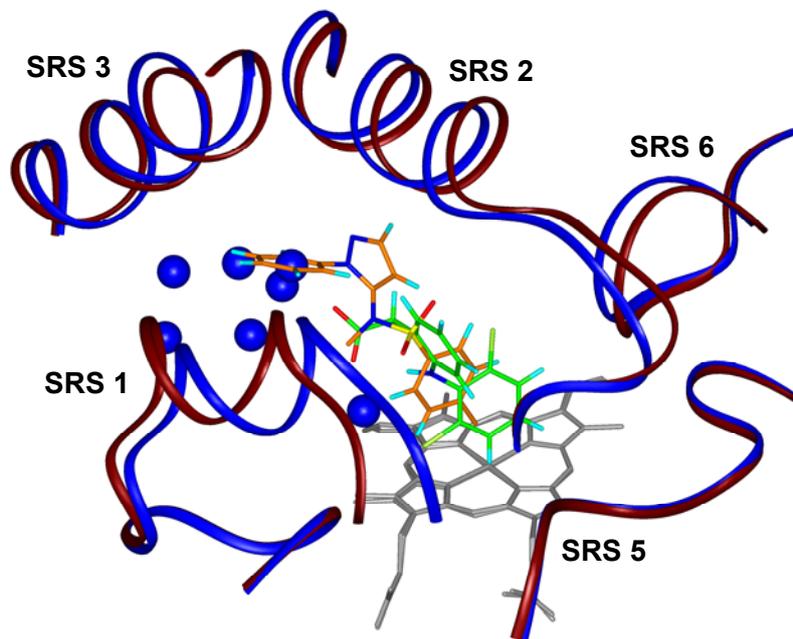


Abbildung 1.27: Überlagerung der Enzym-Substrat-Komplexe von CYP2C5 mit DMZ (1N6B, Proteinrückgrat dunkelrot und Substrat orange) und mit Diclofenac (1NR6, Proteinrückgrat blau und Substrat grün). SRS4, deren Konformation in beiden Komplexen identisch ist, wurde aus Gründen der Übersichtlichkeit weggelassen. Die blauen Kugeln repräsentieren Wassermoleküle aus der Struktur 1NR6.

Erstmals konnte auch anhand der Kristallstruktur eines Säugetiercytochroms die Existenz verschiedener Bindungsmodi nachgewiesen werden. In einem Enzym-Ligand-Komplex von CYP2C9 mit (*S*)-Warfarin befindet sich das Warfarin etwa 10 Å weit vom aktiven Zentrum entfernt in einer Seitentasche des aktiven Zentrums, die von Helix B', Helix F' und SRS5 gebildet wird [94]. Unmittelbar über dem Häm ist Platz für ein zweites Substratmolekül in einer katalytisch aktiven Position. Denkbar ist nun einerseits, dass das Warfarin erst im Verlauf der katalytischen Reaktion die für eine Umsetzung günstige Position einnimmt, oder dass ein weiteres Substrat simultan in einer katalytisch aktiven Position gebunden wird. Im letzteren Fall könnte das in der Kristallstruktur erfasste Warfarin einen allosterischen Modulator der CYP2C9-Aktivität darstellen.

Abschließend soll erwähnt werden, dass das Hämeisen in den bisher aufgeklärten substratgebundenen Kristallstrukturen von Säugetiercytochromen unterschiedliche Spin-Zustände aufweist (vgl. Abschnitt 1.3.1, Schritt 1.). Im Komplex von CYP2C9 mit Flurbiprofen ist die sechste Koordinationsstelle des Eisens trotz der Substratbindung mit einem Wassermolekül besetzt, und spektroskopische Messungen bestätigten, dass das Eisen hier nur teilweise in die *high-spin*-Konfiguration übergeht [95]. Dagegen wurden im Komplex von CYP2C5 mit DMZ ebenfalls nur 20-30 % *high-spin*-Eisen gemessen, ein Wassermolekül über dem Häm konnte aber nicht detektiert werden. Wechselwirkungen des Eisens mit dem Substrat selbst sind in diesem Fall allerdings auch nicht denkbar. Lediglich im Komplex von CYP2C5 mit Diclofenac wurde die typische Spin-Verschiebung zum *high-spin*-Zustand beobachtet, und die sechste Koordinationsstelle des Hämeisens ist in der Kristallstruktur unbesetzt.

2 PROBLEMSTELLUNG UND ZIELSETZUNG

Als Vertreter der in der Leber lokalisierten menschlichen P450-Isoenzyme spielt das humane CYP1A2 eine wesentliche Rolle beim Metabolismus von Arzneistoffen. Neben einer Vielzahl von Arzneistoffen, die selbst durch das Isoenzym umgesetzt werden, ist auch eine Reihe von Inhibitoren bekannt, die dessen Aktivität selektiv hemmen. Aus der simultanen Verabreichung mehrerer CYP1A2-Substrate oder der Kombination von Substraten und Inhibitoren in der Arzneimitteltherapie resultierten in der Vergangenheit bisweilen schwerwiegende Intoxikationen. Die Abschätzung des Interaktionspotentials eines Arzneistoffes aufgrund seiner Affinität zu CYP1A2 bereits in der Entwicklungsphase ist daher von großem Interesse.

Darüber hinaus wird CYP1A2 verschiedentlich mit der Entstehung von Blasen- und Dickdarmkrebs in Verbindung gebracht, da das Isoenzym an der metabolischen Aktivierung zahlreicher Prokarzinogene beteiligt ist. Hier wäre das frühzeitige Erkennen des Gefährdungspotentials eines potentiellen Karzinogens wünschenswert. Auch eine gezielte Entwicklung selektiver CYP1A2-Inhibitoren zur Vermeidung von Aktivierungsreaktionen kann in diesem Fall sinnvoll sein.

Im Rahmen der vorliegenden Arbeit sollten der Bindungsmodus von Substraten und Inhibitoren mithilfe computergestützter theoretischer Methoden auf molekularer Ebene untersucht und Struktur-Wirkungsbeziehungen hergeleitet werden.

Da die dreidimensionale Struktur des CYP1A2-Isoenzymen bisher nicht bekannt ist, sollte im ersten Teil der Arbeit ein Homologiemodell von CYP1A2 auf der Grundlage bereits existierender Kristallstrukturen von Säugetiercytochromen erstellt werden. Die Validierung des Proteinmodells sollte mithilfe von Moleküldynamiksimulationen (MDS) erfolgen. Dabei sollte das Modell auf sein dynamisches Verhalten und seine Stabilität hin analysiert werden. Weitere Validierungskriterien sollten die strukturelle Qualität und das Vorhandensein essentieller Strukturmerkmale darstellen.

Daran anschließend sollte der zweite Teil der Arbeit in einer Untersuchung von Enzym-Substrat-Komplexen zur weiteren Modellvalidierung und zur Erklärung des Bindungsmodus ausgewählter Substrate bestehen. Wichtige Protein-Substrat-Interaktionen sollten identifiziert und mit experimentellen Daten zum Bindungsverhalten der Substrate in Verbindung gebracht werden.

Die Anwendung des Homologiemodells sollte im dritten Teil der Arbeit mit der Durchführung einer 3D-QSAR-Studie erfolgen. Ein Datensatz, der aus kompetitiven CYP1A2-Inhibitoren bestand, sollte mittels automatischen Dockings im aktiven Zentrum der Proteinstruktur überlagert werden. Auf der Grundlage der Überlagerung sollte eine Korrelation mit der biologischen Aktivität der Inhibitoren und damit eine quantitative Beschreibung des Bindungsverhaltens kompetitiver Inhibitoren erzielt werden.

3.1 Analyse der Proteinsequenz

3.1.1 Automatisches Sequenzalignment

Vergleiche von Aminosäuresequenzen in Form von Sequenzalignments dienen einerseits dazu, aus einer Datenbank eine zur Target-Sequenz (Zielsequenz) homologe Sequenz zu finden, andererseits können mit ihrer Hilfe Übereinstimmungen zwischen Target- und Templat-Sequenz gefunden werden. Auf deren Grundlage werden später bei der Konstruktion des Modells die strukturkonservierten Bereiche festgelegt.

Das optimale Alignment zweier Sequenzen über ihre gesamte Länge (globales Alignment) berechnet der Needleman-Wunsch-Algorithmus [96], der auf der Methode der dynamischen Programmierung basiert. Eine Modifikation dieses Algorithmus, der Smith-Waterman-Algorithmus [97], erstellt in sehr ähnlicher Weise optimale lokale Alignments einzelner Sequenzabschnitte. Aus diesen sehr rechenaufwändigen Verfahren, mit denen eine Datenbanksuche kaum möglich ist, wurden so genannte heuristische Ansätze entwickelt, die zwar theoretisch nicht unbedingt das optimale Alignment finden, praktisch aber meist sehr gute Ergebnisse liefern. Dazu gehört die BLAST-Methode (*Basic Local Alignment Search Tool*) [98] mit den Weiterentwicklungen Gapped-BLAST und PSI-BLAST (*Position Specific Iterated BLAST*) [99], die einen heuristischen Ansatz zum Auffinden von möglichst optimalen lokalen paarweisen Alignments verwendet. Sie eignet sich damit zur Suche nach homologen Sequenzen aus einer Datenbank. BLAST 2 SEQUENCES [100] ist ein von BLAST abgeleitetes Programm zum schnellen groben Vergleich zweier Proteinsequenzen.

Ein möglichst genaues Alignment vieler verwandter Aminosäuresequenzen (multiples Sequenzalignment) erreicht man mit der progressiven Alignment-Methode, wie sie beispielsweise das Programm CLUSTAL W [101] verwendet. Dieser Ansatz geht davon aus, dass richtige Alignments auf evolutionärer Verwandtschaft basieren. Ein multiples Alignment wird demnach aufgrund eines phylogenetischen Baumes berechnet. Das Alignment wird dabei in drei Schritten erstellt (Abbildung 3.1). Im ersten Schritt werden alle Sequenzen paarweise miteinander verglichen und es wird eine Distanzmatrix aufgestellt. Hierbei kann man wählen, ob das Alignment mit dynamischer Programmierung oder mit einer heuristischen Methode erstellt werden soll. Als zweites wird ein phylogenetischer Baum aus der Matrix konstruiert. Nach diesem Baum richten sich die Gewichtung der Sequenzen und die Reihenfolge für das eigentliche Alignment, das im letzten Schritt entsprechend dem Baumdiagramm jeweils paarweise von den Spitzen bis zur Wurzel verlaufend durchgeführt wird.

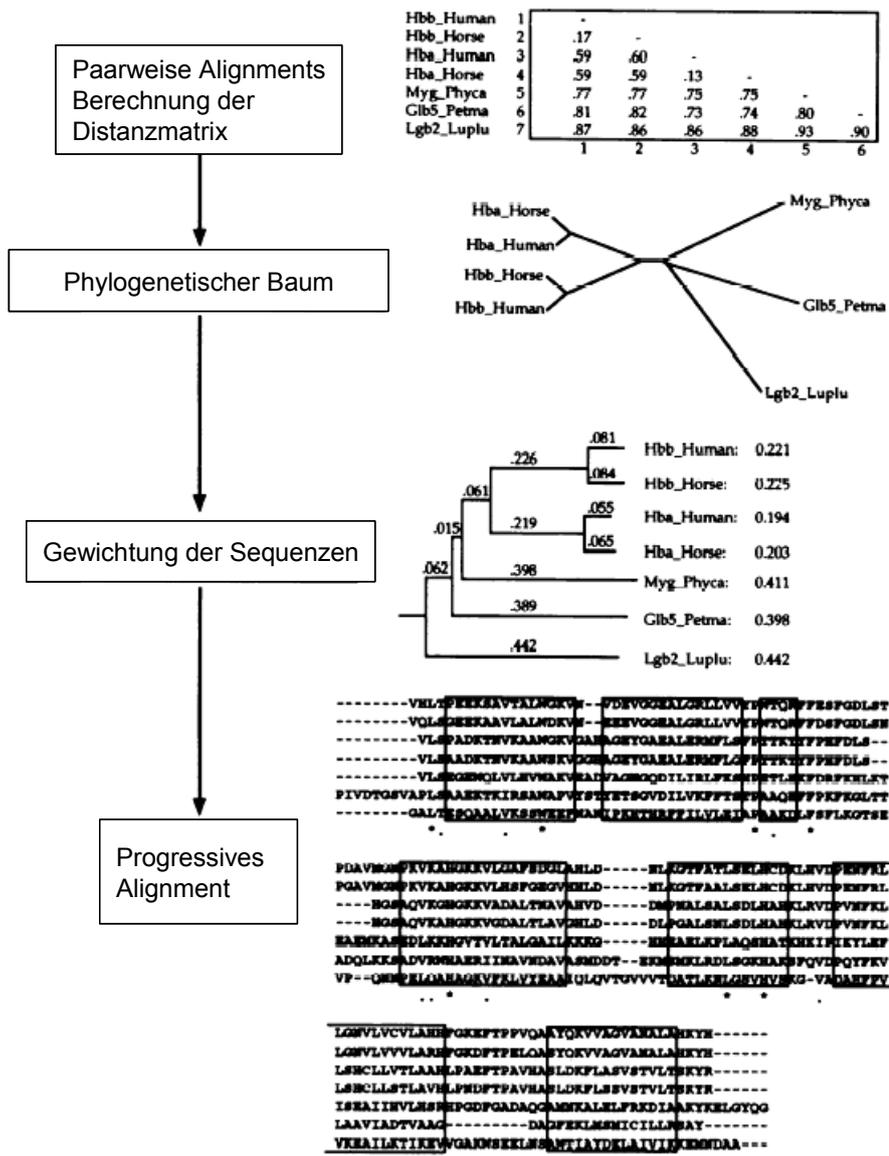


Abbildung 3.1: Vorgehensweise bei der Erstellung eines progressiven Alignments mit CLUSTAL W am Beispiel von sieben Globin-Sequenzen [Referenz s. Text]

Dieser Alignment-Vorgang wird mithilfe der dynamischen Programmierung durchgeführt. Die Sequenzen werden unterschiedlich gewichtet, um eine Überbewertung nahe verwandter Sequenzen zu verhindern. Außerdem werden Strafpunkte, so genannte *gap opening penalties*, die zum Öffnen einer Lücke im Alignment vergeben werden, während des Alignment-Vorganges variiert. Beispielsweise wird die Strafe zur Erzeugung einer Lücke in einer Abfolge von mindestens fünf hydrophilen Aminosäuren herabgesetzt, weil derartige Sequenzabschnitte meist in Loop-Regionen vorkommen. Im Gegensatz dazu wird die Strafe für eine Lücke, deren Abstand weniger als acht Aminosäuren zu einer schon bestehenden Lücke beträgt, erhöht, um das Auftreten von Lücken zwischen den Loop-Regionen zu vermeiden. Auf diesen Verfeinerungen im progressiven Alignment beruht der Erfolg der Methode.

Für alle Alignment-Methoden stehen verschiedene Bewertungsmatrices zur Verfügung, die jeweils vom Benutzer gewählt werden können. Die einfachste Matrix ist die Identitäts-Matrix, bei der lediglich allen identischen Aminosäuren ein Wert von eins und allen anderen ein Wert von null zugeordnet wird. Am häufigsten verwendet werden Mutationsmatrices, wie zum Beispiel die PAM250- [102] oder die BLOSUM62-Matrix [103]. Sie bewerten die Wahrscheinlichkeit, mit der eine Aminosäure gegen eine andere durch zufällige Mutation ausgetauscht werden kann. Für die in dieser Arbeit durchgeführten Alignments wurde jeweils die BLOSUM62-Matrix verwendet. Sie ist die Standardeinstellung in BLAST 2 SEQUENCES und lieferte für CLUSTAL W die besten Ergebnisse.

Ein weiterer Parameter, der vom Benutzer gewählt werden kann, ist die *gap opening penalty* (s.o.). Für BLAST 2 SEQUENCES wurde hier die vorgegebene Standardeinstellung von 11 verwendet, in CLUSTAL W wurde sie auf 30 erhöht, um zu viele Lücken im Alignment zu vermeiden.

3.1.2 Sekundärstrukturvorhersagen

Die Sekundärstruktur von Proteinen, deren dreidimensionale Struktur noch unbekannt ist, kann anhand ihrer Aminosäuresequenz vorhergesagt werden. Neben stereochemischen Methoden, die auf der Auswertung physikochemischer Eigenschaften der Aminosäureseitenketten beruhen [104], und statistischen Methoden, die die statistische Wahrscheinlichkeit analysieren, dass eine bestimmte Aminosäure in einem bestimmten Sekundärstrukturelement vorkommt, [105], [106], sind heute vor allem Methoden von Bedeutung, die von multiplen Sequenz-Alignments abgeleitete Informationen als Eingabe für neuronale Netze benutzen [107]. Für die vorliegende Arbeit wurde nach Evaluierung mehrerer Sekundärstrukturvorhersage-Programme das Programm PSIPRED verwendet [108], [109]. Bei der hier zugrunde liegenden Methode wird zwar nicht explizit ein multiples Sequenz-Alignment erstellt, aber es werden so genannte Sequenz-Profile als Grundlage für die Sekundärstrukturvorhersage eingesetzt, die von dem Alignment-Programm PSI-BLAST (Abschnitt 3.1.1) im Verlauf eines Suchprozesses als Zwischenergebnis erstellt werden. Aus den Profilen wird eine positionsspezifische Scoring-Matrix erhalten, die dann als Eingabe für ein neuronales Netz dient. Wie in anderen Sekundärstrukturvorhersage-Programmen auch werden hier zwei neuronale Netze hintereinander geschaltet, was die Effektivität der Prozedur nochmals steigert. Der Vorteil dieser neueren Methoden liegt darin, dass neuronale Netze trainierbar sind, und dadurch eine sehr hohe Vorhersagegenauigkeit von 76-78 % erreichen.

3.1.3 Erkennung von Sequenz-Struktur-Homologien

Zur Vorhersage der dreidimensionalen Struktur eines Proteins können darüber hinaus so genannte *threading* oder auch *fold recognition* Methoden angewendet werden, bei denen zu der vorhandenen Proteinsequenz homologe Strukturen aus einer Strukturdatenbank gesucht werden, um daraus eine strukturelle oder auch eine funktionale Verwandtschaft abzuleiten. Während derartige Sequenz-Struktur-Alignments in erster Linie dazu dienen, Informationen über die räumliche Struktur von Proteinen zu gewinnen, von denen keine in der Primärsequenz verwandten Strukturen aufgeklärt sind, wurde dieser Ansatz in der vorliegenden Arbeit lediglich dazu benutzt, um einen weiteren Anhaltspunkt für die Sekundärstruktur in einem gering konservierten Loop des zu modellierenden Enzyms zu erhalten. Die Scoring-Matrix dieser Methode, die aus Alignments homologer Proteinstrukturen berechnet wird, enthält nicht wie die klassischen Scoring-Matrices ausschließlich Mutationswahrscheinlichkeiten, sondern sie gibt außerdem wieder, welche Aminosäure in welcher Umgebung (Sekundärstruktur, Hydrophilie, Wasserstoffbrückenmuster) mit welcher Häufigkeit vorkommt. Basierend auf dieser Scoring-Matrix und ebenfalls strukturabhängigen *gap penalties* wird eine Datenbank aus 3D-Profilen für alle bekannten Proteinstrukturen und Struktur-Alignments erstellt. Die Sequenz, zu der eine homologe Proteinstruktur gefunden werden soll, wird dann letztlich mit den 3D-Profilen überlagert. Eine Besonderheit des hier verwendeten FUGUE-Servers [110] ist, dass von der zu untersuchenden Sequenz zuvor ein PSI-BLAST Alignment erstellt wird, sodass nicht nur eine einzelne Sequenz für die Struktursuche verwendet wird, sondern ein ganzes Alignment.

3.2 Homologie-Modelling

Das Homologie-Modelling wurde in der vorliegenden Arbeit mit dem HOMOLOGY-Modul des Software-Paketes INSIGHT II [111] durchgeführt.

3.2.1 Behandlung der strukturkonservierten Bereiche

Die strukturkonservierten Bereiche homologer Proteine sind Regionen im Proteininneren mit hoher Sequenzidentität bzw. Sequenzähnlichkeit. Dazu gehören beispielsweise die Sekundärstrukturelemente, die das Proteingerüst bilden, oder das aktive Zentrum [112]. Die Koordinaten für das Proteinrückgrat dieser strukturkonservierten Bereiche können direkt vom Templat auf das Modell übertragen werden. Handelt es sich um identische Aminosäuren, können die Koordinaten der Seitenkettenatome ebenfalls unverändert übernommen werden. Andernfalls wird die Aminosäure des Modells mit der ursprünglichen Aminosäure des Templats überlagert, sodass deren Konformation zumindest weitgehend

übernommen wird. Hat die neue Aminosäure mehr Atome als die Aminosäure des Templats, wird den zusätzlichen Atomen eine gestreckte Konformation zugewiesen.

3.2.2 Loop Search Routine

Loop-Regionen, welche die Sekundärstrukturelemente miteinander verbinden, unterscheiden sich auch unter homologen Proteinen häufig deutlich in Länge und Struktur voneinander. Sie werden daher auch als strukturvariable Regionen bezeichnet. Neben der Möglichkeit, solche Loops *de novo* zu konstruieren, gibt es die so genannte *loop search* Routine, mit der die Protein Datenbank (PDB) [113] nach passenden Proteinfragmenten durchsucht wird. Die letztgenannte Methode wurde in dieser Arbeit angewendet, da sie den Vorteil hat, dass die aus der Datenbank entnommenen Loops tatsächlich in Proteinen vorkommen und daher eine realistische Konformation mit einem relativ niedrigen Energieinhalt aufweisen. Abgesehen davon, dass die Anzahl der Aminosäuren vorgegeben wird, aus denen das Fragment bestehen soll, wird die Suche ausschließlich nach geometrischen Kriterien durchgeführt. Das Programm erstellt eine Distanzmatrix aus den Abständen der C α -Atome der Aminosäuren, die zu den angrenzenden strukturkonservierten Bereichen gehören (*preflex* (pe) und *postflex* (po) Aminosäuren):

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}, i, j \in pe, po \quad (3.1)$$

Wie viele Aminosäuren jeweils in die Matrix mit einbezogen werden (Anzahl N), kann der Benutzer selbst wählen. Das Programm schlägt einen Standardwert von fünf Aminosäuren von jedem der beiden Enden vor. Wählt man zu wenige Aminosäuren, passt der Loop schlechter, wählt man zu viele, ist das geometrische Kriterium so eng, dass man unter Umständen gar keinen passenden Loop findet. Die so erhaltene Distanzmatrix wird dann mit den entsprechenden Distanzmatrices aller möglichen Loops aus der Datenbank verglichen ($r_c(i)$ und $r_c(j)$ stehen für die Koordinaten der Loop-Kandidaten):

$$D_c^2 = \frac{2}{N(N+1)} \sum_{i \in pe, po} \sum_{j \in pe, po, j > i} (d_{ij} - d_{r_c(i)r_c(j)})^2, i, j \in pe, po \quad (3.2)$$

Die zehn besten Ergebnisse der Suche werden gespeichert und können anschließend visuell ausgewertet werden. Als zusätzliches Auswahlkriterium wird der RMSD-Wert der Überlagerung der Loop-Enden mit den angrenzenden strukturkonservierten Bereichen angegeben. Bei dieser Überlagerung können entweder nur die beiden Endaminosäuren berücksichtigt werden, oder alle *preflex* und *postflex* Aminosäuren (*tails* Option). Die Ausrichtung der Loops kann dabei je nach der Überlagerungsmethode stark variieren. Koordinaten für den gewählten Loop werden schließlich analog dem Verfahren für die strukturkonservierten Bereiche zugewiesen.

3.3 Quantenmechanische Berechnung von Atomeigenschaften

Quantenmechanische Methoden betrachten explizit die Elektronen in einem System, so dass es mit ihrer Hilfe möglich ist, Atom- und Moleküleigenschaften zu bestimmen, die von der Elektronendichteverteilung abhängig sind, oder chemische Reaktionen nachzuvollziehen, in deren Verlauf Bindungen gebrochen und neu ausgebildet werden. Im Rahmen dieser Arbeit wurden quantenchemische Rechnungen mit den einzelnen Liganden durchgeführt, denen bei der Aufstellung einer Molekültopologie für Kraftfeldrechnungen Partialladungen zugewiesen werden müssen. Mittels der Quantenchemie ganze Protein-Ligand-Komplexe oder andere makromolekulare Systeme zu berechnen, ist aufgrund des enormen Rechenaufwandes bis heute nicht möglich.

Die Grundlage der quantenmechanischen Beschreibung eines Systems ist die Schrödinger-Gleichung.

$$\begin{aligned} \bar{H} &: \text{Hamilton - Operator} \\ \bar{H}\Psi = E\Psi & \quad E : \text{Eigenwert der Energie} \\ \psi &: \text{Gesamtwellenfunktion des Systems} \end{aligned} \quad (3.3)$$

Die Lösung dieser Gleichung ergibt die Wellenfunktion des Systems, aus der seine Eigenschaften abgeleitet werden können. Sie lässt sich jedoch nur für Einteilchenprobleme exakt lösen, d.h. in der Quantenchemie für das Wasserstoffatom. Für Mehrteilchenprobleme müssen Näherungen eingeführt werden.

Wichtig für alle quantenmechanischen Methoden ist die Born-Oppenheimer-Näherung. Da die Atomkerne bedeutend schwerer als die Elektronen sind, ist ihre Geschwindigkeit im Vergleich zu den Elektronen vernachlässigbar klein. Man kann somit die Elektronen als bewegte Teilchen in einem Feld von festen Atomkernen annehmen und die Schrödinger-Gleichung in einen Ausdruck für die Elektronen und in einen für die Kerne auftrennen.

$$H = T_{\text{Kern}} + T_{\text{El}} + E_{\text{Kern-Kern}} + E_{\text{El-Kern}} + E_{\text{El-El}} \quad (3.4)$$

In der elektronischen Schrödinger-Gleichung wird dann die kinetische Energie der Atomkerne (T_{Kern}) gleich Null gesetzt und die Abstoßung zwischen den Atomkernen ($E_{\text{Kern-Kern}}$) als konstant betrachtet.

Zur Lösung dieser vereinfachten Schrödinger-Gleichung gibt es verschiedene Näherungsansätze. Einer der bedeutendsten ist die Hartree-Fock-Näherung, welche die Gesamtwellenfunktion des Systems aus Wellenfunktionen für jedes einzelne Elektron zusammensetzt und somit das Mehrteilchenproblem als Summe von

Einteilchenproblemen behandelt. Der Nachteil dieser Methode besteht darin, dass nur die Austausch-Energie (Wechselwirkungen zwischen Elektronen mit gleichem Spin) berücksichtigt werden kann, nicht aber die Korrelations-Energie (Wechselwirkungen zwischen Elektronen mit entgegengesetztem Spin).

Der Ansatz, der in dieser Arbeit verwendet wurde, ist die **Dichte-Funktional-Theorie (DFT)**, bei der ein Mehrteilchensystem lediglich durch Einteilchendichten beschrieben wird, ohne dass einzelne Wellenfunktionen explizit berechnet werden. Nach dem Hohenberg-Kohn-Theorem [114] reicht die Kenntnis der Elektronendichte aus, um die Energie eines Systems im Grundzustand eindeutig zu bestimmen. Ein Lösungsansatz für dieses Energiefunktional wurde von Kohn und Sham [115] vorgeschlagen und bildet die Basis für die DFT-Methode (s. auch [117]).

Alle quantenchemischen Berechnungen im Rahmen dieser Arbeit wurden mit dem Programm JAGUAR [116] durchgeführt.

3.4 Molekülmechanik

Im Gegensatz zur Quantenchemie werden Elektronenbewegungen in der Molekülmechanik vernachlässigt, und die Energie eines Systems wird lediglich als Funktion der Positionen von Atomkernen betrachtet. Folglich können mittels molekülmechanischer Methoden zwar keine Aussagen über das elektronische Verhalten eines Systems getroffen werden, aber Molekülgeometrien und relative Energien können auch für Systeme von Biopolymeren schnell und zuverlässig bestimmt werden. Damit ist die Molekülmechanik für die vorliegende Fragestellung, nämlich die Simulation eines Proteinmodells in einer Lösungsumgebung, die angemessene Methode [112], [117].

Bei der Molekülmechanik handelt es sich um ein empirisches Verfahren, das Wechselwirkungen zwischen den Atomen mithilfe der klassischen Mechanik beschreibt. Atome werden dabei als Kugeln mit einer bestimmten Masse betrachtet und sind über Federn, welche die chemischen Bindungen darstellen, verbunden. Die mathematische Formel für die elastische Verformung einer Feder (Hookesches Gesetz) wird zur Beschreibung der Bindungen herangezogen.

Mit diesem Ansatz wird die sterische Energie einer Struktur berechnet, also der Energieinhalt, der aus einer bestimmten Konformation resultiert. Dieser Energiewert kommt durch abweichende Bindungslängen, Bindungswinkel und Torsionswinkel von empirischen Referenzparametern für bestimmte Atom- und Bindungstypen zustande. Außerdem fließen empirische Werte für nicht-kovalente Wechselwirkungen mit ein. Die berechnete sterische Energie hat keinerlei physikalische Bedeutung, weshalb nur die Auswertung von Energiedifferenzen sinnvoll ist, aber auch nur dann, wenn die zu vergleichenden Energiewerte aus Berechnungen mit demselben Kraftfeld stammen.

Die Potentialfunktionen für die Berechnung der einzelnen Energieterme sowie die notwendigen Referenzparameter für Bindungslängen, Bindungswinkel, Torsionswinkel und nicht-kovalente Wechselwirkungen basieren auf quantenchemischen Berechnungen und sind in einem Kraftfeld hinterlegt. Alle in dieser Arbeit beschriebenen Kraftfeld-Rechnungen wurden mit dem Programm GROMACS [118], [119] und dem darin implementierten GROMOS96-Kraftfeld [120] durchgeführt. Das GROMACS-Programmpaket wurde ausgewählt, weil es ein besonders vielseitiges und leistungsstarkes Simulationsprogramm darstellt, das ursprünglich speziell für die Simulation großer biologischer Systeme wie Proteine oder Lipidmembranen entwickelt wurde.

3.4.1 Das GROMOS96-Kraftfeld

Wie in allen gängigen Kraftfeldern setzt sich die Gesamtenergie E_{tot} zusammen aus den Termen für kovalente Wechselwirkungen, das sind die Bindungslängenenergie E_b , die Bindungswinkelenergie E_a , die „improper dihedral“-Energie E_{id} und die Torsionsenergie E_d , und den Termen für die nicht-kovalenten Wechselwirkungen, das sind die Dispersionsenergie E_{LJ} und die elektrostatische Energie E_C .

$$E_{\text{tot}} = E_b + E_a + E_{id} + E_d + E_{LJ} + E_C \quad (3.5)$$

Wasserstoffbrücken ergeben sich alleine aus dem elektrostatischen Term. Ein spezieller Term, wie er in einigen anderen Kraftfeldern vorkommt, ist dafür nicht vorgesehen. Auch Kopplungen einzelner Energieterme (Kreuzterme) werden hier nicht berücksichtigt.

Tabelle 3.1 listet alle im GROMOS96-Kraftfeld verwendeten Energieterme auf. Es handelt sich meist um einfache harmonische Potentiale, die sich von Kraftfeld zu Kraftfeld nur geringfügig unterscheiden. Trotzdem weist das Kraftfeld einige Besonderheiten auf, die im Folgenden kurz erläutert werden.

Die Energie für die Bindungslängen beispielsweise wird im GROMOS96-Kraftfeld nicht durch ein harmonisches Potential beschrieben, sondern als Fourth-Power-Potential. Dieses Potential ist einerseits schneller zu berechnen, hat aber andererseits den Nachteil, dass es kein harmonisches Potential darstellt. Für diese Arbeit spielt die Berechnung der Bindungslängenenergie allerdings ohnehin keine Rolle, da alle Bindungslängen während der MD-Simulationen konstant gehalten wurden (s. Abschnitt 3.4.3.2).

Für die Berechnung der Bindungswinkelenergie wird ebenfalls zur Erhöhung der Rechengeschwindigkeit eine vereinfachte Funktion benutzt, und zwar eine Cosinus-basierte Funktion.

Als Improper-Dihedral-Energie wird der Energiebetrag bezeichnet, der aus der Abweichung von Ringsystemen und anderen planaren Gruppen von der Planarität und der Konfiguration von Chiralitätszentren von der vorgegebenen Konfiguration resultiert. Der Improper-Dihedral-Winkel ξ ist dabei definiert als der Winkel zwischen der Ebene aus den Atomen i,j,k und der Ebene aus den Atomen j,k,l, wobei i das Zentralatom ist und j,k,l drei seiner Substituenten sind, bzw. in Ringsystemen die drei darauf folgenden Ringatome. Das sich ergebende Potential ist harmonisch.

Wenn wie in diesem Fall die elektrostatischen Wechselwirkungen langer Reichweite mit der PME-Methode berücksichtigt werden sollen, muss auch für die Berechnung der elektrostatischen Wechselwirkungen kurzer Reichweite ein modifizierter Coulomb-Term eingesetzt werden (s. nächster Abschnitt).

Tabelle 3.1: Energierterme des GROMOS96-Krauffeldes

Bindungslängenenergie (Fourth-Power-Potential)	$E_b(r) = \frac{1}{4} k_r (r^2 - r_0^2)^2$	k_r : Kraftkonstante r : tatsächliche Bindungslänge r_0 : Referenzbindungslänge
Bindungswinkelenergie (cos-basiertes Potential)	$E_a(\theta) = \frac{1}{2} k_\theta (\cos(\theta) - \cos(\theta_0))^2$	k_θ : Kraftkonstante θ : tatsächlicher Winkel θ_0 : Referenzwinkel
Improper-Dihedral-Energie	$E_{id}(\xi) = k_\xi (\xi - \xi_0)^2$	k_ξ : Kraftkonstante ξ : tatsächlicher Winkel ξ_0 : Referenzwinkel
Torsionsenergie	$E_d(\phi) = k_\phi (1 + \cos(n\phi - \phi_0))$	k_ϕ : Torsionsbarriere n : Periodizität ϕ : tatsächlicher Torsionswinkel ϕ_0 : Referenztorsionswinkel
Dispersionsenergie (Lennard-Jones-Potential)	$E_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^6}$	r_{ij} : Abstand zwischen i und j $C_{ij}^{(12)}$: Repulsionskoeffizient $C_{ij}^{(6)}$: Attraktionskoeffizient
Elektrostatische Energie (mod. Coulomb-Potential)	$E_c(r_{ij}) = \frac{1}{4\pi\epsilon_0} \frac{\text{erfc}(\beta r_{ij})}{r_{ij}} q_i q_j$	r_{ij} : Abstand zwischen i und j ϵ_0 : Dielektrizitätskonstante (Vakuum) β : Wichtungsfaktor erfc : Fehlerfunktion q_i, q_j : Punktladungen von i und j

3.4.1.1 Die Particle-Mesh-Ewald-Methode (PME-Methode)

Da die elektrostatischen Wechselwirkungen im Vergleich zu den Dispersionswechselwirkungen mit zunehmender Entfernung vom Massepunkt nur sehr langsam abnehmen, führt eine Vernachlässigung dieser Wechselwirkungen ab einem bestimmten *Cut-off*-Wert oft zu strukturellen und dynamischen Artefakten. Deshalb ist es sinnvoll, neben den elektrostatischen Wechselwirkungen kurzer

Reichweite innerhalb eines *Cut-offs* auch die elektrostatischen Wechselwirkungen langer Reichweite mit einzubeziehen. Eine Möglichkeit dafür bietet die Ewald-Summe [121] und eine weniger rechenaufwändige Variante, die Particle-Mesh-Ewald-Methode (PME-Methode) [122]. Zwar kann es auch hier zur Bildung von Artefakten kommen, weil dem System eine nicht vorhandene Periodizität zugeschrieben wird, doch dieser Umstand kann in der Regel vernachlässigt werden.

Die gesamte elektrostatische Energie eines Systems wird berechnet nach

$$E = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{n_x, n_y, n_z} \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{r_{ij,n}} \quad (3.6)$$

Dabei ist $(n_x, n_y, n_z) = \mathbf{n}$ der Ursprungsvektor der Einheitszelle. Der Stern zeigt an, dass der Term $i = j$ vernachlässigt wird, wenn $(n_x, n_y, n_z) = (0, 0, 0)$ ist. Der Abstand $r_{ij,n}$ ist der wahre Abstand zwischen den Ladungen q_i und q_j . Diese nur sehr langsam konvergierende Summe wird nun durch zwei schnell konvergierende Terme (E_{dir} und E_{rec}) und eine Konstante E_0 ersetzt.

$$E = E_{dir} + E_{rec} + E_0 \quad (3.7)$$

Die direkte Summe E_{dir} beschreibt exakt die kurzreichweitigen Coulomb-Wechselwirkungen zwischen benachbarten Atomen innerhalb eines *Cut-off*-Radius.

$$E_{dir} = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{i,j} \sum_{n_x, n_y, n_z} q_i q_j \frac{\text{erfc}(\beta r_{ij,n})}{r_{ij,n}} \quad (3.8)$$

Die reziproke Summe E_{rec} gibt die langreichweitigen Wechselwirkungen mit dem restlichen System mithilfe einer Fourier-Reihe näherungsweise wieder.

$$E_{rec} = \frac{1}{2\pi V} \frac{1}{4\pi\epsilon_0} \sum_{i,j} q_i q_j \sum_{m_x, m_y, m_z} \frac{\exp(-(\pi \mathbf{m} / \beta)^2 + 2\pi i \mathbf{m}(\mathbf{r}_i - \mathbf{r}_j))}{m^2} \quad (3.9)$$

Der konstante Term E_0 eliminiert die Energie, die durch die Wechselwirkungen eines Partikels mit sich selbst zustande kommt.

$$E_0 = -\frac{1}{4\pi\epsilon_0} \frac{\beta}{\sqrt{\pi}} \sum_i q_i^2 \quad (3.10)$$

$\mathbf{m} = (m_x, m_y, m_z)$ und β legt jeweils die relative Wichtigkeit der direkten und der reziproken Summe fest. Die Fehlerfunktion *erf* ist definiert als das Integral über eine Gaußverteilung. Dementsprechend ist *erfc* die komplementäre Fehlerfunktion.

Anstatt entsprechend der Ewald-Summe die Vektoren der reziproken Summe direkt zu addieren, werden die Ladungen bei der PME-Methode auf ein Gitternetz übertragen. Das elektrostatische Potential jedes Gitterpunktes wird über einen 3D-

FFT-Algorithmus (*Fast Fourier Transformation*) berechnet. Anschließend wird das elektrische Feld an jedem Gitterpunkt durch Ableitung des elektrostatischen Potentials bestimmt. Die Kraft, die auf das einzelne Partikel i wirkt, wird daraufhin durch Interpolation erhalten. Der Vorteil dieser Variante liegt darin, dass der Rechenaufwand für die reziproke Summe für N Atome nicht wie bei der Ewald-Summe N^2 beträgt, sondern nur $N \cdot \log(N)$. Das macht die Methode auch für Systeme mit $N > 10^4$ Atomen anwendbar.

3.4.1.2 Parametersatz

Wie viele andere Kraftfelder, die für die Simulation von Proteinen eingesetzt werden, ist auch das GROMOS96-Kraftfeld ein *united-atom*-Kraftfeld. Das bedeutet, dass nur die polaren Wasserstoffe explizit berücksichtigt werden, die unpolare Wasserstoffe werden mit den schweren Atomen zusammengefasst, an die sie gebunden sind. Beispielsweise wird eine CH_2 -Gruppe als ein Massenpunkt mit einer Atommasse von 14 und einem entsprechend größeren Radius definiert. Aromatische CH -Gruppen werden jedoch explizit berücksichtigt, um π - π -Interaktionen besser wiedergeben zu können [123].

Die Parametrisierung des GROMOS96-Kraftfeldes eignet sich besonders gut für die Beschreibung von Proteinen und kleinen Molekülen [124]. Um die Eigenschaften des in Cytochromen enthaltenen Häm richtig wiederzugeben, mussten jedoch einige spezielle Parameter hinzugefügt werden. Dazu wurden die von R. Wade veröffentlichten Partialladungen und die Parameter für die kovalenten Wechselwirkungen [127] aus den Arbeiten von A. Schappach [125] und G. Stahl [126] übernommen (Tabelle 3.2, Tabelle 3.3 und Abbildung 3.2). Für die Beschreibung der Dispersionswechselwirkungen wurde das Lennard-Jones-Potential für Zink, das im GROMOS96-Kraftfeld vorhanden ist, auf das Eisen übertragen.

Tabelle 3.2: Zusätzliche kovalente Parameter für das GROMOS96-Kraftfeld zur Verbindung des Hämeisens mit dem kovalent gebundenen Cystein. Zur Bezeichnung der Atomtypen s. Abbildung 3.2 und Strukturformel des Cysteins neben Tabelle 3.3

Bindung	r_0 (nm)	k_r ($\text{kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$)
SG – FE	0,234	$2,68 \cdot 10^{-6}$
Bindungswinkel	θ_0 ($^\circ$)	k_θ ($\text{kJ} \cdot \text{mol}^{-1} \cdot \text{rad}^{-2}$)
CB – SG – FE	100,0	475,0
SG – FE – NR	90,0	420,0
Torsionswinkel	Φ_0 ($^\circ$)	k_Φ ($\text{kJ} \cdot \text{mol}^{-1} \cdot \text{rad}^{-2}$)
CA – CB – SG – FE	0	2,93
CB – SG – FE – NR	0	0

Tabelle 3.3: Partialladungen des kovalent an das Häm gebundenen Cysteins. Rechts ist die Strukturformel des Cysteins mit den Bezeichnungen der Atomtypen beigelegt.

Atomtyp	q (e)
N	-0,280
H	0,280
CA	0,000
CB	-0,100
SG	-0,500
C	0,380
O	-0,380

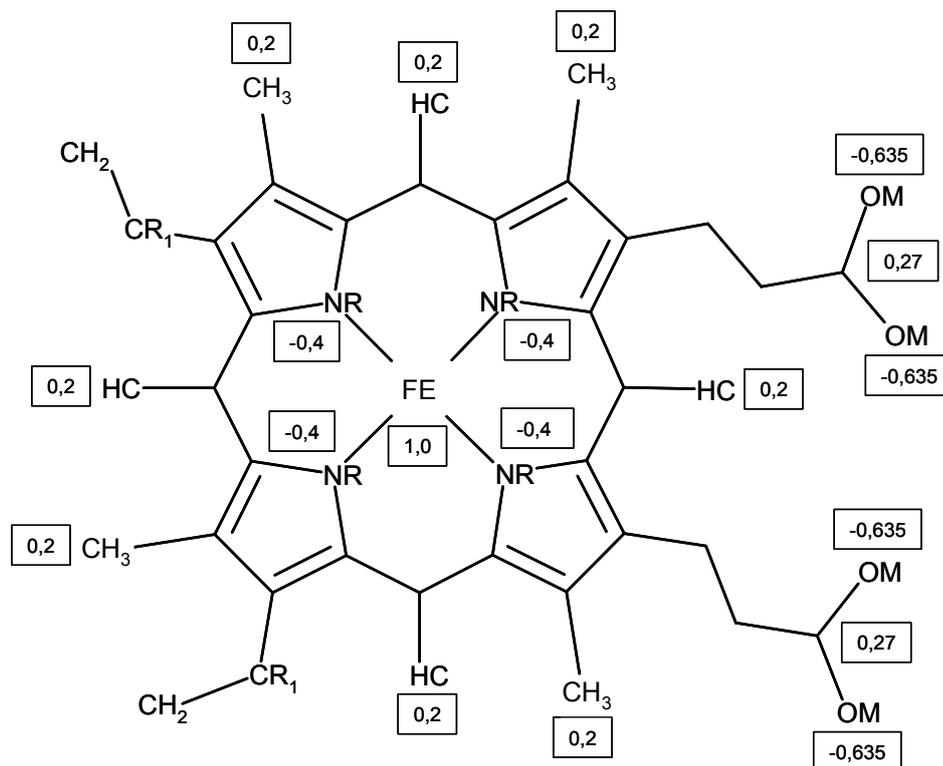
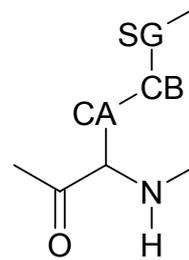


Abbildung 3.2: Atomtypen und Partialladungen des Häm im GROMOS96-Kraftfeld. Mit Ausnahme des Atomtyps „C“ sind alle Atomtypen angegeben. Alle Atome, für die keine Partialladung angegeben ist, sind ungeladen.

3.4.2 Geometrieoptimierung

Die Optimierung von energetisch ungünstigen, verzerrten Geometrien, die beispielsweise aus der Konstruktion eines Homologiemodells resultieren, oder als Momentaufnahmen aus einer MD-Simulation entnommen wurden, wird mithilfe einer Minimierung der potentiellen Energie erreicht. Zwei gängige Verfahren, die beide die erste Ableitung der Energiefunktion zur Ermittlung eines lokalen Minimums verwenden, stellt das Programm GROMACS dafür zur Verfügung.

3.4.2.1 Steepest Descent

Die Anfangskoordinaten werden zunächst um eine festgelegte Schrittweite (z.B. 0,01 nm) in Richtung der negativen Steigung der Energiefunktion verschoben. Dieser Vorgang wird wiederholt, solange sich die potentielle Energie verringert. Die Schrittweite wird dabei stetig erhöht. Folglich kann man mit dieser Methode schnell und effektiv in die Nähe eines lokalen Minimums gelangen, auch wenn man von einer schlechten Startgeometrie ausgegangen ist. Sobald die Energie nach einer Iteration angestiegen ist, wird die Iteration rückgängig gemacht und die Schrittweite wird verringert. Das führt dazu, dass die Methode in der Nähe des Minimums nur sehr langsam konvergiert und unter Umständen um das Minimum oszilliert.

3.4.2.2 Conjugate Gradient

Bei dieser Methode wird jeweils die Steigung des vorangegangenen Minimierungsschrittes zur Festlegung der neuen Schrittweite herangezogen. Dadurch erreicht man, dass die Methode in der Nähe des Minimums besser konvergiert.

Entsprechend dem unterschiedlichen Charakter der beiden Minimierungsverfahren ist es üblich, mit der *steepest descent* Methode zunächst für eine grobe Minimierung der Startstruktur zu beginnen, und daran je nach der Fragestellung eine *conjugate gradient* Minimierung bis zur Konvergenz anzuschließen [112].

3.4.3 Moleküldynamiksimulationen

Da eine Energieminimierung nur ein nahe der Ausgangsstruktur gelegenes lokales Minimum findet, ist es sinnvoll, den Konformationsraum des Modells mit einer effektiveren Methode weiter zu durchsuchen. Die Methode der Wahl für die Konformationsanalyse eines Proteins sind Moleküldynamiksimulationen (MDS), die dazu dienen, molekulare Bewegungen eines Systems über einen bestimmten Zeitraum zu berechnen.

Die anfängliche Equilibrierungsphase einer MDS kann man als Fortsetzung des Modellierungsprozesses bezeichnen, weil hier zunächst eine weitere Entspannung und Verfeinerung der Struktur, bzw. eine Anpassung der Geometrie an die Gegebenheiten der Targetsequenz stattfindet. Hat das System einen Gleichgewichtszustand erreicht, kann man den weiteren Verlauf der MDS zum Absuchen des Konformationsraumes nach einer möglichst repräsentativen Auswahl energetisch günstiger Konformationen nutzen. Andererseits ist es auch möglich, das dynamische Verhalten des Systems zu beobachten, auszuwerten, und mit den Ergebnissen experimenteller Untersuchungen zu vergleichen.

Die Grundlage für die Berechnung molekularer Bewegungen ist die Newtonsche Bewegungsgleichung.

$$F_i = m_i \cdot a_i \quad (3.11)$$

Mit ihr kann man die Beschleunigung a_i eines Atoms i mit der Masse m_i zu einem Zeitpunkt t bestimmen. Die Kraft F_i erhält man dabei aus der Ableitung der potentiellen Energie V , die aus dem Kraftfeld bestimmt wird, nach der Ortskoordinate r_i .

$$F_i = -\frac{\partial E}{\partial r_i} \quad (3.12)$$

Über den Zusammenhang zwischen Beschleunigung und Geschwindigkeit wird als nächstes die Geschwindigkeit des Atoms i über einen Zeitraum Δt berechnet, und über die Geschwindigkeit schließlich die neue Ortskoordinate zum Zeitpunkt Δt . Üblicherweise wird der *leap frog* Algorithmus [128] für die Integration angewendet. Dieser Algorithmus berechnet jeweils die neuen Ortskoordinaten zum Zeitpunkt Δt und die neuen Geschwindigkeiten zum Zeitpunkt $\frac{1}{2} \Delta t$. Die Geschwindigkeit zum Zeitpunkt $-\frac{1}{2} \Delta t$, die bei der Berechnung der neuen Geschwindigkeit berücksichtigt wird, stammt zu Beginn der Simulation aus einer Maxwell-Boltzmann-Verteilung für die vorgegebene Simulationstemperatur T .

3.4.3.1 Allgemeine Simulationsbedingungen

Alle beschriebenen Simulationen wurden wie allgemein üblich mit periodischen Randbedingungen (*periodic boundary conditions*, PBC) durchgeführt. Das bedeutet, dass die Simulationsbox mit identischen Kopien dieser Box umgeben wird, um Artefakte zu vermeiden, die durch Randeffekte entstehen. Man geht dabei davon aus, dass der Fehler, der durch die artifizielle Periodizität entsteht, geringer ist, als der Fehler, der entstehen würde, wenn man die Box mit Vakuum umgeben würde.

Die Simulationstemperatur wurde mit dem Berendsen-Algorithmus [131] kontrolliert. Temperaturabweichungen werden mit diesem Algorithmus innerhalb eines bestimmten Zeitintervalls korrigiert und die Temperatur wird langsam wieder an die gewählte Simulationstemperatur angeglichen. Da die Simulationsbox als NPT-Ensemble (konstante Anzahl an Atomen N , konstante Temperatur T und konstanter Druck P) behandelt wurde, wurde auch der Druck des Simulationssystems mithilfe des Berendsen-Algorithmus auf die gleiche Weise reguliert. Mit einem NPT-Ensemble versucht man, möglichst realistische Simulationsbedingungen zu schaffen, weshalb die Vergleichbarkeit mit experimentellen Daten hier am ehesten gegeben ist.

3.4.3.2 Wahl des Zeitschrittes

Ein wichtiger Schritt vor der Durchführung einer MDS ist die Wahl eines für die Simulation angemessenen und sinnvollen Zeitschrittes Δt . Ein kleiner Zeitschritt

bedeutet eine nur sehr langsame und unvollständige Abdeckung des Konformationsraumes bei einem extrem hohen Rechenaufwand. Andererseits führt ein zu großer Zeitschritt zu Fehlern bei der Integration [117]. Man sollte daher immer einen Zeitschritt wählen, der etwa ein Fünftel bis ein Zehntel der schnellsten Bewegung im Molekül beträgt. In der Regel limitieren dann die Bindungsstreckschwingungen mit einer Schwingungsperiode von ca. 10 fs den Zeitschritt auf 1 fs. Da diese hochfrequenten Schwingungen aber nur einen sehr geringen Einfluss auf das Gesamtverhalten des Systems haben und zudem nur mit quantenchemischen Methoden wirklich korrekt wiedergegeben werden können [129], ist es sinnvoll, alle Bindungen mithilfe von *constraints* auf ihrer Referenzlänge festzuhalten. In der vorliegenden Arbeit wurde dazu der LINCS-Algorithmus [130] verwendet. Gegenüber dem ebenfalls in GROMACS verfügbaren SHAKE-Algorithmus [132] hat dieser Algorithmus den Vorteil, dass er stabiler und erheblich schneller ist. Der Zeitschritt wurde dadurch von 1 fs auf 2 fs erhöht.

3.5 Analyse des Proteinmodells

3.5.1 Strukturvalidierung mit PROCHECK

Der erste Schritt zur Validierung des erstellten Proteinmodells ist die Evaluierung der stereochemischen Qualität. Bevor spezifische Validierungskriterien wie beispielsweise die Nachvollziehbarkeit von Mutationsstudien oder Ligandbindungsaffinitäten herangezogen werden, sollte zunächst die Zuverlässigkeit der Proteinstruktur als solche sichergestellt sein. Man überprüft daher, ob das Proteinrückgrat und die Seitenketten der Aminosäuren überhaupt in einer sinnvollen, energetisch erlaubten Konformation vorliegen.

Auf der Grundlage von hoch aufgelösten Kristallstrukturen ließen sich für bestimmte stereochemische Parameter allgemeingültige Idealwerte ableiten [133], wie z.B. für die Verteilung der Φ -, ψ - und χ_1 -Torsionswinkel, die als Maß für die strukturelle Qualität gelten. Ein Vergleich dieser Idealwerte mit den entsprechenden Parametern der zu untersuchenden Proteinstrukturen kann mit dem Programmpaket PROCHECK [134] automatisch im Rahmen einer umfangreichen Analyse der stereochemischen Eigenschaften durchgeführt werden. Im Einzelnen werden folgende Werte untersucht:

- Verteilung der Φ - und ψ -Winkel (Ramachandran-Plot)
- Verteilung der χ_1 - und χ_2 -Winkel
- Planarität der Peptidbindungen (ω -Winkel)
- Chiralität des C_α

- Sekundärstruktur und Lösungsmittelzugänglichkeit
- Verteilung von Bindungslängen und Bindungswinkeln
- Wasserstoffbrückenbindungsenergie

Die Ergebnisse werden entweder graphisch dargestellt (Ramachandran-Plot, Sekundärstruktur und Lösungsmittelzugänglichkeit) oder als Zahlenwerte in Form von G-Faktoren angeben. Die G-Faktoren sollten möglichst einen Wert $\geq -0,5$ annehmen.

3.5.2 SURFNET-Kavitäten

Zur Untersuchung von Größe und Form der Bindetasche bzw. des aktiven Zentrums ist das Programm SURFNET [135] geeignet, das alle Hohlräume und Aussparungen eines Proteins innerhalb einer bestimmten Größenordnung berechnet. Das Programm legt Kugeln mit einem vorher gewählten Maximalradius in die Hohlräume und reduziert deren Größe so lange, bis es keine van-der-Waals-Überlappungen zu benachbarten Atomen mehr gibt (Abbildung 3.3). Unterschreitet der Kugelradius einen vorher gewählten Minimalradius, wird die Kugel verworfen. Den Umriss der Kugeln, mit denen die Hohlräume ausgefüllt worden sind, kann man schließlich als Oberfläche graphisch darstellen.

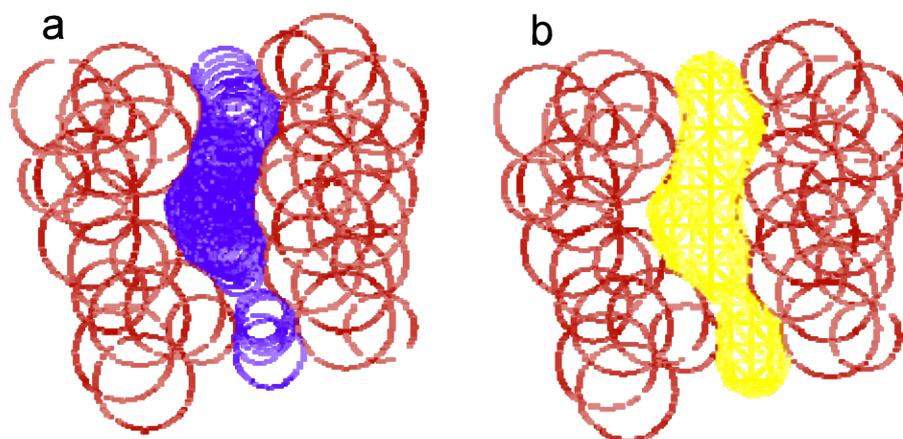


Abbildung 3.3: Berechnung einer SURFNET-Kavität. Die roten Kreise repräsentieren die Atomradien des Proteins, die blauen Kreise stellen die Kugeln dar, mit denen die Kavität ausgefüllt wird. Die resultierende Oberfläche ist gelb dargestellt.

3.5.3 Clusteranalyse mit NMRCLUST

Um eine repräsentative Proteinstruktur aus einer MDS zu erhalten, kann man aus allen Proteinkonformationen aus einem bestimmten Zeitabschnitt der MDS eine Durchschnittsstruktur generieren lassen. Sinnvoller ist es jedoch, die Protein-

konformationen in Cluster einzuteilen und eine Struktur aus der Mitte des größten Clusters zu extrahieren. Eine solche Clusteranalyse kann mit dem Programm NMRCLUST [136] vorgenommen werden.

Zunächst werden alle Konformationen paarweise miteinander überlagert und der RMS-Wert (*root mean square*, mittlere quadratische Abweichung) wird jeweils bestimmt. Im Gegensatz zu anderen Programmen wird die anschließende Clustereinteilung nicht nach einem vorher vom Benutzer als *cut-off* festgelegten RMS-Wert durchgeführt, sondern das Programm errechnet einen als *cut-off* geeigneten RMS-Wert selbst. Zur Ermittlung des *cut-off* werden zwei Parameter gleichzeitig optimiert: so soll sowohl die Anzahl an Clustern als auch die Streuung innerhalb der Cluster so gering wie möglich sein. Bei diesem Verfahren ist die Wahrscheinlichkeit größer, die Konformationen so einzuteilen, wie es für das vorliegende Ensemble am sinnvollsten ist.

3.6 3D-QSAR Untersuchungen

Die dreidimensionale Struktur von Verbindungen eines Datensatzes lässt sich über ein statistisches Modell mit den biologischen Aktivitätsdaten dieser Verbindungen, wie z.B. Bindungsaffinitäten, IC_{50} -Werte usw., an einem Zielprotein in Beziehung setzen. Ein solches 3D-QSAR-Modell dient zur Erklärung der unterschiedlichen Wirksamkeit der einzelnen Verbindungen des Datensatzes und zur Vorhersage der Wirksamkeit unbekannter Strukturen, die nicht Bestandteil des Datensatzes sind.

Voraussetzung für die Erstellung des Modells ist eine Überlagerung der Ausgangsverbindungen. Neben einer direkten Überlagerung der Ausgangsverbindungen (ligandbasierter Ansatz) gibt es auch die Möglichkeit, Informationen über die Struktur des Zielproteins zu nutzen und die Verbindungen mithilfe einer Docking-Prozedur im Protein zu überlagern (proteinbasierter Ansatz) [137], [138]. Die letztgenannte Methode wurde in der vorliegenden Arbeit auf der Grundlage des zuvor erstellten Homologiemodells angewendet. Da sich die von den Docking-Programmen berechneten Wechselwirkungsenergien in der Regel nicht für eine Korrelation mit Aktivitätsdaten eignen, werden stattdessen häufig molekulare Wechselwirkungsfelder als Deskriptoren eingesetzt und in einer PLS-Analyse (s. Abschnitt 3.6.3.2) mit den biologischen Daten korreliert.

3.6.1 Protein-Ligand-Docking

Verfahren zur Vorhersage des Bindungsmodus von Liganden in einer gegebenen Proteinstruktur (Protein-Ligand-Docking) umfassen zwei Komponenten. Dies sind zum einen die Generierung möglicher Protein-Ligand-Komplexe mithilfe eines möglichst effizienten Suchalgorithmus und zum anderen die Bewertung der

resultierenden Komplexe mit dem Ziel, den Komplex mit dem experimentellen Bindungsmodus herauszufiltern [139].

Die Gemeinsamkeit aller neueren Suchalgorithmen ist, dass zwar nur eine Proteinkonformation berücksichtigt werden kann, dass die Flexibilität der Liganden allerdings mit einbezogen wird. Dies erreicht man beispielsweise mit deterministischen Methoden wie der fragment-basierten Methode, bei der die Liganden in einzelne Fragmente zerlegt und die Fragmente inkrementell im Protein wieder zusammengefügt werden [140]. Daneben kommen aber häufig auch stochastische Suchalgorithmen zum Einsatz. Ein Beispiel dafür ist die zufällige Strukturgenerierung mit der Monte-Carlo-Methode in Verbindung mit einer *simulated annealing* Minimierung [141]. In letzter Zeit werden verstärkt genetische Algorithmen zur Komplexgenerierung angewendet, die ebenfalls zu den stochastischen Ansätzen zählen. In diese Kategorie gehört auch das Programm GOLD [142] (Genetic Optimisation for Ligand Docking), das in der vorliegenden Arbeit benutzt wurde.

3.6.1.1 GOLD – Genetischer Docking-Algorithmus

Beim genetischen Algorithmus wird in Anlehnung an die biologische Evolutionstheorie eine bestimmte Anzahl von Ligandkonformationen (Population) durch die Anwendung genetischer Operationen (Migration, Mutation und Crossover) auf den Scoring-Wert (Fitness) optimiert. Die als Gene bezeichneten Variablen der einzelnen Ligandkonformationen sind in diesem Fall die Torsionswinkel der drehbaren Bindungen und die Anzahl der Wasserstoffbrücken des Liganden. Diese Variablen werden im Verlauf von genetischen Operationen zufällig verändert. Dies kann durch eine Mutation geschehen, also eine willkürliche Änderung des Wertes für eine Variable, oder durch ein Crossover, das ist der Austausch des Wertes für eine Variable zwischen zwei Liganden, oder auch durch eine Migration, die einen Wechsel einer Ligandkonformation von einer Subpopulation in eine andere beinhaltet. Der schematische Ablauf der Docking-Prozedur ist in (Abbildung 3.4) dargestellt. Um eine Konvergenz zu erreichen, wird bei der Auswahl der Konformationen nach einer genetischen Operation eine Art Selektionsdruck ausgeübt, der dadurch entsteht, dass die Konformationen mit höherem Scoring-Wert mit einer größeren Wahrscheinlichkeit in die neue Population übernommen werden.

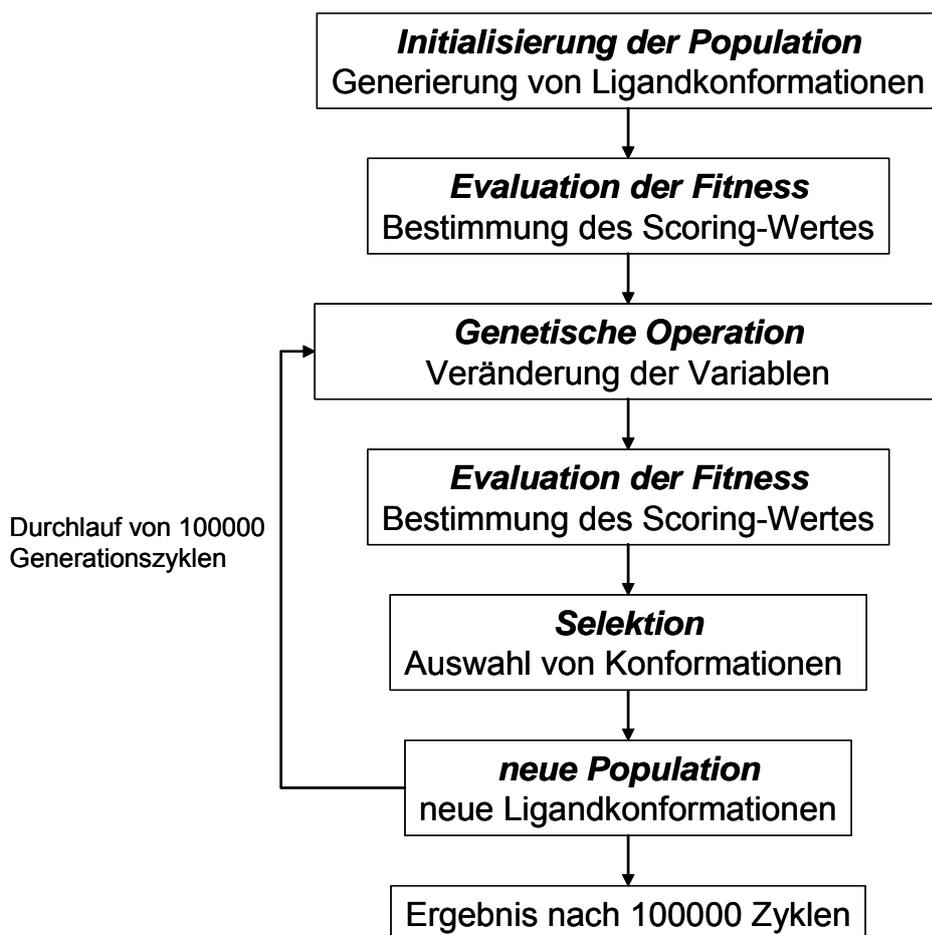


Abbildung 3.4: Schematischer Ablauf des Dockings bei der Anwendung eines genetischen Algorithmus

3.6.1.2 Behandlung von Metallatomen

Die häufig in Proteinen vorkommenden Metallatome Mg, Mn, Zn, Fe und Ca sind in GOLD parametrisiert. Sie können vom Programm erkannt werden, wenn sie in einer tetraedrischen oder in einer oktaedrischen Komplexgeometrie vorliegen. Zur Festlegung der Komplexgeometrie wird die Proteinumgebung des Metallatoms nach Koordinationspartnern abgesucht. Entsprechend der Anordnung der Koordinationspartner werden zum einen die Komplexgeometrie und zum anderen freie Koordinationsstellen, die an der Ligandbindung beteiligt sein können, bestimmt. Metall-Ligand-Wechselwirkungen werden dann als Pseudo-Wasserstoffbrücken behandelt. Folglich binden Metallatome bevorzugt an Wasserstoffbrücken-Akzeptoren und konkurrieren mit den Wasserstoffbrücken-Donatoren des Proteins um die Bindung.

3.6.1.3 Scoring-Funktionen

Zur Bewertung der resultierenden Konformationen während des Dockings und der endgültigen Rangfolge der Konformationen nach dem Docking stehen in GOLD zwei empirische Scoring-Funktionen zur Verfügung.

Der speziell für das Programm entwickelte **GOLDScore** kommt durch die Summe aus drei Energietermen zustande, die für die Ligandbindung wichtig sind. Dies sind die Wasserstoffbrückenbindungsenergie (H_Bond_Energy), die intermolekulare Dispersionsenergie zwischen Protein und Ligand ($Complex_Energy$), und die intramolekulare Torsions- und Dispersionsenergie des Liganden ($Internal_Energy$).

$$Fitness = -H_Bond_Energy - (Internal_Energy + Complex_Energy) \quad (3.13)$$

Die Berechnung der einzelnen Terme basiert auf empirischen Parametern wie Wasserstoffbrückenbindungsenergien, Atomradien, Polarisierbarkeiten usw.

Im Gegensatz dazu wird mit der ebenfalls implementierten **ChemScore**-Funktion [143] die physikalische Bindungsaffinität der Protein-Ligand-Komplexe vorhergesagt. Die freie Bindungsenergie setzt sich dabei aus mehreren Termen zusammen, welche die Bindungsaffinität ausmachen und auch physikalisch interpretierbar sind.

$$\Delta G_{binding} = \Delta G_0 + \Delta G_{hbond} + \Delta G_{metal} + \Delta G_{lipo} + \Delta G_{rot} \quad (3.14)$$

Die einzelnen ΔG -Koeffizienten für die Wasserstoffbrückenbindungsenergie ΔG_{hbond} , die Metall-Ligand-Bindung ΔG_{metal} , hydrophobe Wechselwirkungen ΔG_{lipo} und ein Strafterm für Ligandflexibilität ΔG_{rot} wurden empirisch ermittelt. Dazu wurde eine lineare Regressionsanalyse mit einem Trainingsdatensatzes von 82 kristallographisch aufgeklärten Protein-Ligand-Komplexen durchgeführt, deren Bindungsaffinitäten bekannt sind.

3.6.2 GRID- Wechselwirkungsfelder

Molekulare Wechselwirkungsfelder geben die Wechselwirkungsenergie zwischen einer zu untersuchenden Molekülstruktur und einer virtuellen Sonde mit bestimmten physikochemischen Eigenschaften wieder. Im Programm GRID [144], [145], [146], [147] wird das Molekül in einen dreidimensionalen Gitterkäfig eingesetzt, und die Wechselwirkungsenergie an jedem einzelnen Gitterpunkt E_{xyz} wird mit einer empirischen Energiefunktion berechnet.

$$E_{xyz} = \sum E_{lj} + \sum E_{el} + \sum E_{hb} \quad (3.15)$$

Die Funktion setzt sich aus einem Term für van der Waals-Wechselwirkungen E_{lj} , einem elektrostatischen Term E_{el} und einem Term für die Wasserstoffbrückenbindung E_{hb} zusammen. Die Sonden, mit denen die Gitterpunkte abgefahren werden, repräsentieren Atomtypen, funktionelle Gruppen oder kleine Moleküle, die für die Protein-Ligand-Bindung von Interesse sind.

Prinzipiell ergeben sich zwei Anwendungsmöglichkeiten für die Wechselwirkungsfelder. Einerseits können sie mit verschiedenen Sonden für eine Proteinstruktur zur

Charakterisierung des aktiven Zentrums oder der Bindetasche berechnet werden. Andererseits dienen sie, wie bereits oben erwähnt, in Verbindung mit dem Programm GOLPE (s. Abschnitt 3.6.3) als Deskriptoren für eine Korrelation mit biologischen Daten. Dann müssen die Wechselwirkungsfelder für alle Liganden des Datensatzes berechnet werden, die zuvor in ihrer bioaktiven Konformation überlagert worden sind. Ein wichtiger Schritt hierbei ist die Auswahl einer oder mehrerer charakteristischer Sonden, welche die Eigenschaften der einzelnen Liganden, auf denen ihre unterschiedliche Wirkstärke begründet ist, möglichst gut wiedergeben.

3.6.3 Statistische Auswertung mit GOLPE

Die berechneten molekularen Wechselwirkungsfelder liefern mehrere hundert oder tausend X-Variablen in Form von Wechselwirkungsenergien an den einzelnen Gitterpunkten, die nun mit jeweils einer einzigen Y-Variable, der biologischen Aktivität, in einem 3D-QSAR-Modell korreliert werden sollen. Dieses statistische Problem lässt sich mithilfe des Programmpakets GOLPE (Generating Optimal Linear PLS Estimation) [148] lösen, das speziell für die Erstellung von 3D-QSAR-Modellen entwickelt wurde.

Der erste wichtige Schritt ist eine Vorbehandlung der Daten, bei der zunächst das Untergrundrauschen entfernt wird. Außerdem werden Variablen eliminiert, die fast ausschließlich denselben Wert annehmen oder nur eine geringe Standardabweichung aufweisen. Diese Variablen haben nur eine geringe Aussagekraft und können sogar das Ergebnis verfälschen.

3.6.3.1 Hauptkomponentenanalyse

Eine Analyse der X-Variablen alleine wird mittels einer Hauptkomponentenanalyse (Principal Component Analysis, PCA) durchgeführt. Hierbei wird die Information der X-Variablen in Hauptkomponenten (PC) zusammengefasst, und damit in eine Form gebracht, die sich wesentlich leichter auswerten lässt als die große Menge der Ursprungsdaten. Die Hauptkomponenten werden mit absteigendem Informationsgehalt extrahiert, so dass die erste Komponente meist schon einen Großteil der Unterschiede zwischen den einzelnen Liganden erklärt. Sie stehen außerdem senkrecht aufeinander, es gibt also keine Korrelation zwischen den einzelnen Hauptkomponenten.

Sinnvoll ist eine PCA, weil man mit ihrer Hilfe die Verteilung der untersuchten Liganden im Raum darstellen, und somit Ausreißer detektieren kann. Auch die Verteilung der einzelnen X-Variablen kann man sich anzeigen lassen und dadurch prüfen, welchen Beitrag sie zur Beschreibung der Liganden leisten.

3.6.3.2 PLS-Analyse

Die PLS-Analyse (*Partial Least Squares*) ist eine Regressionstechnik, die es ermöglicht, eine abhängige Y-Variable mit einer großen Anzahl miteinander korrelierender X-Variablen zu beschreiben. Sie ist deshalb die einzige Regressions-technik, die zur Korrelation der GRID-Wechselwirkungsenergien der einzelnen Liganden mit ihrer biologischen Aktivität geeignet ist.

Ähnlich der PCA (s.o.) werden aus den X-Variablen wenige Vektoren extrahiert, die hier latente Variablen genannt werden und den Hauptkomponenten der PCA entsprechen. Sie stehen ebenfalls orthogonal zueinander und auch hier hat die erste latente Variable den höchsten Informationsgehalt. Der einzige wesentliche Unterschied ist, dass die Hauptkomponenten nur die Struktur der X-Matrix repräsentieren und die latenten Variablen die Struktur der X- und der Y-Matrix.

Die Vorhersagekraft eines PLS-Modells wird mithilfe einer **Kreuzvalidierung** überprüft. Eine Kreuzvalidierung dient aber nicht nur dazu, die generelle Qualität des Modells zu evaluieren, sondern sie wird auch durchgeführt, um die optimale Anzahl an latenten Variablen zu ermitteln, die in das Modell einbezogen werden sollen. Man entnimmt dem Datensatz eine bestimmte Anzahl an Verbindungen, erstellt mit den verbliebenen Verbindungen ein neues PLS-Modell und sagt damit die Aktivität der ausgelassenen Verbindungen voraus. Ein strengeres Validierungskriterium als die herkömmliche *leave one out* Methode (Auslassen einer Verbindung) ist die *random groups* Methode, bei der die Verbindungen zufällig in fünf Gruppen eingeteilt werden, und jede Gruppe einmal entnommen und vorhergesagt wird. Diese Prozedur wird zwanzigmal wiederholt.

Als Maß für die Vorhersagekraft eines PLS-Modells dienen der Korrelationskoeffizient q^2 und die Standardabweichung des Fehlers *SDEP* der Vorhersage.

$$q^2 = 1 - \left[\frac{\sum (Y - Y')^2}{\sum (Y - \bar{Y})^2} \right] \quad (3.16)$$

$$SDEP = \sqrt{\sum \frac{(Y - Y')^2}{N}} \quad (3.17)$$

Y steht für die experimentelle, Y' für die vorhergesagte und \bar{Y} für die durchschnittliche Aktivität. N ist die Anzahl der Liganden. Ein zuverlässiges Modell zeichnet sich durch $q^2 > 0,6$ und eine möglichst kleine Standardabweichung aus.

3.6.3.3 Variablenselektion

Die Besonderheit des Programms GOLPE besteht darin, dass dort Methoden zur Variablenselektion implementiert sind, die es ermöglichen, zwischen wichtigen X-Variablen und solchen, die nicht in Zusammenhang mit der biologischen Aktivität der Verbindungen stehen und dadurch die Vorhersagekraft des Modells sogar verschlechtern, zu unterscheiden [149], [150].

Geht man davon aus, dass sich Unterschiede in der Molekülstruktur nicht in der Veränderung einer einzelnen X-Variable äußern, sondern eher in der Veränderung einer ganzen Gruppe benachbarter X-Variablen, so ist es sinnvoll, die X-Variablen in Gruppen mit gleichen strukturellen und statistischen Informationen einzuteilen. Das kann man mit einer *smart region definition* [151] erreichen. Der Vorteil der Gruppen besteht darin, dass sich der Effekt einer einzelnen Variablen leichter ermitteln lässt, wenn man alle Variablen gleichzeitig betrachtet, die dieselbe Information beinhalten. Die Gefahr, dass eine ganze Gruppe von Variablen falsch bewertet wird, ist relativ gering.

Um die Relevanz einzelner Variablen oder Variablengruppen zu ermitteln, kann man ein *fractional factorial design (FFD selection)* durchführen. Einzelne X-Variablen oder Variablengruppen werden dabei weggelassen, und so werden für verschiedene Variablenkombinationen PLS-Modelle aufgestellt und der Einfluss der einzelnen Variablen auf das PLS-Modell getestet. Anschließend werden alle Variablen eliminiert, die die Vorhersagekraft des Modells negativ beeinflussen. Nach diesem Selektionsschritt haben sich die Qualität und die Interpretierbarkeit des 3D-QSAR-Modells in der Regel deutlich verbessert.

4 DURCHFÜHRUNG UND ERGEBNISSE

4.1 Konstruktion des Homologiemodells

4.1.1 Auswahl der Templatstruktur

In der vorliegenden Arbeit wurde die erste publizierte Kristallstruktur eines vollständigen Säugetiercytochroms im substratgebundenen Zustand (PDB-Code 1N6B) für die Konstruktion eines Homologiemodells von CYP1A2 verwendet. Es handelt sich hierbei um die bereits in Abschnitt 1.1 ausführlich vorgestellte Struktur des Kaninchencytochroms CYP2C5 im Komplex mit dem Sulfaphenazol-Derivat DMZ (4-Methyl-*N*-methyl-*N*-(2-phenyl-2*H*-pyrazol-3-yl)benzensulfonamid) [84].

Im Vergleich mit den kurze Zeit später veröffentlichten weiteren Säugetiercytochrom-Strukturen (vgl. Tabelle 1.6) stellte sich 1N6B als gut geeignetes Templat für die Modellierung von CYP1A2 heraus. Das wichtigste Kriterium für die Auswahl einer Templatstruktur ist die Sequenzidentität zwischen der Modellsequenz CYP1A2 und den Aminosäuresequenzen der vorhandenen Kristallstrukturen. Paarweise automatische Alignments von jeder Kristallstruktur-Sequenz mit der CYP1A2-Sequenz, die mithilfe der Routine „BLAST 2 Sequences“ der WWW-Seite des NCBI [100] erstellt wurden, gaben zwischen CYP1A2 und CYP3A4 eine Sequenzidentität von 23 % an, während die Sequenzidentität zu den Cytochromen der Familie 2 (CYP2B4, CYP2C5, CYP2C8, CYP2C9) einheitlich bei 29 % lag. Damit erwiesen sich die Kristallstrukturen von CYP3A4 für die Modellierung von CYP1A2 als weniger gut geeignete Template. CYP3A4 ist zudem auch entwicklungsgeschichtlich wesentlich weiter von CYP1A2 entfernt als die Cytochrome der Familie 2, was wiederum auf eine geringere strukturelle Ähnlichkeit schließen lässt (vgl. Abbildung 1.1).

Auch die Konformation der Templatstruktur ist für den Modellbauvorgang von Bedeutung. Zwar kann sich im Verlauf einer MDS die Konformation eines Modells ändern und anpassen, eine Bewegung von einer offenen in eine substratgebundene geschlossene Konformation ist aber sicher im Rahmen einer MDS nicht möglich. Außerdem kann eine Equilibrierung umso schneller erreicht werden, desto ähnlicher die Ausgangskonformation des Templats der Konformation des Modells ist. Damit erweisen sich die Strukturen 1PO5 (CYP2B4 in der offenen Konformation), 1PQ2 (CYP2C8 mit Palmitinsäure im Substratkanal gebunden), 1OG2 (CYP2C9 mit Wasser als sechstem Hämliganden) und 1OG5 (CYP2C9 mit Warfarin in einer Seitentasche gebunden) ebenfalls als ungeeignet. Die Struktur 1SUO ist ein ungünstiges Templat, weil hier der Inhibitor Chlorophenylimidazol koordinativ an das Hämeisen gebunden ist. Die koordinative Bindung eines Inhibitors wirkt sich auch auf die Gesamtkonformation des Enzyms aus, vor allem auf die Konformation der Helix I, die das katalytische Zentrum enthält. Übrig bleiben somit die gewählte Kristallstruktur 1N6B, sowie die PDB-Einträge 1NR6 (CYP2C5 im Komplex mit Diclofenac) und

1R9O (CYP2C9 im Komplex mit Flurbiprofen), die beide eine aromatische Hydroxylierung katalysieren und abgesehen von geringen Unterschieden in der Auflösung alle gut als Templat für die Erstellung eines Modells von CYP1A2 geeignet sind.

4.1.2 Sekundärstrukturvorhersagen

Bei der Erstellung und Optimierung eines manuellen Sequenz-Alignments ist die Überlagerung der sich entsprechenden Sekundärstrukturen ein wichtiges Kriterium. Deshalb sollte zuerst eine Sekundärstrukturvorhersage für die Sequenz von CYP1A2 durchgeführt werden, um diese Informationen später in das Alignment zu übernehmen zu können. Zunächst wurden verschiedene Vorhersageprogramme, die über den WWW-Meta-Server PredictProtein [152] zugänglich sind, anhand der Kristallstruktur auf ihre Vorhersagekraft hin getestet. Dazu wurden die einzelnen Vorhersagen mit den Angaben aus der PDBsum-Datenbank [153] über die tatsächliche Lage der Sekundärstrukturen im Kristall (Abbildung 4.5) verglichen.

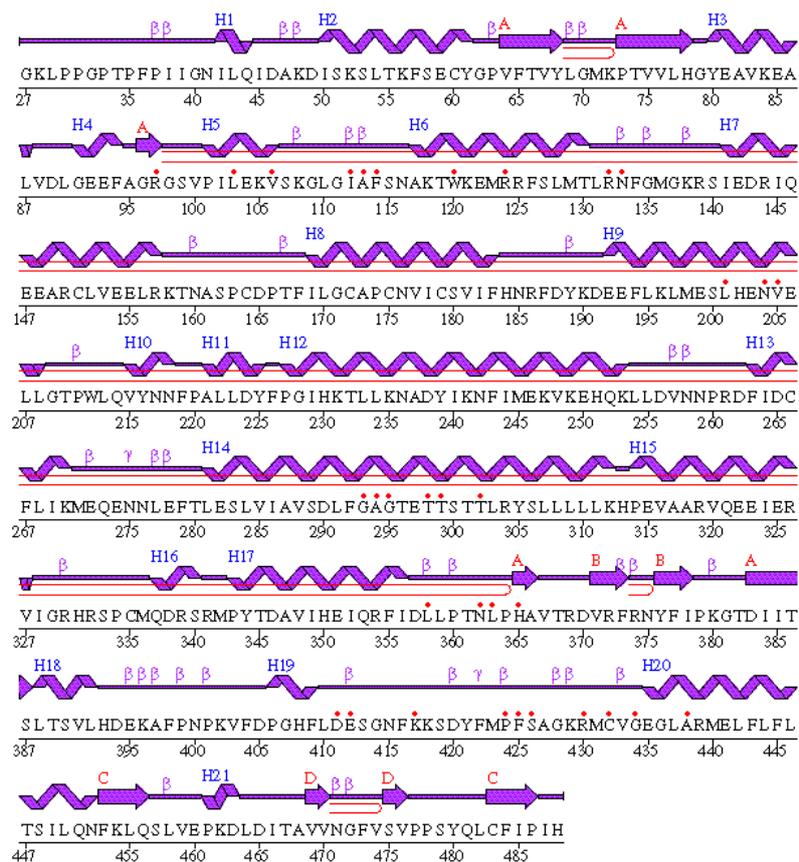


Abbildung 4.5: Sekundärstruktur der Kristallstruktur 1N6B nach dem Eintrag in der PDBsum-Datenbank. Die mit H1-H21 durchnummerierten Spiralen sind die Helices, die mit den Buchstaben A-D gekennzeichneten Blockpfeile stellen β -Faltblätter dar. Turnstrukturen sind durch ein „ β “ bzw. „ γ “ markiert.

Das Programm PSIPRED lieferte dabei die besten Ergebnisse und wurde für diese Arbeit weiter verwendet (Abbildung 4.6).

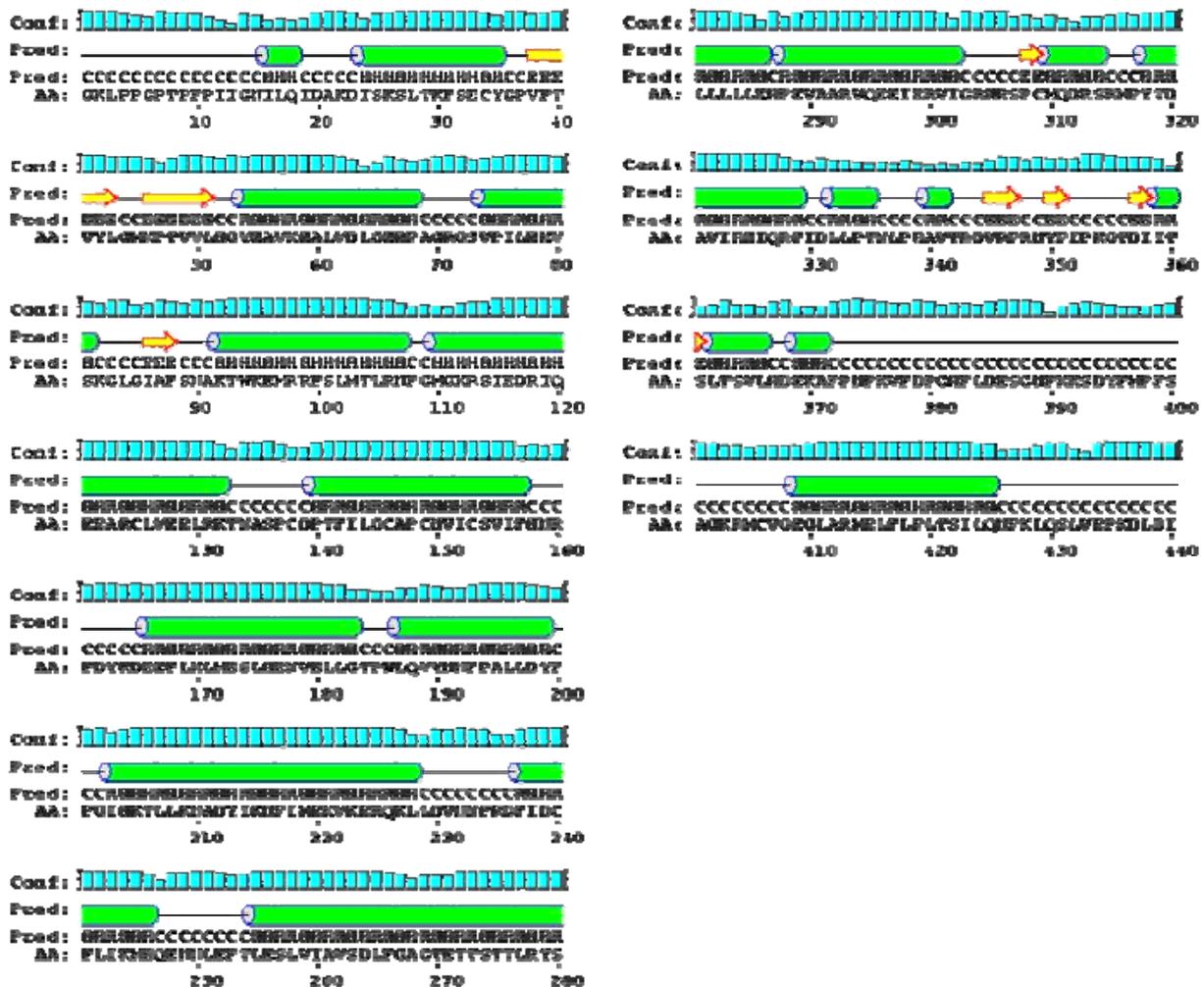


Abbildung 4.6: Sekundärstrukturvorhersage nach PSIPRED für die Kristallstruktur 1N6B. AA: Aminosäuresequenz der Kristallstruktur. Pred: Sekundärstrukturvorhersage; Helices (H) sind als grüne Zylinder dargestellt, β -Faltblätter (E) als gelbe Blockpfeile. Die türkis gefärbten Histogrammsäulen (Conf) geben jeweils die Zuverlässigkeit für die Vorhersage an dieser Stelle an.

Abbildung 4.7 zeigt einen Vergleich der Sekundärstrukturen nach Angabe der PDBsum-Datenbank mit den von PSIPRED vorhergesagten Sekundärstrukturen. In einigen Fällen werden β -Turns als helikal vorhergesagt. Sie sind in der Abbildung mit „ β “ gekennzeichnet. Beim Übergang von Helix B in Helix B', sowie beim Übergang von Helix F' zu Helix G' werden durchgehende Helices vorhergesagt. Die ungenaue Vorhersage für den Bereich von Helix F bis Helix G lässt sich dadurch erklären, dass sich hier der Eingang zum Substratkanal befindet. Dieser hochflexible Bereich legt sich im geschlossenen Zustand wie ein Deckel über das aktive Zentrum und ist im offenen Zustand des Enzyms nach oben geklappt (Abbildung 1.25). In einigen Kristallstrukturen fehlt diese Region ganz, in einigen anderen sind Helix F' oder Helix G' nicht eindeutig als Helix zu erkennen. Da das Programm PSIPRED wie bereits erläutert (Abschnitt 3.1.2) auf einem neuronalen Netz beruht, also aus vorhandenen Kristallstrukturen „lernt“, kann für diese Proteinregion auch keine eindeutige

Vorhersage getroffen werden. Helix B' ist ebenfalls in der Mehrzahl der Kristallstrukturen nicht als Helix ausgebildet, nur eine Helix B'', welche die SRS⁶ bildet, ist immer vorhanden. Dies erklärt, dass auch die Sekundärstruktur dieses Sequenzabschnittes nur ungenau vorhergesagt werden kann. Bei zwei weiteren kurzen Helices, die in der Vorhersage völlig fehlen, handelt es sich nicht um klassische α -Helices, sondern um 3_{10} -Helices, die auch in den Kristallstrukturen nur selten deutlich ausgeprägt sind.

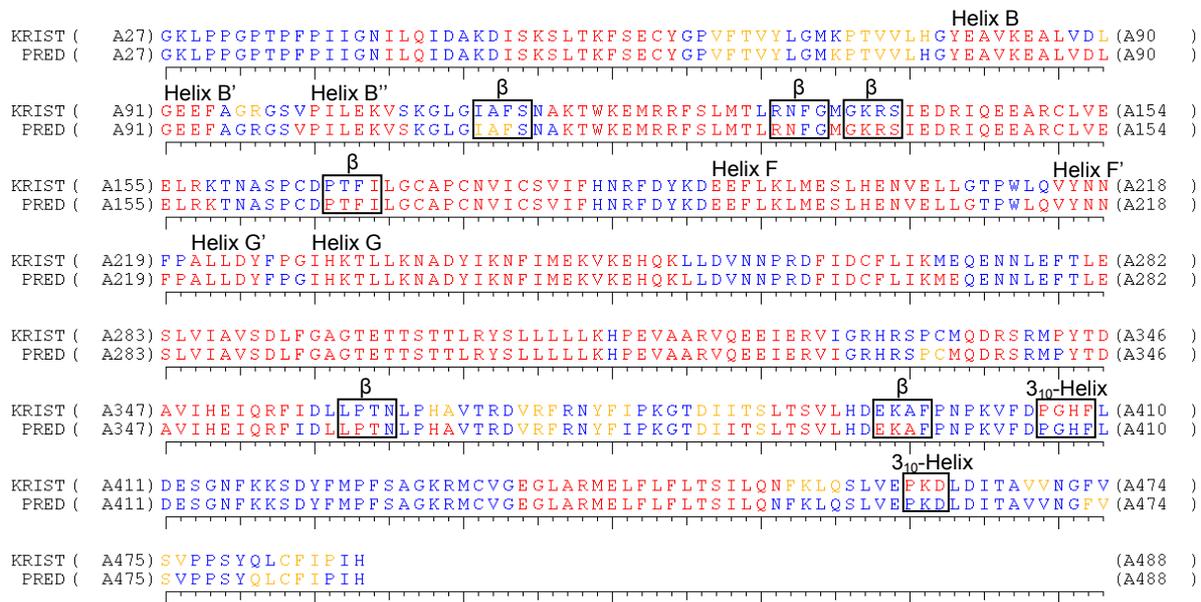


Abbildung 4.7: Vergleich der Sekundärstruktur der Kristallstruktur 1N6B nach Angabe der PDBsum (KRIST) mit der Sekundärstrukturvorhersage des Programms PSIPRED (PRED). Aminosäuren, die zu einer Helix gehören, sind rot gefärbt, Aminosäuren mit β -Faltblattstruktur gelb. β -Turns, die von PSIPRED ganz oder teilweise als Helices vorhergesagt werden, sind mit „ β “ gekennzeichnet.

Von den elf im Enzym vorhandenen β -Faltblattsträngen werden lediglich drei von PSIPRED nicht gefunden und zwei weitere werden fälschlicherweise als helikal vorhergesagt. Alle Haupthelices sind in der Vorhersage vorhanden und ihre Länge unterscheidet sich meist nur um ein bis zwei Aminosäuren von der PDBsum-Angabe. Insgesamt wird durch den Vergleich deutlich, dass die Sekundärstrukturvorhersage in den meisten Fällen sehr gut mit den Angaben aus der PDBsum übereinstimmt. Somit kann man davon ausgehen, dass auch die Sekundärstruktur des Targets CYP1A2 zuverlässig vorhergesagt wird (Abbildung 4.8).

⁶ SRS: Substraterkennungsregion (*Substrate Recognition Site*), vgl. (Einleitung)

3. alle Säugetiersequenzen der Familie CYP 2 und CYP1A2 (98 Sequenzen)
4. alle Säugetiersequenzen der Familie CYP1 und CYP2C5 (18 Sequenzen).

Tabelle 4.4 zeigt die Sequenzen aus der SwissProt-Sequenzdatenbank [18], die in die multiplen Alignments einbezogen wurden.

Tabelle 4.4: Übersicht über die für die multiplen Sequenzalignments verwendeten Sequenzen

	Familie	Anzahl	
1. humane Sequenzen	CYP1	3	
	CYP2	13	
	CYP3	5	
	CYP4	7	
	CYP7-CYP51	19	gesamt: 47
	Unterfamilie	Anzahl	
2.-4. Säugetiersequenzen	CYP1A	14	
	CYP1B	3	gesamt: 17
	CYP2A	13	
	CYP2B	12	
	CYP2C	34	
	CYP2D	19	
	CYP2E	8	
	CYP2F	4	
	CYP2G	2	
	CYP2J	5	gesamt: 97

Bei der Auswertung der Alignments wurde darauf geachtet, dass möglichst wenige Sekundärstrukturen durch Lücken im Alignment unterbrochen wurden, und dass Lücken in potentiellen Loop-Regionen an Stellen mit niedriger Homologie eingefügt wurden. Alle vier Sequenzkombinationen lieferten gute Alignments. Die besten Ergebnisse erhält man jedoch, wenn man entweder ein breites Spektrum von verschiedenen Informationen in das Alignment einfließen lässt (Möglichkeit 1), oder möglichst viele Informationen aus der Familie des Targets und des Templats (Möglichkeit 2). Das zuletzt genannte Alignment wurde als Grundlage für den weiteren Modellbau verwendet und ist auszugsweise in Abbildung 4.10 dargestellt. Aus Gründen der Übersichtlichkeit sind hier nur die Sequenzen der Kristallstrukturen und je ein Vertreter aus jeder humanen CYP-Unterfamilie abgebildet.

Bei diesem Alignment ist nur eine einzige Helix, und zwar die Helix G, durch eine Lücke unterbrochen. Alle anderen Lücken sind in Loop-Regionen, und auch dort sind nur wenige manuelle Verbesserungen des Alignments nötig.

```

sp|P10632|CPC8_HUMAN-----MEPFVVLVLCCLFMLLFLSWRQSCRRLKPPGPTPLPII
sp|P11712|CPC9_HUMAN-----MDSLVLVLCCLCLLLLSLWRQSSGRKLPKPPGPTPLPVI
sp|P33261|CPCJ_HUMAN-----MDPFVVLVLCCLCLLLLSLWRQSSGRKLPKPPGPTPLPVI
sp|P00179|CPC5_RABIT-----MDPVVVLVLCCLCLLLLSLWKQNSGRKLPKPPGPTPLPPII
sp|P05181|CPE1_HUMAN-----MSALGVTVALLVWAAFLLLVSMMRWQVHSSWNLLPPEFLPII
sp|P24903|C2F1_HUMAN-----MDSISTAILLLLLLALVCLLLTSSRDKGKLPKPPGPRPLSIL
sp|P11509|CPA6_HUMAN-----MLASGMLLVALLVCLTMVVLMSVWQQRKSKGKLPKPPGPTPLPFI
sp|P00178|CPB4_RABIT-----MEFSLLLLAFLAGLLLLLLFR-GHPKAHGRLLPPGPSPLPVL
sp|P20813|CPB6_HUMAN-----MELSVLLFLALLTGLLLLLVQ-RHPNTHDRLLPPEPRPLPLL
sp|P51589|CPJ2_HUMAN-----MLAAMGSLAAALWAVVHPRITLLLTGTVAFLLAADFLKRRRPNKPYPPGPWRLLPFL
sp|P10635|CPD6_HUMAN-----MGLEALVPLAVIVAIFLLLVDLMHRRQRWAARYPPGGLPLPGL
sp|P04798|CP11_HUMAN-----MLFPISMSATEFLLASVIFCLVFWVIRASRPQVPKGLKNPPGPWGWPLI
sp|P05177|CP12_HUMAN-----MALSQSVFSAATELLASAIIFCLVFWVLKGLRPRVKGLKSPPEPWGPLL
sp|Q16678|CP1B_HUMANMGTSLSFNDFWPLNPLSIQQTLLLLLSVLATVHVGGQRLLRQRRLRSAPPGEFAWPLI

sp|P10632|CPC8_HUMANGNMLQIDVKDICKSFTNFSKVYGPVFTVYFGMNPVVVHGYEAVKEALIDNGEEFSGRGN
sp|P11712|CPC9_HUMANGNILQIGIKDISKSLTNLSKVYGPVFTLYFGLKPIVVLHGYEAVKEALIDLGEFFSGRGI
sp|P33261|CPCJ_HUMANGNILQIDIKDVSKSLTNLSKIYGPVFTLYFGLERMVVLHGYEVVKEALIDLGEFFSGRGI
sp|P00179|CPC5_RABITGNILQIDAKDISKSLTKFSECYGPVFTVYLGMPKPTVVLHGYEAVKEALVDLGEFFAGTGS
sp|P05181|CPE1_HUMANGNLFLQELKNIKPSFTRLAQRFQVFTVYVGSQRMVVMHGYKAVKEALLDYKDEFSGRGD
sp|P24903|C2F1_HUMANGNLLLLCSQDMLTSLTKLSKEYGSMYTVHLGPRRVVLSGYQAVKEALVDQGEFFSGRGD
sp|P11509|CPA6_HUMANGNYLQLNTEQMYNSLMKISERYGPVFTIHLGPRRVVLCGHDAVREALVDQAEFFSGRGE
sp|P00178|CPB4_RABITGNLLQMDRKGKLLRSFLRLREKYGDVFTVYLGSEPVVLCGTDAREALVDQAEAFSGRGK
sp|P20813|CPB6_HUMANGNLLQMDRRGKLLRSFLRFREKYGDVFTVHLGPRRVVLCGVAIREALVDKAEAFSGRGK
sp|P51589|CPJ2_HUMANGNFFLVDFEQSHLEVQLFVKYGNLPSLELGDISAVLITGLPLIKEALIHMDQNFGRNPV
sp|P10635|CPD6_HUMANGNLLHVDFQNTPYCFDQLRRRFGDVFSLQLAWTPVVVVLNGLAAVREALVTHGEDTADRPP
sp|P04798|CP11_HUMANGHMLTLGK-NPHLALSRRMSQQYGDVLRIGSTPVVVLSGLDTIRQALVRQGDFFKGRPD
sp|P05177|CP12_HUMANGHVLTGK-NPHLALSRRMSQRYGDVLRIGSTPVLVLSRLDTIRQALVRQGDFFKGRPD
sp|Q16678|CP1B_HUMANGNAAAVGQ-AAHLSFARLARRYGDVLRIGSCPIVVLNGERAIHQALVQQGSFAFADRA

sp|P10632|CPC8_HUMANSPISQRITKGL----GIISNGKRWKEIRRFSLTTLRNF GMGKRS-----IEDRVQEE
sp|P11712|CPC9_HUMANFPLAERANRGF----GIVFSNGKWKKEIRRFSLMTRLNFGMGKRS-----IEDRVQEE
sp|P33261|CPCJ_HUMANFPLAERANRGF----GIVFSNGKWKKEIRRFSLMTRLNFGMGKRS-----IEDRVQEE
sp|P00179|CPC5_RABITVPILKRVSKGL----GIAPSNAKTWKEMRRFSLMTRLNFGMGKRS-----IEDRIQEE
sp|P05181|CPE1_HUMANLEPAFHHRDR----GIIFNNGPTWKDIRRFSLTTLRNYGMGKQG-----NESRIQRE
sp|P24903|C2F1_HUMANYPAFNFNFTKGN----GIAPSSGDRWKVLRQFSIQILRNF GMGKRS-----IEERILEE
sp|P11509|CPA6_HUMANQATFDWVFKGY----GVVFSNGERAKQLRRFSIATLRDFGVGKRG-----IEERIQEE
sp|P00178|CPB4_RABITIAVVDPIFQGY----GVIFANGERWRALRRFSLATMRDFGMGKRS-----VEERIQEE
sp|P20813|CPB6_HUMANIAMVDPFFRGY----GVIFANGNRWVLRFRFSVTMRDFGMGKRS-----VEERIQEE
sp|P51589|CPJ2_HUMANTPMREHIFKKN----GLIMSSGQAWKEQRRFSLTALRNFGLGKKS-----LEERIQEE
sp|P10635|CPD6_HUMANVPITQILGFGPRSQQGVFLARYGPVWREQRRFSVSTLRNLGLGKKS-----LEQWVTEE
sp|P04798|CP11_HUMANLYTFTLLSNGQ--SMSFSPDSGPVWAARRRLAQNGLKSFSIASDPASSSSCYLEHVSKE
sp|P05177|CP12_HUMANLYTSTLITDQG--SLTFSQDSGPVWAARRRLAQNALNTPFSIASDPASSSSCYLEHVSKE
sp|Q16678|CP1B_HUMANFASF RVVSGGR---SMAFG-HYSEHWKVQRRRAHSMRNFPTRQP---RSRQVLEGHVLS

sp|P10632|CPC8_HUMANAHCLVEELRKT KASP--CDPTFILGCAPCNVICSVVVFQKRFDYKQDFLTLMKRFNENFR
sp|P11712|CPC9_HUMANARCLVEELRKT KASP--CDPTFILGCAPCNVICSIIFHKRFDYKQDFLNLMEKLNENIK
sp|P33261|CPCJ_HUMANARCLVEELRKT KASP--CDPTFILGCAPCNVICSIIFQKRFDYKQDFLNLMEKLNENIR
sp|P00179|CPC5_RABITARCLVEELRKTNASP--CDPTFILGCAPCNVICSVIFHNRFDYKDEFFLKLMSLNENVR
sp|P05181|CPE1_HUMANAHFLLALRKTQGP--FDPTFILGCAPCNVIADILFRKHF DYNDEKFLRLMYLFNENFH
sp|P24903|C2F1_HUMANGSFLLADV RKTGEPE--FDPTFVLSRSVSNII CSVLF GSRFDYDDELLTII RLINDNFQ
sp|P11509|CPA6_HUMANAGFLIDAHRTGGAN--IDPTFFLSRTVSNVSIIVFGDRFDYKDEFFLSLLRMLLGI FQ
sp|P00178|CPB4_RABITARCLVEELRKS K GAL--LDNTLLFHSITSNII CSIVFGKRF DYKDPVFLRLLLDFFQSFS
sp|P20813|CPB6_HUMANAQCLIEELRKS K GAL--MDPTFLFQSITANII CSIVFGKRFHYQDFLKLMLNLFYQTF S
sp|P51589|CPJ2_HUMANAQHLTEAIKEENGQF--FDPHF KINNAVSNIICSI TFGERFEYQDSWFQQLLKLLEDEVTY
sp|P10635|CPD6_HUMANAAACCAAFANHSGRP--FRNGLLDKAVSNVIASLTCGRF FEYDDPRFLRLLLDAQEGLK
sp|P04798|CP11_HUMANAEVLISLQELMAGEGHFNPYRVVSVTNVICAI CFGRYDHNHQELLSLVNLNPNFGE
sp|P05177|CP12_HUMANAKALISRLQELMAGEGHFDPNYQVVSVANVIGAMCFGQHFPESSDEMLSLVKNTHFVE
sp|Q16678|CP1B_HUMANARELVALLVRGSDGAF LDRPRLTVAVANVMSAVCFGC RYSHDDPEFRELLSHNEEFGR

```

Abbildung 4.9: Auszug aus dem multiplen Sequenzalignment aller Säugetiersequenzen der Familien CYP1 und CYP2, erstellt mit CLUSTAL W. Aus Gründen der Übersichtlichkeit sind nur einige repräsentative Sequenzen abgebildet. Dies sind in der Reihenfolge von oben nach unten: CYP2C8 human, CYP2C9 human, CYP2C19 human, CYP2C5 Kaninchen, CYP2E1 human, CYP2F1 human, CYP2A6 human, CYP2B4 Kaninchen, CYP2B6 human, CYP2J2 human, CYP2D6 human, CYP1A1 human, CYP1A2 human und CYP1B1 human.

sp P10632 CPC8_HUMAN	ILNSPWIQVCNNFPL----LIDCFPGTH----NKVLKNVALTRSYIREKVKHQASLDVN
sp P11712 CPC9_HUMAN	ILSSPWIQICNNFSP----IIDYFPGTH----NKLLKNVAFMKSYILEKVKHQESMDMN
sp P33261 CPCJ_HUMAN	IVSTPWIQICNNFPT----IIDYFPGTH----NKLLKNLAFMESDILEKVKHQESMDIN
sp P00179 CPC5_RABIT	ILSSPWLQVYNNFPA----LLDYFPGIH----KTLLKNADYIKNFIMEKVKHEKLLDVN
sp P05181 CPE1_HUMAN	LLSTPWLQLYNNFPS----FLHYLPGSH----RKVIKNVAEVKEYVSEKVKHHQSLDPN
sp P24903 C2F1_HUMAN	IMSSPWGELYDILDRFPFSLLDWVPGPH----QRIFQNFKCLRDLIAHSVHHDQASS---
sp P11509 CPA6_HUMAN	FTSTSTGQLYEMFSS----VMKHLPGPQ----QQAFQLLQGLEDFIAKKVEHNQRTLDPN
sp P00178 CPB4_RABIT	LISFSSQVPELFPG----FLKHPGTH----RQIYRNLQEIINTFIGQSVEKHRATLDPS
sp P20813 CPB6_HUMAN	LISVFGQLPELFSG----FLKYFPGAH----RQVYKNLQEIINAYIGHSVKHKRETLDPS
sp P51589 CPJ2_HUMAN	LEASKTCQLYNVFP----WIMKFLPGPH----QTLFSNWKLLKLFVSHMIDKHKRDWNP
sp P10635 CPD6_HUMAN	EESGFLREVLNVP----VLLHIPALA----GKVLRFQKAFLTQLDELLEHRTWDPA
sp P04798 CP11_HUMAN	VVGS--GNPADFIP----ILRYLNPNS---LNAPKDLNKFYSFMQKMKVHEHYKTFEK-
sp P05177 CP12_HUMAN	TASS--GNPLDFFP----ILRYLENPA---LQRFKAFNRQLWFLQRTVQEHYQDFDK-
sp Q16678 CP1B_HUMAN	TVGA--GSLVDVMP----WLQYFPNVRTVFRFEQLNRNFSNFIIDKFLRHCESLRPG
sp P10632 CPC8_HUMAN	N-PRDFIDCFLIKMEQEKD----NQKSEFNIEENLVGTADLVFVAGTETTSTTLRYGLLLL
sp P11712 CPC9_HUMAN	N-PQDFIDCFLMKMEKEKH----NQPSEFTIESLENTAVDLFGAGTETTSTTLRYALLLL
sp P33261 CPCJ_HUMAN	N-PRDFIDCFLIKMEKEKQ----NQOSEFTIENLVITAADLLGAGTETTSTTLRYALLLL
sp P00179 CPC5_RABIT	N-PRDFIDCFLIKMEQENN-----LEFTLESLVIAVDLFGAGTETTSTTLRYGLLLL
sp P05181 CPE1_HUMAN	C-PRDLTDCLLVMEKEKH----SAERLYTMDGTVTVADLVFAGTETTSTTLRYGLLLL
sp P24903 C2F1_HUMAN	--PRDFIQCFLLTKMAEKEK----DPLSHFHMDTLLMTHNLFLFGTETVSTTLHHAFLL
sp P11509 CPA6_HUMAN	S-PRDFIDSFLIRMQEEEK----NPNTFYLKKNLVMTTLLNLFIGGTETVSTTLRYGFLLL
sp P00178 CPB4_RABIT	N-PRDFIDVYLLRMEKDKS----DPSSEFHHQNLITLVLSLFFAGTETTSTTLRYGFLLM
sp P20813 CPB6_HUMAN	A-PKDLIDTYLLHMEKEKS----NAHSEFSSHQNLNLTLSLFFAGTETTSTTLRYGFLLM
sp P51589 CPJ2_HUMAN	E-TRDFIDAYLKEMSKHTG----NPTSSFHEENLICSTLDDFFAGTETTSTTLRWALLYM
sp P10635 CPD6_HUMAN	QPPRDLTEAPLAEMEKAKG----NPESFPDENLRIVVADLFSAGMVTSTTTLAWGLLM
sp P04798 CP11_HUMAN	GHIRDITDSLIEHCQEKQ----LDENANVQLSDEKINIIVLDLFGAGFTVTTALSWSLMYL
sp P05177 CP12_HUMAN	NSVRDITGALFKHKKKGP--RASGNLIPQ--EKIVNLVNDIFGAGFTVTTALSWSLMYL
sp Q16678 CP1B_HUMAN	AAPRDMMDAFILSAEKKAAGDSHGGGARLDLENVPATITDIFGASQDTLSTALQWLLLLL
sp P10632 CPC8_HUMAN	LKHPEVTAKVQEEIDHVIGRHRSPCMQDRSHMPYTDVAVVHEIQRYIS--DLVPTGVPHAVTT
sp P11712 CPC9_HUMAN	LKHEVTAKEVQEEIERVIGRNRSPCMQDRSHMPYTDVAVVHEVQRYI--DLLPTSLPHAVTC
sp P33261 CPCJ_HUMAN	LKHEVTAKEVQEEIERVIGRNRSPCMQDRSHMPYTDVAVVHEVQRYI--DLIPTSLPHAVTC
sp P00179 CPC5_RABIT	LKHPEVAARVQEEIERVIGRHRSPCMQDRSRMPYTDAVIHEIQRFI--DLLPTNLPHAVTR
sp P05181 CPE1_HUMAN	MKYPEIEEKLHEIEDRVI GPSRIPAIKDRQEMPYMDAVVHEIQRFI--TLVPSNLPHAVTR
sp P24903 C2F1_HUMAN	MKYPKVQARVQEEIDLVVGRARLPALKDRAMPYTDAVIHEVQRFI--DIIPMLNPHAVTR
sp P11509 CPA6_HUMAN	MKHEVEAKVHEIEDRVI GKNRQPKFEDRAKMPYMEAVIHEIQRFI--DVIPMSLARRVKK
sp P00178 CPB4_RABIT	LKYPHVTERVQKEIEQVIGSHRPPALDDRAKMPYTDAVIHEIQRLG--DLIPFGVPHAVTK
sp P20813 CPB6_HUMAN	LKYPHVAERVREIEQVIGPHRPPALHRAKMPYTEAVIYEQRFIS--DLLPMGVPHAVTK
sp P51589 CPJ2_HUMAN	ALYPEIQEKVQAEIDRVI GGGQSPSTAARESMPTNAVIHEVQRMG--NIIPLNVPREVTV
sp P10635 CPD6_HUMAN	ILHPDVQRRVQEEIDVDVIGQVRRPEMGOAHMPYTTAVIHEVQRFI--DIVPLGTMHTSR
sp P04798 CP11_HUMAN	VMNPRVQRKIQKELDTVIGRSRRPRLSDRSHLPYMEAFIETFRHS--SFVPTTIPHSTTR
sp P05177 CP12_HUMAN	VTKEPIQRKIQKELDTVIGRERRPRLSDRQPLPYLEAFIETFRHS--SFVPTTIPHSTTR
sp Q16678 CP1B_HUMAN	TRYPDVQTRVQAEIQVGRDRRLPCMGDQENLPLVLAFLYEMRFS--SFVPTTIPHATTA
sp P10632 CPC8_HUMAN	DTKFRNYLIPKGTTIMALLTSLVHDDKEFPNPNIFDPGHFLDKNG--NFKKSDYFMPFS
sp P11712 CPC9_HUMAN	DIKFRNYLIPKGTITLISLTSVLHDKNEFPNPEMFDPHHFLDEGG--NFKKSKYFMPFS
sp P33261 CPCJ_HUMAN	DVKFRNYLIPKGTITLISLTSVLHDKNEFPNPEMFDPRHFLEDEGG--NFKKSNYFMPFS
sp P00179 CPC5_RABIT	DVRFNRNYFIPKGTDIITSLTSVLHDEKAFENPKVFDPGHFLDESG--NFKKSDYFMPFS
sp P05181 CPE1_HUMAN	DTIFRGYLIIPKGTIVVPTLDSVLYDNQEPDPEKFKPEHFLNENG--FKKYSDFYKPFPS
sp P24903 C2F1_HUMAN	DTAFRGFLIPKGTIVVPTLLNTVHVDPSQFLTPQEFNPEHFLDANQ--SFKKSPAFMPPFS
sp P11509 CPA6_HUMAN	DTKFRDFFLIPKGTVEVYMLGSLRDRPSFFSNPQDFNPQHFLENEK--QFKKSDAFVPPFS
sp P00178 CPB4_RABIT	DTQFRGYVIIPKNTVEVYVLLSSALHDPYFETPNTFNPGHFLDANG--ALKRNEGFMPPFS
sp P20813 CPB6_HUMAN	HTSFRGYIIPKDETEVFLILSTALHDPHYFEKPDAPNPDHFLEANG--ALKKTEAFIPFS
sp P51589 CPJ2_HUMAN	DTTLAGYHLIPKGTIMLITNLTAHLDPTWATPDTFNPDHFLENG--QFKKREAFMPPFS
sp P10635 CPD6_HUMAN	DIEVQGFRIIPKGTITLITNLSSVLKDEAVWEKPFHFPEHFLDAQG--NIIPLNVPREVTV
sp P04798 CP11_HUMAN	DTSLKGFYIIPKGRCVFNQVQVNHQKLVVNPSEFLPERFLTPDG--ADKVLSEKVIIFG
sp P05177 CP12_HUMAN	DTTLNGFYIIPKCCVFNQVQVNHDPLEWEDPSEFRPERFLTDAGTAINKPLSEKMLFG
sp Q16678 CP1B_HUMAN	NTSVLGYHIPKDTVVFNQVSNHDPVKWPNPENFDPAFLDKDG--LINKDLTSRVMIFFS
sp P10632 CPC8_HUMAN	AG-----KRICAGEGLARMELFLFLTTILQNFNLKSVDDLKLNLTAVTKGIVSLP
sp P11712 CPC9_HUMAN	AG-----KRICVGEALAGMELFLFLTSLQNFNLKSLVDPKLNLTTPVNVNGFASVP
sp P33261 CPCJ_HUMAN	AG-----KRICVGEGLARMELFLFLTFLILQNFNLKSLDLPKDLTTPVNVNGFASVP
sp P00179 CPC5_RABIT	AG-----KRMCVGEGLARMELFLFLTSLQNFKLQSLVEPKDLDTAVNVNGFVSV
sp P05181 CPE1_HUMAN	TG-----KRVCSAGEGLARMELFLLLCAILQHFNKLPVDPKIDLSPIHGFGCIP
sp P24903 C2F1_HUMAN	AG-----RRLCLGELLARMELFLYLTAILQSFSLQPLGAPEDIDLTPLSGGLGNLP
sp P11509 CPA6_HUMAN	IG-----KRNCGEGLARMELFLFFTVMQNFRLKSSQSPKDDIVSPKHVGFATIP
sp P00178 CPB4_RABIT	LG-----KRICLGEIARTEFLFLFTTILQNFASVPEPIDLTPRESGVGNVP
sp P20813 CPB6_HUMAN	LG-----KRICLGEIARTEFLFLFTTILQNFASVPEPIDLTPRESGVGNVP
sp P51589 CPJ2_HUMAN	IG-----KRACLGEQLARTEFLFIPTSLMQKFTFRPPNNEK--LSLKRMGITIS
sp P10635 CPD6_HUMAN	AG-----RRACLGEPLARTEFLFFTSLQHFSSVPTGQPR--PSHHGVFAFLVSP
sp P04798 CP11_HUMAN	MG-----KRKCIGETIARWEVFLFLAILLQRFVFSVPLGVK--VDMTPIYGLTMKHA
sp P05177 CP12_HUMAN	MG-----KRKCIGEVLARWEVFLFLAILLQQLQLEFSVPPGVK--VDLTPYIYGLTMKHA
sp Q16678 CP1B_HUMAN	VG-----KRRCTGEELSRMQLFLFLSILAHQCDFRANPNEP--ARMNFSYGLTIKPK
sp P10632 CPC8_HUMAN	PSYQICFIPV-----
sp P11712 CPC9_HUMAN	PFYQLCFIPV-----
sp P33261 CPCJ_HUMAN	PFYQLCFIPV-----
sp P00179 CPC5_RABIT	PSYQLCFIPI-----
sp P05181 CPE1_HUMAN	PRYKLCVIRPS-----
sp P24903 C2F1_HUMAN	RPFQLCLRR-----
sp P11509 CPA6_HUMAN	RNYTMSFLPR-----
sp P00178 CPB4_RABIT	PSYQIRFLAR-----
sp P20813 CPB6_HUMAN	PTYQIRFLPR-----
sp P51589 CPJ2_HUMAN	VSHRLCAVPQV-----
sp P10635 CPD6_HUMAN	SPYELCAVPR-----
sp P04798 CP11_HUMAN	CCEHFQMQLRS-----
sp P05177 CP12_HUMAN	RCEHVQAR-RFSIN-----
sp Q16678 CP1B_HUMAN	SFKVNVTLRESMELLSAVQNLQAKETCQ

Abbildung 4.10: Auszug aus dem multiplen Sequenzalignment (Fortsetzung)

4.1.4 Festlegung der strukturkonservierten Bereiche

Auf der Grundlage des oben beschriebenen Sequenzalignments und unter Zuhilfenahme der Sekundärstrukturvorhersage wurden im nächsten Schritt die strukturkonservierten Bereiche festgelegt und die entsprechenden Koordinaten auf das Target CYP1A2 übertragen (Abbildung 4.11). Als strukturkonserviert angenommen wurden neben den Sekundärstrukturelementen auch alle Loops, die sich in ihrer Länge nicht von den entsprechenden Loops im Templat unterscheiden. Der Membrananker, der sich am N-terminalen Ende des Proteins befindet, fehlt in der Kristallstruktur, und wurde deshalb auch im Modell nicht berücksichtigt. Nach der Originalnummerierung müsste das Alignment daher mit Aminosäure 39 beginnen, die Nummerierung im Alignment weicht somit von der Originalnummerierung ab.

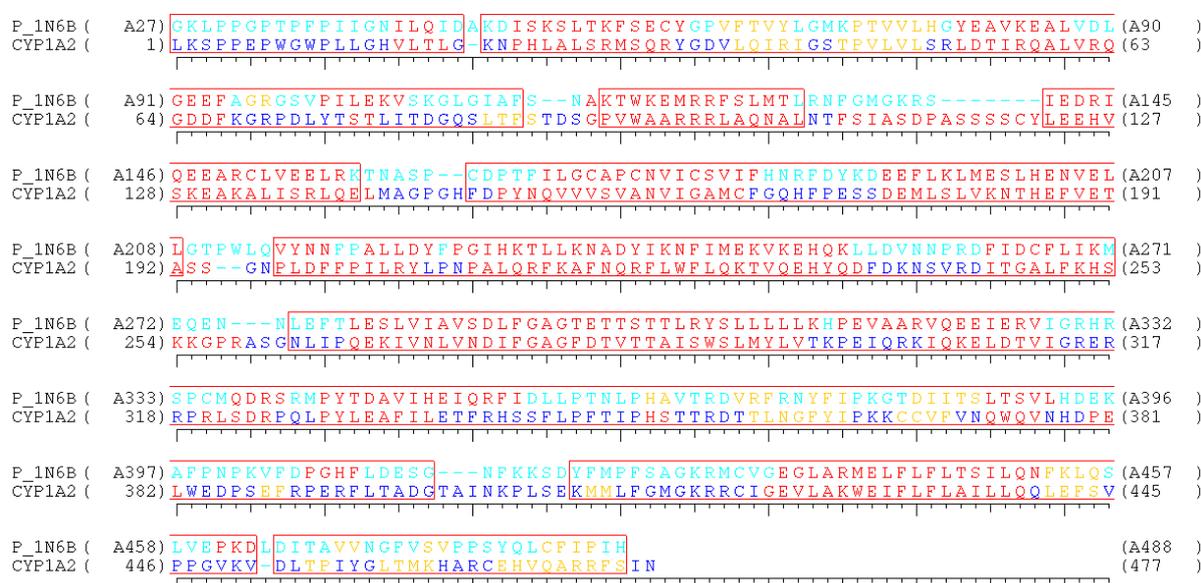


Abbildung 4.11: Strukturkonservierte Bereiche (rot eingerahmt) im manuellen Alignment der Kristallstruktur (P_1N6B) mit dem Target CYP1A2. Aminosäuren, die zu einer Helix gehören, sind rot gefärbt, Aminosäuren mit β -Faltblattstruktur gelb.

Leicht verändert wurde das Alignment im Loop zwischen Helix B' und Helix C. An dieser Stelle ist ein korrektes Alignment besonders wichtig, weil sich dort die SRS1 befindet. Der β -Turn „IAFS“ wurde abweichend vom automatischen Alignment mit dem Sequenzabschnitt „LTFS“ der Targetsequenz überlagert. Dies erschien sinnvoll, weil dadurch einerseits das einzufügende Bruchstück außerhalb der Substraterkennungsregion unmittelbar vor der Helix C liegt, und weil es sich hier andererseits sicher um homologe Sequenzabschnitte handelt, die das automatische Alignment nicht erkannt hat. Dafür spricht, dass der Sequenzabschnitt „LTFS“ ebenso wie der β -Turn „IAFS“ von PSIPRED fälschlicherweise als β -Faltblattstruktur vorhergesagt wird. Nach der Korrektur nimmt Phe 125 dieselbe Position ein wie Phe 114 der Templatstruktur. Dieses Phenylalanin ist in der Familie 2 konserviert und kommt auch in allen Sequenzen der Familie 1 vor (vgl. Abbildung 4.10). In sämtlichen Kristallstrukturen geht dieses Phenylalanin π - π -Interaktionen mit dem jeweiligen

Substrat ein. Da die Substrate von CYP1A2 überwiegend aromatisch sind, wird vermutet, dass dieser Kontakt auch hier eine wichtige Rolle spielt.

Problematisch war die Modellierung des Bereichs zwischen Helix F und Helix G mit den Helices F' und G'. Eine möglichst gute Modellierung dieses Bereichs ist wichtig, da sich hier, wie bereits in Abschnitt 4.1.2 erläutert, der Eingang zum Substratkanal befindet, der im substratgebundenen Zustand von den Helices F' und G' verschlossen wird. Die von CLUSTALW vorgeschlagene Überlagerung konnte hier ebenfalls nicht übernommen werden, weil dadurch zu viele Lücken und Bruchstücke entstehen, welche die Sekundärstruktur in diesem Bereich völlig zerstören. Stattdessen wurde versucht, unter Vernachlässigung der exakten Sequenzüberlagerung die für CYP1A2 vorhergesagte Sekundärstruktur an die in der Kristallstruktur vorhandene anzupassen, um die Ausprägung der Helices F' und G' sicherzustellen. Die Sekundärstrukturvorhersage gibt für die Helices F' und G' eine durchgehende Helix an. Anhand des Prolins, das häufig eine Unterbrechung der helikalen Struktur signalisiert, lässt sich die Trennung zwischen den Helices F' und G' aber gut zuordnen. Die beiden Lücken zu Beginn der Helix G' und in Helix G, die sich nur jeweils über eine Aminosäure erstrecken, wurden durch Ineinanderschieben der Sequenzen geschlossen. Dadurch ergibt sich in Helix G' und zu Beginn von Helix G eine Verschiebung um eine Aminosäure. Helix G' wird aber trotz des von der Homologie abweichenden Alignments korrekt ausgebildet. Bezüglich des Alignments in Helix G ist zu beachten, dass sich in der Mitte der Helix SRS3 befindet. Da die Sequenzverschiebung aber nur die ersten vier Aminosäuren betrifft, ist sicher, dass sich die an der Substratbindung beteiligten Aminosäuren auf der richtigen Position relativ zum aktiven Zentrum befinden.

4.1.5 Modellierung der strukturvariablen Bereiche

Nach Festlegung der strukturkonservierten Bereiche wurden die strukturvariablen Bereiche mithilfe der *loop search* Methode eingefügt (vgl. Abschnitt 3.2.2). Insertionen aufgrund abweichender Länge von Loop-Regionen waren nötig zwischen den Helices B'' und C, C und D, D und E, F und F', H und I, sowie zwischen Mäander-Loop und Cystein-Tasche vor Helix L (Abbildung 4.11). Zwischen den Helices A' und A am N-Terminus und zwischen der zweiten 3_{10} -Helix und dem Faltblattstrang β_{4-1} am C-Terminus ist die Sequenz von CYP1A2 jeweils um eine Aminosäure kürzer als das Templat. Hier war kein *loop search* notwendig, sondern die beiden Enden der strukturkonservierten Bereiche konnten jeweils durch eine lange Bindung miteinander verbunden werden. Eine Übersicht über alle beim Modellbau eingefügten Loops gibt Tabelle 4.5. Um die Zuordnung zu erleichtern, beginnt die Nummerierung der Aminosäuren wie im Alignment in Abbildung 4.11 mit 1.

Tabelle 4.5: Übersicht über alle beim Modellbau eingefügten Loops. SVR: strukturvariable Region; Flex: Anzahl der eingefügten Aminosäuren; preflex und postflex: Anzahl der Aminosäuren aus den angrenzenden strukturkonservierten Regionen, die mit dem Loop überlagert wurden; Tails: Überlagerung der gesamten preflex- und postflex-Bereiche; PDB-Code und Start-Aminosäure des eingefügten Fragmentes; RMSD-Wert der Überlagerung in Å.

SVR	Position	Flex	Pre-Flex	Post-Flex	Tails	PDB-Code	Start-Aminosäure	RMSD
Helix B'' – C	87-93	5	5	5	ja	1BU7	A84	1,00
Helix C – D	106-123	16	4	5	nein	1FCD	B120	0,70
Helix D – E	140-148	7	5	5	ja	1ECP	B24	1,44
Helix F – F''	192-197	4	5	5	ja	1PEH	236	1,24
Helix H – I	253-262	8	5	5	ja	1RHS	253	1,92
Helix K' – L	399-409	9	5	5	ja	1NBC	B114	2,28

Abgesehen von einer Ausnahme wurden jeweils fünf Aminosäuren aus den angrenzenden strukturkonservierten Bereichen in den *loop search* mit einbezogen und diese *preflex* und *postflex* Bereiche wurden auch mit überlagert.

Bei den Loops zwischen den Helices B'' und C, D und E sowie dem Loop in der Hämbinderegion handelte es sich im Vergleich zur Kristallstruktur jeweils um Erweiterungen der Loops um zwei bis drei Reste. Die per *loop search* ausgewählten Teilstücke konnten ohne Schwierigkeiten eingefügt werden und verliefen weitgehend parallel zu den entsprechenden Abschnitten der Kristallstruktur.

Zwischen den Helices C und D musste ein Teilstück gefunden werden, das sieben Aminosäuren länger ist als der entsprechende Loop der Kristallstruktur. Laut Sekundärstrukturvorhersage ist abgesehen von zwei Aminosäuren der gesamte Bereich zwischen Helix C und Helix D helikal. Betrachtet man in der Kristallstruktur die Lage der Helices C und D zueinander, stellt man fest, dass diese Vorhersage aus sterischen Gründen nicht stimmen kann. Auf diese Weise könnte keine Verbindung zwischen den beiden Helices gefunden werden. Da sich in der Mitte des Loops ein Prolin befindet, ist es wahrscheinlich, dass dort eine Helix beginnt oder endet. Denkbar wäre nun einerseits, dass der Loop eine Helix C' enthält, die mit dem Prolin endet, oder dass eine Verlängerung von Helix D enthalten ist, die mit dem Prolin beginnt. Die Sekundärstrukturvorhersage für die Kristallstruktur sagt den β -Turn „GKRS“ ebenfalls fälschlicherweise als Helix voraus. Deshalb wurde vermutet, dass die Vorhersage für CYP1A2 in diesem Bereich ebenfalls falsch ist, und dass stattdessen nach einer Verlängerung der Helix D gesucht werden muss. Diese Vermutung wurde durch eine Struktursuche mit dem FUGUE-Server (s. Abschnitt 3.1.3) bestätigt. Mit der Sequenz von CYP1A2 als Eingabe wurde neben einem Strukturalignment vieler bekannter Cytochromstrukturen (CYP2C5, CYP51, CYP102, CYP108) auch die 2003 veröffentlichte Kristallstruktur von CYP152A1 als homolog gefunden. Es handelt sich hierbei um eine Fettsäurehydroxylase aus dem Bakterium *Bacillus subtilis*. Zwar ist die Gesamthomologie zu CYP1A2 natürlich relativ gering, weshalb diese Struktur auf

räumliche Anordnung der beiden Helices F und F' mit dem eingefügten Loop in Abbildung 4.14, wird deutlich, dass eine Verlängerung der Helix F, wie sie die Sekundärstrukturvorhersage angibt, mit diesem Templat nicht realisiert werden kann, und dass daher nur nach der kürzesten möglichen Verbindung der beiden Helixenden gesucht werden konnte.

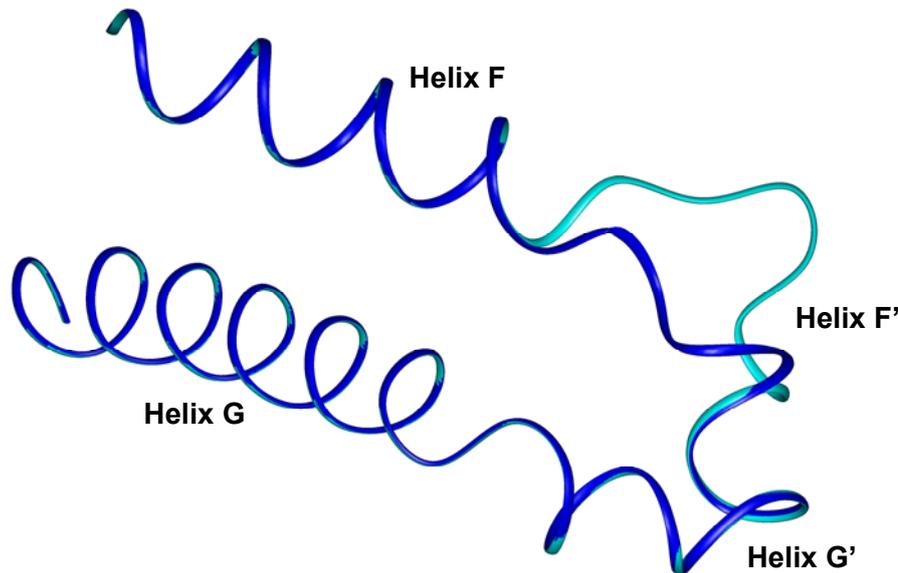


Abbildung 4.14: Der Loop zwischen den Helices F und F' der Kristallstruktur (türkis) im Vergleich mit dem per *loop search* eingefügten Teilstück im Modell (blau)

Eine weitere Verlängerung einer Helix entsprechend der Sekundärstrukturvorhersage für CYP1A2 (Abbildung 4.11) wurde im Loop zwischen den Helices H und I eingebracht. Dort wurde die Helix H um eine Helixwindung verlängert. Auch bei diesem Loop war eine sorgfältige Modellierung wichtig, da Helix I einerseits die Substraterkennungsregion SRS4 beinhaltet, aber andererseits auch Teil des katalytischen Zentrums des Enzyms ist.

Zuletzt wurden den beiden C-terminalen Aminosäuren, für die keine Koordinaten aus der Kristallstruktur übernommen werden konnten, mit der *end repair* Routine Koordinaten zugewiesen. Diese Routine entnimmt die entsprechenden Aminosäuren einer Templatbibliothek, wobei immer eine gestreckte Konformation angenommen wird.

4.1.6 Energieminimierung

Zur gezielten Minimierung der zuvor eingefügten strukturvariablen Bereiche stehen in INSIGHT II in Verbindung mit dem DISCOVER-Modul zwei Routinen zur Verfügung. Die Minimierungen basieren auf einer modifizierten Version des CVFF-Krafftfelds [154], die Parameter zur Beschreibung von Hämproteinen enthält [125].

Mit der *splice repair* Routine wird eine Energieminimierung der Peptidbindungen durchgeführt, die den Übergang zwischen einem strukturkonservierten Bereich und einem per *loop search* eingefügten Fragment bilden. Trotz einer sorgfältigen Überlagerung der angrenzenden Teilstücke kann die Geometrie dieser Peptidbindungen verzerrt sein, weshalb eine Angleichung der Bindungslängen und vor allem der Torsionswinkel wichtig ist. In die Minimierung einbezogen werden lediglich die beiden benachbarten Aminosäuren des Loops und die Peptidbindung selbst. Mit dieser Routine wurden zunächst die Torsionswinkel der ω -Bindungen aller an einem Loopübergang beteiligten Peptidbindungen entsprechend den Standardparametern der *splice repair* Routine mit einer Kraftkonstante von $50 \text{ kcal}\cdot\text{mol}^{-1}$ über 100 Schritte Steepest-Descent-Minimierung auf 180° gezwungen.

Die *relax* Routine ermöglicht eine selektive Energieminimierung einzelner Bereiche des Proteinmodells zur Beseitigung von Überlappungen mutierter Proteinseitenketten oder schlecht eingepasster Loop-Bereiche. Der Rest des Proteins bleibt dabei jeweils fixiert. Folgende Teilbereiche des Modells wurden auf diese Weise nacheinander entspannt:

1. der mit *end repair* angefügte C-Terminus
2. alle Loop-Seitenketten
3. alle Loop-Atome
4. alle mutierten Seitenketten der strukturkonservierten Regionen
5. alle Seitenketten der strukturkonservierten Regionen
6. alle Atome aus 1., 3. und 5. gleichzeitig

Es wurde jeweils eine Steepest-Descent-Minimierung über 100 Schritte durchgeführt, bei der alle beteiligten Atome mit einer leichten Rückhaltekraft von $100 \text{ kcal}\cdot\text{\AA}^{-2}$ belegt wurden. Die Rückhaltekraft verhindert unrealistisch heftige Ausweichbewegungen von Atomen, die beim Aufheben von van-der-Waals-Überlappungen entstehen können.

4.1.7 Einfügen des Häm in das Modell

Die Koordinaten für das Häm konnten direkt aus dem Templat übertragen werden. Da das Eisenatom, das Zentralatom des Häm, über das Schwefelatom von Cys 458 kovalent an das Protein gebunden ist, wurde zwischen diesen beiden Atomen eine kovalente Bindung der Länge $2,3 \text{ \AA}$ definiert.

4.2 Strukturoptimierung und Validierung des Modells

Die weitere Verfeinerung der Proteinstruktur und die anschließende Validierung des Modells erfolgten im Rahmen von Moleküldynamiksimulationen (MDS, s. Abschnitt 3.4.3) mit dem Simulationspaket GROMACS. Neben der bereits in Abschnitt 3.4.1.2 erwähnten für die Simulation von Proteinen gut geeigneten Parametrisierung des GROMOS96-Krauffeldes bietet dieses Programm auch die Möglichkeit, das dynamische Verhalten von Proteinen mit einem vertretbaren Zeitaufwand unter physiologischen Bedingungen zu simulieren. Das heißt, dass die Proteinmodelle in eine Wasserbox mit einer physiologischen Kochsalzkonzentration eingesetzt werden. Die Berücksichtigung der Lösungsumgebung ist deshalb besonders wichtig, weil die Proteinkonformation vor allem in den lösungsmittel-exponierten Bereichen stark von der Solvatation durch das Lösungsmittel beeinflusst wird. Wie die kürzlich veröffentlichten Kristallstrukturen gezeigt haben, spielt das Lösungsmittel aber auch bei der Substratbindung im aktiven Zentrum eine große Rolle. Häufig werden keine direkten Wasserstoffbrückenbindungen zwischen Protein und Substrat ausgebildet, sondern hydrophile Gruppen des Substrates werden in Wasserclustern solvatisiert, die eine Verbindung zu Aminosäuren im aktiven Zentrum des Cytochroms herstellen [93] (s. auch Abschnitt 1.4.2.4).

Zur Überprüfung der Simulationsbedingungen wurde die als Templat verwendete Kristallstruktur 1N6B ebenfalls einer MDS unterzogen. Das mitkristallisierte Substrat DMZ wurde entfernt, um zunächst nur das Verhalten des leeren Enzyms mit dem leeren Modell zu vergleichen.

4.2.1 Durchführung und Simulationsbedingungen

Zuerst wurden die zu untersuchenden Proteinstrukturen in eine quaderförmige Simulationsbox eingesetzt. Die Größe der Box und die Position des Proteins in der Box wurden dabei so gewählt, dass der Abstand zwischen jedem Proteinatom und der Boxwand mindestens 0,6 nm betrug, um Interaktionen des Proteins mit seinen Nachbarn zu vermeiden.

Im nächsten Schritt wurde die Box mit *simple point charge* (SPC) Wassermolekülen [155] aufgefüllt. Das SPC-Modell ist ein rigides Wassermodell, das von GROMACS standardmäßig als Lösungsmittel verwendet wird. Es enthält drei Interaktionspunkte für elektrostatische Wechselwirkungen, nämlich das Sauerstoffatom und die zwei Wasserstoffatome, deren Ladung insgesamt null ergibt. Van-der-Waals-Wechselwirkungen werden nur für die Sauerstoffatome berechnet, der van-der-Waals-Radius der Wasserstoffatome wird vernachlässigt. Der Abstand zwischen dem Sauerstoff und den Wasserstoffen, sowie der Abstand zwischen den beiden

Wasserstoffen werden mit dem SETTLE-Algorithmus [156], einer speziell für Wassermoleküle modifizierten Variante des SHAKE-Algorithmus, konstant gehalten.

Bei Verwendung der PME-Methode zur Berechnung der elektrostatischen Wechselwirkungen muss das Simulationssystem insgesamt elektronisch neutral sein [122]. Deshalb muss die Ladung des Proteins durch das Einfügen von Gegenionen in das Simulationssystem ausgeglichen werden. Es ist an dieser Stelle sinnvoll, zusätzlich zum Ladungsausgleich eine physiologische Ionenkonzentration einzustellen, um eine gleichmäßige Verteilung von Ionenladungen in der Simulationsbox zu gewährleisten [157]. Dazu wurde eine Programmroutine benutzt, die nach dem Zufallsprinzip eine festgelegte Anzahl Wassermoleküle durch Na^+ - und Cl^- -Ionen ersetzt. Die Anzahl der benötigten Ionen wurde zuvor jeweils für die Anzahl der Wassermoleküle in der Box berechnet. Die vollständige Zusammensetzung beider Simulationssysteme ist in Tabelle 4.6 aufgeführt. Abbildung 4.15 zeigt das Modell CYP1A2 in der Simulationsbox.

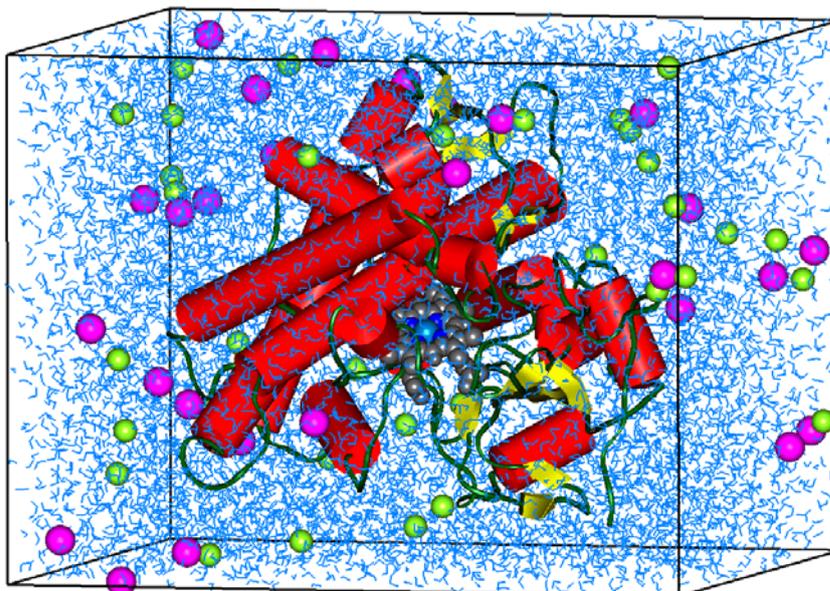


Abbildung 4.15: Proteinmodell von CYP1A2 in der Simulationsbox. Das Proteinrückgrat ist als dunkelgrünes Band dargestellt. Die roten Zylinder symbolisieren die Helices des Proteins, die gelben Blockpfeile die Faltblätter. Das Häm ist als graues Kalottenmodell dargestellt. Die Wassermoleküle sind hellblau eingefärbt. Die magentafarbenen Kugeln repräsentieren die Na^+ -Ionen, die hellgrünen Kugeln die Cl^- -Ionen.

Tabelle 4.6: Zusammensetzung der Simulationssysteme

	Kristallstruktur	Modell
Kantenlänge der Box (nm)	8,176 x 5,821 x 7,469	7,923 x 6,134 x 7,470
Gesamtladung Protein (e)	-7	+5
Anzahl Wassermoleküle	9.412	9.513
Anzahl Na^+ -Ionen	37	32
Anzahl Cl^- -Ionen	30	37

Da nach der Erstellung des Homologiemodells bisher nur eine grobe Energieminimierung einzelner Bereiche des Modells stattgefunden hat (s. Abschnitt 4.1.6), wurde als nächstes eine Energieminimierung des gesamten Proteins in der zuvor geschaffenen Simulationsumgebung angeschlossen. Auch die Kristallstruktur wurde in der Simulationsbox einer Energieminimierung unterzogen, um ungünstige van-der-Waals-Kontakte zum einen innerhalb der Proteinstruktur und zum anderen zwischen Protein und Lösungsmittelmolekülen oder Ionen aufzuheben. Es wurde hierzu eine *steepest descent* Minimierung bis zu einer Konvergenz von $100 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-1}$ durchgeführt. Die Fortsetzung mit einer *conjugate gradient* Minimierung oder die Wahl eines niedrigeren Konvergenzkriteriums ist nicht sinnvoll, da es hier lediglich darum geht, eine gute Startstruktur für eine anschließende MDS zu generieren.

Die weitere Equilibrierung des Systems nach der Energieminimierung wurde mit MDS erreicht. Eine erste Equilibrierungsdynamik über 200 ps diente dazu, die regelmäßige Struktur der Wassermoleküle aufzuheben, die Ionen gleichmäßig in der Box zu verteilen und die Proteinseitenketten nach intramolekularen Wechselwirkungen oder Wechselwirkungen mit Lösungsmittelmolekülen auszurichten. Um in diesem frühen Stadium der Equilibrierung größere Bewegungen des Proteinrückgrats zu vermeiden, die zu unerwünschten Veränderungen der Proteinstruktur bis hin zur Auflösung von Sekundärstrukturelementen führen können, wurden die Atome des Proteinrückgrats und des Häm mithilfe von *position restraints* auf ihren Ausgangspositionen fixiert. Die Kraftkonstante der *position restraints* betrug $1000 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$ für alle drei Raumrichtungen. In Anlehnung an die von B. Rupp angewendete Equilibrierungsmethode [158] wurden per *loop search* eingefügte Loopbereiche, in denen sich nach dem Ramachandran Plot Aminosäuren in strukturell ungünstigen Regionen befanden, von der Fixierung ausgenommen. In diesen eng begrenzten Bereichen wurde bereits auch das Proteinrückgrat freigelassen, um eine effektive Strukturverbesserung zu ermöglichen. Zur Validierung dieser Equilibrierungsmethode wurden auch in der Kristallstruktur kurze Loopbereiche mit schlechter struktureller Qualität freigelassen.

Eine freie MDS, in deren Verlauf das gesamte Simulationssystem, also auch das Rückgrat des Proteinmodells, frei beweglich war, bildete die zweite Equilibrierungsphase. Hier wurde auch das Proteinrückgrat in die Strukturoptimierung mit einbezogen. Sobald die Proteinstruktur einen Gleichgewichtszustand erreicht hatte, was hier nach 1,5 ns der Fall war, schloss sich eine Auswertungsphase von 1 ns an. Über diesen Simulationszeitraum wurde alle 5 ps die Proteinkonformation extrahiert, um einen Eindruck vom Konformationsraum des Proteinmodells zu gewinnen. Einen Überblick über das gesamte Simulationsprotokoll gibt Abbildung 4.16.

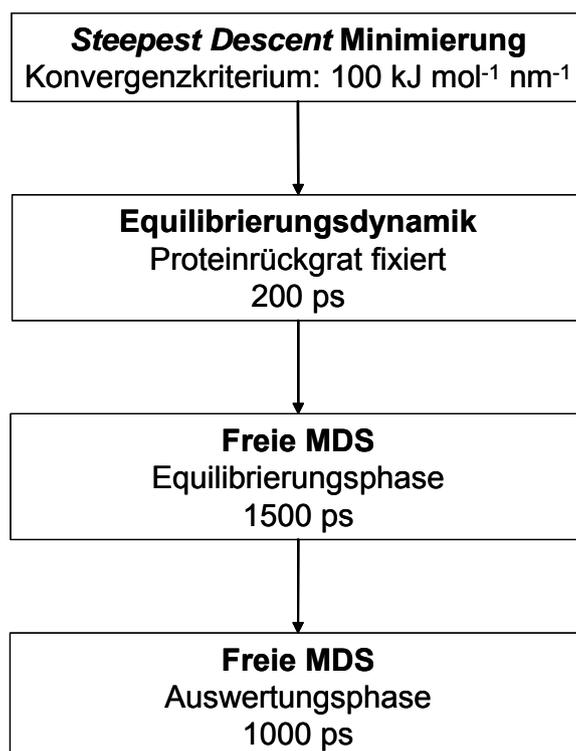


Abbildung 4.16: Simulationsprotokoll der MDS

Alle MD-Simulationen wurden bei einer Temperatur von 310 K und einem Druck von 1 bar durchgeführt. Wie bereits in Abschnitt 3.4.1.1 erläutert, wurden die elektrostatischen Wechselwirkungen langer Reichweite nach der PME-Methode berechnet. Für die elektrostatischen Wechselwirkungen kurzer Reichweite und für die van-der-Waals-Wechselwirkungen wurde ein *cutoff* von 0,9 nm verwendet. Um einen Integrationszeitschritt von 2 fs verwenden zu können, wurden alle Bindungslängen mit dem LINCS-Algorithmus fixiert. Auch darauf wurde bereits in Abschnitt 3.4.3.2 ausführlicher eingegangen. Ein vollständiges Input-File mit allen verwendeten Parametern ist im Anhang beigelegt.

4.2.2 Equilibrierung der Kristallstruktur

4.2.2.1 Strukturelle Entwicklung

Zur Validierung der oben erläuterten Equilibrationsmethode wurde zunächst die strukturelle Qualität der Kristallstruktur im Verlauf der Equilibrierung mit dem Programm PROCHECK verfolgt. Es stellte sich heraus, dass die Energieminimierung keine Strukturverbesserung ergab, sondern eher eine leichte Verschlechterung (Tabelle 4.7). Das liegt daran, dass die Struktur nicht bis zur Konvergenz minimiert wurde, sondern nur bis zu einem relativ groben Konvergenzkriterium, um eine Startstruktur für die MDS zu erhalten. Von den sechs Aminosäuren, die nach der Minimierung in den beiden ungünstigeren Konformationsbereichen lagen (Abbildung 4.17 a), befinden sich zwei am N-Terminus des Enzyms (Lys 28, Ile 38), eine im Loop zwischen den Helices B' und C (Ala 117) und eine im Loop zwischen den

Helices H und I (Glu 272). Deshalb wurden während der Equilibrierungsdynamik der N-Terminus von Gly 27 bis Asn 41, der Loop B''C von Gly 109 bis Lys 118 und der Loop HI von Glu 272 bis Thr 280 komplett, also inklusive des Proteinrückgrates, freigelassen.

Tabelle 4.7: Auswertung des Ramachandran Plots im Verlauf der Equilibrierung der Kristallstruktur 1N6B. Die Verteilung der Aminosäurekonformationen auf die Konformationsbereiche ist in Prozent angegeben. Für den zulässigen und den unerlaubten Konformationsbereich ist die absolute Anzahl der darin enthaltenen Aminosäuren in Klammern aufgeführt.

Konformationsbereich	Kristallstruktur	nach Minimierung	nach Equilibrierung
bevorzugter Bereich	85,7	82,8	85,5
günstiger Bereich	12,8	15,8	12,6
zulässiger Bereich	0,7 (3)	0,5 (2)	1,0 (4)
unerlaubter Bereich	0,7 (3)	1,0 (4)	1,0 (4)

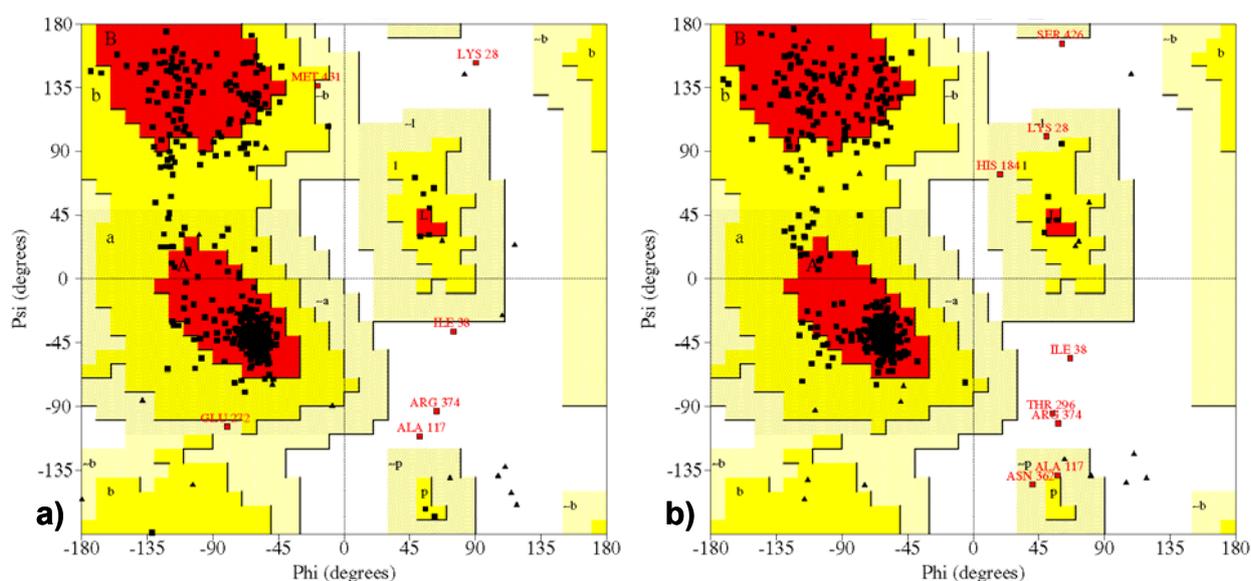


Abbildung 4.17: Ramachandran Plot a) nach einer Minimierung der Kristallstruktur und b) nach der Equilibrierungsdynamik.

Wie aus dem Ramachandran Plot der Kristallstruktur nach der Equilibrierungsdynamik ersichtlich ist (Abbildung 4.17 b) und Tabelle 4.7), hat der Anteil an Aminosäuren im bevorzugten Konformationsbereich wieder zugenommen. Folglich ist eine Entspannung der Struktur und eine Solvatation des Proteins durch das Lösungsmittel erfolgt. Glu 272 (Loop HI) ist aus den ungünstigen Konformationsbereichen verschwunden, und Lys 28 (N-Terminus) und Ala 117 (Loop B''C) sind von einer unerlaubten Konformation in eine zulässige Konformation übergegangen. Somit scheint das Freilassen von strukturell schlechten Loops zur Entspannung dieser Regionen beigetragen zu haben. Insgesamt gesehen befinden sich nun jedoch acht anstatt vorher sechs Aminosäuren in den ungünstigen Konformationsbereichen. Zusätzlich zu den bereits erwähnten Loop-Aminosäuren nehmen Asn 362 und Ser 426, zwei Aminosäuren, die über Wasserstoffbrücken zu den Propionatseitenketten

an der Hämbindung beteiligt sind, ungünstige Konformationen ein. Ser 426 liegt aber bereits in der Kristallstruktur nur im zulässigen Bereich, was darauf schließen lässt, dass an der Hämbindung beteiligte Aminosäuren häufiger in ungewöhnlichen Konformationen vorliegen. Eine detaillierte Auswertung der stereochemischen Qualität ist aber erst nach der sich anschließenden freien MDS sinnvoll.

4.2.2.2 Potentielle Energie und RMSD-Wert

Der Verlauf der potentiellen Energie während der freien MDS der Kristallstruktur zeigt, dass keine Energiedrift stattfindet, sondern dass das Simulationssystem über den gesamten Simulationszeitraum stabil bleibt (Abbildung 4.18). Nach einer Equilibrierungszeit von 500 ps schwankt die potentielle Energie um einen konstanten Wert von $-463.403,3 \text{ kJ}\cdot\text{mol}^{-1}$.

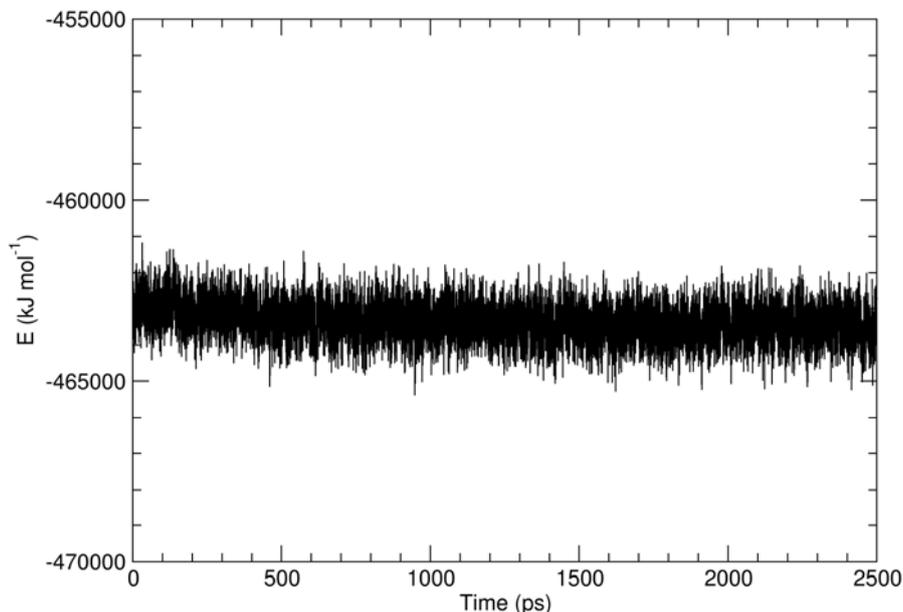


Abbildung 4.18: Verlauf der potentiellen Energie des Simulationssystems während der freien MDS.

Als Maß für die Equilibrierung der Proteinstruktur wurde der RMSD-Wert des Proteinrückgrates der Proteinkonformationen im Verlauf der Simulation bezogen auf das Proteinrückgrat der Ausgangsstruktur herangezogen. Der RMSD-Wert (*root mean square deviation*, mittlere quadratische Standardabweichung) einer Atomposition r_i zum Zeitpunkt t_1 in Bezug auf die Atomposition der Referenzstruktur zum Zeitpunkt t_0 wird dabei nach einer Überlagerung der Strukturen berechnet nach:

$$RMSD(t_1, t_0) = \left[\frac{1}{N} \sum_{i=1}^N \|r_i(t_1) - r_i(t_0)\|^2 \right]^{\frac{1}{2}} \quad (4.1)$$

Wie aus Abbildung 4.19 hervorgeht, ist ab 1500 ps ein Gleichgewichtszustand der Proteinstruktur erreicht. Der RMSD-Wert pendelt um einen Mittelwert von 0,21 nm, eine weitere Entfernung von der Startstruktur findet nicht mehr statt. In einer 2003

veröffentlichten MDS-Studie untersuchten H. Fan und A.E. Mark den RMSD-Wert von 19 Kristallstrukturen nach einer 5 ns langen MDS bezogen auf die Startstruktur [159]. Die Studie wurde ebenfalls mit dem Programm GROMACS und dem GROMOS96-Kraftfeld durchgeführt, und auch die Simulationsbedingungen waren mit den hier gewählten vergleichbar. Die Autoren fanden heraus, dass neun der Durchschnittsstrukturen aus den MDS einen RMSD-Wert $\leq 0,2$ nm und acht Durchschnittsstrukturen einen RMSD-Wert zwischen 0,2 und 0,3 nm aufwiesen. Der hier berechnete mittlere RMSD-Wert von 0,21 nm stimmt somit mit dem Ergebnis der Studie überein und scheint für das Programm und die Simulationsbedingungen typisch zu sein.

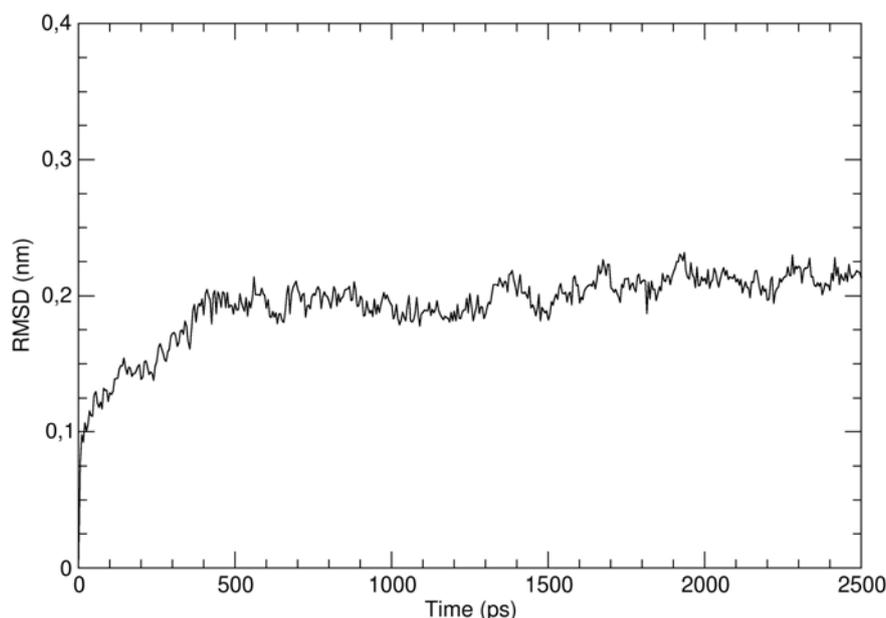


Abbildung 4.19: RMSD-Wert des Proteinrückgrates bezogen auf die Startstruktur im Verlauf der freien MDS.

Der equilibrierte Zeitabschnitt der Trajektorie von 1500 bis 2500 ps wird im Folgenden ausführlich ausgewertet.

4.2.3 Auswertung der MDS mit der Kristallstruktur

4.2.3.1 RMS-Fluktuation

Im Gegensatz zum RMSD-Wert, der die strukturellen Veränderungen des gesamten Proteins im Verlauf der Simulationszeit wiedergibt, kann mithilfe der RMS-Fluktuation (RMSF-Wert) die Beweglichkeit einzelner Proteinabschnitte gemittelt über die equilibrierte Phase der MDS verfolgt werden. Hier wurden die Fluktuationen für alle Atome des Proteinrückgrats berechnet und anschließend wurde für jede Aminosäure ein Mittelwert gebildet. Die ersten Maxima im Diagramm der RMS-Fluktuation (Abbildung 4.20) lassen sich eindeutig den Loop-Übergängen zwischen den Helices vom N-Terminus bis zur Helix J zuordnen. Das höchste Maximum kommt dabei am

Übergang zwischen den Helices C und D zustande. Dieser Loop ist mit neun Aminosäuren der längste Loop.

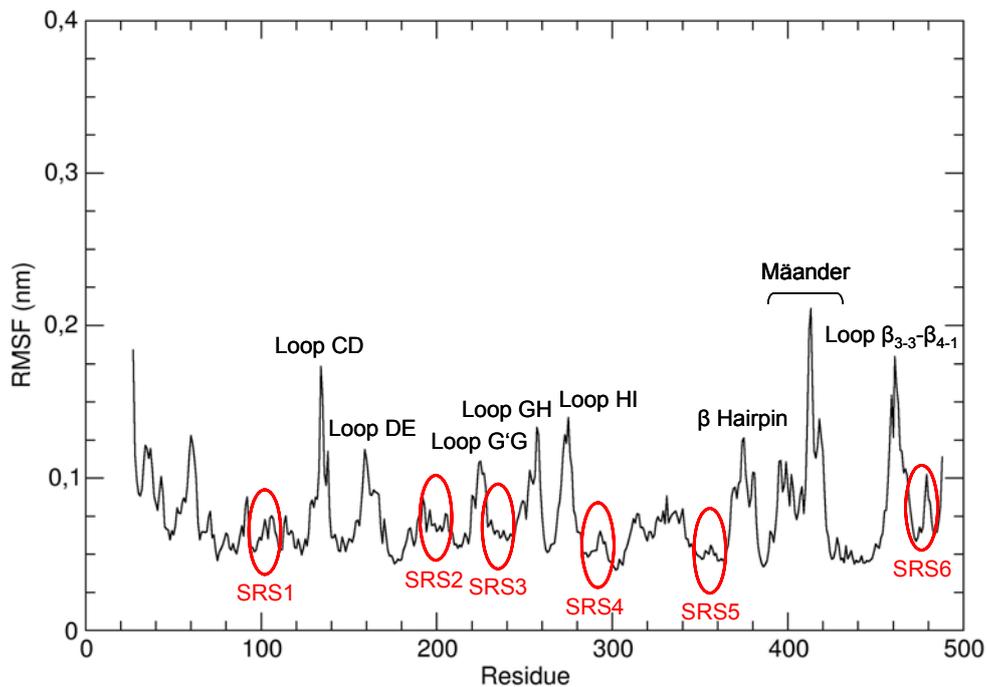


Abbildung 4.20: RMS-Fluktuation der einzelnen Aminosäuren zwischen 1500 und 2500 ps der freien MDS. Die roten Markierungen umfassen die Substratbinderegionen SRS1-6.

Erstaunlich ist die relativ geringe Flexibilität des Bereiches Helix F – Helix F' – Helix G' – Helix G. Da wie bereits in Abschnitt 4.1.2 erwähnt, der Loop zwischen den Helices F und G mit den Nebenhelices F' und G' eine Art Deckel des aktiven Zentrums bildet, der je nach Bedarf geöffnet oder geschlossen werden kann, wurde für den gesamten Proteinabschnitt zwischen den Helices F und G eigentlich eine viel höhere Flexibilität erwartet. Dass sich dieser Bereich in der hier durchgeführten MDS nur mäßig bewegt, könnte darin begründet sein, dass hier zwar einerseits ein leeres Enzym vorliegt, das sich aber andererseits im geschlossenen Zustand befindet. Ein Übergang vom geschlossenen in den geöffneten Zustand lässt sich offensichtlich im Rahmen einer MDS von nur 2500 ps auch mit einem leeren Enzym nicht nachvollziehen. Stattdessen bleibt der geschlossene Zustand stabil.

In den Faltblattregionen des Cytochroms sind ebenfalls zwei Turn- bzw. Loopstrukturen lokalisiert, die ein Maximum im RMSF-Plot verursachen.

Das höchste Maximum der RMS-Fluktuation wird durch die Mäander-Region hervorgerufen. Hierbei handelt es sich um den Bereich zwischen der Helix K' und der so genannten Cystein-Tasche, die das konservierte Cystein enthält, das die kovalente Bindung zwischen Häm und Protein ausbildet. Die Mäander-Region umfasst einen Bereich von etwa 20 Aminosäuren, der in scheinbar strukturlosen Schleifen verläuft und keinerlei Sekundärstrukturelemente aufweist [83]. Die einzige

strukturstabilisierende Komponente dieser Region ist die ERR-Triade, eine konservierte Dreiergruppe aus den Aminosäuren Glu-Arg-Arg/His, deren Seitenketten Wasserstoff- bzw. Salzbrücken mit dem Proteinrückgrat der Mäander-Region eingehen. Diese Interaktionen sorgen dann für die typische mäanderförmige Konformation dieses Proteinabschnittes. Man kann sich somit gut vorstellen, dass dieser Bereich und der sich daran anschließende nochmals etwa 10 Aminosäuren umfassende Loop-Übergang in die Cystein-Tasche von einer erhöhten Beweglichkeit geprägt sind.

Erwartungsgemäß befinden sich die sechs an der Substratbindung beteiligten Regionen SRS1-6 in eher bewegungsarmen Bereichen im Proteininneren. Bemerkenswert ist die kleine Spitze in SRS4. Sie markiert genau die Stelle, an der das regelmäßige Wasserstoffbrückenmuster der Helix I aufgehoben ist, und ein für alle Cytochrome charakteristischer Knick in der Helix entsteht. SRS6 liegt in einem β -Turn zwischen den Faltblättern 4-1 und 4-2. Dieser Turn ist neben dem Loop FG am Abschluss des Substratkanals beteiligt, was auch hier eine gewisse Beweglichkeit erklärt.

Der hier beobachtete Verlauf der RMS-Fluktuation stimmt weitgehend mit den Ergebnissen einer freien MDS mit der Kristallstruktur 1BU7 von CYP102 (P450 BM-3) über 14 ns in wässriger Umgebung überein [160]. In dieser MDS-Studie zeigte sich ebenfalls das höchste Maximum im Bereich der Mäander-Region. Deutlich abweichend verhielt sich allerdings der Loop FG, der dort das zweithöchste Maximum verursachte. Grund dafür ist, dass in Bakterien Cytochromen die Nebenhelices F' und G' fehlen, die in den Säugetiercytochromen für eine Stabilisierung dieses Loops sorgen.

4.2.3.2 Protein-Häm-Interaktionen

Ein weiteres wichtiges Kriterium für die Qualität der Proteinstruktur von Cytochromen sind die Protein-Häm-Wechselwirkungen. Neben der kovalenten Bindung von Cys 432 an das Hämeisen existieren eine Reihe von über die gesamte Superfamilie mehr oder weniger konservierten Salz- und Wasserstoffbrücken zwischen den Propionatseitenketten der Hämringe A und D und Aminosäureseitenketten des Proteins, die die Lage des Häm im aktiven Zentrum stabilisieren (s. Abschnitt 1.4.2.1). Abbildung 4.21 zeigt die aufgetretenen Interaktionen und ihre Stabilität im Verlauf der MDS. Es stellte sich heraus, dass alle Salz- und Wasserstoffbrücken, die man in der Kristallstruktur findet, auch in der MDS auftraten, und über die gesamte Dauer stabil blieben. Zusätzlich zu den in der Kristallstruktur erkennbaren Interaktionen zu Aminosäureseitenketten kommt es in der MDS zur Ausbildung einer Wasserstoffbrücke zur Amidfunktion des Proteinrückgrats von Leu 363. Diese Interaktion kommt aufgrund einer leicht gedrehten Orientierung der Propionatgruppe

des Hämrings A zustande. Sie führt zu keiner Verzerrung der Proteinstruktur und kann somit hingenommen werden.

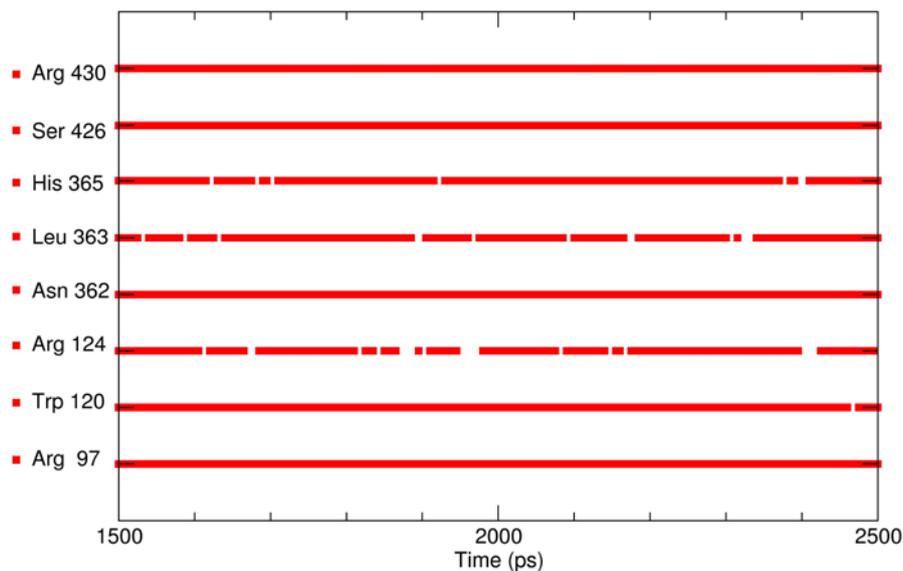


Abbildung 4.21: Protein-Häm-Wechselwirkungen im Verlauf der equilibrierten Phase der MDS mit der leeren Kristallstruktur.

4.2.3.3 Wasserstoffbrückenmuster in Helix I

Ein besonderes Strukturmerkmal der Cytochrome ist wie bereits erwähnt ein charakteristischer Knick in der Helix I, im katalytischen Zentrum direkt über dem Häm. Im Rahmen der Auswertung der MDS mit der Kristallstruktur wurde untersucht, ob dieses charakteristische Strukturmerkmal erhalten bleibt.

Die Kristallstruktur 1N6B enthält zwei Wassermoleküle in der aufgeweiteten Helixwindung in Helix I, die Wasserstoffbrücken zum Proteinerückgrat und zur Seitenkette von Thr 298 innerhalb der Helixwindung ausbilden. Im Verlauf der MDS befindet sich nur noch ein Wassermolekül in der Helixwindung (Abbildung 4.22). Die Wasserstoffbrücke zwischen Ala 294 und der Seitenkette von Thr 298 ist nun direkt, sodass die Aufweitung trotzdem bestehen bleibt. Eine sehr ähnliche Konstellation findet man in der anderen von CYP2C5 bekannten Kristallstruktur 1NR6. Auch dort existiert eine direkte Wasserstoffbrücke zwischen Ala 294 und der Hydroxylgruppe von Thr 298, und auch dort befindet sich nur ein Wassermolekül in der Helixwindung. Dies zeigt, dass das Wasserstoffbrückenmuster dieses Strukturmerkmals durchaus variieren kann, auch wenn es sich um dasselbe Isoenzym handelt.

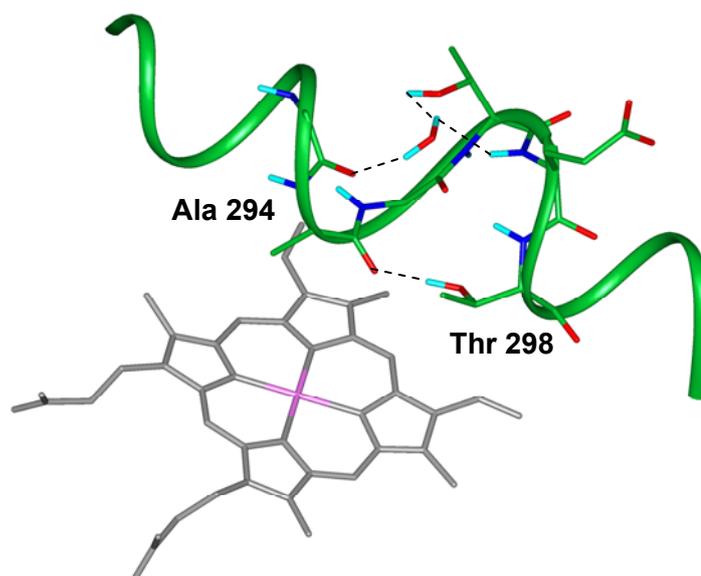


Abbildung 4.22: Wasserstoffbrückenmuster in Helix I am Beispiel der repräsentativen Struktur aus der MDS mit der Kristallstruktur.

4.2.3.4 Proteinstruktur

Zur Beurteilung der strukturellen Qualität der Kristallstruktur nach der freien MDS wurde mit allen 200 Strukturen aus der equilibrierten Phase eine Clusteranalyse durchgeführt. Dabei wurden zur Überlagerung der Strukturen alle Proteinrückgrat- atome der Helices benutzt und für die anschließende Clustereinteilung die Proteinrückgrat- atome aller Aminosäuren. Auf diese Weise bekommt man einen Überblick über den gesamten Konformationsraum des Proteins.

Tabelle 4.8 listet die fünf größten Cluster mit einer PROCHECK-Auswertung der jeweils repräsentativen Struktur auf. Der Zeitpunkt, von dem die repräsentative Struktur jedes Clusters stammt, ist ebenfalls mit angegeben. Er bezieht sich auf den Simulationszeitraum der equilibrierten Phase (1500-2500 ps). Man erkennt, dass die großen Cluster gleichmäßig über den Simulationszeitraum der MDS verteilt sind. Der repräsentative Vertreter des größten Clusters weist gleichzeitig die beste strukturelle Qualität auf.

Tabelle 4.8: Übersicht über die fünf größten Cluster der NMRCLUST-Analyse mit Zeitpunkt und struktureller Qualität der jeweils repräsentativen Struktur.

Nr.	Anzahl Strukturen	Zeitpunkt (ps)	Konformationsbereiche nach PROCHECK			
			bevorzugt (%)	günstig (%)	zulässig (%)	unerlaubt (%)
1	13	2000	83,0	15,8	0,5 (2)	0,7 (3)
2	12	1550	81,5	16,5	0,5 (2)	1,5 (6)
3	9	2135	82,0	16,5	0,5 (2)	1,0 (4)
4	9	1650	83,0	14,8	1,0 (4)	1,2 (5)
5	9	2260	84,0	14,4	0,5 (2)	1,2 (5)

Nur fünf Aminosäuren befinden sich in den ungünstigen Konformationsregionen. Im zulässigen Bereich sind das Ile 38 am N-Terminus und Val 474 am Ende der SRS6. Val 474 wird durch das benachbarte Phe 473 in eine leicht gespannte Konformation gezwungen, da dieses Phenylalanin aus Platzgründen ins aktive Zentrum hineinragt, wenn das Enzym leer ist. In den MDS mit den Substratkomplexen der Kristallstruktur geht Phe 473 π - π -Interaktionen mit dem Substrat ein und nimmt dadurch eine andere Position ein. Val 474 ist dann ebenfalls an der Substratbindung beteiligt und liegt in einer bevorzugten Konformation vor. Eine unerlaubte Konformation nehmen Thr 296, Arg 374 und Ser 426 ein. Ser 426 ist an der Hämbindung beteiligt und befand sich bereits in der Kristallstruktur und während der Equilibrierung in einem ungünstigen Konformationsbereich (s. Abschnitt 4.2.2). Arg 374 liegt in der Spitze des β -Hairpin-Turns von Faltblatt β_2 . Auch diese Aminosäure wurde bereits aus der Kristallstruktur in einer unerlaubten Konformation übernommen. Der Turn liegt weit entfernt von SRS5 in der Peripherie des Enzyms. Thr 296 ist Bestandteil des für Cytochrome charakteristischen Knicks in der Mitte der Helix I. Es ist durchaus vorstellbar, dass in diesem außergewöhnlichen Strukturmerkmal Aminosäuren in einer ungewöhnlichen Konformation vorliegen, in der man sie in idealen Helices nicht vorfindet.

Für die repräsentative Struktur des größten Clusters sind in Tabelle 4.9 die G-Faktoren für die Verteilung der Torsionswinkel sowie der Bindungslängen und der Bindungswinkel des Proteinrückgrats angegeben. Wenn wie in diesem Fall eine Kristallstruktur einer MDS unterzogen wurde, sind die G-Faktoren vor allem ein Maß dafür, wie gut die stereochemischen Gegebenheiten des Proteins vom verwendeten Kraftfeld wiedergegeben werden können.

Tabelle 4.9: Stereochemische Qualität der repräsentativen Struktur aus der MDS mit der Kristallstruktur, angegeben in Form von G-Faktoren

Stereochemischer Parameter	G-Faktor
Torsionswinkel	
Φ - ψ -Winkel	-0,34
χ_1 - χ_2 -Winkel	-0,14
ω -Winkel	-0,80
kovalente Parameter	
Bindungslängen	0,26
Bindungswinkel	0,02
gesamt	-0,16

Mit Ausnahme des Wertes für die ω -Winkel liegen alle G-Faktoren im Bereich für eine gute strukturelle Qualität über dem Grenzwert von -0,5. Die Werte für die kovalenten Parameter liegen sogar im positiven Bereich. Bezüglich der ω -Winkel lässt das Kraftfeld offensichtlich zu große Abweichungen der Peptidbindungen von der Planarität zu. Dieses Phänomen kann man häufig auch bei planaren

Ringsystemen beobachten. Da der Wert größer als -1 ist, liegt er aber immer noch innerhalb der Toleranzgrenze und muss als Schwäche des Kraftfeldes akzeptiert werden.

Abbildung 4.23 zeigt eine Überlagerung der repräsentativen Struktur des größten Clusters bei 2000 ps mit der ursprünglichen Kristallstruktur. Der RMSD-Wert der beiden Strukturen beträgt 0,207 nm und entspricht damit in etwa der durchschnittlichen RMS-Abweichung der Strukturen während der MDS im Bezug auf die Startstruktur (vgl. Abbildung 4.19).

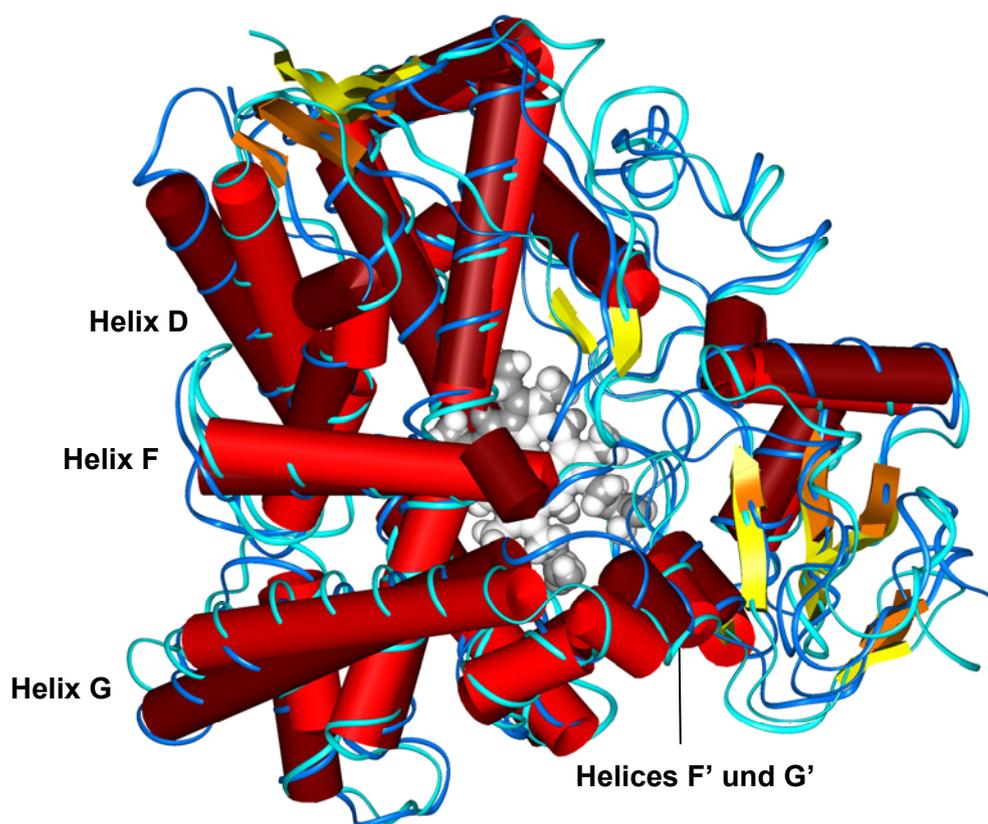


Abbildung 4.23: Überlagerung einer repräsentativen Struktur der MDS mit der Kristallstruktur in der Aufsicht. Farbcodierung der Kristallstruktur: Das Häm ist als hellgraues Karlotenmodell dargestellt, das Proteinrückgrat als türkisfarbenes Band, Helices als hellrote Zylinder und Faltblätter als gelbe Blockpfeile. Farbcodierung der Struktur aus der MDS: Das Häm ist als dunkelgraues Karlotenmodell dargestellt, das Proteinrückgrat als blaues Band, Helices als rotbraune Zylinder und Faltblätter als orangefarbene Blockpfeile.

Die Lage und die Länge der Helices unterscheiden sich im Allgemeinen kaum. Auffällig ist jedoch die unterschiedliche Ausrichtung der Helix D. Das liegt vermutlich daran, dass in der MDS der Faltblattstrang β_{3-1} , der in der Kristallstruktur fehlt, deutlich ausgebildet ist. Dazu muss sich der Loop DE, auf dem sich dieser Faltblattstrang befindet, an den benachbarten antiparallelen β_{3-2} -Strang anlagern. Aus dieser Interaktion resultiert die Verschiebung des C-terminalen Endes der Helix D. Die Nebenhelices F' und G', die den Eingang zum Substratkanal verschließen und

damit ohnehin einen flexiblen Bereich des Enzyms darstellen, unterscheiden sich ebenfalls in ihrer Lage von der Ausgangsstruktur. Auch die angrenzenden Haupthelices F und G können aus diesem Grund in ihrer Orientierung leicht variieren.

Zusammenfassend lässt sich sagen, dass die MDS mit den hier gewählten Simulationsbedingungen in der Lage war, das dynamische Verhalten der Kristallstruktur im leeren, geschlossenen Zustand in Übereinstimmung mit experimentellen Daten und in einer guten strukturellen Qualität wiederzugeben. Die Simulationsbedingungen sollten folglich auch für die Berechnung des CYP1A2-Modells gut geeignet sein.

4.2.4 Equilibrierung des Proteinmodells

4.2.4.1 Strukturelle Entwicklung

Auch die Equilibrierung des Proteinmodells wurde zunächst mithilfe des Programms PROCHECK überprüft.

Tabelle 4.10: Auswertung des Ramachandran Plots im Verlauf der Equilibrierung der Kristallstruktur 1N6B. Die Verteilung der Aminosäurekonformationen auf die Konformationsbereiche ist in Prozent angegeben. Für den zulässigen und den unerlaubten Konformationsbereich ist die absolute Anzahl der darin enthaltenen Aminosäuren in Klammern aufgeführt.

Konformationsbereich	nach Minimierung	nach Equilibrierung
bevorzugter Bereich	80,3	83,9
günstiger Bereich	16,8	14,4
zulässiger Bereich	1,4 (6)	0,7 (3)
unerlaubter Bereich	1,4 (6)	1,0 (4)

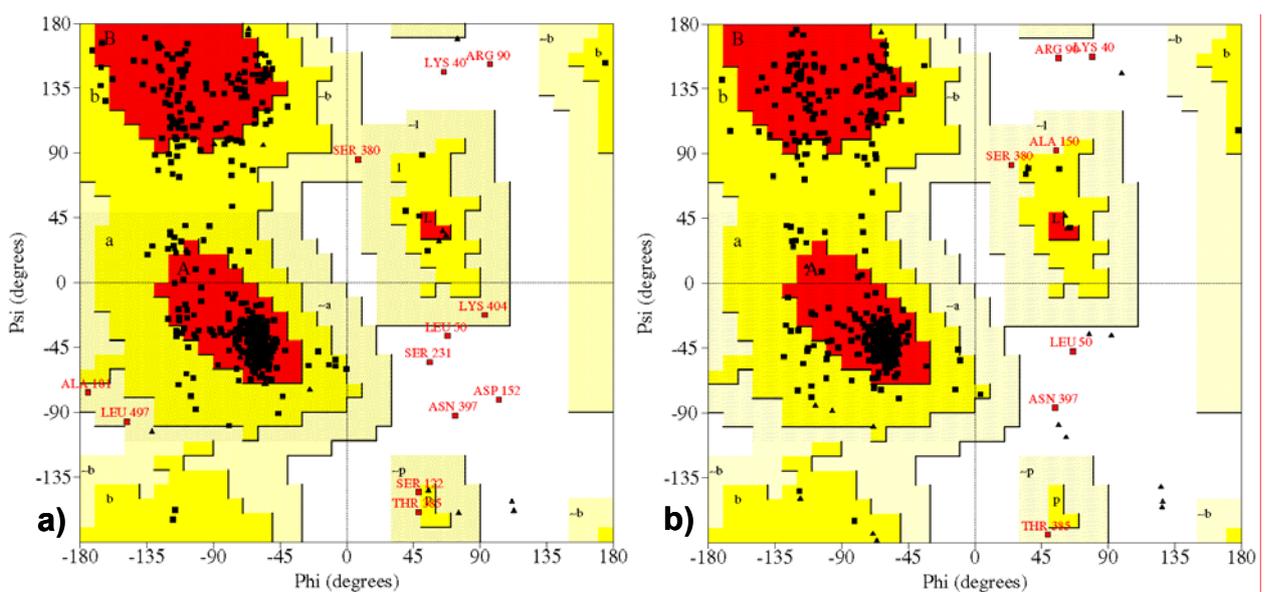


Abbildung 4.24: Ramachandran Plot a) nach einer Minimierung des Proteinmodells und b) nach der Equilibrierungsdynamik

Nach der Minimierung befanden sich insgesamt zwölf Aminosäuren in den ungünstigen Konformationsbereichen des Ramachandran Plots (Abbildung 4.24 a) und Tabelle 4.10). Davon liegen zwei am N-Terminus (Lys 40, Leu 50) und vier weitere in den per *loop search* eingefügten Bereichen des Modells (Ser 122, Asp 152, Ala 181, Ser 231)⁷.

Um eine effektive Entspannung dieser Bereiche während der Equilibrierungsdynamik zu erreichen, wurden die beiden Termini und die per *loop search* eingefügten Abschnitte, wie in Tabelle 4.11 aufgelistet, komplett frei gelassen.

Tabelle 4.11: Proteinregionen mit Aminosäuren, die während der Equilibrierungsdynamik frei gelassen wurden.

Proteinregion	frei gelassene Aminosäuren
N-Terminus	39-53
Loop B''C	119-130
Loop CD	144-153
Loop DE	180-188
Loop FF'	230-232
Loop HI	296-303
Loop β_{3-3} - β_{4-1}	488-491
C-Terminus	512-515

Nach der Equilibrierungsdynamik hatte sich die strukturelle Qualität des Modells deutlich verbessert. In den Loops B''C, DE und FF' befanden sich keine Aminosäuren mehr, die eine ungünstige Konformation aufweisen und in Loop CD lag nur noch eine Aminosäure im zulässigen Konformationsbereich. Lediglich im N-Terminus kam es zu keiner deutlichen Entspannung der Struktur.

4.2.4.2 Potentielle Energie und RMSD-Wert

Abbildung 4.25 zeigt den Verlauf der potentiellen Energie während der freien MDS. Auch hier erreicht die potentielle Energie nach 500 ps ein Plateau. Der Mittelwert beträgt $-466.676,8 \text{ kJ}\cdot\text{mol}^{-1}$ und unterscheidet sich damit kaum von der MDS mit der Kristallstruktur ($-463.403,3 \text{ kJ}\cdot\text{mol}^{-1}$). Das Simulationssystem bleibt somit auch hier bis zum Ende der MDS stabil.

Analog der MDS mit der Kristallstruktur ist auch bei der MDS mit dem Proteinmodell nach 1500 ps die Equilibrierung der Proteinstruktur erfolgt (Abbildung 4.26). Der RMSD-Wert beträgt in der equilibrierten Phase der MDS 0,25 nm. Im Vergleich mit dem Verlauf des RMSD-Wertes der MDS mit der Kristallstruktur fällt zum einen auf,

⁷ In diesem und in den folgenden Kapiteln bezieht sich die Nummerierung der Aminosäuren stets auf die Originalnummerierung der Proteinsequenz, um den Vergleich mit Literaturdaten zu erleichtern.

dass die Schwankungen des RMSD-Wertes beim Proteinmodell stärker ausgeprägt sind, und zum anderen, dass der RMSD-Wert des Proteinmodells insgesamt um etwa 0,04 nm höher liegt als der der Kristallstruktur. Beides ist auf den Modellierungsprozess zurückzuführen.

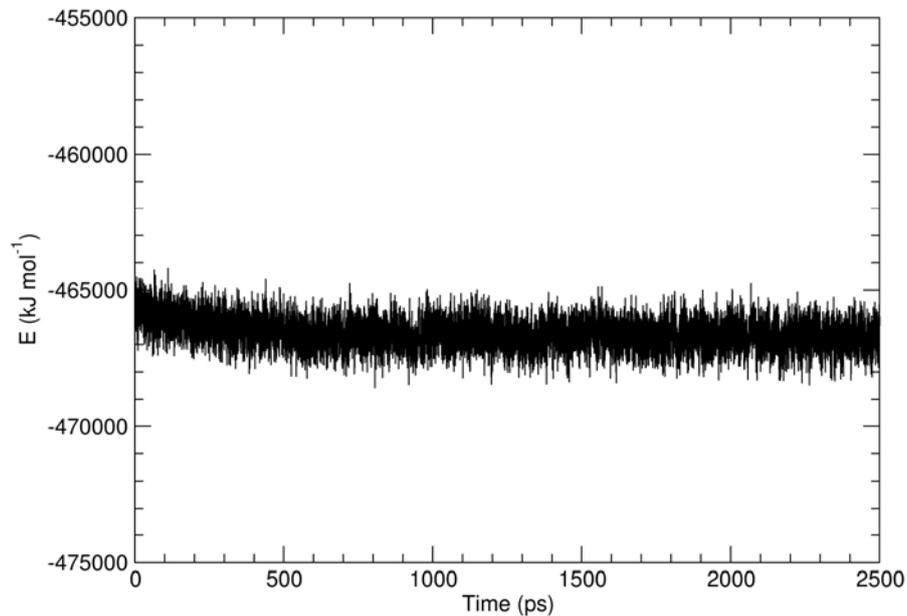


Abbildung 4.25: Verlauf der potentiellen Energie des Simulationssystems während der freien MDS.

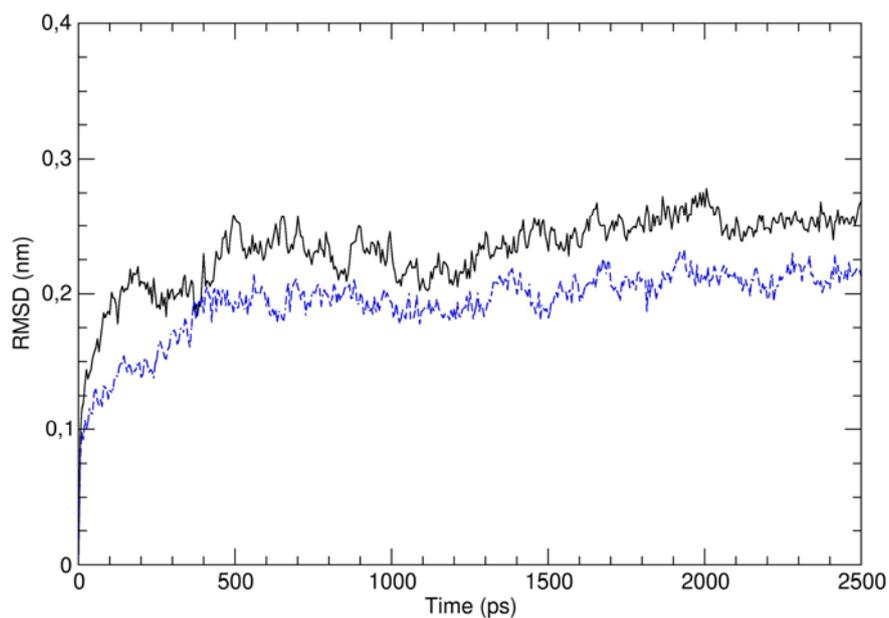


Abbildung 4.26: RMSD-Wert des Proteinrückgrates bezogen auf die Startstruktur im Verlauf der freien MDS mit dem Proteinmodell (schwarz). Zum Vergleich ist der RMSD-Wert der MDS mit der Kristallstruktur eingetragen (blau und gestrichelt).

4.2.5 Auswertung der MDS mit dem Modell

4.2.5.1 RMS-Fluktuation

Wie bereits bei der MDS mit der Kristallstruktur wurde auch hier die Beweglichkeit einzelner Proteinbereiche während der equilibrierten Phase der MDS zur Auswertung herangezogen (Abbildung 4.27). Im direkten Vergleich mit der Kristallstruktur fällt zunächst auf, dass die Bewegungen des CYP1A2-Modells vor allem in der helixreichen Region des Proteins grundsätzlich stärker sind als die der Kristallstruktur. Das Einfügen von Proteinsegmenten mittels *loop search* als Verbindung zwischen den Helices bringt erwartungsgemäß eine gewisse Instabilität des Modells mit sich. Außerdem weist die N-terminale Hälfte der Cytochrome P450 bis zur Helix I in der Regel generell eine deutlich niedrigere Sequenzidentität auf als die C-terminale Hälfte [83] (s. auch Abschnitt 1.4.2), was ebenfalls die Stabilität eines auf Sequenzhomologien beruhenden Modells in diesem Proteinbereich herabsetzt. Tatsächlich kann man deutlich erkennen, dass der Bereich vor Helix I bis etwa zur Aminosäure 300 auf einem höheren Bewegungsniveau liegt, als der Proteinabschnitt danach.

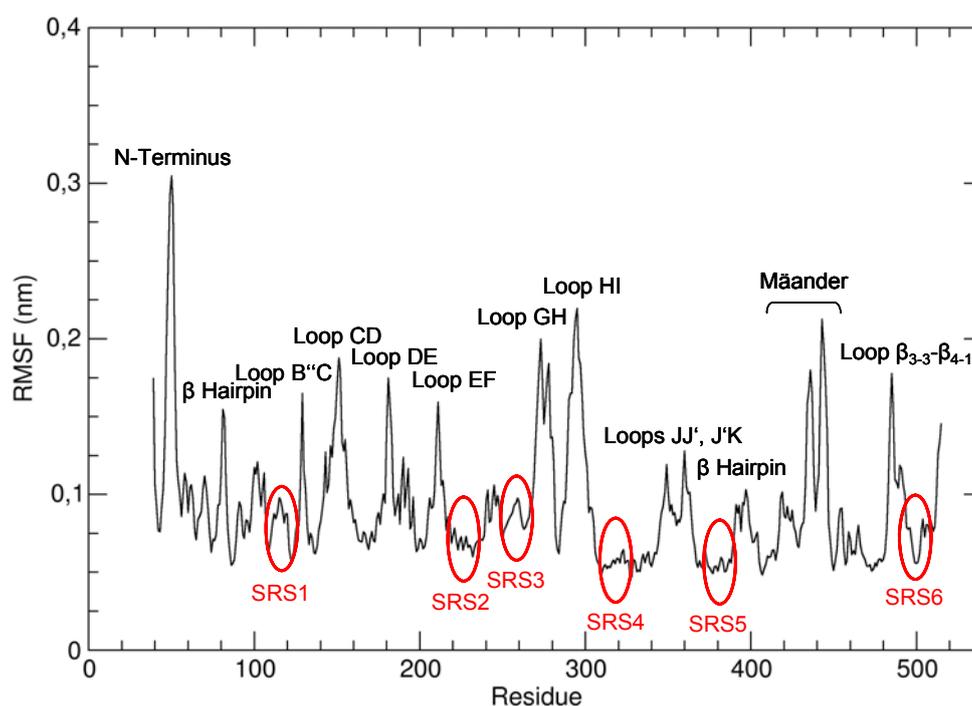


Abbildung 4.27: RMS-Fluktuation der einzelnen Aminosäuren zwischen 1500 und 2500 ps der freien MDS. Die roten Markierungen umfassen die Substratbinderegionen SRS1-6.

Das höchste Maximum und damit der beweglichste bzw. instabilste Bereich des Modells befindet sich am N-Terminus. Hierbei handelt es sich mit Sicherheit um ein Artefakt des Modells, das sich einerseits dadurch erklären lässt, dass der N-Terminus einer der am geringsten konservierten Bereiche der Proteinsequenz ist. Andererseits weist auch die Kristallstruktur in diesem Bereich eine gewisse

Instabilität auf, was sich damit begründen lässt, dass der Membrananker zur Kristallisation entfernt wurde, und daher sowohl in der MDS mit der Kristallstruktur als auch in der MDS mit dem Modell als strukturgebendes Element fehlt. Laut Sekundärstrukturvorhersage besteht beispielsweise der Membrananker von CYP1A2 aus einer 28 Aminosäuren langen Helix (vgl. Abbildung 4.8).

Im weiteren Verlauf der RMS-Fluktuation lassen sich klare Parallelen zu den entsprechenden RMSF-Werten der Kristallstruktur erkennen. So sind die nachfolgenden Maxima in der Helixdomäne wie auch bei der Kristallstruktur eindeutig den Loops zwischen den Helices zuordnen. Obwohl der Loop CD mit 16 Aminosäuren am längsten ist, verursacht er nicht das höchste Maximum, weil hier eine Verlängerung der Helix D eingefügt wurde (vgl. Abbildung 4.13), die den Loop offensichtlich gut stabilisiert. Aus der nur mäßigen Beweglichkeit des Bereichs Helix F – Helix F' – Helix G' – Helix G zwischen den Regionen SRS2 und SRS3 kann man erkennen, dass auch das CYP1A2-Modell über den gesamten Verlauf der MDS in einem leeren aber geschlossenen Zustand vorliegt.

Analog zur Kristallstruktur befinden sich deutlich ausgeprägte Maxima im Bereich der Mäander-Region, die sich aufgrund derselben Strukturmerkmale ergeben (s. Abschnitt 4.2.3). Auch in der Faltblattdomäne ist der Verlauf der RMS-Fluktuation ähnlich dem der Kristallstruktur.

Die Substratbinderegion SRS1 befindet sich bei CYP1A2 in der Helix B''. Diese kurze Helix gehört wie der Bereich von Helix F bis Helix G und SRS6 zum Eingang des Substratkanals und ist deshalb relativ beweglich. Dass SRS6 sich hier eher in einem Minimum der RMS-Fluktuation befindet, liegt daran, dass Thr 498 aus SRS6 im leeren Enzym eine Wasserstoffbrücke zum benachbarten Ser 231 der Helix F' ausbildet, die die Beweglichkeit dieses Bereiches einschränkt. Sobald sich allerdings ein Substrat im aktiven Zentrum des Modells befindet, wird diese Wasserstoffbrücke gebrochen. In SRS4 ist wiederum der charakteristische Knick in der Mitte der Helix I als kleine Spitze zu erkennen.

4.2.5.2 Protein-Häm-Interaktionen

Die Mehrzahl der an der Hämbindung beteiligten Aminosäuren ist hochkonserviert. Deshalb kann man die meisten der in der equilibrierten Phase der MDS mit dem CYP1A2-Modell aufgetretenen Salz- und Wasserstoffbrücken zwischen dem Protein und den Propionatseitenketten der Hämringe A und D entsprechenden Wechselwirkungen in der Kristallstruktur zuordnen. Arg 137, Trp 133 und Arg 456 des Modells bilden Salz- bzw. Wasserstoffbrücken zur Propionatseitenkette des Hämringes D (Abbildung 4.28) in Analogie zu Arg 124, Trp 120 und Arg 430 der Kristallstruktur. Die Wasserstoffbrücke zu Trp 133 ist dabei wesentlich schwächer als die entsprechende Wasserstoffbrücke in der Kristallstruktur (Abbildung 4.29). Das liegt daran, dass sich Trp 133 während der MDS nach oben biegt und über den

Indolstickstoff eine Wasserstoffbrücke zum Carbonylsauerstoff des Proteinrückgrats von Leu 123 ausbildet. Dadurch schiebt sich der Phenylring von Trp 133 vor das Propionat und verhindert so den gleichzeitigen Erhalt der Wasserstoffbrücke zum Häm. Diese strukturelle Störung am N-terminalen Ende der Helix C ist vermutlich auf die unmittelbar vor und nach Helix C per *loop search* eingefügten Proteinsegmente zurückzuführen.

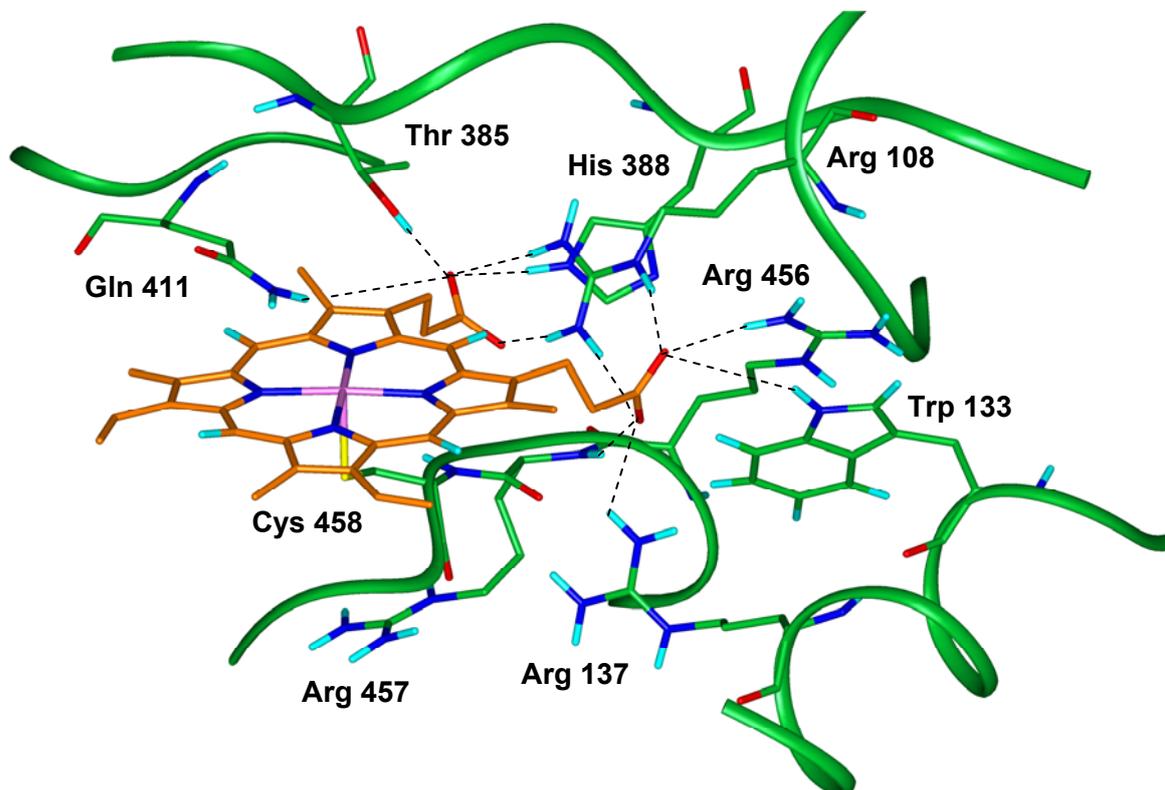


Abbildung 4.28: Protein-Häm-Interaktionen im Proteinmodell von CYP1A2 am Beispiel einer repräsentativen Struktur. Das Protein ist in grün dargestellt, das Häm in orange. Alle Heteroatome sind atomkodierte eingefärbt.

Arg 108 entspricht Arg 97 der Kristallstruktur und bildet Salzbrücken zu beiden Propionatseitenketten des Häm. Die Propionatseitenkette des Hämrings A ist außerdem über His 388 (His 365 in der Kristallstruktur) und Thr 385 mit dem Protein verbunden. Thr 385 befindet sich im CYP1A2-Modell an derselben Stelle wie Asn 362 in der Kristallstruktur, das ebenfalls eine stabile Wasserstoffbrücke zum Propionat des Hämrings A bildet. Zusätzlich zu diesen konservierten Interaktionen kommt es zur Bildung einer Wasserstoffbrücke zwischen der Amidfunktion des Proteinrückgrats von Arg 457 und dem Propionat des Hämrings D. Dafür fehlt im Modell die Interaktion zum Proteinrückgrat von Ile 386. Ein zusätzliches Artefakt der Simulation ist eine schwach ausgeprägte Wasserstoffbrücke von Gln 411, das nicht in einer der Hämbinderegionen liegt, zum Propionat des Hämrings A. Die Wasserstoffbrücke zu Ser 426 oder einer entsprechenden Aminosäure fehlt völlig, da in CYP1A2 an dieser Stelle ein Glycin liegt. Alle umliegenden Aminosäuren sind lipophil und enthalten keine Wasserstoffbrückendonatoren. Ser 426 kommt zwar in

allen Isoenzymen der Familie 2 vor (s. Abbildung 4.10), ist aber darüber hinaus nicht konserviert [83].

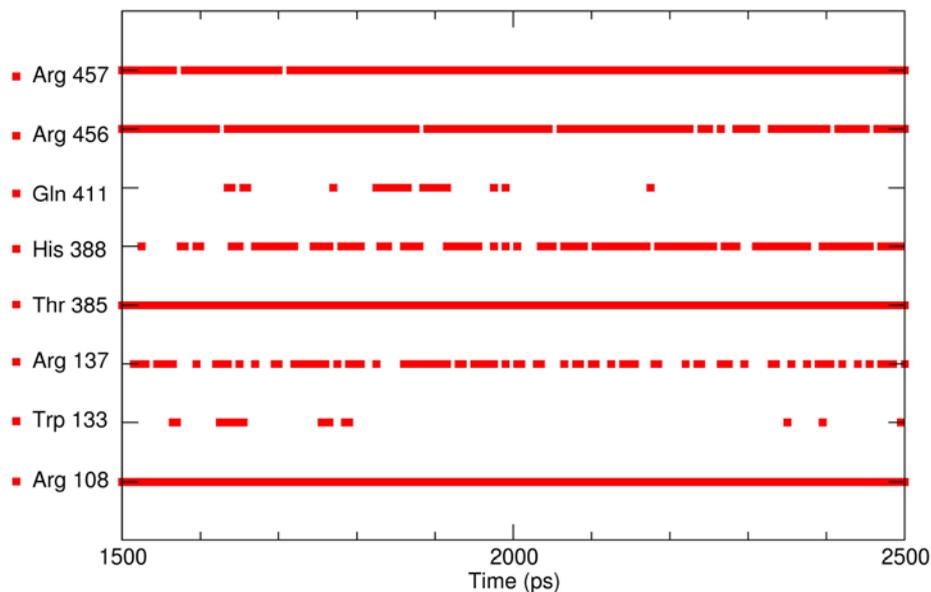


Abbildung 4.29: Protein-Häm-Wechselwirkungen des CYP1A2-Modells im Verlauf der equilibrierten Phase der MDS.

4.2.5.3 Wasserstoffbrückenmuster in Helix I

Im Gegensatz zu den Kristallstrukturen 1N6B und 1NR6 befindet sich im Verlauf der MDS mit dem CYP1A2-Modell kein Wasser in der Helixwindung über dem katalytischen Zentrum (Abbildung 4.30).

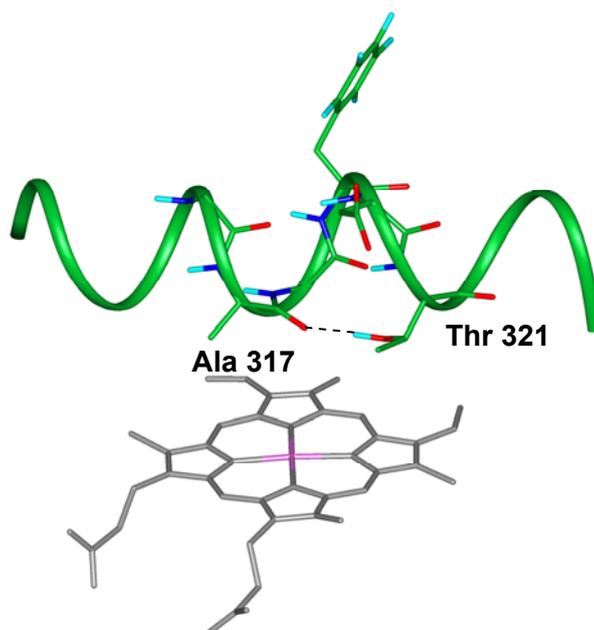


Abbildung 4.30: Irreguläre Wasserstoffbrücke in Helix I am Beispiel der repräsentativen Struktur aus der MDS mit dem CYP1A2-Modell.

Trotzdem bleibt die charakteristische Aufweitung der Helixwindung erhalten, die hier durch eine direkte Wasserstoffbrücke zwischen Ala 317 und der Hydroxylgruppe von Thr 321 entsteht. Zieht man weitere Kristallstrukturen zum Vergleich heran, stellt man fest, dass in der Kristallstruktur 1PQ2 (CYP2C8 mit Palmitinsäure) und in zwei weiteren Kristallstrukturen von CYP3A4 (1W0E und 1W0G) ebenfalls kein Wasser in der Helixwindung über dem Hämeisen enthalten ist. Allen Strukturen gemeinsam ist jedoch die Unterbrechung des regelmäßigen Wasserstoffbrückenmusters durch eine Wasserstoffbrücke zur Seitenkette des konservierten Threonins, wie sie auch im vorliegenden Modell beobachtet wird.

4.2.5.4 Proteinstruktur

Der Konformationsraum der Proteinstruktur in der equilibrierten Phase der MDS (1500-2500 ps) wurde wiederum mithilfe einer Clustereinteilung analysiert (Tabelle 4.12). Auch hier sind die großen Cluster gleichmäßig über den Simulationszeitraum verteilt, worin sich eine gute Equilibrierung der Struktur widerspiegelt. Erwartungsgemäß ist die Qualität der Proteinstruktur bei einem Homologiemodell etwas schlechter als diejenige der Kristallstruktur im Verlauf der MDS. Lagen bei der Kristallstruktur 81,5-84 % der Aminosäuren im bevorzugten Konformationsraum und nur 5-9 Aminosäuren in den ungünstigen Konformationsbereichen, so sind es beim Modell 78,1-81,5 % bzw. 7-14 Aminosäuren.

Tabelle 4.12: Übersicht über die fünf größten Cluster der NMRCLUST-Analyse mit Zeitpunkt und struktureller Qualität der jeweils repräsentativen Struktur.

Nr.	Anzahl Strukturen	Zeitpunkt (ps)	Konformationsbereiche nach PROCHECK			
			bevorzugt (%)	günstig (%)	zulässig (%)	unerlaubt (%)
1	18	2390	78,1	20,2	0,5 (2)	1,2 (5)
2	15	1780	78,6	19,0	0,7 (3)	1,7 (7)
3	13	2445	79,8	18,0	1,0 (4)	1,2 (5)
4	13	1580	81,5	16,1	1,4 (6)	1,0 (4)
5	12	1690	79,8	16,8	1,7 (7)	1,7 (7)

In Bezug auf die Anzahl der Aminosäuren in den ungünstigen Konformationsbereichen ist auch hier die repräsentative Struktur des größten Clusters die beste Struktur. Abbildung 4.31 gibt einen Überblick über die Lage der Aminosäuren in den ungünstigen Konformationsbereichen für diese Struktur. Lys 106 und Asn 142 weisen hier eine zulässige Konformation auf. Lys 106 liegt im Übergang zwischen Helix B und dem Faltblatt β_{1-5} und Asn 142 gehört zum Loop zwischen den Helices C und D, der per *loop search* eingefügt wurde. Von den fünf im unerlaubten Konformationsbereich liegenden Aminosäuren befinden sich ebenfalls drei Aminosäuren in Loops, für die ein *loop search* durchgeführt wurde. Das sind Ile 440 im Loop zwischen Mäanderregion und Cystein-Tasche, Leu 301 im Loop HI und Ser 232 im Loop FF'. Die ungewöhnliche Konformation von Ser 232 kommt durch eine

Wasserstoffbrücke der Amidfunktion des Proteinrückgrats mit der Seitenkette von Tyr 112 der Helix B'' zustande. Eine ähnliche Interaktion findet man in der Kristallstruktur 1PQ2 (CYP2C8). Dort befindet sich an der Stelle des Serins ein Asparagin, dessen Seitenkette eine Wasserstoffbrücke zur NH-Gruppe des Proteinrückgrats von Val 477 im Turn zwischen Faltblatt β_{4-1} und β_{4-2} ausbildet. Offensichtlich tragen derartige Wasserstoffbrücken zwischen einer Aminosäureseitenkette und dem Proteinrückgrat zum Verschluss des Substratkanals bei. Im Verlauf der MDS mit Substraten im aktiven Zentrum des Modells verändert sich dieses Interaktionsmuster teilweise, und Ser 232 geht dann in eine zulässige Konformation über. Im Übrigen sollte man bei der Betrachtung des Loops FF' stets berücksichtigen, dass sich in diesem Bereich auch in der Kristallstruktur mutierte Aminosäuren befinden, um die Kristallisation zu ermöglichen. Diese Mutationen können sich durchaus auf die Konformation und die Stabilität des C-terminalen Endes von Helix F und den daran angrenzenden Loop auswirken. Die übrigen beiden Aminosäuren befinden sich am N-terminalen Ende des Modells. Leu 50 wurde bereits aus der Kristallstruktur in einer unerlaubten Konformation übernommen (Ile 38), und Arg 90 liegt im Übergang zwischen Faltblatt β_{1-2} und Helix B.

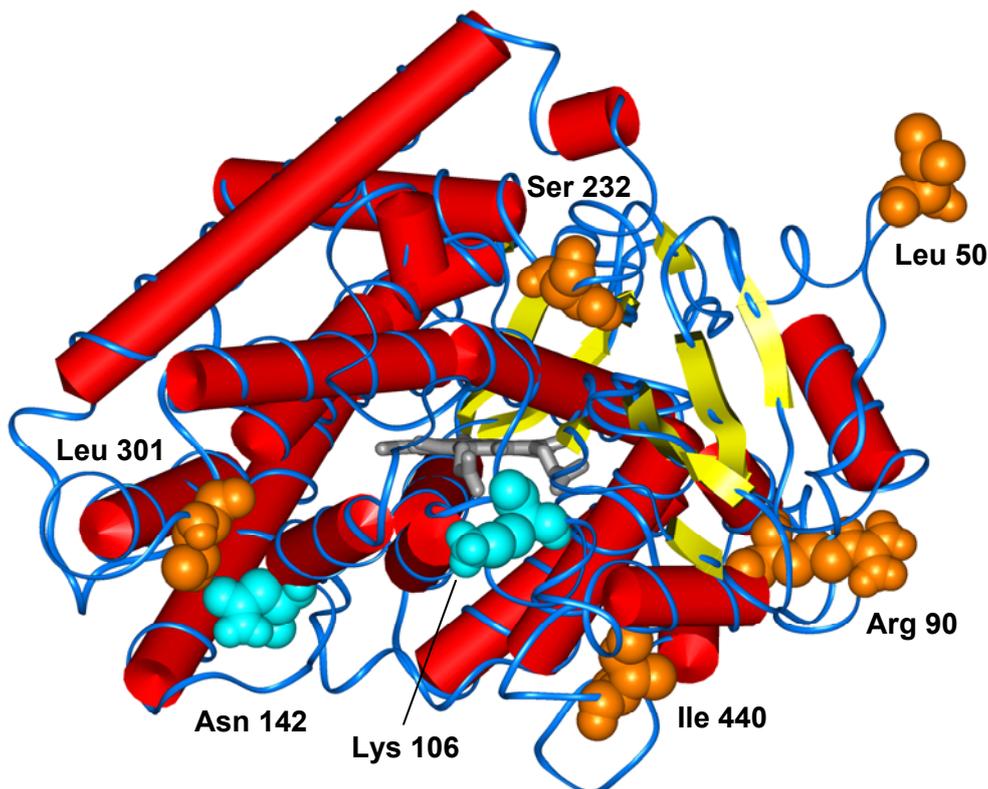


Abbildung 4.31: PROCHECK-Auswertung der repräsentativen Struktur aus der MDS. Aminosäuren in zulässiger Konformation sind türkis eingefärbt und Aminosäuren in unerlaubter Konformation orange.

Die G-Faktoren, die wiederum für die repräsentative Struktur des größten Clusters exemplarisch berechnet wurden (Tabelle 4.13), unterscheiden sich nur wenig von den entsprechenden Werten für die repräsentative Struktur aus der MDS mit der

Kristallstruktur (vgl. Tabelle 4.9). Das bedeutet, dass sich der Modellierungsprozess nur wenig auf die stereochemische Qualität ausgewirkt hat. Geringfügig schlechtere Werte findet man für die Verteilung der Bindungswinkel und für die Verteilung der Φ - und ψ -Winkel (s. ausführliche Auswertung oben). Insgesamt liegt das Modell in einer guten stereochemischen Qualität vor.

Tabelle 4.13: Stereochemische Qualität der repräsentativen Struktur des Modells aus der MDS, angegeben in Form von G-Faktoren

Stereochemischer Parameter	G-Faktor
Torsionswinkel	
Φ - ψ -Winkel	-0,46
χ_1 - χ_2 -Winkel	-0,14
ω -Winkel	-0,76
kovalente Parameter	
Bindungslängen	0,25
Bindungswinkel	-0,08
gesamt	-0,19

Die Überlagerung der repräsentativen Struktur aus der MDS mit der Kristallstruktur in Abbildung 4.32 zeigt, dass im Wesentlichen alle Sekundärstrukturelemente erhalten geblieben sind. Die einzige Ausnahme bildet die Nebenhelix F'. Sie wird nicht mehr explizit als Helix angezeigt. Zwar sollte sich laut der Sekundärstrukturvorhersage dort eine Helix befinden, aber wie bereits in Abschnitt 4.1.2 erläutert, kann die Sekundärstruktur gerade in diesem Bereich nicht unbedingt korrekt vorhergesagt werden. Darüber hinaus gibt es auch Kristallstrukturen, in denen eine der beiden Nebenhelices F' oder G' nicht deutlich als Helix ausgeprägt sind. Das gilt beispielsweise für die Helix F' in der Struktur 1NR6 (CYP2C5) und für die Helix G' in der Struktur 1PQ2 (CYP2C8). Da sich diese Helices an der Oberfläche des Proteins befinden, könnte das allerdings auch auf die Kristallisation zurückzuführen sein. Im Verlauf der MDS des Modells mit Substraten im aktiven Zentrum treten teilweise beide Nebenhelices F' und G' auf. Es wäre also denkbar, dass ein Substratkomplex stabilisierend auf diesen Bereich des Modells wirkt. Wie schon in der MDS mit der Kristallstruktur bildet sich auch hier während der MDS der Faltblattstrang β_{3-1} , der in der Kristallstruktur fehlt. Des Weiteren hat sich die Länge einiger Helices verändert, was zum größten Teil in Übereinstimmung mit der Sekundärstrukturvorhersage bzw. mit den eingefügten Loopstrukturen steht. Helix B'' zum Beispiel ist im Modell deutlich kürzer als in der Kristallstruktur und Helix E ist deutlich länger (vgl. Abbildung 4.11). Für die Helices D und H wurden per *loop search* Helixverlängerungen eingefügt, die auch stabil geblieben sind.

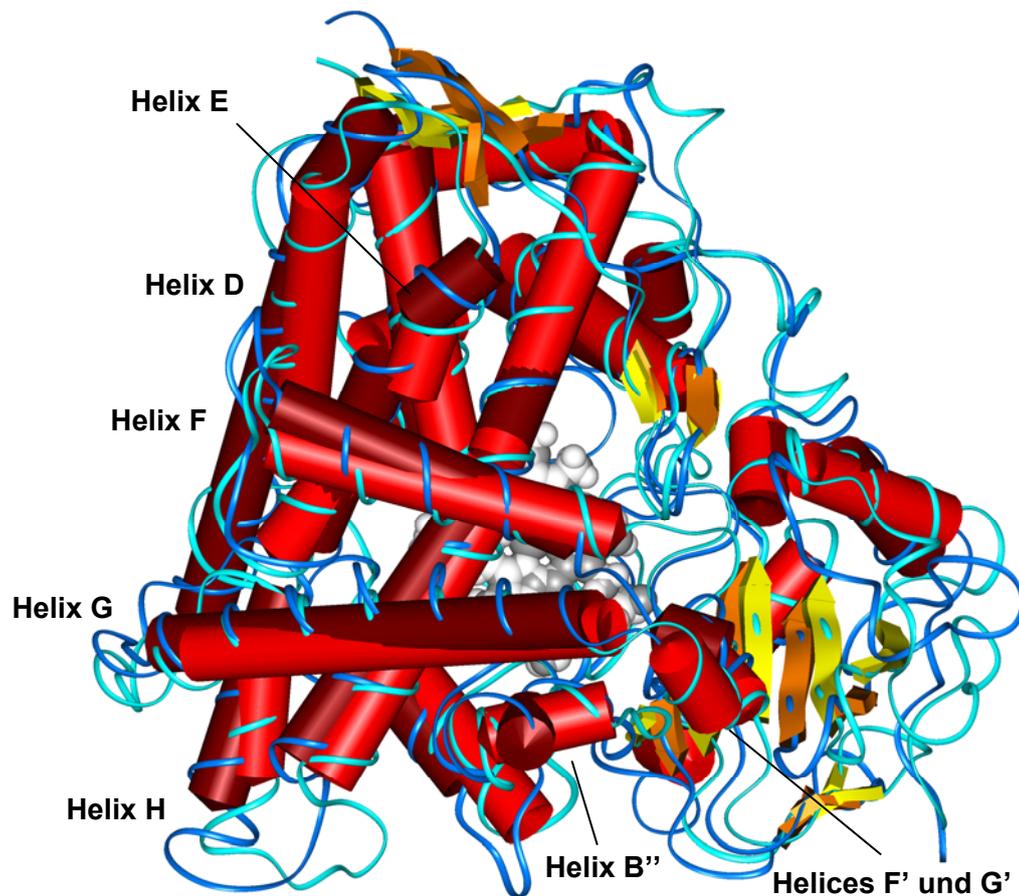


Abbildung 4.32: Überlagerung der repräsentativen Struktur des CYP1A2-Modells aus der MDS mit der Kristallstruktur in der Aufsicht. Farbcodierung der Kristallstruktur: Das Häm ist als hellgraues Karlotenmodell dargestellt, das Proteinerückgrat als türkisfarbenes Band, Helices als hellrote Zylinder und Faltblätter als gelbe Blockpfeile. Farbcodierung der Struktur aus der MDS: Das Häm ist als dunkelgraues Karlotenmodell dargestellt, das Proteinerückgrat als blaues Band, Helices als rotbraune Zylinder und Faltblätter als orangefarbene Blockpfeile.

4.3 Untersuchung von Enzym-Substrat-Komplexen

4.3.1 Auswahl von Substraten und Positionierung in das Modell

Eine weitere wichtige Validierung des Modells besteht in der Untersuchung von Enzym-Substrat-Interaktionen. Hierzu sollten Substrate ausgewählt werden, für die es eine möglichst umfangreiche experimentelle Grundlage gibt, die dazu geeignet ist, anhand einer MDS nachvollzogen zu werden. Diese experimentellen Daten können einerseits Mutationsstudien sein, deren Auswirkung auf die Bindung bestimmter Substrate man mithilfe eines Homologiemodells erklären kann, oder andererseits Informationen über Regioselektivitäten oder Umsetzungsraten verschiedener Substrate.

Im Rahmen der vorliegenden Arbeit wurden die bereits in der Einleitung vorgestellten Verbindungen MeIQ (Abschnitt 1.2.1.1), 7-Ethoxyresorufin und Coffein (Abschnitt 1.2.1.3) als repräsentative Substrate für eine detaillierte Untersuchung mittels MDS ausgesucht (Abbildung 4.33). Für MeIQ und Ethoxyresorufin existieren Mutationen von Aminosäuren der SRS1-5 des humanen Isoenzym, die die katalytische Aktivität signifikant verändern und somit vermutlich eine wichtige Rolle bei der Substratbindung spielen. Von Coffein weiß man, dass durch Umsetzung mit CYP1A2 verschiedene Metabolite gebildet werden können, und man kennt deren Bildungsraten. In diesem Fall ist es interessant, anhand einer Berechnung verschiedener Substratkomplexe eine Erklärung für die unterschiedlich häufige Bildung der Metabolite zu finden. In der vorliegenden Arbeit wurden daher zwei MDS mit Coffein durchgeführt, einmal in der Position zur Umsetzung zum Hauptmetaboliten Paraxanthin und einmal in der Position zur Umsetzung zum Nebenmetaboliten Theobromin, der zu ca. 12 % gebildet wird.

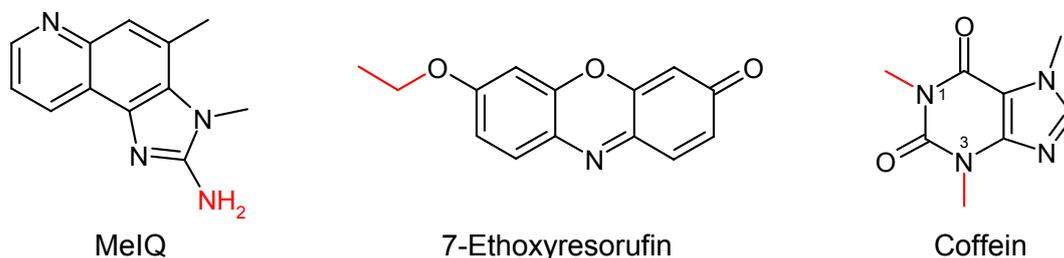


Abbildung 4.33: Strukturformeln und Hydroxylierungsstellen der drei zur Berechnung von Substratkomplexen mit CYP1A2 verwendeten Substrate (vgl. auch Einleitung)

Als Ausgangsstruktur des Enzyms wurde die bereits in Abschnitt 4.2.5.4 gründlich untersuchte repräsentative Struktur aus der MDS mit dem leeren Proteinmodell eingesetzt. Von den drei Substratmolekülen konnte Coffein direkt als Kristallstruktur aus der *Cambridge Structural Database* [162] entnommen werden, 7-Ethoxyresorufin basiert auf einer Kristallstruktur von 7-Methoxyresorufin und MeIQ wurde mit der *sketch molecule* Routine des Programms SYBYL [161] neu konstruiert. Auf der

Grundlage der Kristallstrukturdaten und den Kraftfeldparametern des Tripos-Kraftfeldes wurden den Verbindungen für die spätere Durchführung von MDS in GROMACS analoge Parameter des GROMOS96-Kraftfeldes zugewiesen. Die Partialladungen der Atome wurden in einer quantenchemischen Berechnung mit dem Programm JAGUAR bestimmt und an die im GROMOS96-Kraftfeld für Aminosäuren verwendeten Partialladungen angeglichen. Die Durchführung einer Konformationsanalyse zur Generierung einer günstigen Startkonformation war nicht nötig, da alle Substrate aus starren Ringsystemen bestehen und nur wenig konformatorischen Spielraum lassen.

Ein wichtiger Schritt ist die Positionierung der Substrate im aktiven Zentrum, um in der anschließenden MDS von einer möglichst guten Startkonformation des Enzym-Substrat-Komplexes ausgehen zu können. Zunächst wurde deshalb die Position der Substrate in den drei als Kristallstrukturen zur Verfügung stehenden Komplexen von CYP2C5 bzw. CYP2C9 mit DMZ, Diclofenac bzw. Flurbiprofen analysiert (Abbildung 4.34). Da die Cytochromfamilien CYP1 und 2 eng miteinander verwandt sind und eine hohe Sequenzidentität untereinander aufweisen, ist auch die Lage der Substrate in den Kristallstrukturen ein guter Anhaltspunkt für die Positionierung von MelQ, 7-Ethoxyresorufin und Coffein im Homologiemodell von CYP1A2.

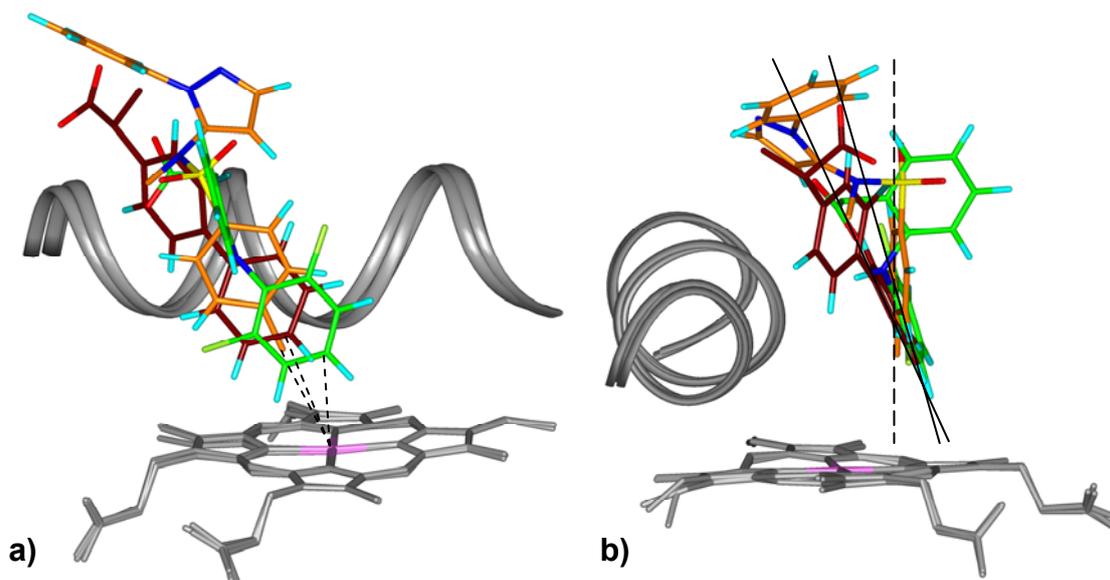


Abbildung 4.34: Position der Substrate aus den Kristallstrukturen über dem Häm. Orange: DMZ, grün: Diclofenac, dunkelrot: Flurbiprofen. Alle Heteroatome sind atomkodiert eingefärbt. Zur Orientierung sind das Häm und der zentrale Abschnitt der Helix I eingeblendet. a) Vorderansicht. Der Abstand zum Hämeisen (jeweils als gestrichelte Linie eingezeichnet) beträgt für DMZ 4,28 Å, für Diclofenac 4,66 Å und für Flurbiprofen 4,91 Å. b) Seitenansicht. Die Neigung zur Hämenebene ist jeweils durch eine Linie dargestellt. Gestrichelte Linie: DMZ (aliphatische Hydroxylierung), durchgezogene Linie: Diclofenac und Flurbiprofen (aromatische Hydroxylierung).

Alle Substrate sind mit ihrer Hydroxylierungsstelle zum Häm hin orientiert, wobei die Hydroxylierungsstelle zwischen 4,3 und 4,9 Å vom Zentralatom Eisen entfernt ist. Darüber hinaus sind alle Substrate in Richtung der Helices B'', F und G geneigt, wo

SRS1-3 lokalisiert sind. Aus der Seitenansicht wird deutlich, dass die beiden Substrate, die einer aromatischen Hydroxylierung unterliegen eine andere Neigung zum Hämeisen hin aufweisen als DMZ, das an einer aliphatischen Seitenkette hydroxyliert wird. Die Positionierung steht dabei aber jeweils im Einklang mit dem der Umsetzung zugrunde liegenden Mechanismus (vgl. Abschnitt 1.3.2.1 und 1.3.2.2). Bei einer aromatischen Hydroxylierung ist die Bindung der Eisen-Oxo-Verbindung an den aromatischen Kohlenstoff der erste Schritt. Dementsprechend ist das Ringsystem so zum Häm hin geneigt, dass dieser elektrophile Angriff gut erfolgen kann und nicht vom Wasserstoffsubstituenten behindert wird. Im Fall der aliphatischen Hydroxylierung erfolgt als erstes die Übertragung eines Wasserstoffatoms vom aliphatischen Kohlenstoff auf die Eisen-Oxo-Verbindung. Die Methylgruppe des DMZ steht daher senkrecht zur Hämebene, sodass eines der substituierten Wasserstoffatome in Richtung Hämeisen zeigen muss und leicht auf den Sauerstoff übertragen werden kann.

Das aktive Zentrum der Startkonformation des Proteinmodells wurde mithilfe der Programme SURFNET und GRID untersucht (Abschnitt 3.5.2 und 3.6.2). Die SURFNET-Parameter für den minimalen und den maximalen Kugelradius wurden so gewählt, dass das größtmögliche zusammenhängende Volumen im aktiven Zentrum entsteht. In diesem Fall betrug der minimale Kugelradius 1,8 Å und der maximale Kugelradius 3,2 Å. Aus Abbildung 4.35 und Abbildung 4.36 kann man entnehmen, dass sich das aktive Zentrum wie auch bei den Kristallstrukturen in die Kavität zwischen den Helices B', F und G ausdehnt, und dass folglich die Orientierung der Substrate im Modell ähnlich ist wie in den Kristallstrukturen. Zur Charakterisierung von möglichen Enzym-Substrat-Interaktionen wurde das aktive Zentrum mit GRID-Sonden abgetastet, die die physikochemischen Eigenschaften von funktionellen Gruppen der Substrate wiedergeben. Das waren zunächst die aromatische Kohlenstoffsonde zur Detektion aromatischer und lipophiler Interaktionsbereiche und die Carbonylsonde in der Funktion als Wasserstoffbrückenakzeptor. Die Kontur der Kohlenstoffsonde gibt im Wesentlichen den Umriss der SURFNET-Kavität wieder, es ist dabei aber deutlicher zu erkennen, dass die Ringsysteme der Substrate wie auch in zwei der drei Kristallstrukturen zur Helix I hin geneigt sein sollten (s. Abbildung 4.36 b). Die Interaktionsfelder der Carbonylsonde zeigen zahlreiche Übereinstimmungen mit der Lokalisation Wasserstoffbrücken donierender Aminosäuren in SRS1, 3 und 6 (Abbildung 4.35 a-c).

Auf der Grundlage dieser Voruntersuchungen wurden die Substrate so ins aktive Zentrum positioniert, dass Orientierung und Abstand der Hydroxylierungsstelle zum Hämeisen mit den experimentellen Daten aus den Kristallstrukturen übereinstimmen, dass das Volumen der Substrate möglichst gut mit der SURFNET-Kavität übereinstimmt, und dass Wasserstoffbrückenakzeptoren der Substrate mit den GRID-Feldern der Carbonylsonde zur Deckung kommen. Beim 7-Ethoxyresorufin (Abbildung 4.35 a) ist der Abstand zwischen dem α -Kohlenstoff des Ethoxyl-

substituenten und dem Hämeisen mit 5,04 Å relativ groß, was auf die voluminöse Ethylgruppe zurückzuführen ist. Im Bereich der Carbonylgruppe des Ethoxyresorufin befindet sich ein GRID-Feld der Carbonylsonde, dass durch Asn 257 aus SRS3 verursacht wird. Weitere Interaktionsfelder decken sich mit dem Sauerstoff und dem Stickstoff des Heterozyklus. Sie sind komplementär zu den Aminosäuren Tyr 112 und Thr 113 aus SRS1. In Abbildung 4.35 b und c sind zwei verschiedene Positionierungen des Coffeins dargestellt, die der Umsetzung zum Hauptmetaboliten Paraxanthin und zum Nebenmetaboliten Theobromin entsprechen.

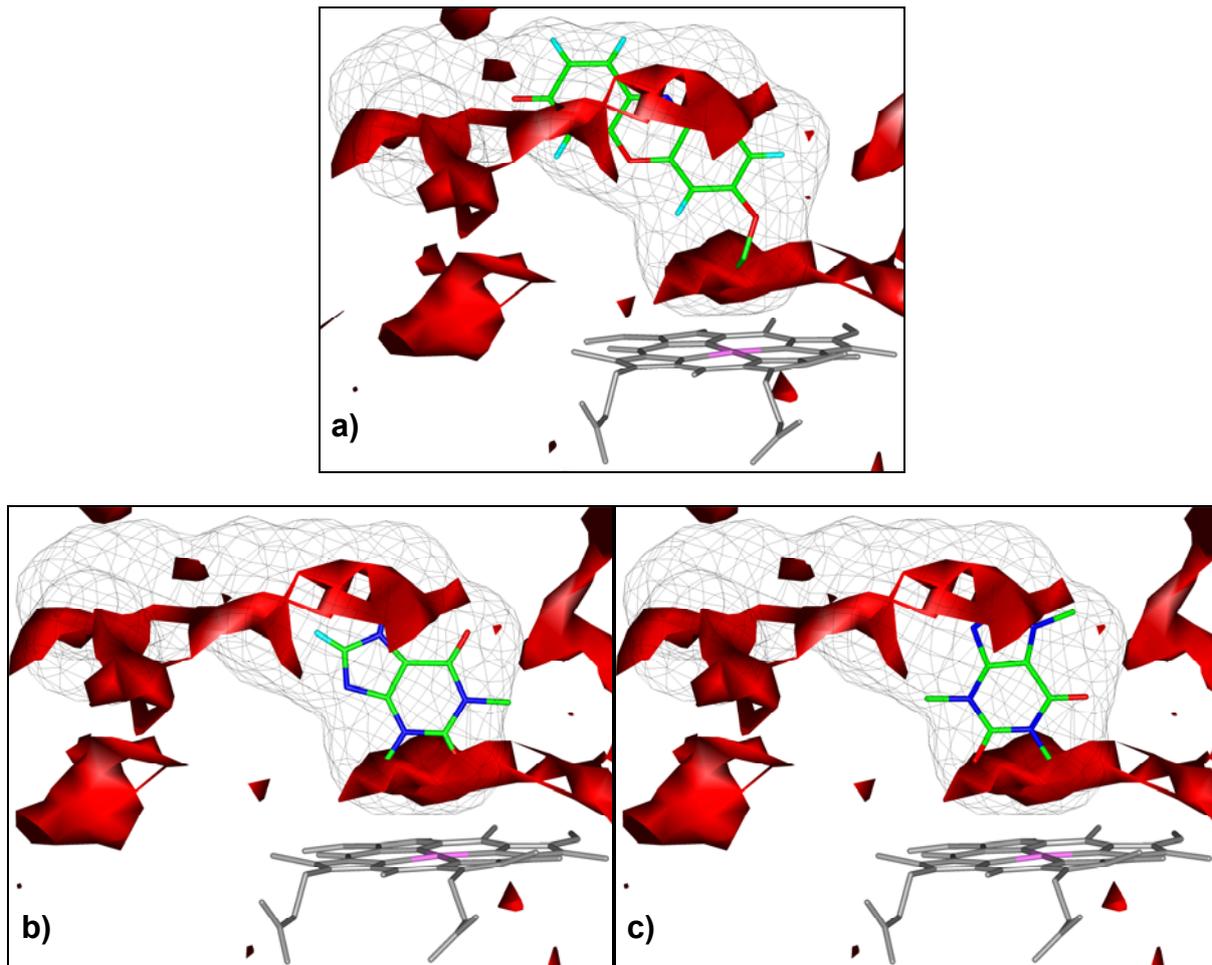


Abbildung 4.35: Positionierung von 7-Ethoxyresorufin und Coffein im aktiven Zentrum von CYP1A2. Das graue Gitternetz stellt die SURFNET-Kavität dar. In rot sind die GRID-Interaktionsfelder der Carbonylsonde mit der Konturierung bei einer Wechselwirkungsenergie von $-4,5 \text{ kcal}\cdot\text{mol}^{-1}$ abgebildet. Zur Orientierung ist das Häm in grau eingezeichnet. Alle Substrate sind atomkodierte eingefärbt. a) 7-Ethoxyresorufin, b) Coffein zur Bildung des Hauptmetaboliten Paraxanthin und c) Coffein zur Bildung von Theobromin.

Befindet sich das Coffein in der Position für die Demethylierung am Stickstoffatom N3 zum Paraxanthin (Abbildung 4.35 b), ist die entsprechende Methylgruppe 4,63 Å vom Hämeisen entfernt und die Carbonylgruppe am C6 liegt zwischen zwei GRID-Feldern der Carbonylsonde, die mögliche Interaktionen mit Ser 231 im Loop FG und Thr 498 in SRS6 anzeigen. Wird Coffein so platziert, dass die Umsetzung zum Theobromin möglich ist (Abbildung 4.35 c), befindet sich die Hydroxylierungsstelle im

Abstand von 4,80 Å zum Eisen. N9 liegt dann in einem GRID-Feld der Carbonylsonde, das durch Aminosäuren aus SRS1 zustande kommt (vgl. auch Abbildung 4.35 a). Während in Abbildung 4.35 a) und b) die gewählte Substratposition aufgrund der SURFNET-Kavität offensichtlich die einzig sinnvolle ist, wäre in c) prinzipiell auch die spiegelbildliche Anordnung denkbar. Sie konnte nach Durchführung einer MDS jedoch ausgeschlossen werden, da das Substrat nicht in einer für eine Umsetzung günstigen Orientierung im aktiven Zentrum blieb.

Das noch verbleibende Substrat MelQ konnte ebenfalls eindeutig aufgrund des SURFNET-Volumens ins aktive Zentrum von CYP1A2 positioniert werden (Abbildung 4.36). Der Abstand vom N1, der Hydroxylierungsstelle, zum Hämeisen beträgt 4,50 Å und der Chinolinstickstoff deckt sich mit dem bereits mehrfach erwähnten GRID-Feld im Bereich von SRS1.

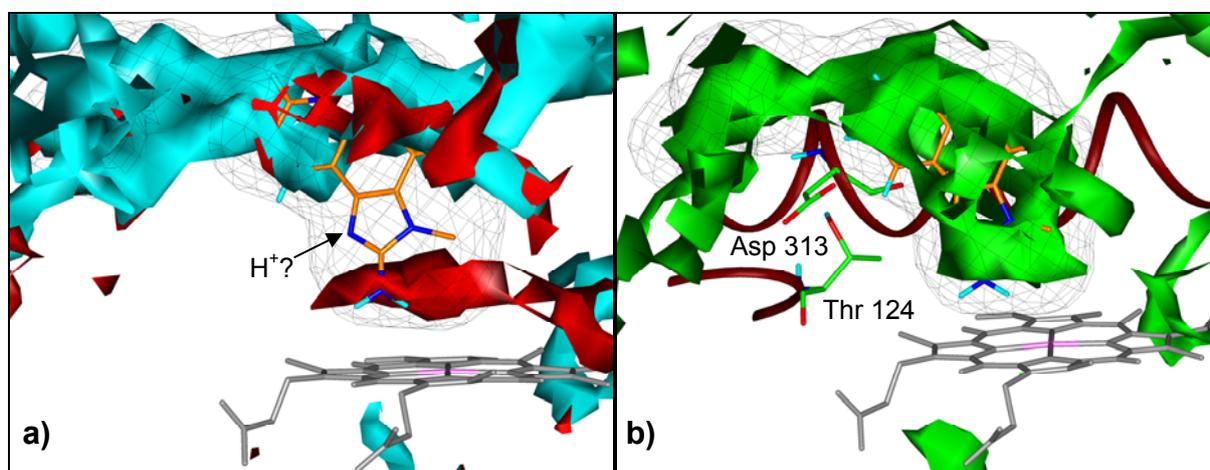


Abbildung 4.36: Positionierung von MelQ im aktiven Zentrum von CYP1A2. Das graue Gitternetz stellt die SURFNET-Kavität dar und zur Orientierung ist das Häm jeweils in grau eingezeichnet. Das Kohlenstoffgerüst von MelQ ist orange eingefärbt. a) GRID-Interaktionsfelder der Carbonylsonde (rot) und der kationischen NH-Sonde (türkis) jeweils mit der Konturierung bei einem Energiewert von $-4,5 \text{ kcal}\cdot\text{mol}^{-1}$. Der protonierbare Stickstoff des MelQ ist mit „H⁺“ gekennzeichnete. b) GRID-Interaktionsfelder der aromatischen Kohlenstoffsonde (grün) mit der Konturierung bei einem Energiewert von $-2 \text{ kcal}\cdot\text{mol}^{-1}$.

Da MelQ aufgrund der Guanidinogruppe eine relativ starke Base mit einem pK_S -Wert von 9,77 ist [163], stellte sich hier die Frage, ob die Verbindung im aktiven Zentrum des Enzyms protoniert oder neutral vorliegt. Dass diese Frage durchaus berechtigt ist, zeigt beispielsweise die Tatsache, dass Diclofenac im aktiven Zentrum von CYP2C5 als freie Säure gebunden wird, im aktiven Zentrum von CYP2C9 aber im deprotonierten Zustand [93]. Daher wurde das CYP1A2-Modell zusätzlich mit der kationischen NH-Sonde abgetastet, um potentielle Interaktionspartner für einen protonierten Stickstoff zu finden. Es zeigte sich, dass in dem Bereich, in dem der protonierbare Stickstoff des MelQ liegen würde, weder für einen Wasserstoffbrückenakzeptor (neutraler Stickstoff) noch für einen protonierten Stickstoff Wechselwirkungsfelder existieren (Abbildung 4.36 a). Stattdessen zeigt die aromatische

Kohlenstoffsonde für diese Region lipophile Interaktionen an. Die zu diesem Feld korrespondierende Aminosäure ist Thr 124 aus SRS1, dessen Methylgruppe in Richtung Substrat zeigt, weil die Hydroxylgruppe eine intramolekulare Wasserstoffbrücke zu Asp 313 aus Helix I bildet. Dass die Konformation von Thr 124 und die intramolekulare Wasserstoffbrücke kein Artefakt der Simulation sind, bestätigt eine Mutationsstudie, in der beobachtet wurde, dass der Austausch des Threonins gegen ein Serin zu einer um 70 % verminderten katalytischen Aktivität gegenüber 7-Ethoxyresorufin führt [50]. Folglich ist die Methylgruppe des Threonins in diesem Proteinbereich tatsächlich an der Ausbildung lipophiler Interaktionen zum Substrat beteiligt, weshalb dementsprechend angenommen wird, dass MeIQ neutral und nicht protoniert im aktiven Zentrum vorliegt. Die endgültige Bestätigung dieser Überlegungen erbrachte eine MDS mit dem protonierten MeIQ. Das protonierte NH drehte sich weg aus der lipophilen Umgebung und bildete eine Salzbrücke zu Asp 320 in Helix I, was zu einer groben Verzerrung dieser Helix führte. Außerdem vermutet man, dass dieses Aspartat beim Protonentransport ins aktive Zentrum eine Rolle spielt (vgl. Abschnitt 1.4.2.2), sodass es eher unwahrscheinlich ist, dass diese Aminosäure über eine Salzbrücke an der Substratbindung beteiligt ist.

Die Kristallstruktur von CYP2C9 mit Flurbiprofen im aktiven Zentrum einerseits und diverse spektroskopische Untersuchungen andererseits haben gezeigt, dass die Bindung eines Substrates nicht zwangsläufig mit der Verdrängung des Wassers als sechstem Hämliganden einhergeht, sondern dass es durchaus vorkommen kann, dass das Wasser gebunden bleibt, obwohl auch ein Substrat gebunden ist, wobei sich das Hämeisen dann weiterhin im *low-spin*-Zustand befindet (reverse Typ-I-Spektren, vgl. Abschnitt 1.3.1). Wie bereits erwähnt stellt CYP1A2 schon im substratfreien Zustand eine Ausnahme dar, weil das Hämeisen in der *high-spin*-Konformation vorliegt und kein Wasser an die sechste Koordinationsstelle gebunden ist (Abschnitt 1.3.1). Bei Ratten-CYP1A2, dessen Hämeisen im substratfreien Zustand ebenfalls zu zwei Drittel im *high-spin*-Zustand vorliegt, konnte man keine Veränderung der Absorptionsspektren nach Zugabe eines Substrates feststellen [164], was bedeuten würde, dass auch während oder nach der Substratbindung kein Wasser gebunden ist. In einer spektroskopischen Studie mit humanem CYP1A2 zum Metabolismus von Coffein konnte man erst nach Zugabe von unphysiologisch hohen Coffeinkonzentrationen (25 mmolar) eine unvollständige Spin-Verschiebung von 28 % zum *low-spin*-Eisen beobachten [54]. Bei dieser Coffeinkonzentration änderten sich allerdings auch die Entstehungsraten der einzelnen Metabolite, sodass dieses Ergebnis wohl nicht als allgemeingültig für physiologische Verhältnisse betrachtet werden kann. Wenn wie von Harris und Loew theoretisch berechnet die Bindung von Wasser an das Hämeisen von den elektronischen Eigenschaften der Proteinumgebung im Cytochrom abhängt (s. Abschnitt 1.3.1), ist es darüber hinaus unwahrscheinlich, dass nach der Bindung eines zumeist lipophilen Substrates zusätzlich ein polares Wassermolekül gebunden wird, nachdem bereits die elektronischen

Verhältnisse im leeren Enzym die Bindung von Wasser nicht zuließen. Somit wird für alle folgenden Berechnungen von Enzym-Substrat-Komplexen mit dem Modell von CYP1A2 davon ausgegangen, dass auch zum Zeitpunkt der Substratbindung kein Wassermolekül an die sechste Koordinationsstelle des Hämeisens gebunden ist.

4.3.2 Durchführung der MDS mit Enzym-Substrat-Komplexen

Analog dem Verfahren bei der Durchführung der MDS mit den leeren Enzymen wurden alle Enzym-Substrat-Komplexe zunächst in eine Wasserbox mit einer physiologischen Ionenkonzentration eingesetzt. Die Simulationssysteme wurden wie schon in Abschnitt 4.2.1 beschrieben einer *steepest descent* Minimierung und einer anschließenden Equilibrierungsdynamik über 200 ps unterzogen. Abweichend von der Equilibrierung des leeren Enzymmodells wurden hierbei allerdings das gesamte Proteinrückgrat sowie das Häm und das Substrat mithilfe von *position restraints* fixiert. Da hier mit einer bereits equilibrierten Proteinstruktur aus der MDS mit dem leeren Proteinmodell gearbeitet wurde, war es nicht mehr nötig, Loopbereiche während der Equilibrierungsdynamik zur Strukturanpassung komplett frei zu lassen. Das Substrat wurde fixiert, um eine Solvation durch Wassermoleküle im aktiven Zentrum und die Ausbildung von Kontakten zu Aminosäuren des aktiven Zentrums unter dem Erhalt der für das jeweilige Substrat vorgesehenen Position zu ermöglichen. Im Anschluss an diese Equilibrierungsdynamik wurde eine freie MDS unter den gleichen Simulationsbedingungen wie zuvor durchgeführt. Die Länge der MDS richtete sich in diesem Fall nicht nur nach der Equilibrierung der Proteinstruktur sondern auch nach der Equilibrierung der Substratposition und der Enzym-Substrat-Interaktionen. Sie betrug daher insgesamt 5 ns, wobei die letzten 2 ns zur Auswertung der Komplexstrukturen und der Enzym-Substrat-Interaktionen herangezogen wurden. Wie schon bei der Validierung des leeren Modells wurden auch hier zum Vergleich MDS mit den beiden als Kristallstruktur zur Verfügung stehenden Enzym-Substrat-Komplexen von CYP2C5 mit dem Sulfaphenazol-Derivat DMZ (1N6B) und mit Diclofenac (1NR6) unter denselben Simulationsbedingungen durchgeführt. Tabelle 4.14 listet die Zusammensetzung der Simulationssysteme für alle durchgeführten MDS auf.

Tabelle 4.14: Zusammensetzung der Simulationssysteme für die einzelnen MDS. 1N6B und 1NR6: Kristallstruktur von CYP2C5 mit DMZ bzw. Diclofenac im aktiven Zentrum; ERO, MIQ, CAF PAX und CAF THB: Proteinmodell von CYP1A2 mit den Substraten 7-Ethoxyresorufin, MeIQ, Coffein zur Bildung von Paraxanthin und Coffein zur Bildung von Theobromin.

	1N6B	1NR6	ERO	MIQ	CAF PAX	CAF THB
Kantenlänge der Box (nm)	8,125 x 5,820 x 7,478	8,307 x 6,257 x 7,619	7,861 x 6,185 x 7,689	7,862 x 6,185 x 7,689	7,923 x 6,134 x 7,470	7,861 x 6,185 x 7,689
Gesamtladung Protein (e)	-7	-7	+5	+5	+5	+5
Anzahl Wassermoleküle	9.294	10.509	9.908	9.925	9.513	9.910
Anzahl Na ⁺ -Ionen	37	41	33	33	32	33
Anzahl Cl ⁻ -Ionen	30	34	38	38	37	38

4.3.3 MDS mit der Kristallstruktur 1N6B

Im Fall der Kristallstruktur 1N6B, die für das vorliegende Homologiemodell als Templat diente und daher auch weiterhin zur Validierung des MDS-Verfahrens herangezogen werden sollte, konnte der Bindungsmodus des Substrates DMZ im aktiven Zentrum von CYP2C5 im Verlauf der MDS nicht erhalten werden.

Die Analyse der Kristallstruktur ergab, dass DMZ vor allem durch lipophile Interaktionen im aktiven Zentrum gebunden wird (Abbildung 4.37). Es sind weder direkte noch wasservermittelte Wasserstoffbrücken zwischen dem Substrat und Aminosäuren des aktiven Zentrums vorhanden. Wassermoleküle existieren lediglich in der bereits häufiger erwähnten charakteristischen Helixwindung der Helix I und zur Solvation von Asp 290 in Helix I. Keines von ihnen steht in Kontakt mit dem Substrat. Auch die beiden hydrophilen Funktionen des DMZ, der Pyrazolstickstoff und die Sulfonylgruppe, sind von lipophilen Aminosäuren umgeben. Auffällig ist, dass einer der beiden Wasserstoffe des Amidstickstoffs der Seitenkette von Asn 204 in SRS2 auf das Zentrum des endständigen Phenylrings des DMZ weist. Der Abstand zwischen dem Amidstickstoff von Asn 204 und dem Zentrum des Phenylrings von DMZ beträgt dabei etwa 3,4 Å. Denkbar wäre hier, dass zusätzlich zu den lipophilen Interaktionen eine schwache Interaktion des π -Elektronensystems des endständigen Phenylrings von DMZ mit der Amidfunktion von Asn 204 existiert. Derartige Pseudo-Wasserstoffbrücken zwischen dem Wasserstoff einer Amidfunktion und einem π -Elektronensystem werden in seltenen Fällen in Kristallstrukturen beobachtet, wie beispielsweise in der Kristallstruktur der bovinen Trypsininhibitors (PDB-Code 4PTI) [165].

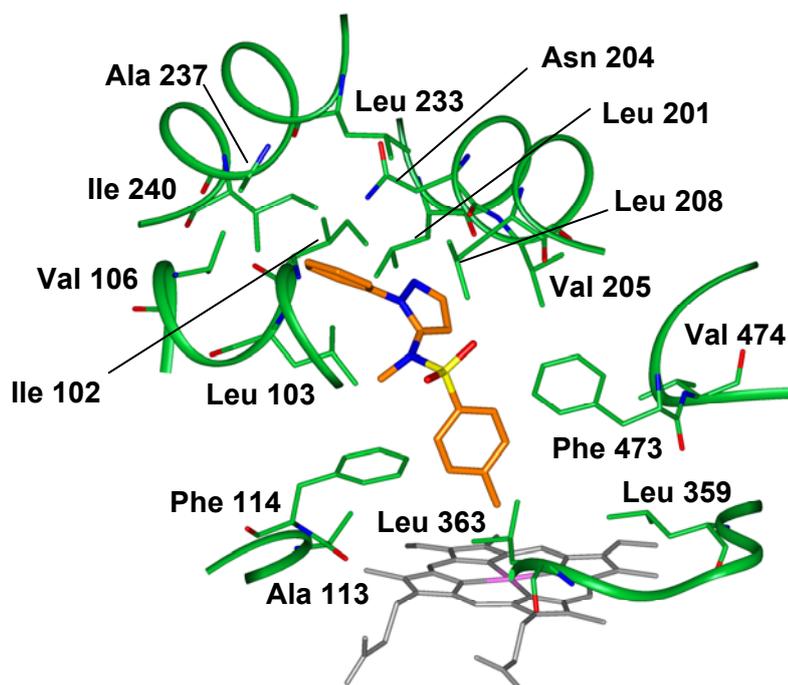


Abbildung 4.37: Protein-Substrat-Interaktionen in der Kristallstruktur 1N6B. Aminosäuren des Enzyms sind grün eingefärbt und das Substrat DMZ orange. Zur Orientierung ist das Häm in grau dargestellt.

Bereits zu Beginn der Equilibrierungsdynamik mit der Kristallstruktur 1N6B, wenn die Proteinseitenketten freigegeben werden, schwingt die Seitenkette von Asn 204 sofort weg vom Phenylring des Substrates, weil das Kraftfeld hier offensichtlich eine van-der-Waals-Abstoßung berechnet. Existierte hier tatsächlich eine Interaktion des π -Elektronensystems des endständigen Phenylrings von DMZ mit der Amidfunktion von Asn 204, so kann sie jedenfalls durch das Kraftfeld nicht wiedergegeben werden, da die empirischen Parametersätze von Kraftfeldern nicht in der Lage sind, die elektronischen Gegebenheiten einer solchen Interaktion korrekt zu berechnen. Dies ist sicherlich nur mit quantenchemischen Rechenverfahren möglich, mit denen Elektronendichteverteilungen explizit berechnet werden können. Im Verlauf der freien MDS rutscht der Phenylring nach unten aus der Seitentasche heraus und durch den Zugangskanal dringt Wasser in diesen Bereich des aktiven Zentrums. Es kommt zur Ausbildung einer wasservermittelten Wasserstoffbrücke zwischen der Amidfunktion von Asn 204 und dem Pyrazolstickstoff von DMZ, die bis zum Ende der MDS bestehen bleibt (Abbildung 4.38).

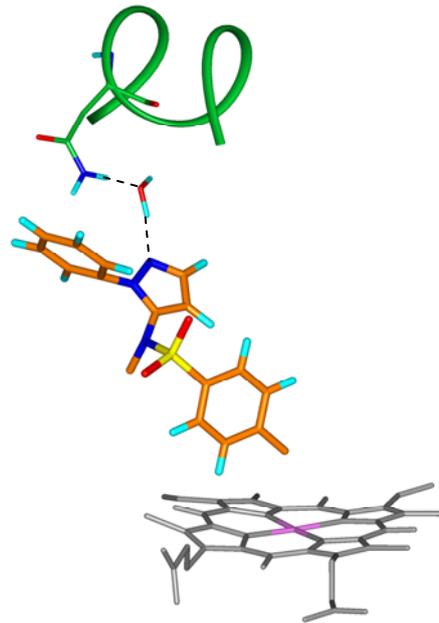


Abbildung 4.38: Wasservermittelte Interaktion zwischen Asn 204 und DMZ im Verlauf einer freien MDS, dargestellt anhand einer repräsentativen Struktur zum Simulationszeitpunkt 3735 ps. Der Proteinabschnitt aus Helix F ist grün eingefärbt und das Substrat DMZ orange. Zur Orientierung ist das Häm in grau dargestellt.

4.3.4 MDS mit der Kristallstruktur 1NR6

Zusätzlich wurde die zweite als Substratkomplex existierende Kristallstruktur von CYP2C5 mit Diclofenac im aktiven Zentrum zur Validierung des MDS-Verfahrens hinzugezogen. Neben dem Protein und dem Substrat wurden auch die mitkristallisierten Wassermoleküle in das Simulationssystem übernommen. Für diese Kristallstruktur konnte zunächst die Vermutung der Autoren, dass Diclofenac protoniert im aktiven Zentrum von CYP2C5 vorliegt, moleküldynamisch bestätigt werden. Während einer MDS mit der deprotonierten Form bildete sich eine Salzbrücke zu einem Lysin im oberen Teil des aktiven Zentrums, die dazu führte, dass sich die Hydroxylierungsstelle des Substrates in einem für eine Umsetzung unmöglichen Abstand von 8,5 Å zum Hämeisen befand und die Proteinstruktur verzerrt wurde. Die MDS mit der freien Säure verlief dagegen stabil und wird im Folgenden ausgewertet.

4.3.4.1 RMSD- und RMSF-Wert

Betrachtet man den Verlauf des RMSD-Wertes des gesamten Proteinrückgrates (schwarzer Graph in Abbildung 4.39), stellt man fest, dass es nach etwa 4000 ps nochmals zu einem Anstieg auf einen Wert von ca. 0,23 nm kommt, dass sich die Proteinstruktur hier also noch einmal ändert. Zieht man dagegen zur Berechnung des RMSD-Wertes nur die Bereiche heran, in denen sich die SRS1-6 befinden oder die an der Hämbindung beteiligt sind, so wird deutlich, dass der Kern des Proteins auch

über einen Simulationszeitraum von 5 ns stabil bleibt (blau gestrichelter Graph) und sich schon nach 1500 ps auf einen RMSD-Wert von 0,14 nm einpendelt.

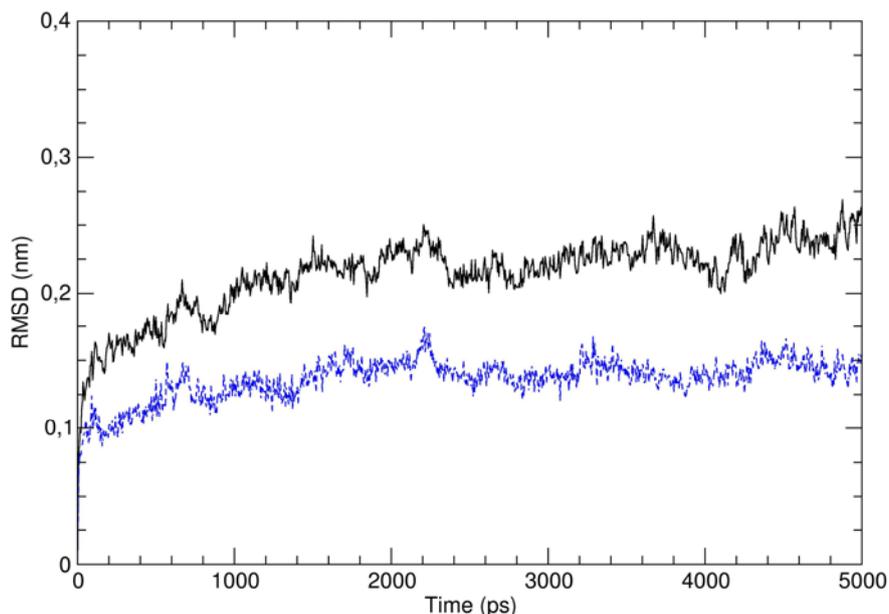


Abbildung 4.39: RMSD-Wert des Proteinrückgrates bezogen auf die Startstruktur im Verlauf der freien MDS mit der Kristallstruktur 1NR6. Aufgetragen ist der RMSD-Wert des ganzen Proteinrückgrates (schwarz) und des Proteinrückgrates des Proteinkerns, d.h. der SRS1-6 und der Hämbinderegionen (blau gestrichelt).

In der bereits in Abschnitt 4.2.2.2 zitierten MDS-Studie über den Verlauf von RMSD-Werten bei diversen Kristallstrukturen zeigte sich, dass die RMSD-Werte der verschiedenen Kristallstrukturen über den Simulationszeitraum unterschiedlich gut konvergierten. Es stellte sich aber in einer 50 ns langen Testsimulation auch heraus, dass nach 5 ns keine starken strukturellen Fluktuationen mehr stattfinden, und dass daher eine Simulationszeit von 5 ns in der Regel ausreichend ist. Ein Beispiel für eine MDS mit einer CYP-Kristallstruktur ist die bereits in Abschnitt 4.2.3.1 zitierte Simulation mit der Struktur 1BU7 von CYP102, bei der ebenfalls immer wieder leichte Fluktuationen des über das gesamte Proteinrückgrat berechneten RMSD-Wertes zu erkennen sind [160]. Beide oben genannten MDS wurden mit GROMACS unter mit dieser Arbeit vergleichbaren Simulationsbedingungen durchgeführt.

Hinweise darauf, welche Proteinbereiche den Anstieg des RMSD-Wertes in der hier betrachteten Kristallstruktur verursachen und warum, gibt die Auswertung der RMS-Fluktuation (Abbildung 4.40). Hier zeigt sich zunächst, dass das Gesamtprofil der RMS-Fluktuation dem der kürzeren MDS mit der leeren Kristallstruktur (blau gestrichelter Graph) sehr ähnlich ist. Bewegungen des Proteinrückgrats finden also wiederum vor allem in den Loopregionen und in der Mäanderregion statt (vgl. Abschnitt 4.2.3.1). Der größte Unterschied im Verlauf der beiden Kurven ergibt sich in der N-terminalen Region bis zur SRS1, die in der längeren MDS mit dem Substratkomplex von deutlich stärkeren Bewegungen im Proteinrückgrat geprägt ist.

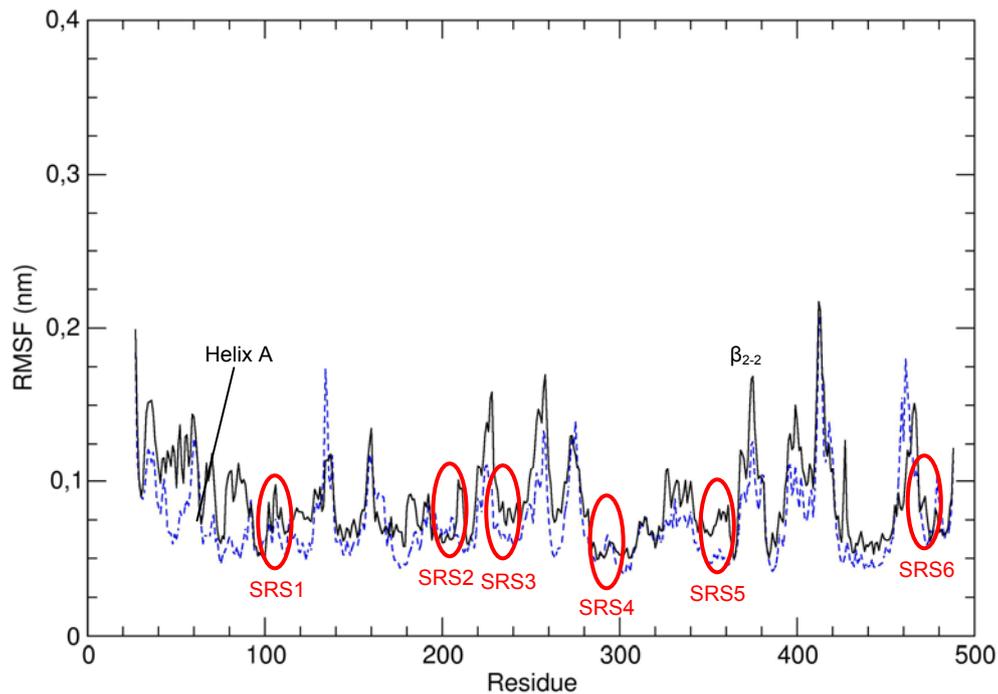


Abbildung 4.40: RMS-Fluktuation der einzelnen Aminosäuren zwischen 3000 und 5000 ps der freien MDS (schwarz). Zum Vergleich ist als blaue gestrichelte Linie der RMSF-Verlauf aus der MDS der leeren Kristallstruktur mit aufgetragen. Die roten Markierungen umfassen SRS1-6.

Die Überlagerung einer Proteinstruktur vom Anfang der equilibrierten Phase mit einer Struktur vom Ende der MDS lässt auch tatsächlich eine deutliche strukturelle Abweichung im Bereich des N-Terminus erkennen. Bereits in Abschnitt 4.2.3.1 und in Abschnitt 4.2.5.1 wurde auf die Instabilität des N-Terminus aufgrund des fehlenden Membranankers hingewiesen, die sich hier mit zunehmender Simulationsdauer noch stärker bemerkbar macht. Die Ausdehnung der Bewegungen bis zur SRS1 lässt sich aufgrund von Antikörperbindungsstudien erklären, die ergaben, dass vermutlich mehrere dem N-Terminus benachbarte hydrophobe Region an der Membranbindung beteiligt sind [81]. Das entspräche den Aminosäuren vor und nach Helix A und dem Faltblattstrang β_{2-2} .

4.3.4.2 Enzym-Substrat-Interaktionen

Bereits zu Beginn der freien MDS stellt sich ein Abstand von etwa 0,42 nm zwischen dem Hämeisen und dem Kohlenstoffatom C4', der Hydroxylierungsstelle des Diclofenac, ein, der über den gesamten Verlauf der MDS bestehen bleibt (Abbildung 4.41). Da dieser Abstand für eine Umsetzung etwa 0,4-0,5 nm betragen sollte, befindet sich das Substrat somit in einem für den Ablauf der Hydroxylierungsreaktion günstigen Abstand zum katalytischen Zentrum. In der Kristallstruktur, die als Ausgangsstruktur eingesetzt wurde, ist dieser Abstand mit 0,47 nm um 0,05 nm größer. Man sollte in diesem Zusammenhang das Zustandekommen der hier verwendeten Kraftfeldparameter, insbesondere der Partialladungen für das Häm berücksichtigen (vgl. Abschnitt 3.4.1.2). Bei der Parametrisierung testeten Helms &

Wade verschiedene Ladungsverteilungen in kurzen MDS, indem sie jeweils den Abstand zwischen dem Hämeisen und einem als sechster Ligand darüber liegenden Wassermolekül maßen und mit Kristallstrukturdaten verglichen. Für den besten Parametersatz ergab sich ein Abstand von 0,201 nm zwischen dem Wassersauerstoff und dem Hämeisen, der in der zugrunde gelegten Kristallstruktur mit 0,228 nm um etwa 0,03 nm größer war [127]. Das zeigt, dass schon die Parametrisierung eines so komplexen Systems wie des Häm mit einem gewissen Fehler behaftet ist. Dennoch ist trotz dieser Abweichung vom experimentellen Wert der Abstand der Hydroxylierungsstelle zum Hämeisen im Verlauf einer MDS grundsätzlich als Anhaltspunkt dafür geeignet, ob sich ein Substrat in einer für die Umsetzung geeigneten Position befindet, und er kann eventuell für den Vergleich verschiedener Substrate herangezogen werden, um eine Aussage über deren Enzymaffinität zu treffen.

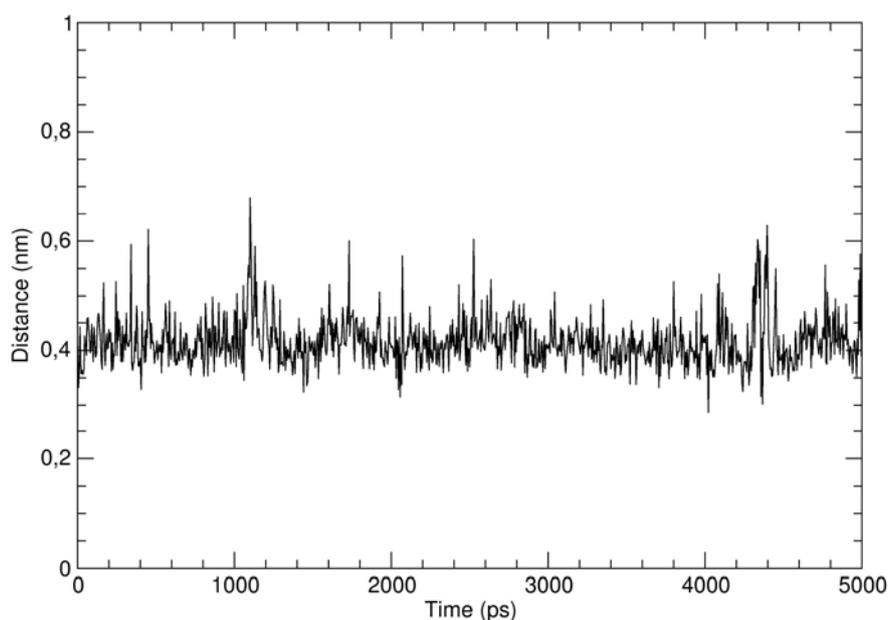


Abbildung 4.41: Abstand der Hydroxylierungsstelle des Diclofenac zum Hämeisen im Verlauf der MDS

In Abbildung 4.42 ist das Auftreten von Wasserstoffbrücken zwischen Diclofenac und Aminosäuren von CYP2C5 im Verlauf der letzten zwei Nanosekunden der freien MDS dargestellt. Das Wasserstoffbrückenmuster verläuft sehr gleichmäßig über den abgebildeten Zeitraum, was wiederum für die Equilibrierung des gesamten Enzym-Substrat-Komplexes spricht. Im Wesentlichen treten zwei verschiedene Wasserstoffbrücken durchgehend auf. Dies sind zum einen eine wasservermittelte Interaktion der Hydroxylgruppe von Ser 289 und zum anderen eine direkte Wasserstoffbrücke der Carboxylatgruppe von Asp 290 jeweils zur Säurefunktion des Diclofenac.

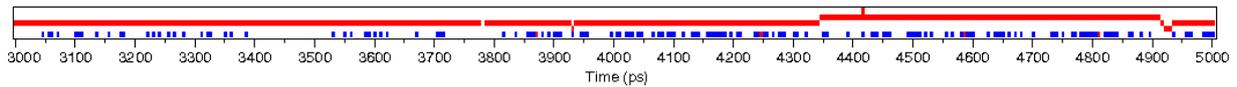


Abbildung 4.42: Matrix des Auftretens von Wasserstoffbrücken zwischen Diclofenac (DIF) und Aminosäuren des Proteins im Simulationszeitraum von 3000 bis 5000 ps. Direkte Wasserstoffbrücken sind in rot eingezeichnet und wasservermittelte in blau. In der Reihenfolge von unten nach oben sind aufgetragen die Wasserstoffbrücken Ser 289 OH – DIF COOH, Ser 289 O – DIF COOH, Asp 290 COO⁻ – DIF COOH (zweimal), Ala 294 N – DIF NH.

Zieht man nun wiederum die Kristallstruktur zum Vergleich heran, findet man dort keine direkten Wasserstoffbrücken zwischen Protein und Substrat, wohl aber tritt die wasservermittelte Interaktion zwischen Ser 289 und der Säurefunktion des Diclofenac auch dort auf. Sie ist Bestandteil eines Clusters bestehend aus sechs Wassermolekülen, das unter Beteiligung der Aminosäuren Asn 204, Lys 241, Ser 289 und Asp 290 die Säuregruppe des Substrates solvatisiert (vgl. Abschnitt 1.4.2.4).

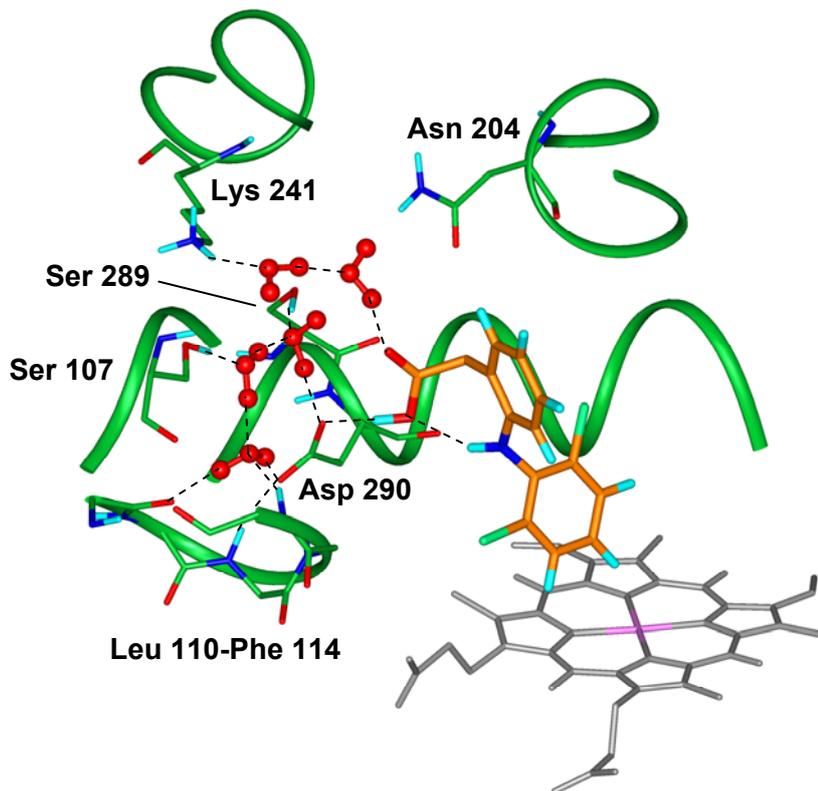


Abbildung 4.43: Repräsentative Struktur aus der MDS der Kristallstruktur 1NR6 nach 3385 ps. Aminosäuren des Proteins sind grün eingefärbt und das Substrat orange. Wassermoleküle sind als rote Kugelmodelle dargestellt. Von den Aminosäuren Leu 110 bis Phe 114 ist nur das Proteinrückgrat eingeblendet. Gestrichelte schwarze Linien symbolisieren Wasserstoffbrücken.

Unter diesem Gesichtspunkt wurden daraufhin die Wasserstoffbrücken zwischen Enzym, Lösungsmittelmolekülen und Diclofenac nochmals anhand von repräsentativen Komplexstrukturen aus der MDS visuell untersucht. Repräsentative Strukturen wurden wiederum mithilfe einer Clusteranalyse erhalten, wobei diesmal anstatt des gesamten Proteinrückgrates alle Atome der Aminosäuren der SRS1-6 für

die Clustereinteilung benutzt wurden, um einen Überblick über den Konformationsraum der an der Substratbindung beteiligten Aminosäuren zu gewinnen. Aus der repräsentativen Struktur des größten Clusters in Abbildung 4.43 ist ersichtlich, dass auch im Verlauf der MDS Kontakte zwischen dem Substrat und Aminosäuren des Proteins über ein Wassercluster zustande kommen. Allerdings besteht es hier nur aus fünf Wassermolekülen und nicht wie in der ursprünglichen Kristallstruktur aus sieben.

Die Überlagerung dieser repräsentativen Struktur mit der Kristallstruktur (Abbildung 4.44 a) lässt erkennen, dass sich das Substrat aufgrund der Ausbildung der direkten Wasserstoffbrücke zwischen der Säurefunktion und der Carboxylatgruppe von Asp 290 entlang der Helix I in Richtung Asp 290 parallel um etwa 2,2 Å verschiebt. Dadurch wird die hydrophile Kavität kleiner und es werden zwei der ursprünglich sieben Wassermoleküle daraus verdrängt. Außerdem ist wegen dieser Verschiebung auch Ser 107, das in der Kristallstruktur nur eine untergeordnete Rolle spielt, im Verlauf der MDS stärker an der Substratbindung beteiligt. Darüber hinaus sind die gleichen Aminosäuren wie auch in der Kristallstruktur in die Bindung des Diclofenac involviert, dies sind Asn 204, Lys 241, Ser 289 und Asp 290. Gegen Ende der MDS (Abbildung 4.44 b) schiebt sich die Hydroxylierungsstelle wieder zurück bis fast in die Ausgangsposition und der Phenylring neigt sich mehr zum Häm hin. Die Säuregruppe sackt dadurch noch tiefer über die Hämebene. Die Wechselwirkungen zwischen Protein und Substrat verändern sich dadurch aber nicht mehr.

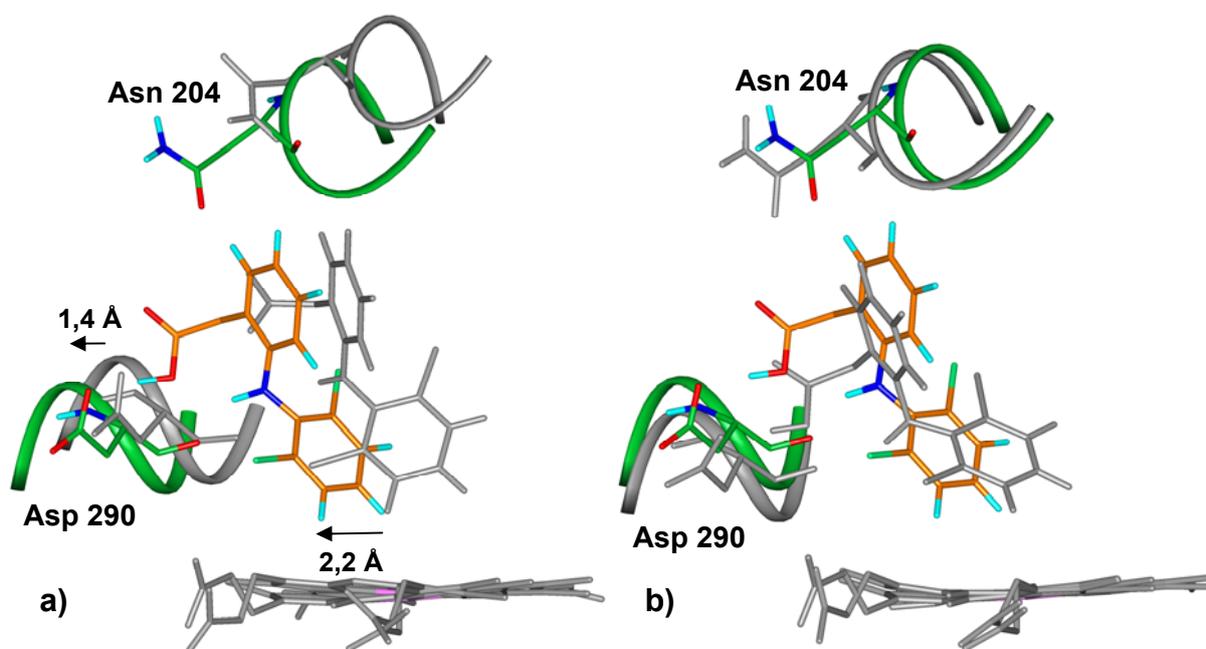


Abbildung 4.44: Überlagerung der repräsentativen Struktur nach 3385 ps (farbig, Farbcodierung wie in Abbildung 4.43) mit a) der Kristallstruktur und b) einer weiteren repräsentativen Struktur nach 4760 ps (jeweils in grau).

Die in Abbildung 4.44 a) eingeblendeten Aminosäuren zeigen, dass auch im Proteinerückgrat deutliche Verschiebungen stattfinden, und zwar im Bereich der Helices F, G, H und dem N-Terminus der Helix I. Während der N-Terminus der Helix I sich dabei nur um etwa 1,4 Å in dieselbe Richtung wie das Substrat verschiebt, kippt Helix G zusätzlich tiefer ins aktive Zentrum. Für aromatische und lipophile Interaktionen zwischen Protein und Substrat gilt prinzipiell das Gleiche wie für die Wasserstoffbrücken. Die Wechselwirkungen sind qualitativ identisch, aber auch hier hat eine Verschiebung der beteiligten Aminosäuren stattgefunden. Dies könnte eine Erklärung für die im Vergleich zur MDS mit der leeren Kristallstruktur leicht erhöhte RMS-Fluktuation im Bereich der SRS1-6 sein.

Insgesamt bleibt die Architektur des aktiven Zentrums jedoch erhalten, wie aus den Interaktionen von Asp 290 mit dem Proteinerückgrat der Aminosäuren Leu 110 bis Phe 114 (Abbildung 4.43) ersichtlich wird. Asp 290 ist in der Kristallstruktur über intramolekulare Wasserstoffbrücken mit dem β -Turn zwischen Leu 110 und Phe 114 verbunden. Dabei existiert eine direkte Wasserstoffbrücke zur Amidfunktion des Proteinerückgrates von Ile 112 und zwei wasservermittelte Wasserstoffbrücken zum Proteinerückgrat von Leu 110 und Phe 114. Diese Wechselwirkungen bleiben auch während der gesamten MDS erhalten.

Zusammenfassend kann man sagen, dass zwar deutliche Unterschiede zwischen der Kristallstruktur und den Komplexen aus dem Verlauf der MDS bestehen, dass die Substratbindung insgesamt aber qualitativ wiedergegeben werden konnte. Somit konnte die Stabilität und das dynamische Verhalten eines Enzym-Substrat-Komplexes über einen Zeitraum von 5 ns verfolgt werden und bietet eine hinreichende Validierung für die Berechnung von Substratkomplexen mit dem Homologiemodell von CYP2C5. Die Bedeutung von Wasserclustern für die Substratbindung wurde hier noch einmal besonders offensichtlich, und sollte auch im weiteren Verlauf der Arbeit berücksichtigt werden.

Für die artifizielle Ausbildung der direkten Wasserstoffbrücke zwischen Diclofenac und Asp 290 gibt es zwei Erklärungsmöglichkeiten. Zum einen ist denkbar, dass es im Kraftfeld bisweilen zu einer Überbewertung von elektrostatischen Interaktionen kommt. Zum anderen sollte überprüft werden, ob nicht auch Asp 290 wie das Substrat Diclofenac protoniert vorliegen könnte. Das würde die Attraktivität des Carboxylatsauerstoffs von Asp 290 für das saure Proton des Diclofenac sicherlich erheblich senken, und eine wasservermittelte Interaktion könnte erhalten bleiben. Zwar gibt es keinerlei experimentelle Anhaltspunkte für unterschiedliche Protonierungszustände im aktiven Zentrum von P450-Isoenzymen, jedoch wäre eine Protonierung des Aspartats in diesem Fall mit der Geometrie der Kristallstruktur vereinbar.

4.3.5 MDS des Modells mit MelQ

4.3.5.1 RMSD- und RMSF-Wert

Wie schon in der vorangegangenen MDS mit der Kristallstruktur steigt der RMSD-Wert des gesamten Proteinrückgrates bis zum Ende der Simulationszeit weiter leicht an, auch wenn die Steigung dabei abflacht (Abbildung 4.45).

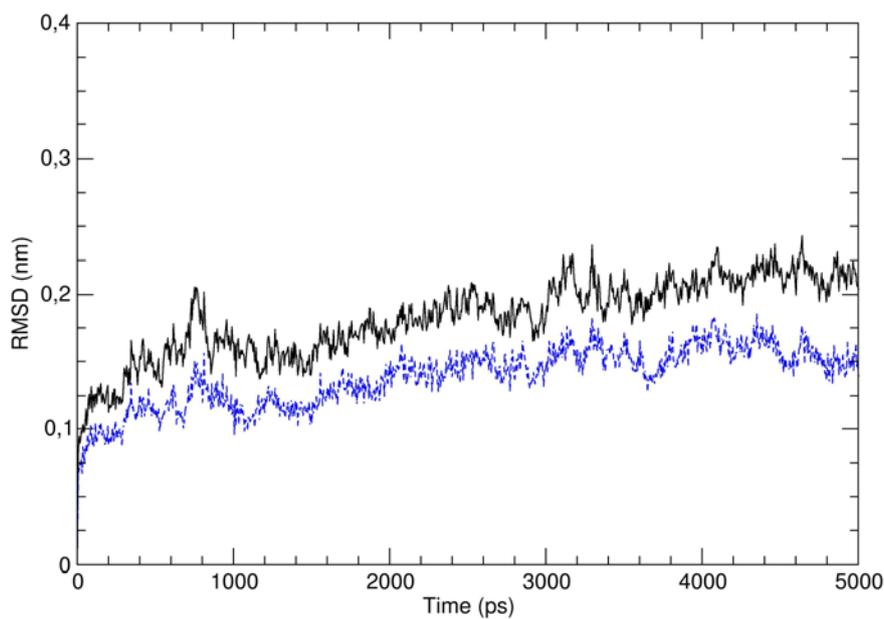


Abbildung 4.45: RMSD-Wert des Proteinrückgrates bezogen auf die Startstruktur im Verlauf der freien MDS des CYP1A2-Modells mit dem Substrat MelQ. Aufgetragen ist der RMSD-Wert des gesamten Proteinrückgrates (schwarz) und des Proteinrückgrates des Proteinkerns, d.h. der SRS1-6 und der Hämbinderegionen (blau gestrichelt).

Daher wird auch für sämtliche MDS von Enzym-Substrat-Komplexen des CYP1A2-Modells der RMSD-Wert des Proteinkerns als zusätzliches Stabilitäts- und Equilibrierungskriterium herangezogen. Als Proteinkern werden wiederum alle Bereiche betrachtet, die zu SRS1-6 gehören oder an der Hämbindung beteiligt sind. Dies sind im Einzelnen der Faltblattstrang β_{1-5} und die anschließende Helix B', der β -Turn vor Helix C und die Helix C selbst, die Helices F, G, I und K und der folgende Faltblattstrang β_{1-4} , die Cysteintasche und die Helix L, sowie das Faltblatt β_4 . Der RMSD-Wert dieser Bereiche liegt für den vorliegenden MelQ-Komplex mit 0,16 nm nur geringfügig höher als der entsprechende Wert aus der MDS mit der Kristallstruktur (0,14 nm), erreicht hier aber erst ab 3000 ps einen Gleichgewichtszustand. Die deutlich spätere Equilibrierung lässt sich dadurch begründen, dass zu Beginn der MDS des Modells mit verschiedenen Substraten zunächst eine Anpassung des Enzyms an das jeweilige Substrat stattfinden muss, die im Fall der Kristallstruktur wegfällt.

Um die Bewegungen innerhalb des CYP1A2-Modells während der equilibrierten Phase der MDS nachvollziehen zu können, wird wiederum die RMS-Fluktuation des

Proteinrückgrates analysiert (Abbildung 4.46). Wie bereits bei der Auswertung der MDS mit der Kristallstruktur ist auch hier zum Vergleich der RMSF-Verlauf aus der MDS mit dem leeren Modell mit aufgetragen. Wie schon im Fall der Kristallstruktur ähneln sich die beiden Kurvenverläufe stark. Bemerkenswert ist, dass der N-Terminus im Komplex von CYP1A2 mit MeIQ ein deutlich niedrigeres Maximum verursacht als im leeren Modell. Im Gegensatz zur Kristallstruktur scheint dies nicht die entscheidende Schwachstelle des Modells zu sein. Eine erhöhte Flexibilität lässt sich stattdessen am C-Terminus, in der Helix J und im Loop FG erkennen.

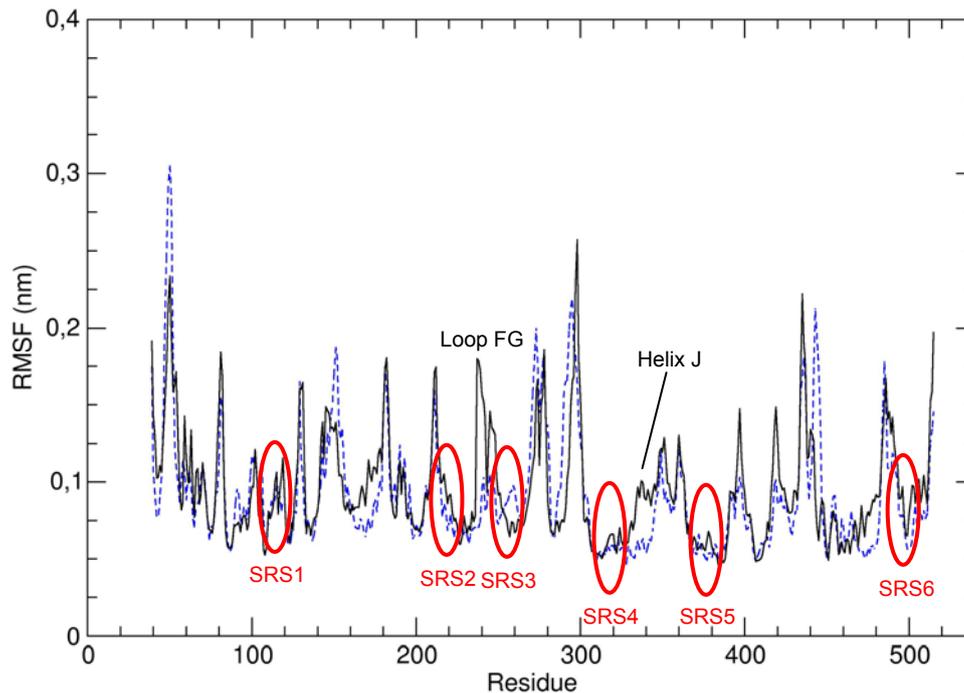


Abbildung 4.46: RMS-Fluktuation der einzelnen Aminosäuren zwischen 3000 und 5000 ps der freien MDS (schwarz). Zum Vergleich ist als blaue gestrichelte Linie der RMSF-Verlauf aus der MDS des leeren Proteinmodells mit aufgetragen. Die roten Markierungen umfassen SRS1-6.

Aus einer Überlagerung von Proteinstrukturen vom Beginn und vom Ende der equilibrierten Phase der MDS geht hervor, dass diese drei Bereiche wohl auch für den Anstieg des RMSD-Wertes verantwortlich sind. Der C-Terminus wurde, da er um zwei Aminosäuren länger ist als das Templat, beim Bau des Modells mit der *end repair* Routine angefügt (vgl. Abschnitt 4.1.5), was eine gewisse Instabilität erklärt. Helix J befindet sich in räumlicher Nachbarschaft zum C-Terminus und wird von dessen Beweglichkeit beeinflusst. Eine im Vergleich zur kürzeren MDS des leeren Enzyms höhere Beweglichkeit des FG-Loops ist auch bei der Kristallstruktur zu erkennen, ist jedoch nicht so ausgeprägt wie im Modell. Wie bereits in Abschnitt 4.2.3.1 erläutert, würde man hier den umgekehrten Fall erwarten, nämlich dass der Loop im leeren Zustand eher beweglich ist, und im substratgebundenen Zustand das aktive Zentrum fest verschließt. Diese Vermutung bestätigt sich zum wiederholten Mal nicht. Es sollte an dieser Stelle noch einmal darauf hingewiesen werden, dass in

diesem Proteinabschnitt in der Kristallstruktur Aminosäuren mutiert wurden, um die Kristallisation zu ermöglichen. Dies könnte ein unerwartetes Verhalten des FG-Loops erklären.

4.3.5.2 Wasserstoffbrückenmuster in Helix I

Im Gegensatz zur MDS mit dem leeren Proteinmodell (vgl. Abschnitt 4.2.5.3) befindet sich während der MDS mit den Substraten auch beim CYP1A2-Modell ein Wassermolekül in der charakteristischen Helixwindung von Helix I. Das resultierende Wasserstoffbrückenmuster ist hier am Beispiel der MDS mit MeIQ exemplarisch dargestellt (Abbildung 4.47).

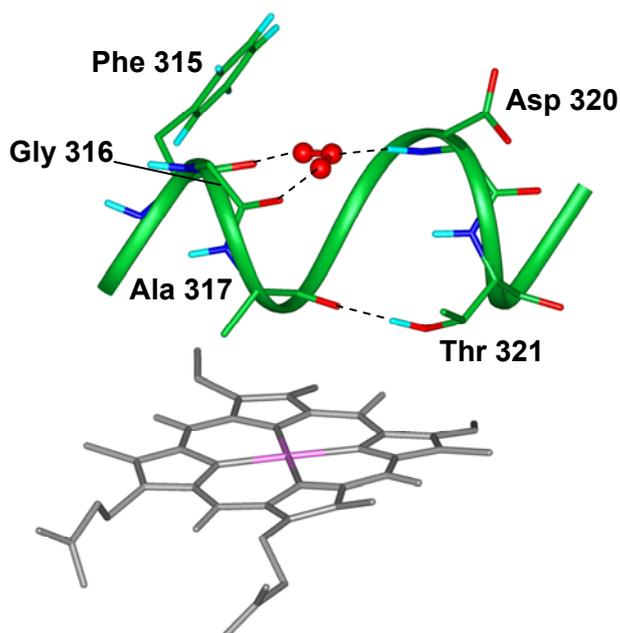


Abbildung 4.47: Wasserstoffbrückenmuster in Helix I am Beispiel einer repräsentativen Struktur aus der MDS des CYP1A2-Modells mit MeIQ.

Ähnlich wie in der für die MDS eines Substratkomplexes zugrunde gelegten Kristallstruktur 1NR6 (Abschnitt 4.3.4) besteht eine direkte Wasserstoffbrücke zwischen Ala 317 und Thr 321. Das Wassermolekül im oberen Teil der Helixwindung wird durch Wasserstoffbrücken zu den Proteinrückgratatomen der Aminosäuren Phe 315, Gly 316 und Asp 320 festgehalten. Daraus lässt sich ableiten, dass das Enzymmodell durch Substrate im aktiven Zentrum weiter in seiner aktiven Konformation stabilisiert wird.

4.3.5.3 Enzym-Substrat-Interaktionen

Im Gegensatz zur MDS von 1N6B mit Diclofenac, bei der der Abstand der Hydroxylierungsstelle zum Hämeisen von Beginn an konstant war, befindet sich die Hydroxylierungsstelle des MeIQ, der exozyklische Stickstoff, erst ab einer Simulationszeit von 2400 ps gleichmäßig im Abstand von etwa 0,42 nm zum Hämeisen (Abbildung 4.48).

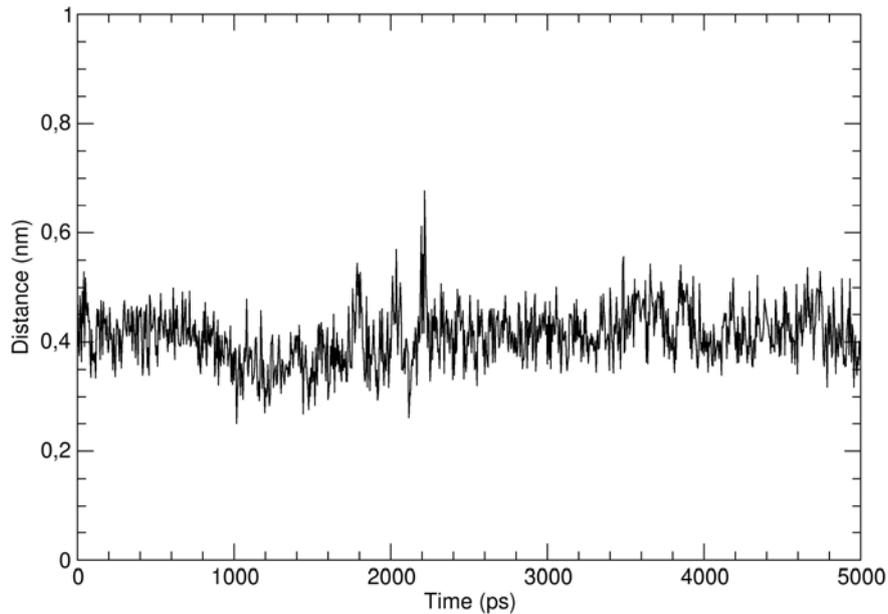


Abbildung 4.48: Abstand der Hydroxylierungsstelle des MeIQ zum Hämeisen im Verlauf der MDS

Zu Beginn der freien MDS variiert das Substrat seine Orientierung, bis sich Enzym und MeIQ aneinander angepasst haben. Der Abstand eines Substrates zum Hämeisen kann somit auch als Maß für die Equilibrierung der Substratorientierung gelten. Darüber hinaus geht aus dem Graphen hervor, dass MeIQ während des gesamten Verlaufs der MDS in einem für eine Umsetzung günstigen Abstand zum katalytischen Zentrum vorliegt.

Abbildung 4.49 zeigt die Matrix der zwischen MeIQ und Aminosäuren des Enzyms im Verlauf der Auswertungsphase aufgetretenen Wasserstoffbrücken. Daraus ist ersichtlich, dass auch die Protein-Substrat-Interaktionen in der Auswertungsphase von 300 bis 5000 ps einen equilibrierten Zustand erreicht haben. Es existiert eine stabile direkte Wasserstoffbrücke zwischen der Hydroxylgruppe von Tyr 112 und dem Chinolinstickstoff des MeIQ, die das Substrat fest im aktiven Zentrum bindet.

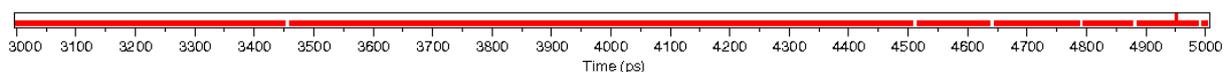


Abbildung 4.49: Matrix des Auftretens direkter Wasserstoffbrücken zwischen MeIQ und Aminosäuren des Proteins im Simulationszeitraum von 3000 bis 5000 ps. In der Reihenfolge von unten nach oben sind aufgetragen die Wasserstoffbrücken Tyr 112 OH – MeIQ Chinolinstickstoff, Thr 124 OH – MeIQ endozyklischer Guanidinstickstoff.

Anhand einer repräsentativen Struktur aus einer Clusteranalyse, die nach den schon in Abschnitt 4.3.4.2 erläuterten Kriterien durchgeführt wurde, ist diese Wasserstoffbrücke in Abbildung 4.50 dargestellt. Man erkennt hier, dass sie durch Ser 231 aus dem Loop FG zusätzlich stabilisiert wird.

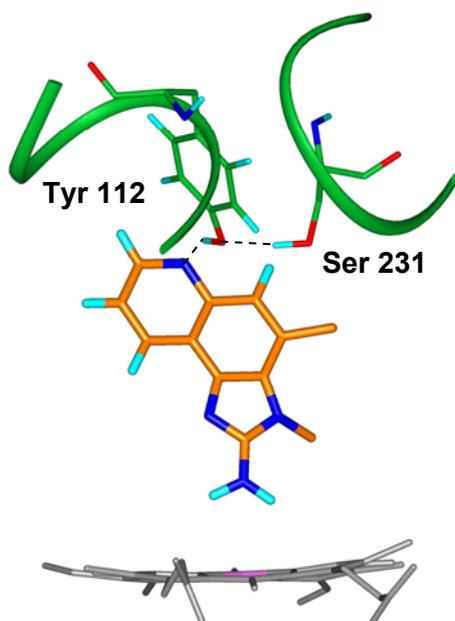


Abbildung 4.50: Wasserstoffbrücke zwischen Tyr 112 und dem Chinolinstickstoff von MeIQ, dargestellt anhand der repräsentativen Struktur des größten Clusters aus der MDS nach 4875 ps. Die Farbcodierung entspricht der in Abbildung 4.43.

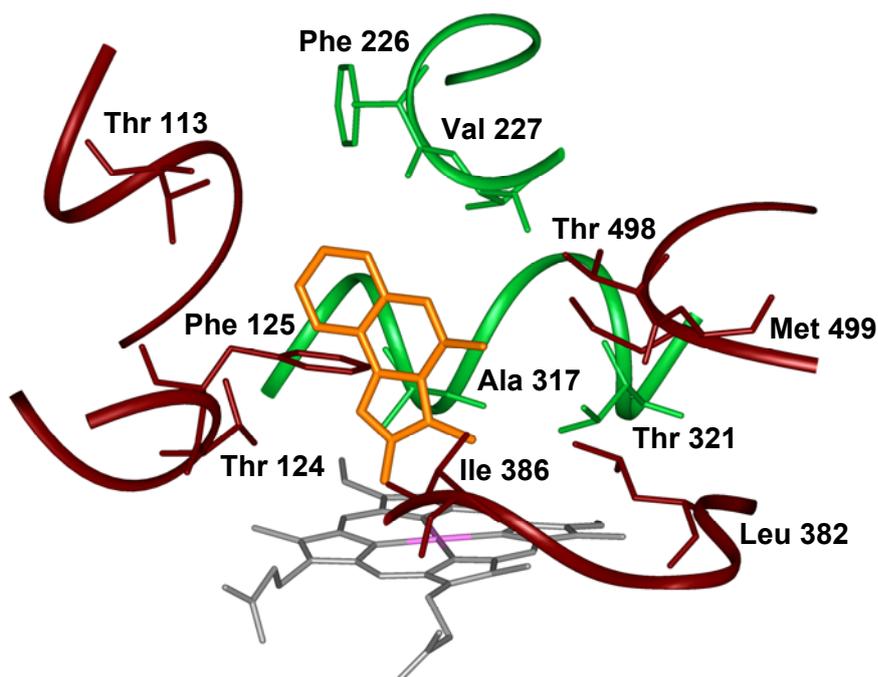


Abbildung 4.51: Lipophile und aromatische Protein-Substrat-Interaktionen von MeIQ im aktiven Zentrum des CYP1A2-Modells, dargestellt anhand der repräsentativen Struktur des größten Clusters aus der MDS nach 4875 ps. Aus Gründen der Übersichtlichkeit sind Aminosäuren der SRS1, 5 und 6 braun eingefärbt und Aminosäuren der SRS2 und 4 grün. Alle Wasserstoffatome wurden ausgeblendet.

Dieselbe repräsentative Struktur des größten Clusters nach 4875 ps Simulationszeit ist in Abbildung 4.51 noch einmal zur Darstellung aromatischer und lipophiler Wechselwirkungen des Substrates mit Aminosäuren von CYP1A2 abgebildet. Aromatische π - π -Interaktionen ergeben sich zwischen dem Ringsystem des MeIQ

und Phe 125 aus SRS1 sowie Phe 226 aus SRS2. Thr 113 und Thr 124 aus SRS1, Thr 321 aus SRS4 und Thr 498 aus SRS6 sind mit ihrer Methylgruppe zum Substrat hin orientiert und spielen so ebenfalls für lipophile Interaktionen eine Rolle. Hinzu kommen außerdem Val 227 aus SRS2, Ala 317 aus SRS4, Leu 382 und Ile 386 aus SRS5 und Met 499 aus SRS6.

4.3.5.4 Vergleich der aufgetretenen Interaktionen mit Mutationsstudien und mit der Kristallstruktur

Wie bereits in Abschnitt 1.2.1.1 kurz erwähnt, findet man in der Literatur mehrere Veröffentlichungen über Mutationsstudien am humanen CYP1A2, bei denen MeIQ als Substrat verwendet wurde. Zwei davon befassen sich mit Mutationen innerhalb der SRS1-5 und werden im Folgenden zum Vergleich mit den im CYP1A2-Modell gefundenen Protein-Substrat-Interaktionen herangezogen (Tabelle 4.15). In beiden Studien wurde die Genotoxizität der aus der Enzymreaktion resultierenden Metabolite als Maß für die Enzymaktivität gegenüber MeIQ in einem quantitativen Genotoxizitätstest (*umu*-Test) ausgewertet.

Tabelle 4.15: Übersicht zu Mutationsstudien am humanen CYP1A2 mit MeIQ als Substrat. Die Mutationen sind nach den SRS, die sie betreffen, sortiert. In der vierten Spalte ist der Effekt bezogen auf den Wildtyp als Aktivitätszunahme (+), Aktivitätsabnahme (-) oder unveränderter Aktivität (0) angegeben. Besonders aussagekräftige Mutationen sind fett hervorgehoben.

SRS	Mutation	CYP2C5	Effekt	Referenz
1	Ser 126 Lys	Ser 115	+	[34]
1	Ser 126 Lys	Ser 115	-	[35]
1	Ser 129 Asp, Glu	Asn 116	+	[34]
1	Ser 129 Glu	Asn 116	-	[35]
2	Glu 225 Ile	Glu 203	++	[34]
2	Glu 225 Asn, Ser	Glu 203	--	[34]
2	Glu 225 Ile, Ser	Glu 203	-	[35]
2	Phe 226 Ile, Thr, Tyr	Asn 204	---	[34], [35]
3	Lys 254 Gly	Leu 234	+	[34]
3	Lys 254 Gly	Leu 234	0	[35]
4	Asp 320 Ala	Glu 297	---	[34]
4	Asp 320 Ser	Glu 297	---	[35]
4	Thr 321 Pro	Thr 298	---	[34]
4	Thr 321 Ser	Thr 298	--	[34]
4	Thr 321 Pro, Ser	Thr 298	-	[35]
4	Val 322 Ala, Ser, Thr	Thr 299	--	[34], [35]
5	Phe 384 Leu	Thr 361	--	[34]
5	Phe 384 Leu	Thr 361	0	[35]
5	Thr 385 Leu	Asn 362	+	[34]
5	Thr 385 Leu	Asn 362	-	[35]
5	Ile 386 Leu, Pro	Leu 363	-	[34]
5	Ile 386 Thr	Leu 363	---	[34]
5	Ile 386 Leu	Leu 363	-	[35]
5	Ile 386 Pro, Thr	Leu 363	--	[35]

Für die beiden in SRS1 durchgeführten Mutationen (Ser126 und 129) konnten nur sehr unspezifische und zwischen den beiden Veröffentlichungen widersprüchliche

Effekte beobachtet werden, woraus geschlossen werden kann, dass sie keine wichtige Rolle für die Substratbindung spielen. Im in der vorliegenden Arbeit erstellten CYP1A2-Modell wird Ser 126 durch Phe 125 abgeschirmt, das eine π - π -Interaktion mit dem MelQ eingeht (vgl. Abbildung 4.51) und Ser 129 befindet sich unmittelbar vor Beginn der Helix C auf Höhe der Hämebene. Damit sind beide Aminosäuren für eine direkte Substrat-Interaktion zu weit vom aktiven Zentrum entfernt und sind auch im Modell nicht für die Substratbindung von Bedeutung.

Ebenfalls von untergeordneter Bedeutung ist die Mutation von Glu 225 in SRS2 zu Ile, Asn oder Ser. Die Aminosäure weist im Modell in die Peripherie und nicht zum aktiven Zentrum. Dagegen ist das benachbarte Phe 226 hochsensitiv gegenüber einem Austausch gegen Ile, Thr oder auch Tyr. Darin stimmen beide Referenzen überein. Aus dem Aktivitätsverlust bei einer Mutation zu Ile und Thr kann man vermuten, dass hier eine π - π -Interaktion für die Substratbindung wichtig ist, wie sie auch im vorliegenden CYP1A2-Modell gefunden wird (Abbildung 4.51). Dort steht Phe 226 senkrecht zum Ringsystem des MelQ, was erklären würde, warum auch die Tyr-Mutante, die nur π - π -Interaktionen eingehen kann, wenn sie parallel zum Ringsystem des Substrates liegt, eine deutlich geringere katalytische Aktivität aufweist. Im Templat CYP2C5 befindet sich an dieser Stelle ein Asparagin, es scheint sich also an dieser Position um eine für die Substratspezifität entscheidende Aminosäure zu handeln.

Die einzige in SRS3 durchgeführte Mutation zeigt wiederum einen unspezifischen Effekt. Im Modell weist Lys 254 in die Peripherie des Enzyms.

Der ausgeprägte Effekt eines Austausches von Asp 320 in SRS4 gegen Ala oder Ser ist wohl weniger auf eine Beteiligung an der Substratbindung zurückzuführen als vielmehr auf die wichtige Funktion einer sauren Aminosäure an dieser Position beim Ablauf des katalytischen Mechanismus (vgl. Abschnitt 0). Diese Annahme wird dadurch unterstützt, dass in beiden Studien die Expressierung dieser Mutanten schwierig war. Demgegenüber ist Thr 321 zwar einerseits auch an der Katalyse beteiligt, indem es durch die irreguläre Wasserstoffbrücke der Hydroxylgruppe zur Carbonylfunktion von Ala 317 den charakteristischen Knick in Helix I verursacht, es scheint aber auch eine gewisse Bedeutung für die Substratbindung zu haben, wie die verminderte Aktivität der Ser-Mutante zeigt, die ja auch eine Hydroxylgruppe enthält. Anhand des CYP1A2-Modells lässt sich nachvollziehen, dass die Methylgruppe des Threonins für den richtigen Abstand des Substrates zur Helix I sorgt. Nicht nachvollziehbar ist der deutliche Aktivitätsverlust einer Mutation von Val 322. Da diese Aminosäure in der Helix I auf Thr 321 folgt, zeigt sie in der aufsteigenden Helixwindung bereits weg vom aktiven Zentrum. Auf der Grundlage des Modells ist hier allenfalls ein indirekter Effekt auf die Substratbindung denkbar, der sich nicht ohne weiteres aus der Modellstruktur ableiten lässt.

Phe 348 aus SRS5 ist im Modell in der Schleife zwischen Leu 382 und Ile 386 lokalisiert und ist damit für eine direkte Substratinteraktion zu weit vom aktiven Zentrum entfernt. Das stimmt mit dem uneinheitlichen Ergebnis der Mutationsstudien überein. Erstaunlich ist der geringe Effekt der Entfernung der Hydroxylgruppe von Thr 385 durch Mutation zu Leu, da die Seitenkette dieser Aminosäure im Modell wie auch das im Templat an dieser Position befindliche Asn 362 an der Hämbindung beteiligt ist. Man könnte vermuten, dass das Häm im wesentlichen durch die starken Salzbrücken zu Arg 108, Arg 137 und Arg 456 stabilisiert wird, sodass die zusätzlich vorhandenen Wasserstoffbrücken nur eine unterstützende Funktion besitzen. Demgegenüber ist eindeutig, dass an der Position von Ile 386 eine flexible lipophile Aminosäure für die Substratbindung entscheidend ist, wie die deutliche Aktivitätsabnahme der Thr- und der Pro-Mutante im Vergleich zum Wildtyp zeigen. Im Templat CYP2C5 befindet sich an dieser Stelle ein Leu, woraus man schließen kann, dass Ile 386 zwar für die Substratbindung verantwortlich ist aber weniger für die Substratspezifität. Im vorliegenden Modell wird auch diese Interaktion wieder gefunden, Ile 386 ist hier an der Positionierung der Hydroxylierungsstelle unmittelbar über der Hämebene beteiligt (Abbildung 4.51).

Anhand von Mutationsstudien nicht überprüft werden kann die Existenz der Wasserstoffbrücke zwischen Tyr 112 und dem Chinolinstickstoff von MeIQ. Es kann hier lediglich die Kristallstruktur zum Vergleich herangezogen werden. Dort befindet sich an der Position des Tyr das Pro 101. Substrat-Interaktionen werden in diesem Bereich der SRS1 zu Val 110 und zu Leu 103 gefunden, also zu zwei benachbarten lipophilen Aminosäuren. Man weiß allerdings aus multiplen Sequenzalignments, dass gerade die Proteinregion, in der sich Helix B' befindet, unter den P450-Isoenzymen von einer relativ geringen Sequenzhomologie geprägt ist [83], so dass es möglich ist, dass hier die Substratspezifität der verschiedenen Isoenzyme determiniert wird. Eine klare Übereinstimmung findet sich demgegenüber in der π - π -Interaktion von Phe 125 mit dem Substrat. Dieser Rest ist konserviert (Phe 114 in der Kristallstruktur) und zeigt in den beiden kristallisierten Substratkomplexen eine sehr ähnliche Wechselwirkung mit Diclofenac bzw. DMZ.

4.3.6 MDS des Modells mit 7-Ethoxyresorufin

4.3.6.1 RMSD- und RMSF-Wert

Wie in Abbildung 4.52 dargestellt ist im Fall der MDS des CYP1A2-Modells mit 7-Ethoxyresorufin im aktiven Zentrum auch der RMSD-Wert des gesamten Proteinerückgrates über den Simulationszeitraum der Auswertungsphase von 3000 bis 5000 ps equilibriert und lässt keine Steigung mehr erkennen. Er schwankt um einen Wert von 0,22 nm. Trotzdem wird zum Vergleich mit den vorangegangenen MDS auch hier der RMSD-Verlauf des Proteinkerns berechnet. Er ist ebenfalls ab 3000 ps

equilibriert und pendelt sich auf einen Mittelwert von 0,14 nm ein wie auch der entsprechende RMSD-Wert der Kristallstruktur.

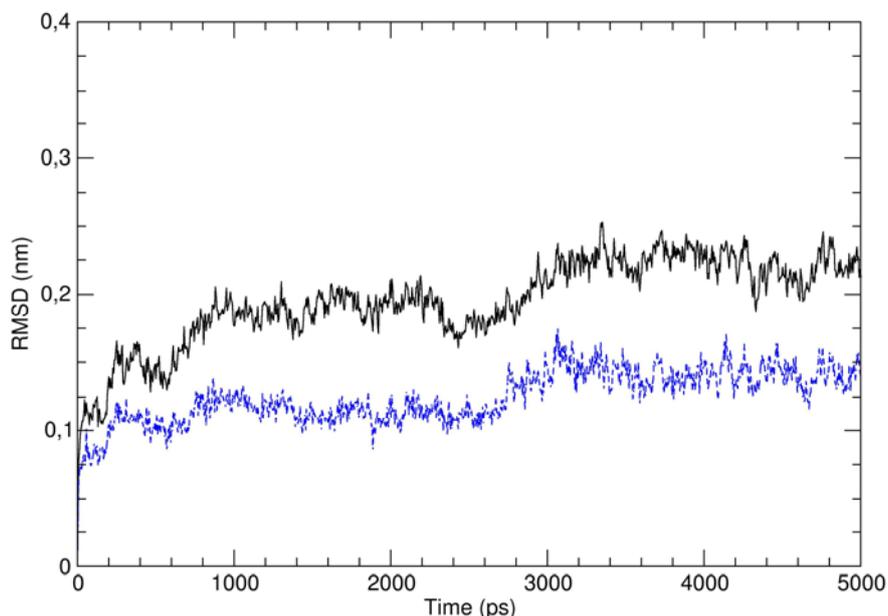


Abbildung 4.52: RMSD-Wert des Proteinrückgrates bezogen auf die Startstruktur im Verlauf der freien MDS des CYP1A2-Modells mit dem Substrat 7-Ethoxyresorufin. Aufgetragen ist der RMSD-Wert des ganzen Proteinrückgrates (schwarz) und des Proteinrückgrates des Proteinkerns, d.h. der SRS1-6 und der Hämbinderegionen (blau gestrichelt).

Der Verlauf der RMS-Fluktuation (Abbildung 4.53) weist im Prinzip das gleiche Profil auf wie in der vorangegangenen MDS des CYP1A2-Modells mit MeIQ.

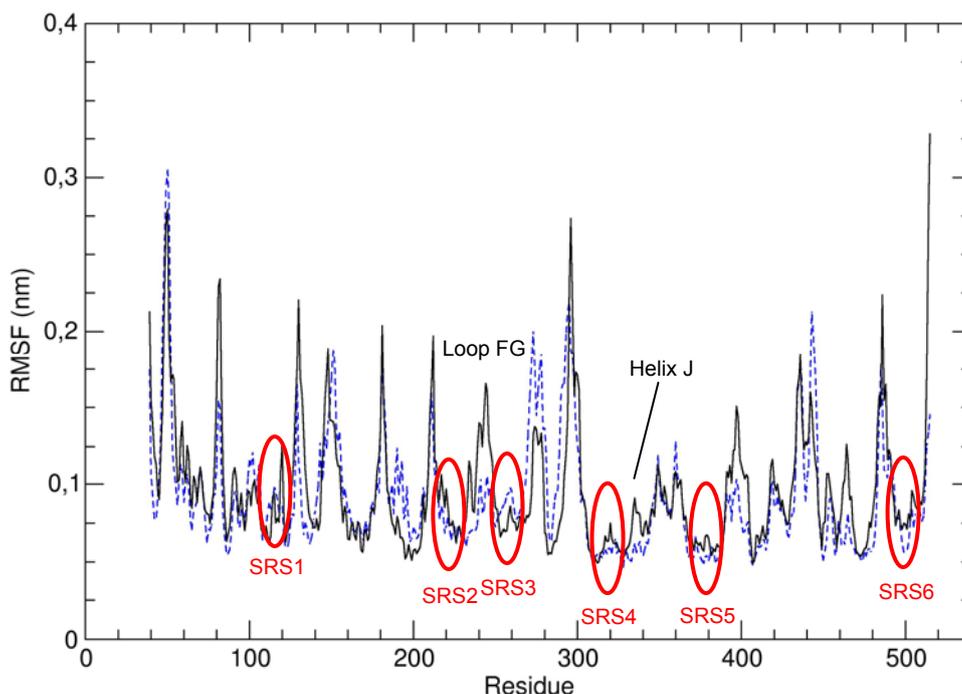


Abbildung 4.53 RMS-Fluktuation der einzelnen Aminosäuren zwischen 3000 und 5000 ps der freien MDS (schwarz). Zum Vergleich ist als blaue gestrichelte Linie der RMSF-Verlauf aus der MDS des leeren Proteinmodells mit aufgetragen. Die roten Markierungen umfassen SRS1-6.

Wiederum sind es der C-Terminus und der Loop FG, die eine deutlich höhere Fluktuation aufweisen als im leeren Modell. Es gelten daher die bereits in Abschnitt 4.3.5.1 angestellten Überlegungen.

4.3.6.2 Protein-Substrat-Interaktionen

Der Abstand zwischen dem Hämeisen und der Hydroxylierungsstelle des Ethoxyresorufins, dem α -Kohlenstoff des Ethoxysubstituenten, ist wie aus Abbildung 4.54 hervorgeht mit durchschnittlich 0,47 nm deutlich größer als der Abstand der Hydroxylierungsstelle zum Eisen in der MDS mit MeIQ (0,42 nm). Außerdem schwankt der Abstand hier stärker als im Fall des MeIQ. Der Grund dafür könnte der Ethoxysubstituent sein, der zum einen voluminöser und zum anderen konformatorisch flexibler ist als die exozyklische NH_2 -Gruppe des MeIQ. Trotz dieser Schwankungen befindet sich das Ethoxyresorufin jedoch im gesamten Verlauf der MDS in einem für eine Umsetzung günstigen Abstand zum Häm.

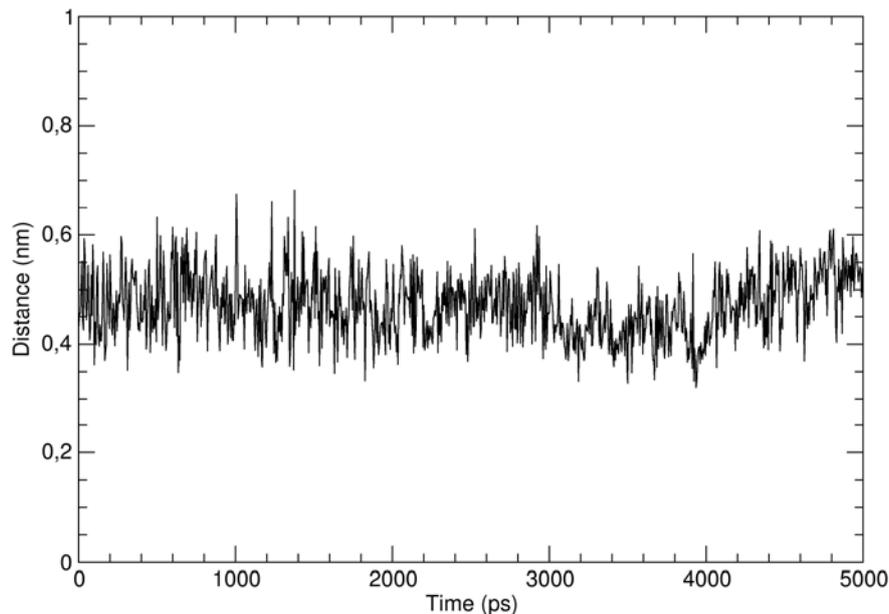


Abbildung 4.54: Abstand der Hydroxylierungsstelle des 7-Ethoxyresorufin zum Hämeisen im Verlauf der MDS

Die Wasserstoffbrückenmatrix in Abbildung 4.55 zeigt, dass im Verlauf der Auswertungsphase der MDS drei wasservermittelte und zwei direkte Wasserstoffbrücken zwischen dem Substrat und Aminosäuren des Proteins auftreten.

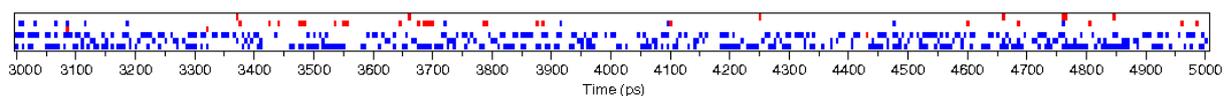


Abbildung 4.55: Matrix des Auftretens von Wasserstoffbrücken zwischen 7-Ethoxyresorufin (ERO) und Aminosäuren des Proteins im Simulationszeitraum von 3000 bis 5000 ps. Direkte Wasserstoffbrücken sind in rot eingezeichnet und wasservermittelte in blau. In der Reihenfolge von unten nach oben sind aufgetragen die Wasserstoffbrücken Tyr 112 OH – ERO N, Thr 113 OH – ERO CO, Thr 118 OH – ERO CO, Asn 257 NH_2 – ERO CO, Asn 312 NH_2 – ERO CO (zweimal).

Ähnlich der Interaktion zwischen Tyr 112 und dem Chinolinstickstoff des MeIQ existiert auch beim Ethoxyresorufin regelmäßig eine Wasserstoffbrücke dieses Tyrosins zum heterozyklischen Stickstoff des Ringsystems. Sie ist hier allerdings nicht direkt, sondern wasservermittelt (s. auch Abbildung 4.57). Die wasservermittelten Interaktionen von Thr 113 und Thr 118 aus SRS1 mit der Carbonylfunktion des Ethoxyresorufins treten alternierend auf. Wie aus Abbildung 4.56 ersichtlich ist, kommt die Wasserstoffbrücke über zwei verschiedene Wassermoleküle zustande, von denen sich jeweils eins in einer Orientierung befindet, in der ein Kontakt zum Substrat entstehen kann. Die beiden Wassermoleküle werden zusätzlich durch Wasserstoffbrücken zu Asp 313 aus SRS4 stabilisiert, das wiederum über intramolekulare Wasserstoffbrücken mit Thr 124 und Ser 122 verbunden ist (vgl. Positionierung von MeIQ).

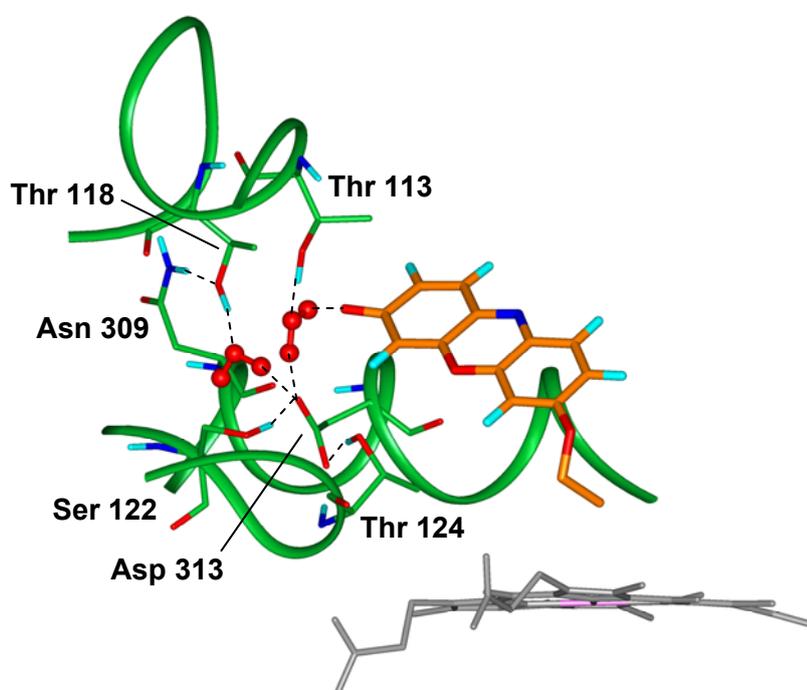


Abbildung 4.56: Wasserstoffbrückenmuster zwischen Thr 113, Thr 118 und der Carbonylfunktion von 7-Ethoxyresorufin, dargestellt anhand einer Struktur aus dem größten Cluster der MDS nach 3385 ps. Die Farbcodierung entspricht der in Abbildung 4.43.

Relativ selten aber doch über den gesamten Simulationszeitraum verteilt und zum Ende der MDS wieder verstärkt auftretend, kann zusätzlich eine direkte Wasserstoffbrücke der Carbonylfunktion von Ethoxyresorufin zur Amidfunktion von Asn 312 aus SRS4 beobachtet werden. Eine direkte Wasserstoffbrücke zu Asn 257 aus SRS3 kommt nur zweimal zu Beginn der equilibrierten Phase der MDS zustande, häufiger ist diese Aminosäure an der Fixierung von Asn 312 beteiligt (Abbildung 4.57).

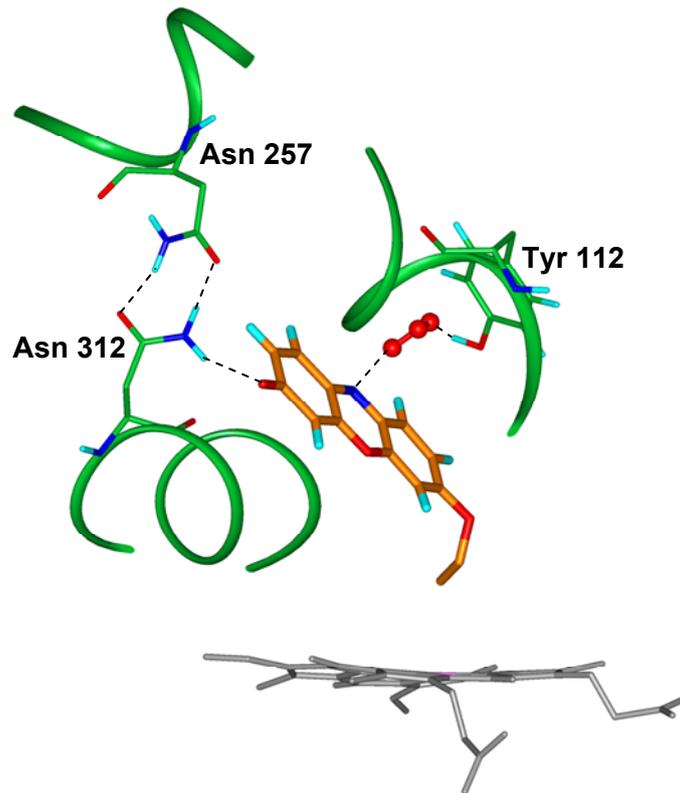


Abbildung 4.57: Wasserstoffbrücken zwischen Asn 312 und der Carbonylfunktion, und zwischen Tyr 112 und dem heterozyklischen Stickstoff von 7-Ethoxyresorufin, dargestellt anhand einer Struktur aus dem größten Cluster der MDS nach 3475 ps. Die Farbcodierung entspricht der in Abbildung 4.43.

Die lipophilen und aromatischen Interaktionen des Ethoxyresorufins mit den Aminosäuren des aktiven Zentrums (Abbildung 4.58) sind vielfach die gleichen wie beim MelQ, da Größe und Form der beiden Substrate zwar nicht gleich, aber doch ähnlich sind. Beide Verbindungen sind relativ groß und lang gestreckt und nehmen eine ähnliche Orientierung im aktiven Zentrum ein.

Zusätzlich zu den bereits in Abschnitt 4.3.5.3 genannten Wechselwirkungen existiert beim Ethoxyresorufin eine π - π -Interaktion zu Phe 260 aus SRS3 und eine lipophile Interaktion zur Methylgruppe von Thr 223 in SRS2, da das Ringsystem des Ethoxyresorufins aus drei Sechsringen nebeneinander besteht und daher tiefer in die Kavität zwischen den Helices F und G hineinreicht als das MelQ. Met 499 weist hier weg vom aktiven Zentrum und ist daher nicht eingeblendet.

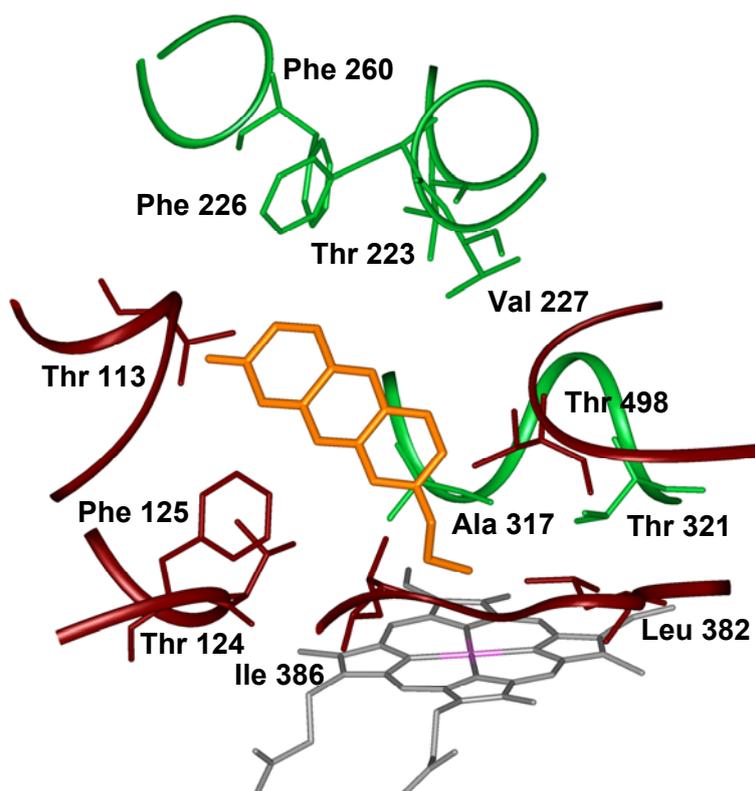


Abbildung 4.58: Lipophile und aromatische Protein-Substrat-Interaktionen von 7-Ethoxyresorufin im aktiven Zentrum des CYP1A2-Modells, dargestellt anhand der repräsentativen Struktur des größten Clusters aus der MDS nach 3430 ps. Die Farbcodierung entspricht der in Abbildung 4.51.

4.3.6.3 Vergleich der aufgetretenen Interaktionen mit Mutationsstudien und mit der Kristallstruktur

Wie auch beim MeIQ gibt es zwei Veröffentlichungen, die sich mit Mutationen im Bereich der SRS1-5 befassen und dabei 7-Ethoxyresorufin als Substrat verwenden. Die Enzymaktivität wurde in beiden Studien durch eine fluorimetrische Messung der Konzentration des aus der Enzymreaktion resultierenden Metaboliten Resorufin mithilfe des bereits in Abschnitt 1.2.1.3 kurz erwähnten EROD-Assays bestimmt. Die in Tabelle 4.16 aufgeführten Mutationen von Parikh et al. wurden im Rahmen derselben Studie durchgeführt, in der auch MeIQ untersucht wurde. Es wurden auch aufgrund der Ähnlichkeit der Substrate häufig analoge Effekte beobachtet, auf die hier nicht mehr ausführlich eingegangen wird (vgl. hierzu Abschnitt 4.3.5.4). Als besonders wichtig für die Substratbindung stellte sich wiederum die Aminosäure Phe 226 heraus, die wie aus Abbildung 4.58 hervorgeht, auch hier eine π - π -Interaktion mit dem Ringsystem des Substrates eingeht. Ein deutlicher Unterschied ergab sich bei der Mutante Ile 386 Thr, die in Bezug auf MeIQ zu einem Aktivitätsverlust führte, beim Ethoxyresorufin aber eine deutliche Aktivitätszunahme hervorruft. Eine mögliche Erklärung dafür ist, dass die Hydroxylgruppe des Threonins sich auf derselben Höhe wie der Ethoxysubstituent des Ethoxyresorufins befinden würde, und dass es dadurch zur Ausbildung einer Wasserstoffbrücke kommen

könnte. Eine derartige zusätzliche Fixierung der Hydroxylierungsstelle über dem aktiven Zentrum kann mit Sicherheit eine gesteigerte Umsetzung bewirken.

Tabelle 4.16: Übersicht zu Mutationsstudien am humanen CYP1A2 mit 7-Ethoxyresorufin als Substrat. Die Mutationen sind nach den SRS, die sie betreffen, sortiert. In der vierten Spalte ist der Effekt bezogen auf den Wildtyp als Aktivitätszunahme (+), Aktivitätsabnahme (-) oder unveränderter Aktivität (0) angegeben. Aussagekräftige Mutationen sind fett hervorgehoben.

SRS	Mutation	CYP2C5	Effekt	Referenz
1	Thr 124 Ser	Ala 113	- -	[50]
1	Ser 126 Lys	Ser 115	-	[34]
1	Ser 129 Glu, Arg	Asn 116	-	[34]
2	Thr 223 Asn	Leu 201	- -	[50]
2	Glu 225 Asn, Ser	Glu 203	+ +	[34]
2	Phe 226 Ile, Thr, Tyr	Asn 204	- - -	[34]
2	Val 227 Gly	Val 205	- -	[50]
4	Asn 312 Leu	Ser 289	- - -	[50]
4	Thr 321 Ser	Thr 298	-	[34]
4	Val 322 Ala	Thr 299	-	[34]
4	Val 322 Ser	Thr 299	+ + +	[34]
5	Leu 382 Ala	Leu 359	- - -	[50]
5	Leu 382 Val	Leu 359	0	[50]
5	Ile 386 Leu	Leu 363	+ +	[34]
5	Ile 386 Thr	Leu 363	+ +	[34]

Weitere aussagekräftige Mutationen stammen aus der Untersuchung von Liu et al. Bereits im Zusammenhang mit der Positionierung des MeIQ ins aktive Zentrum wurde die Mutation Thr 124 Ser in SRS1 erwähnt (Abschnitt 4.3.1). Es zeigt sich nun, dass in den MDS beider Substratkomplexe des CYP1A2-Modells die Orientierung der Methylgruppe in Richtung Substrat erhalten geblieben ist (Abbildung 4.51 und Abbildung 4.58), und dass somit auch das Modell die Bedeutung von Thr 124 für die Substratbindung wiedergeben kann.

Ähnlich wie Thr 124 weist auch Thr 223 aus SRS2 im Modell mit seiner Methylgruppe ins aktive Zentrum und ist somit an der Bindung des Ethoxyresorufins über eine lipophile Interaktion beteiligt (Abbildung 4.58). Dies steht in Übereinstimmung mit der Aktivitätsabnahme beim Austausch gegen eine hydrophile Aminosäure wie Asparagin. Ebenfalls in SRS2, eine Helixwindung weiter und unmittelbar über dem aktiven Zentrum neben Phe 226 befindet sich im CYP1A2-Modell Val 227. Über eine lipophile Interaktion mit dem Ringsystem des MeIQ und des Ethoxyresorufin schließt es in beiden MDS das aktive Zentrum nach oben ab. Befindet sich an dieser Position nur ein Glycin, könnte das eine Vergrößerung des Abstands vom Substrat zum Hämeisen zur Folge haben, was eine Umsetzung erschwert.

Einen gravierenden Effekt auf die CYP1A2-Aktivität hat ein Austausch von Asn 312 in SRS4 gegen ein Leucin, wodurch die EROD-Aktivität auf 1,1 % der Wildtyp-Aktivität sinkt. Dieser starke Aktivitätsverlust und die Tatsache, dass hier eine

Aminosäure, die in der Lage ist, Wasserstoffbrücken sowohl zu donieren als auch zu akzeptieren, gegen eine lipophile ersetzt wurde, machen die Existenz einer Wasserstoffbrücke zwischen Ethoxyresorufin und Asn 312 wahrscheinlich. Das steht wiederum im Einklang mit der MDS des CYP1A2-Modells mit Ethoxyresorufin, bei der eine Wasserstoffbrücke zwischen der Carbonylfunktion des Substrates und der Amidfunktion des Asn 312 zwar nicht durchgehend, aber mehrfach über den gesamten Verlauf der MDS beobachtet werden konnte (Abbildung 4.55 und Abbildung 4.57).

In SRS5 konnte im Rahmen der Mutationsstudie zusätzlich zu Ile 386 ein Einfluss von Leu 382 auf die Substratbindung festgestellt werden. Ein Austausch gegen ein Alanin führte hier zu einer Aktivitätsabnahme. Im Proteinmodell ist Leu 382 in beiden Substratkomplexen wie Ile 386 auch an der Fixierung der Hydroxylierungsstelle über dem Eisen beteiligt (Abbildung 4.51 und Abbildung 4.58). Eine Mutation des Leucins zum kürzeren Alanin würde demnach keine so exakte Positionierung der Hydroxylierungsstelle über dem katalytischen Zentrum mehr erlauben.

4.3.7 MDS des Modells mit Coffein zur Umsetzung zum Paraxanthin

4.3.7.1 RMSD- und RMSF-Wert

Der RMSD-Wert des gesamten Proteinrückgrates (Abbildung 4.59) erreicht zwar nach 3000 ps ein Plateau, steigt jedoch ab 4000 ps nochmals an bis auf etwa 0,24 nm.

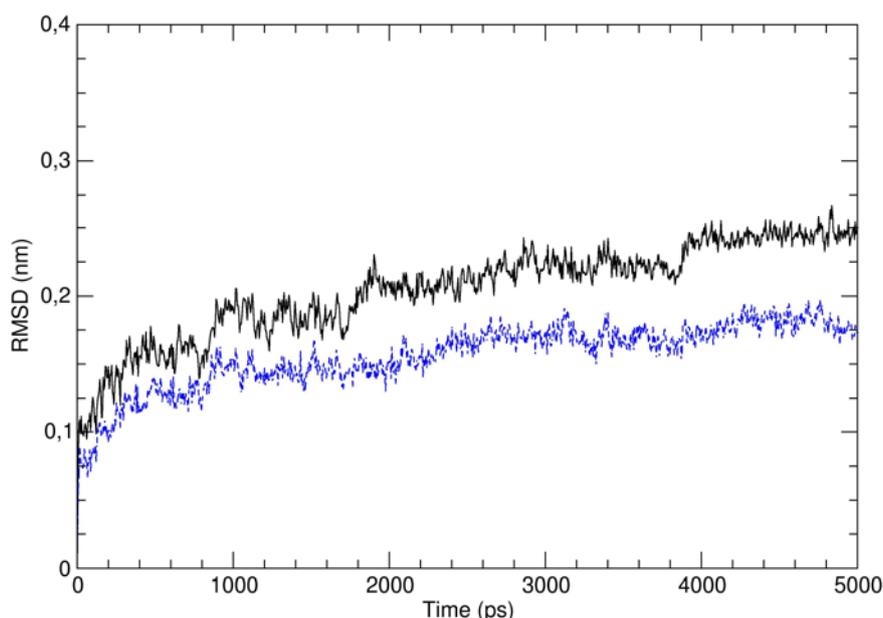


Abbildung 4.59: RMSD-Wert des Proteinrückgrates bezogen auf die Startstruktur im Verlauf der freien MDS des CYP1A2-Modells mit Coffein zur Umsetzung zum Paraxanthin. Aufgetragen ist der RMSD-Wert des ganzen Proteinrückgrates (schwarz) und des Proteinrückgrates des Proteinkerns, d.h. der SRS1-6 und der Hämbinderegionen (blau gestrichelt).

Der RMSD-Wert des Proteinkerns dagegen ist in der Auswertungsphase equilibriert und beträgt durchschnittlich 0,17 nm. Auch hier erkennt man zwar ab 4000 ps eine leichte Steigung, die Kurve fällt aber ab 4600 ps wieder ab, sodass man davon ausgehen kann, dass der RMSD-Wert um einen Gleichgewichtswert pendelt. Im Vergleich mit den RMSD-Werten im Verlauf der MDS der Substratkomplexe mit MelQ (0,16 nm) und 7-Ethoxyresorufin (0,14 nm) verläuft der RMSD-Wert des Coffein-Komplexes etwas höher.

Die Ursache für den sprunghaften Anstieg des RMSD-Wertes des gesamten Proteinrückgrates bei etwa 4000 ps lässt sich anhand der RMS-Fluktuation während der Auswertungsphase gut lokalisieren. Aus Abbildung 4.60 geht hervor, dass die Mäanderregion ein ungewöhnliches hohes Maximum im Profil der RMS-Fluktuation verursacht. Dieses Maximum erreicht einen Wert von fast 0,5 nm. Eine Überlagerung von Strukturen verschiedener Simulationszeitpunkte der MDS zeigt, dass der Mäander-Loop bei etwa 3850 ps in eine andere Konformation umklappt. Bereits in Abschnitt 4.2.3.1 wurde auf die mehr oder weniger strukturlose Beschaffenheit des Mäander-Loops eingegangen. Da sich dieser Bereich aber unterhalb der Hämebene, also nicht auf derselben Seite wie das aktive Zentrum befindet, wird die Konformation des aktiven Zentrums dadurch nicht beeinträchtigt. Dies kann man wiederum daran erkennen, dass das Profil der RMS-Fluktuation für das restliche Protein dem Profil der beiden anderen Substratkomplexe mit dem CYP1A2-Modell stark gleicht.

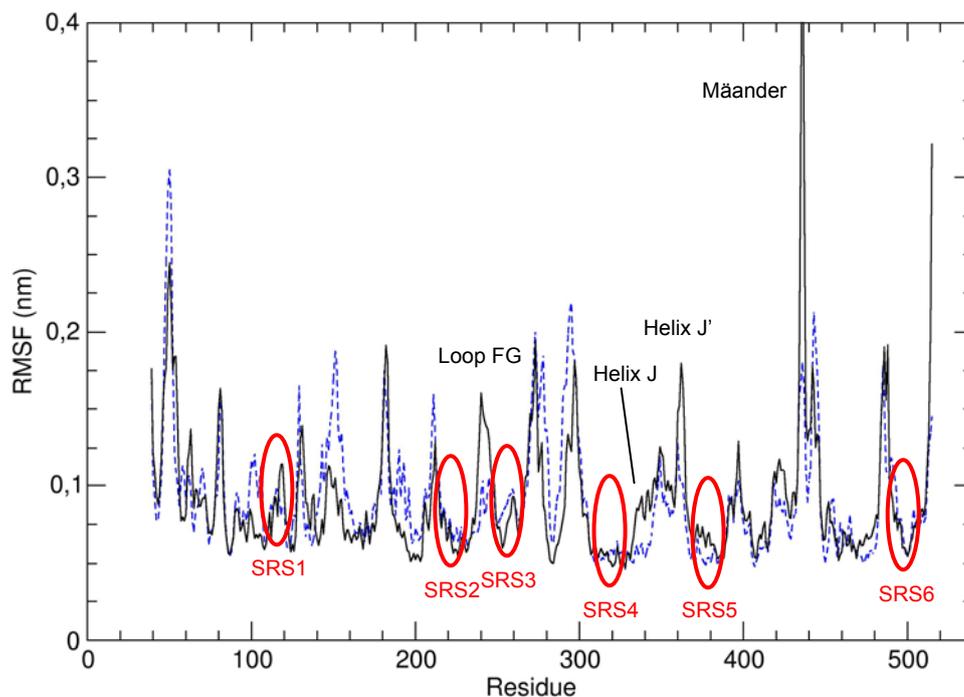


Abbildung 4.60: RMS-Fluktuation der einzelnen Aminosäuren zwischen 3000 und 5000 ps der freien MDS (schwarz). Zum Vergleich ist als blaue gestrichelte Linie der RMSF-Verlauf aus der MDS des leeren Proteinmodells mit aufgetragen. Die roten Markierungen umfassen SRS1-6.

4.3.7.2 Protein-Substrat-Interaktionen

In Abbildung 4.61 ist der Abstand zweier potentieller Hydroxylierungsstellen des Coffeins, der N3-Methylgruppe und der N1-Methylgruppe, zum Hämeisen im Verlauf der MDS mit Coffein in der Position zur Umsetzung zum Paraxanthin (Hydroxylierung der N3-Methylgruppe) aufgetragen. Aus der Tatsache, dass sich zunächst die N1-Methylgruppe näher am Hämeisen befindet als die N3-Methylgruppe, ist ersichtlich, dass sich das Substrat unmittelbar nach dem Start der freien MDS zunächst in die Position zur Bildung des Nebenmetaboliten Theobromin (Hydroxylierung der N1-Methylgruppe) dreht. Im Simulationszeitraum von 700 bis 1000 ps sind kurzzeitig beide Hydroxylierungsstellen gleich weit vom Hämeisen entfernt, bis dann ab etwa 1000 ps das Substrat seine endgültige Orientierung mit der N3-Methylgruppe auf das Hämeisen gerichtet erreicht hat, die es bis zum Ende der MDS nach 5000 ps beibehält. In der Auswertungsphase der MDS beträgt der Abstand der N3-Methylgruppe zum Hämeisen gleichmäßig 0,39 nm, womit diese Methylgruppe in einer für die Umsetzung günstigen Orientierung vorliegt, und die N1-Methylgruppe ist 0,74 nm vom Hämeisen entfernt, also zu weit für eine Umsetzung. Ein ähnliches Hin- und Herdrehen des Substrates wird auch in der Auswertungsphase der zweiten MDS mit Coffein beobachtet, wo dieses Phänomen näher erläutert wird (Abschnitt 4.3.8).

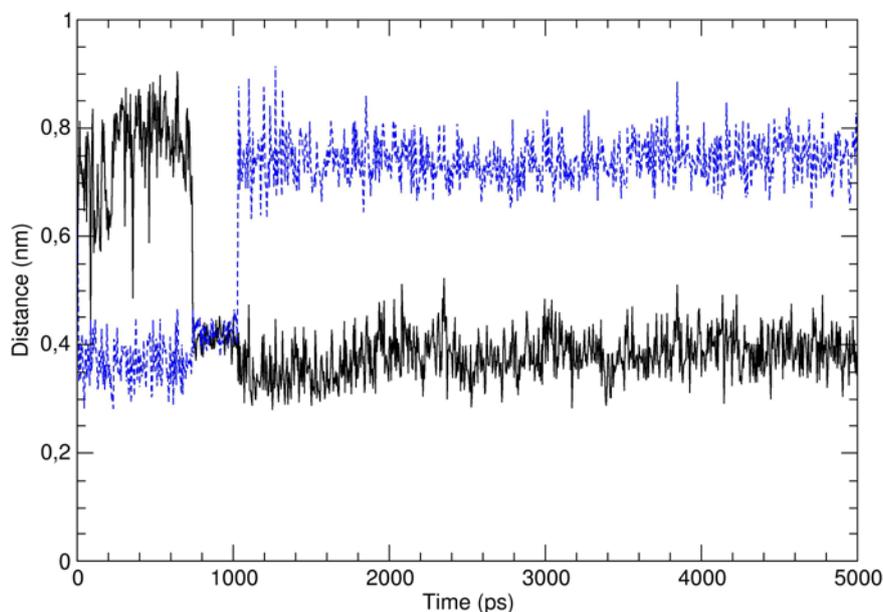


Abbildung 4.61: Abstand der beiden potentiellen Hydroxylierungsstellen N3-Methylgruppe (schwarz) und N1-Methylgruppe (blau gestrichelt) des Coffeins zum Hämeisen im Verlauf der MDS mit Coffein zur Umsetzung zum Paraxanthin.

Im Verlauf der equilibrierten Phase der MDS treten zwei verschiedene Wasserstoffbrücken zwischen dem Substrat Coffein und Aminosäuren des aktiven Zentrums auf (Abbildung 4.62). Relativ selten, aber über den gesamten Simulationszeitraum verstreut, wird eine direkte oder wasservermittelte Wasserstoffbrücke der C6-Carbonylgruppe des Coffeins zu Ser 231 im Loop FG detektiert. Diese Wasserstoffbrücke ist Teil eines Wasserclusters, das sich wie in Abbildung 4.63

dargestellt unter Beteiligung von Thr 124, Asp 313, Tyr 112, Ser 231 und Thr 498 über den oberen Teil des aktiven Zentrums zwischen Helix B'' und dem Loop FG erstreckt.

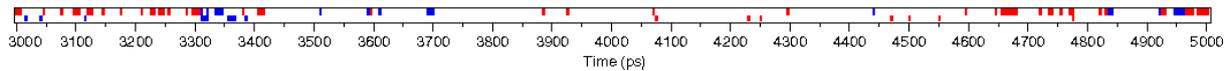


Abbildung 4.62: Matrix des Auftretens von Wasserstoffbrücken zwischen Coffein und Aminosäuren des Proteins im Simulationszeitraum von 3000 bis 5000 ps. Direkte Wasserstoffbrücken sind in rot eingezeichnet und wasservermittelte in blau. In der Reihenfolge von unten nach oben sind aufgetragen die Wasserstoffbrücken Ser 231 OH ($-H_2O$) und Thr 498 OH zur Carbonylfunktion am C6 des Coffeins.

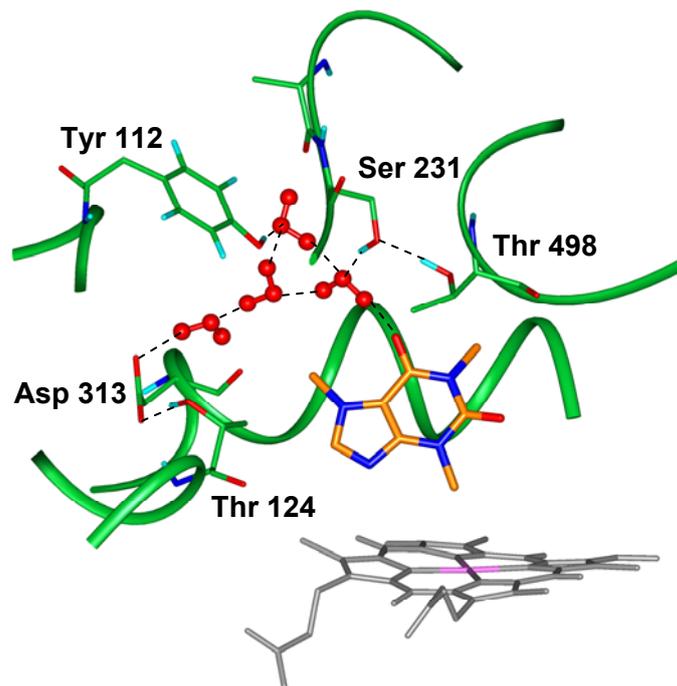


Abbildung 4.63: Wasserstoffbrückenmuster zwischen Aminosäuren des aktiven Zentrums und der Carbonylfunktion am C6 des Coffeins, dargestellt anhand einer Struktur aus dem zweitgrößten Cluster der MDS nach 3385 ps. Die Farbcodierung entspricht der in Abbildung 4.43.

Besonders am Anfang und am Ende der Auswertungsphase besteht häufig eine direkte Wasserstoffbrücke zwischen Thr 498 aus SRS6 und der C6-Carbonylgruppe (Abbildung 4.64). In diesem Fall werden Ser 231 und das Wassermolekül vom Carbonylsauerstoff verdrängt und der Kontakt des Coffeins zum Wassercluster geht verloren. Das Wassercluster bleibt trotzdem weiter bestehen und füllt die Kavität zwischen Helix B'' und dem Loop FG aus, in die das im Vergleich zu MelQ und 7-Ethoxyresorufin relativ kleine Coffein nicht hineinreicht.

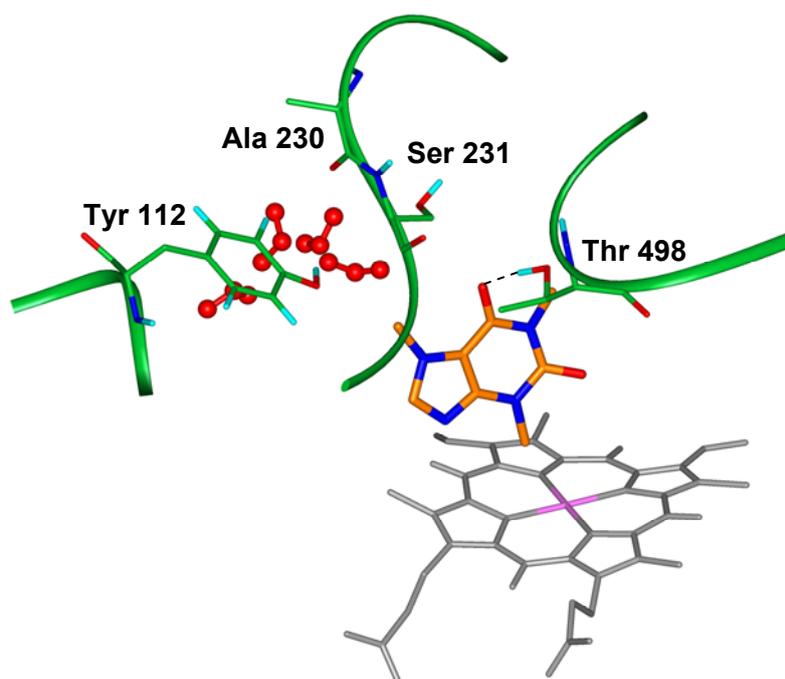


Abbildung 4.64: Wasserstoffbrücke zwischen Thr 498 und der Carbonylfunktion am C6 des Coffeins, dargestellt anhand einer repräsentativen Clusterstruktur der MDS nach 3405 ps. Die Farbcodierung entspricht der in Abbildung 4.43.

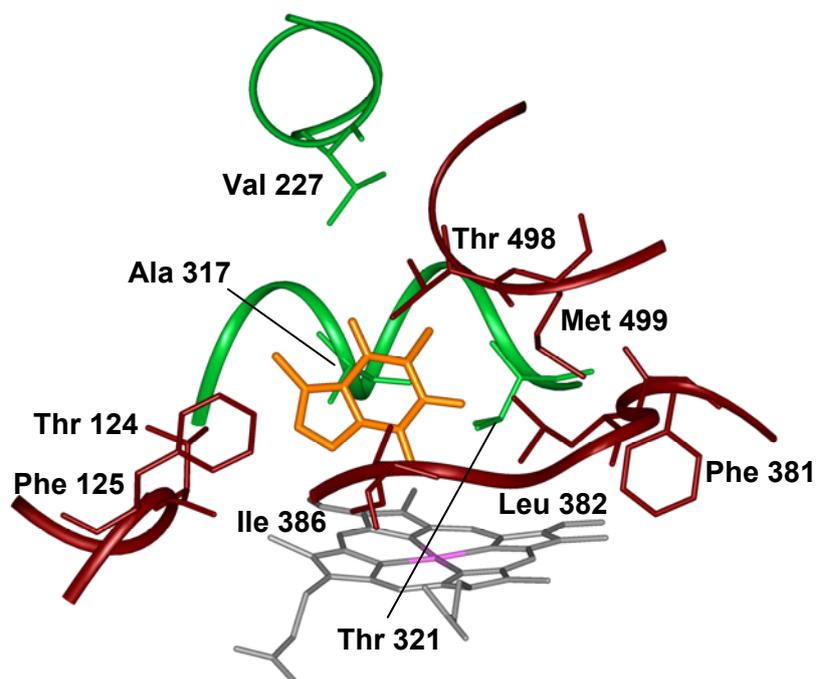


Abbildung 4.65: Lipophile und aromatische Protein-Substrat-Interaktionen von Coffein im aktiven Zentrum des CYP1A2-Modells, dargestellt anhand der repräsentativen Struktur des größten Clusters aus der MDS nach 3950 ps. Die Farbcodierung entspricht der in Abbildung 4.51.

An lipophilen und aromatischen Kontakten zum Coffein (Abbildung 4.65) sind prinzipiell dieselben Aminosäuren beteiligt wie in den beiden bereits beschriebenen MDS mit Substratkomplexen des CYP1A2-Modells. Es fehlen hier lediglich Interaktionen mit relativ weit vom Häm entfernten Aminosäuren wie dem Thr 113 aus

SRS1, dem Thr 223 und dem Phe 226 aus SRS2 und dem Phe 260 aus SRS3. Dafür spielt teilweise Phe 381, das am Beginn der SRS5 lokalisiert ist, für die Substratbindung eine Rolle, da das Coffein im Gegensatz zu MelQ und 7-Ethoxyresorufin aufgrund der Wasserstoffbrücke zu Thr 498 in SRS6 etwas mehr in diesen Teil des aktiven Zentrums verschoben ist.

4.3.8 MDS des Modells mit Coffein zur Umsetzung zum Theobromin

4.3.8.1 RMSD- und RMSF-Wert

Im Vergleich zu den vorangegangenen MDS mit Substratkomplexen des CYP1A2-Modells nimmt der Verlauf des RMSD-Wertes hier erstaunlicherweise insgesamt nur relativ niedrige Werte von bis zu 0,2 nm bezogen auf das gesamte Proteinrückgrat an (Abbildung 4.66). Der RMSD-Wert erreicht ein Plateau nach etwa 2000 ps, steigt aber in der Auswertungsphase bei etwa 4000 ps nochmals leicht an. Abweichend von allen bisher analysierten MDS verläuft der RMSD-Wert des Proteinkerns in diesem Fall parallel zum RMSD-Wert des gesamten Proteinrückgrates. Auch er steigt also nach etwa 4000 ps Simulationszeit wieder leicht an. Das bedeutet, dass das Protein noch nicht vollständig equilibriert ist, und dass auch im aktiven Zentrum noch geringfügige Konformationsänderungen stattfinden.

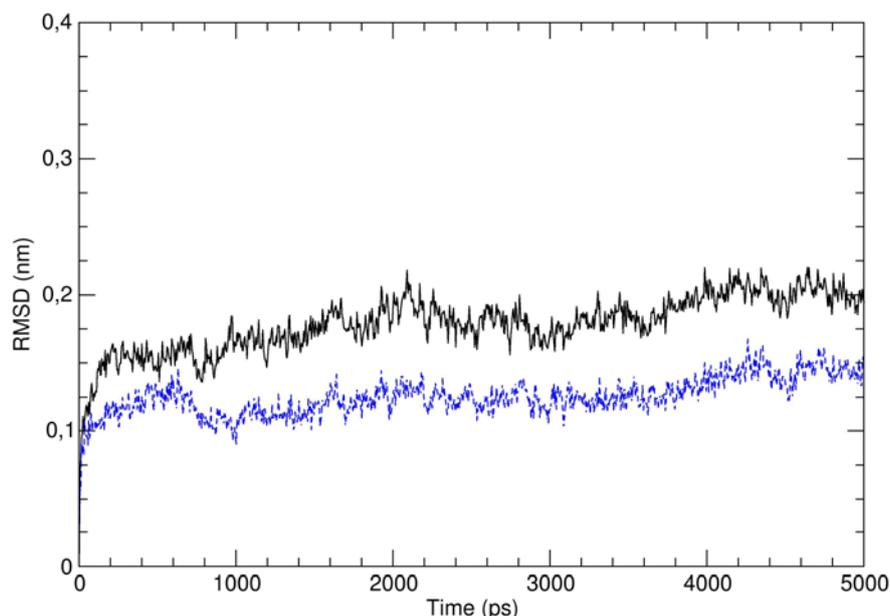


Abbildung 4.66: RMSD-Wert des Proteinrückgrates bezogen auf die Startstruktur im Verlauf der freien MDS des CYP1A2-Modells mit Coffein zur Umsetzung zum Theobromin. Aufgetragen ist der RMSD-Wert des ganzen Proteinrückgrates (schwarz) und des Proteinrückgrates des Proteinkerns, d.h. der SRS1-6 und der Hämbinderegionen (blau gestrichelt).

Betrachtet man das Profil der RMS-Fluktuation in Abbildung 4.67, so fällt wie auch bei der vorangegangenen MDS mit Coffein ein hohes Maximum von etwa 0,4 nm auf.

Es ist hier aber nicht in der Mäander-Region lokalisiert sondern im Bereich des Loops HI. Im weiteren Verlauf der Kurve ist in der Mitte von Helix I (SRS4) eine deutlich höhere Spitze als bisher zu erkennen. Auch in SRS1 und in SRS5 ist eine erhöhte Beweglichkeit im Vergleich zur MDS des leeren Proteinmodells und auch im Vergleich mit den bisher dargestellten RMSF-Profilen von Substratkomplexen offensichtlich.

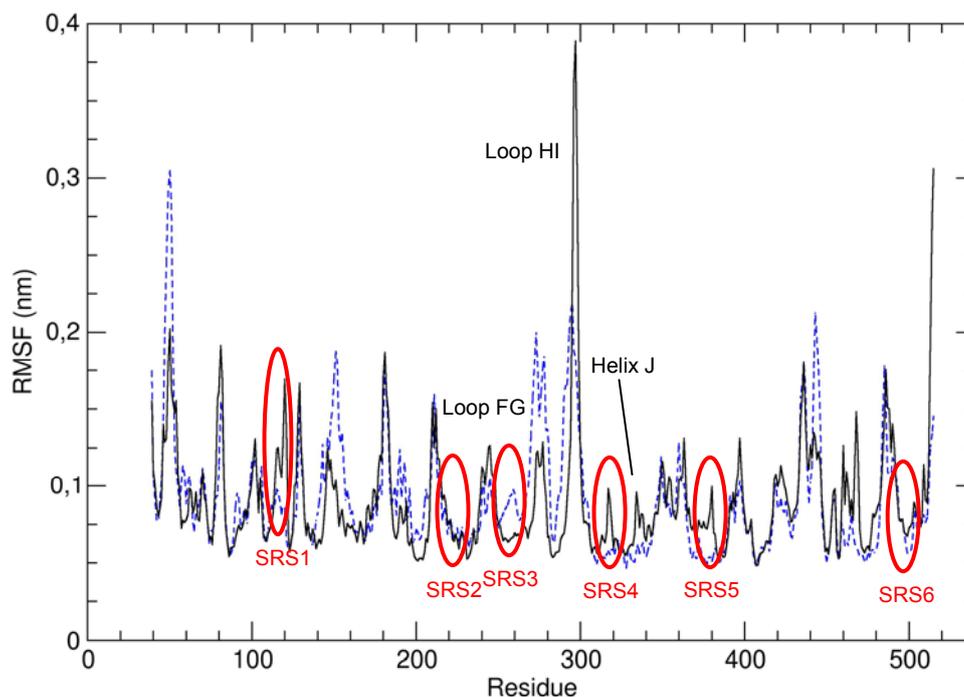


Abbildung 4.67: RMS-Fluktuation der einzelnen Aminosäuren zwischen 3000 und 5000 ps der freien MDS (schwarz). Zum Vergleich ist als blaue gestrichelte Linie der RMSF-Verlauf aus der MDS des leeren Proteinmodells mit aufgetragen. Die roten Markierungen umfassen SRS1-6.

4.3.8.2 Protein-Substrat-Interaktionen

Wie schon in der bereits beschriebenen MDS mit Coffein in der Position für die Umsetzung zum Paraxanthin wurde auch bei dieser MDS der Abstand der N1- und der N3-Methylgruppe zum Hämeisen über den gesamten Simulationsverlauf verfolgt (Abbildung 4.68). Es zeigt sich, dass sich die N1-Methylgruppe mit einem Abstand von 0,39 nm durchgehend in einer für eine Umsetzung günstigen Orientierung zum Eisen befindet. Gegen Ende der MDS steigt der Abstand dieser Methylgruppe bis auf etwa 0,4 nm ganz leicht an. Die N3-Methylgruppe hingegen, die sich zu Beginn der MDS im Abstand von durchschnittlich etwa 0,75 nm vom Hämeisen befindet, nähert sich ab etwa 3000 ps, also ab dem Beginn der Auswertungsphase, der Hämebene langsam bis auf etwa 0,68 nm an. Folglich deutet sich während der letzten zwei Nanosekunden der MDS die Tendenz des Coffeins an, sich von der Positionierung zur Bildung des Nebenmetaboliten Theobromin in die Position zur Bildung des Hauptmetaboliten Paraxanthin zu drehen. Das würde sowohl den erneuten Anstieg

des RMSD-Wertes im Proteinkern als auch die erhöhte RMS-Fluktuation im Bereich der SRS1, 4 und 5 und dem vor SRS4 liegenden Loop HI erklären.

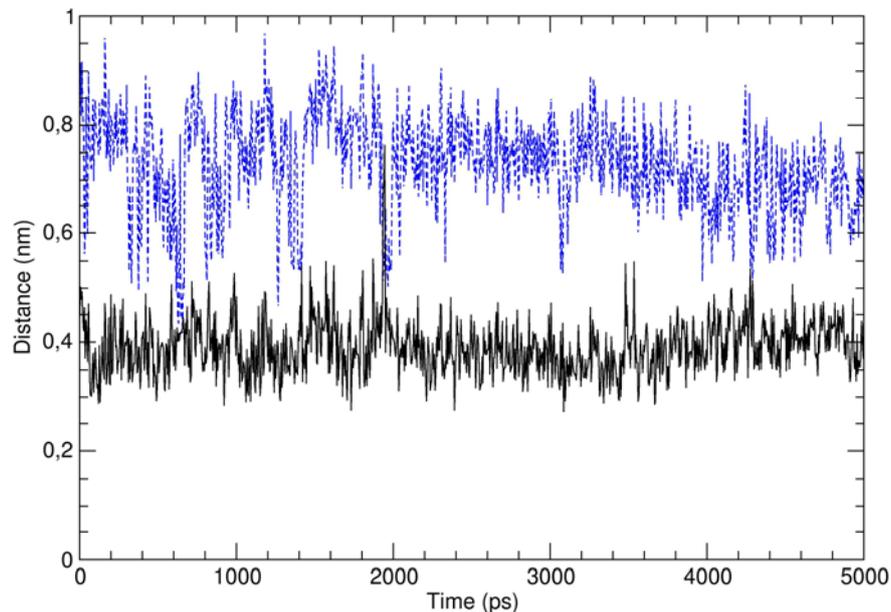


Abbildung 4.68: Abstand der beiden potentiellen Hydroxylierungsstellen des Coffeins, der N1-Methylgruppe (schwarz) und der N3-Methylgruppe (blau gestrichelt), zum Hämeisen im Verlauf der MDS mit Coffein zur Umsetzung zum Theobromin.

Die in Abbildung 4.69 abgebildete Matrix zeigt im Verlauf der letzten zwei Nanosekunden der MDS ebenfalls Veränderungen im Wasserstoffbrückenmuster an. Zu Beginn der Auswertungsphase tritt häufig eine wasservermittelte Interaktion zwischen der Hydroxylgruppe von Thr 112 aus SRS1 und dem nicht methylierten Imidazolstickstoff N9 des Coffeins auf. Darüber hinaus existiert eine direkte Wasserstoffbrücke der Hydroxylgruppe von Thr 321 aus SRS4 mit der Carbonylfunktion am C6 des Coffeins.

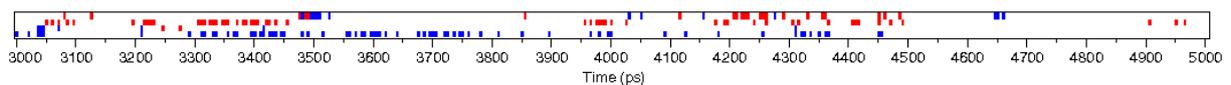


Abbildung 4.69: Matrix des Auftretens von Wasserstoffbrücken zwischen Coffein und Aminosäuren des Proteins im Simulationszeitraum von 3000 bis 5000 ps. Direkte Wasserstoffbrücken sind in rot eingezeichnet und wasservermittelte in blau. In der Reihenfolge von unten nach oben sind aufgetragen die Wasserstoffbrücken Tyr 112 OH – H₂O und Ser 231 OH (– H₂O) zum N9 des Coffeins, sowie Thr 321 OH und Thr 498 OH zur Carbonylfunktion am C6 des Coffeins.

Abbildung 4.70 stellt dieses Wasserstoffbrückenmuster zu Beginn der MDS anhand der repräsentativen Struktur des größten Strukturclusters zum Simulationszeitpunkt von 3650 ps dar. Die wasservermittelte Interaktion zwischen Tyr 112 und N9 des Coffeins ist wiederum Teil eines Wasserclusters, das unter Beteiligung von Thr 113 und Thr 124 aus SRS1, Asn 257 aus SRS3 und Asn 312 aus SRS4 die Kavität zwischen Helix B'' und dem Loop FG ausfüllt. Die Wasserstoffbrücke des Carbonylsauerstoffs am C6 des Coffeins wird nicht immer direkt zur Hydroxylgruppe

von Thr 321 ausgebildet, sondern verläuft wie in der Abbildung teilweise auch über das in der Helixwindung eingelagerte Wassermolekül zu den Proteinrückgratatomen dieser Helixwindung. Der wasservermittelte Kontakt des Coffeins zu Tyr 112 bricht nach etwa 4500 ps ab, und auch die direkte Wasserstoffbrücke zum Thr 321 wird danach nur noch in drei Strukturen detektiert.

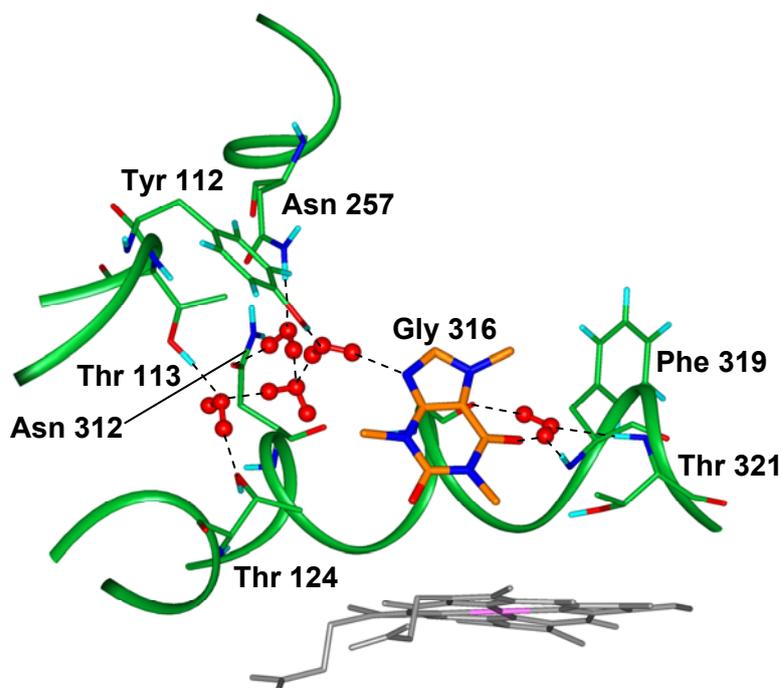


Abbildung 4.70: Wasserstoffbrückenmuster zu Beginn der MDS mit Coffein in der Position zur Umsetzung zum Theobromin, dargestellt anhand einer repräsentativen Struktur nach 3650 ps. Die Farbcodierung entspricht der in Abbildung 4.43.

Stattdessen tritt ab etwa 4000 ps verstärkt eine wasservermittelte oder direkte Wasserstoffbrücke zwischen Thr 498 aus SRS6 und dem C6-Carbonylsauerstoff auf, die auch in der ersten MDS mit Coffein häufiger zu beobachten war. Dass dieser Kontakt auch bis zum Ende der MDS nicht völlig abreißt, sondern letztlich bestehen bleibt, geht aus Abbildung 4.71 hervor, die eine repräsentative Clusterstruktur zum Simulationszeitpunkt 4940 ps zeigt. Dort ist der Carbonylsauerstoff des Coffeins über zwei Wassermoleküle mit der Hydroxylgruppe von Thr 498 verbunden. Die Position der Threonins wird zusätzlich durch eine Wasserstoffbrücke zu Ser 231 des Loops FG stabilisiert. Darüber hinaus erkennt man in der Abbildung, dass der Kontakt des C6-Carbonylsauerstoffes vom Coffein zur SRS4 unterbrochen ist. Die Helixwindung weist wieder die für Cytochrome charakteristische Form auf. Auch eine Wasserstoffbrücke des Imidazolstickstoffes von Coffein zu Tyr 112 ist nicht mehr möglich, das Substrat hat sich aufgrund des Kontaktes zu Thr 498 weggedreht. Vergleicht man die Orientierung des Coffeins in zu Beginn der Auswertungsphase in Abbildung 4.70 mit der gegen Ende der MDS in Abbildung 4.71, so wird auch die Annäherung der N3-Methylgruppe an das Hämeisen nachvollziehbar.

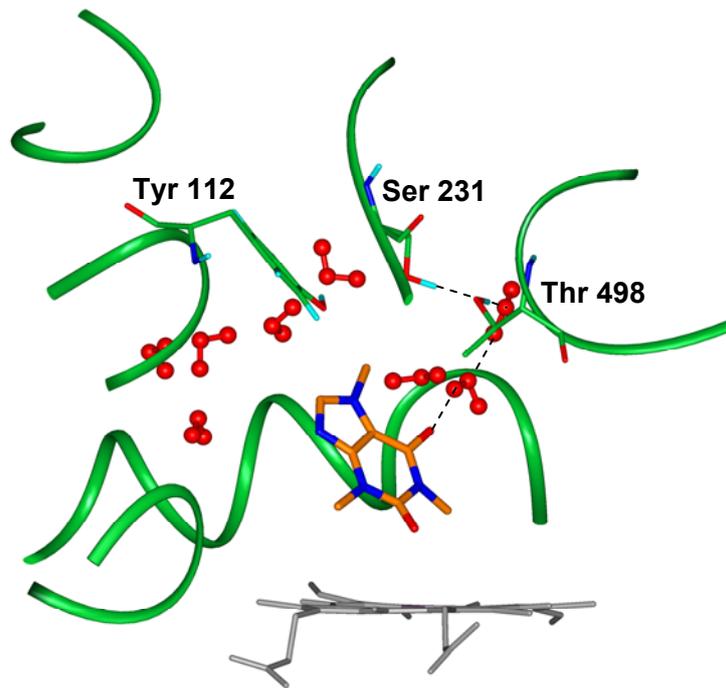


Abbildung 4.71: Wasserstoffbrückenmuster am Ende der MDS mit Coffein in der Position zur Umsetzung zum Theobromin, dargestellt anhand einer repräsentativen Struktur nach 4940 ps. Die Farbcodierung entspricht der in Abbildung 4.43. Aus Gründen der Übersichtlichkeit sind die am Wassercluster beteiligten Aminosäuren und die Aminosäuren in Helix I, die das katalytische Wasser fixieren, ausgeblendet.

Auch Abbildung 4.72 verdeutlicht noch einmal die Drehung des Coffeins im Verlauf der MDS, die eine Tendenz zur Orientierung der N3-Methylgruppe zum Hämeisen und damit zur Bildung des Hauptmetaboliten Paraxanthin zeigt.

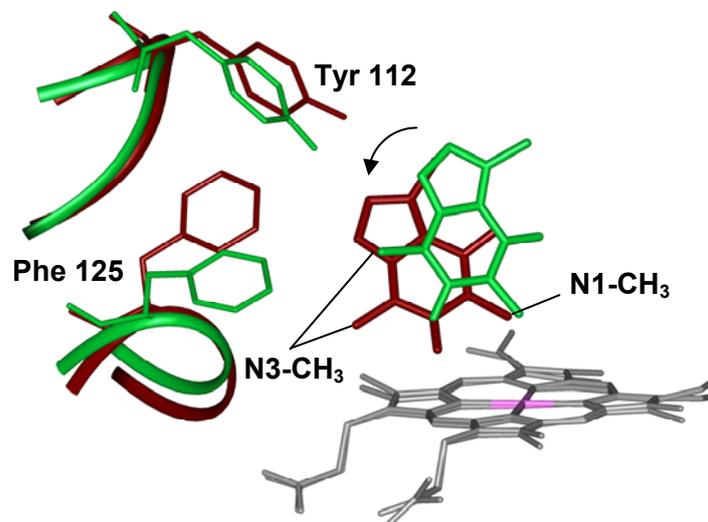


Abbildung 4.72: Wichtige Unterschiede der lipophilen und aromatischen Protein-Substrat-Interaktionen zwischen dem Beginn und dem Ende der MDS mit Coffein zur Umsetzung zum Theobromin, dargestellt anhand zweier repräsentativer Strukturen nach 3650 ps (grün) und nach 4940 ps (braun).

Bezüglich der aromatischen und lipophilen Wechselwirkungen zwischen Protein und Substrat gibt es im Verlauf der MDS zwei Veränderungen. Sowohl Tyr 112 als auch

Phe 125 richten sich so aus, dass eine Polarisationswechselwirkung ihrer Ringsysteme mit dem Imidazolring des Coffeins möglich wird. Zu Beginn der MDS dagegen ist Tyr 112 wie bereits erläutert über seine Hydroxylfunktion am Wasserstoffbrückenmuster beteiligt und Phe 125 steht über eine lipophile Interaktion zur N3-Methylgruppe im Kontakt mit dem Substrat. Ansonsten werden dieselben lipophilen Interaktionen beobachtet wie auch in der ersten MDS mit Coffein.

4.3.9 Vergleich der beiden MDS mit Coffein

Zunächst ist festzustellen, dass die beiden mit Coffein durchgeführten MDS zu einem übereinstimmenden Ergebnis kommen. In der ersten MDS, in der die N3-Methylgruppe zur Bildung des Hauptmetaboliten Paraxanthin auf das Hämeisen gerichtet war, fand zwar zu Beginn eine spontane Drehung des Coffeins mit Orientierung der N1-Methylgruppe zum katalytischen Zentrum statt, doch bereits innerhalb der ersten Nanosekunde der Simulationszeit wurde die ursprüngliche Orientierung wieder hergestellt und blieb bis zum Ende der MDS nach 5 ns stabil. Der Grund für die anfängliche Umorientierung liegt darin, dass zunächst wie in der zweiten MDS mit Coffein, bei der von Beginn an die N1-Methylgruppe auf das Häm gerichtet war, eine wasservermittelte Interaktion vom N9 des Coffeins zu Tyr 112 zustande kam. Die Ursache dafür könnte eine ungünstige Anordnung der Lösungsmittelmoleküle im aktiven Zentrum der Startstruktur sein. Im weiteren Verlauf der MDS kommt nach kurzer Zeit zuerst eine Wasserstoffbrücke zwischen der C6-Carbonylfunktion des Coffeins und Thr 498 zustande und zum Schluss eine Polarisationswechselwirkung von Phe 125 mit dem Imidazolring des Coffeins, die die endgültige Rückorientierung der N3-Methylgruppe auf das Hämeisen bewirkt. Genau derselbe Ablauf ist tendenziell und viel langsamer auch über den Zeitraum der Auswertungsphase der zweiten MDS mit Coffein zu beobachten.

Betrachtet man die Bilanz der Protein-Substrat-Interaktionen, so ist die bevorzugte Orientierung der N3-Methylgruppe zum Hämeisen zunächst nicht zu verstehen. Ist die N3-Methylgruppe auf das Eisen hin orientiert, wird das Substrat durch eine Wasserstoffbrückeninteraktion am Carbonylsauerstoff und durch eine Polarisationswechselwirkung mit dem Imidazolring in seiner Position fixiert. Im Fall einer Orientierung der N1-Methylgruppe zum Hämeisen existieren zwei Wasserstoffbrückeninteraktionen, und zwar einmal zum Imidazolstickstoff N9 und einmal zum Carbonylsauerstoff. Aufgrund ihrer Wechselwirkungsenergie müssten zwei Wasserstoffbrücken ein Substrat fester binden als eine Wasserstoffbrücke und eine Polarisationswechselwirkung. Betrachtet man jedoch die Wasserstoffbrückenpartner der C6-Carbonylgruppe im Fall der zweiten MDS mit Coffein, so fällt auf, dass es sich hierbei entweder um das für den Ablauf der Katalyse wichtige Wassermolekül in Helix I handelt, das vermutlich beim Protonentransport eine Rolle spielt, oder um das ebenfalls am katalytischen Mechanismus beteiligte hochkonservierte Thr 321 in

derselben Helixwindung. Dass dieser Bereich an der Substratbindung beteiligt ist, ist also eher unwahrscheinlich. Man könnte daher vermuten, dass diese Interaktion entweder ein Artefakt der MDS ist, das es überhaupt ermöglicht, dass das Substrat so lange in dieser Orientierung verbleibt, oder dass diese Interaktion zumindest ungünstig ist und eine Umsetzung behindert, weshalb eine Substratorientierung bevorzugt wird, bei der die Carbonylgruppe die Konformation der an der Katalyse beteiligten Helixwindung nicht stört.

4.3.10 Vergleich der drei untersuchten Substratkomplexe

Da es sich bei den drei im Komplex mit dem CYP1A2-Modell untersuchten Substraten um Verbindungen handelt, die nicht nur strukturell unterschiedlich sind, sondern auch eine unterschiedliche Affinität zu CYP1A2 aufweisen und nach verschiedenen Reaktionsmechanismen umgesetzt werden, sollen die drei Substratkomplexe hier abschließend miteinander verglichen werden.

Im Zusammenhang mit der Positionierung der drei Substrate nach dem Vorbild der Kristallstrukturen wurde erwähnt, dass die verschiedenen Substrate in den Kristallstrukturen je nach Umsetzungsreaktion eine unterschiedliche Neigung zur Hämebene aufweisen (Abschnitt 4.3.1). Ein ähnliches Phänomen kann man auch bei den Substratkomplexen des CYP1A2-Modells beobachten (Abbildung 4.73). MeIQ wird von CYP1A2 in Form einer N-Oxidation des exozyklischen Stickstoffs umgesetzt. Bei einer N-Oxidation ist die Übertragung eines Elektrons vom Stickstoff auf die Eisen-Oxo-Verbindung der erste Schritt (Abschnitt 1.3.2.3). Folglich ist der Stickstoff im Substratkomplex so in Richtung Häm geneigt, dass eine Elektronenübertragung gut erfolgen kann (Abbildung 4.73 a). Beim 7-Ethoxyresorufin ist das Ringsystem zwar auch zum Häm geneigt, die Ethoxylseitenkette, die hier in einer O-Desalkylierungsreaktion oxidativ entfernt wird, steht aber nahezu senkrecht über dem Häm, sie ist sogar leicht vom Häm weg geneigt (Abbildung 4.73 b). Auch dies steht im Einklang mit dem Reaktionsmechanismus, da bei O-Desalkylierungsreaktionen ähnlich wie bei der aliphatischen Hydroxylierung zuerst die Übertragung eines Wasserstoffatoms vom Alkylrest auf die Eisen-Oxo-Verbindung erfolgt. Steht der Alkylrest senkrecht zur Hämebene, zeigt eines der beiden Wasserstoffatome in Richtung Hämeisen und kann so leicht übertragen werden. Das Coffein weist wiederum eine deutliche Neigung zur Hämebene auf (Abbildung 4.73 c). In diesem Fall ist ein Rückschluss auf den Reaktionsmechanismus allerdings nicht eindeutig möglich, da bezüglich der N-Desalkylierungen zwei Erklärungsmodelle für den Ablauf der Reaktion existieren. Hier ist sowohl die Übertragung eines Elektrons vom Stickstoff als auch die Übertragung eines Wasserstoffatoms vom Alkylrest als erster Reaktionsschritt denkbar (vgl. Abschnitt 1.3.2.4). In Analogie zum MeIQ würde die Neigung des Coffeins zum Häm in der Proteinstruktur allerdings eher auf eine Elektronenübertragung hindeuten als auf eine Wasserstoffübertragung.

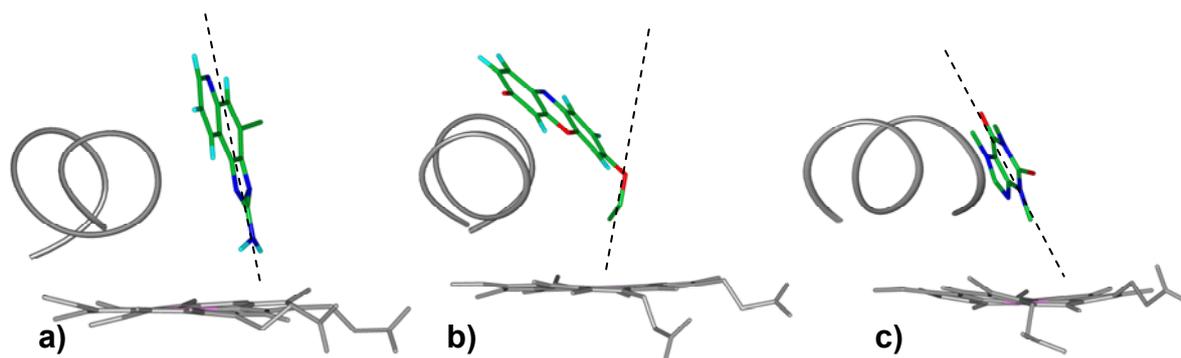


Abbildung 4.73: Unterschiedliche Neigung der Substrate zur Hämebene, dargestellt jeweils anhand einer repräsentativen Clusterstruktur. a) MelQ, b) 7-Ethoxyresorufin und c) Coffein. Zur Orientierung sind jeweils das Häm und der zentrale Abschnitt der Helix I eingeblendet.

Die Affinität der drei Substrate zum CYP1A2 unterscheidet sich jeweils um eine Zehnerpotenz, wie die K_m -Werte in Tabelle 4.17 zeigen. Coffein ist das kleinste der untersuchten Substrate. Es nimmt nicht das gesamte Volumen des aktiven Zentrums ein, sodass ein Teil des aktiven Zentrums in diesem Substratkomplex mit Wasser gefüllt ist. Dadurch kommen weniger lipophile und aromatische Kontakte zwischen Substrat und Protein zustande als in den übrigen beiden Substratkomplexen. Coffein besitzt demnach die geringste Affinität zum CYP1A2. MelQ ist zwar wie das Coffein auch nur über eine Wasserstoffbrücke an das Protein gebunden, füllt aber aufgrund seiner Molekülgröße einen größeren Teil des aktiven Zentrums aus und bildet folglich mehr lipophile Kontakte zum Protein aus. Beispielsweise kommt beim MelQ zusätzlich eine π - π -Interaktion zu Phe 226 zustande, die beim Coffein fehlt. Das MelQ weist somit eine um eine Zehnerpotenz höhere Affinität zu CYP1A2 auf als das Coffein. Die höchste Affinität besitzt jedoch das 7-Ethoxyresorufin. Es ragt zum einen aufgrund des aus drei Ringen bestehenden Ringsystems noch tiefer in die Kavität zwischen den Helices F und G hinein und bildet zusätzliche Kontakte zu Phe 260 und Thr 223 aus. Zum anderen ist es aber auch im Gegensatz zu den beiden übrigen Substraten über zwei Wasserstoffbrücken an das Protein gebunden, nämlich einmal vom Carbonylsauerstoff ausgehend und einmal über den heterozyklischen Stickstoff. Diese Überlegungen zeigen, dass auf der Grundlage der unterschiedlichen beobachteten Protein-Substrat-Interaktionen Abstufungen in der Enzymaffinität zumindest grob qualitativ nachvollzogen werden können.

Tabelle 4.17: K_m -Werte der untersuchten Substrate (Lewis 2003)

7-Ethoxyresorufin	MelQ	Coffein
$K_m = 1,7 \mu\text{mol}$	$K_m = 13 \mu\text{mol}$	$K_m = 180 \mu\text{mol}$

4.4 3D-QSAR-Studie

Das in den vorangegangenen Kapiteln erstellte und validierte Proteinmodell von CYP1A2 wurde im Anschluss in einer 3D-QSAR-Studie für die Überlagerung verwendet. Wie bereits in Abschnitt 3.6 erwähnt, sollten die Verbindungen eines dafür zur Verfügung stehenden Datensatzes, der sich hauptsächlich aus Naphthalin- und Laktonstrukturen zusammensetzt, mithilfe eines Docking-Programms im aktiven Zentrum des Proteins überlagert werden. Zu diesem Zweck wurde verschiedene Programme getestet und das Programm GOLD ausgewählt, weil es die Hämumgebung am besten beschreiben konnte.

Die für das Docking benötigte Proteinstruktur wurde aus der equilibrierten Phase der MDS mit dem leeren Proteinmodell entnommen (s. Abschnitt 4.2.5). Hierzu wurde wiederum eine Clusteranalyse durchgeführt, bei der zur Clustereinteilung wie schon im Rahmen der Auswertung der Protein-Substrat-Komplexe alle Atome der Aminosäuren der SRS1-6 verwendet wurden, um einen Überblick über verschiedene Konformationen des aktiven Zentrums zu erlangen. Ausgewählt wurde daraufhin eine repräsentative Struktur aus einem der größten Cluster zum Simulationszeitpunkt 1840 ps, die eine gute strukturelle Qualität aufwies. Bevor die Struktur für das Docking eingesetzt wurde, wurde sie zuerst mit dem *steepest descent* Algorithmus und anschließend mit dem *conjugate gradient* Algorithmus minimiert.

Zwar erwies sich von den zur Verfügung stehenden Docking-Programmen das Programm GOLD für die vorliegende Fragestellung als am besten geeignet, doch es bestanden auch bei diesem Programm Schwierigkeiten bei der korrekten Behandlung des Häm. Es musste daher eine Modifikation der Eisenparameter eingeführt werden und das Docking-Verfahren wurde anschließend zunächst anhand eines Substratdatensatzes validiert.

4.4.1 Validierung des Dockings

In einem ersten Testlauf, in dem die drei für die MDS verwendeten Substrate MeIQ, 7-Ethoxyresorufin und Coffein in das aktive Zentrum des Proteinmodells positioniert wurden, stellte sich heraus, dass zunächst keines der drei Substrate in einer für eine Umsetzung geeigneten Orientierung über das Häm platziert wurde. Für Coffein wurden mehrere Orientierungen gefunden, bei denen sich das Molekül aber stets im oberen Teil des aktiven Zentrums befand, die potentielle Hydroxylierungsstelle also viel zu weit vom Hämeisen entfernt war. Im Fall des MeIQ konnte auch unter Verwendung eines *distance constraints*, der den exozyklischen Guanidinstickstoff in Richtung Hämeisen zwingen sollte, nie eine Orientierung des Stickstoffs zum Eisen hin erreicht werden. Für 7-Ethoxyresorufin schließlich wurde eine Komplexierung der Carbonylfunktion mit dem Hämeisen an der unbesetzten sechsten Koordinations-

stelle von GOLD vorgeschlagen. Diese Beobachtungen stimmen mit den Ergebnissen einer kürzlich veröffentlichten Studie überein, in der die Vorhersagekraft des Programms GOLD in Bezug auf Cytochrome P450 und andere Hämproteine anhand von Kristallstrukturen getestet wurde [166]. Dort fanden die Autoren heraus, dass Carbonylfunktionen von Liganden fast immer eine Koordination mit dem Hämeisen eingingen. Dass dies offensichtlich auf einer Überbewertung von Eisen-Sauerstoff-Interaktionen durch das Programm beruht, zeigte die Tatsache, dass in der CSD keine einzige Komplexverbindung existiert, bei der ein Carbonylsauerstoff an ein Eisenatom gebunden ist. Der Grund für diese Überbewertung liegt in der Definition von Metall-Ligand-Wechselwirkungen als Pseudo-Wasserstoffbrücken (s. Abschnitt 3.6.1.2), die dazu führt, dass Interaktionen zwischen dem Eisen und potentiellen Akzeptoren überschätzt werden. Als weitere Unzulänglichkeit des Programms stellte sich heraus, dass planare Stickstoffe nicht als lipophil behandelt werden, und somit niemals korrekt über der extrem lipophilen Hämebene platziert werden können. Kirton et al. schlugen daraufhin eine Reparametrisierung von Eisen-Akzeptor-Interaktionen und eine Klassifizierung von planaren Stickstoffen (abgesehen von Amidstickstoffen) als lipophile Atome vor. Diese Modifikationen erfordern einen Eingriff in den *source code* des Programms und konnten daher für diese Arbeit leider nicht übernommen werden. Da der Datensatz, der hier für die QSAR-Analyse verwendet werden soll, zwar keine Verbindungen mit exozyklischen lipophilen Stickstoffen, wohl aber Laktonstrukturen enthält, musste zumindest für die korrekte Behandlung sauerstoffhaltiger Verbindungen eine Lösung gefunden werden. Geht man davon aus, dass die Enzym-Ligand-Interaktionen bei den Cytochromen im wesentlichen durch die Aminosäuren des Proteins und durch das lipophile Hämringsystem bestimmt werden, und dass das Eisen für die Ligandbindung in der Regel nur eine untergeordnete Rolle spielt, so ist es möglich, das Eisen als *dummy atom*⁸ zu definieren und diesem *dummy atom* lediglich die van-der-Waals-Parameter des Eisens zuzuweisen. Dies hat zur Folge, dass die Donor-Eigenschaften des Eisens wegfallen, während das Atomvolumen weiterhin bestehen bleibt.

Ob diese Veränderung zulässig ist und eine Verbesserung der Vorhersage sauerstoffhaltiger Verbindungen bewirkt, wurde zunächst anhand eines Testdatensatzes typischer CYP1A2-Substrate getestet, der in Anlehnung an Fassbender 1997 [167] und De Rienzo et al. 2000 [47] zusammengestellt wurde. Er besteht aus 12 Verbindungen, die bereits in der Einleitung kurz vorgestellt wurden (Abschnitt 1.2.1.3, Abbildung 4.74). Sie sind größtenteils aromatisch, enthalten vielfach Carbonyl- oder Etherfunktionen und weisen damit eine gewisse Ähnlichkeit zum Inhibitor Datensatz für die QSAR-Analyse auf (vgl. nächster Abschnitt). Beispiele für N-Hydroxylierungen

⁸ dummy, engl.: Attrappe. Ein *dummy atom* erfüllt nur eine geometrische Funktion, besitzt aber keine physikochemischen Eigenschaften.

wurden weggelassen, weil sie nach wie vor nicht korrekt wiedergegeben werden können und auch im QSAR-Datensatz nicht vorkommen.

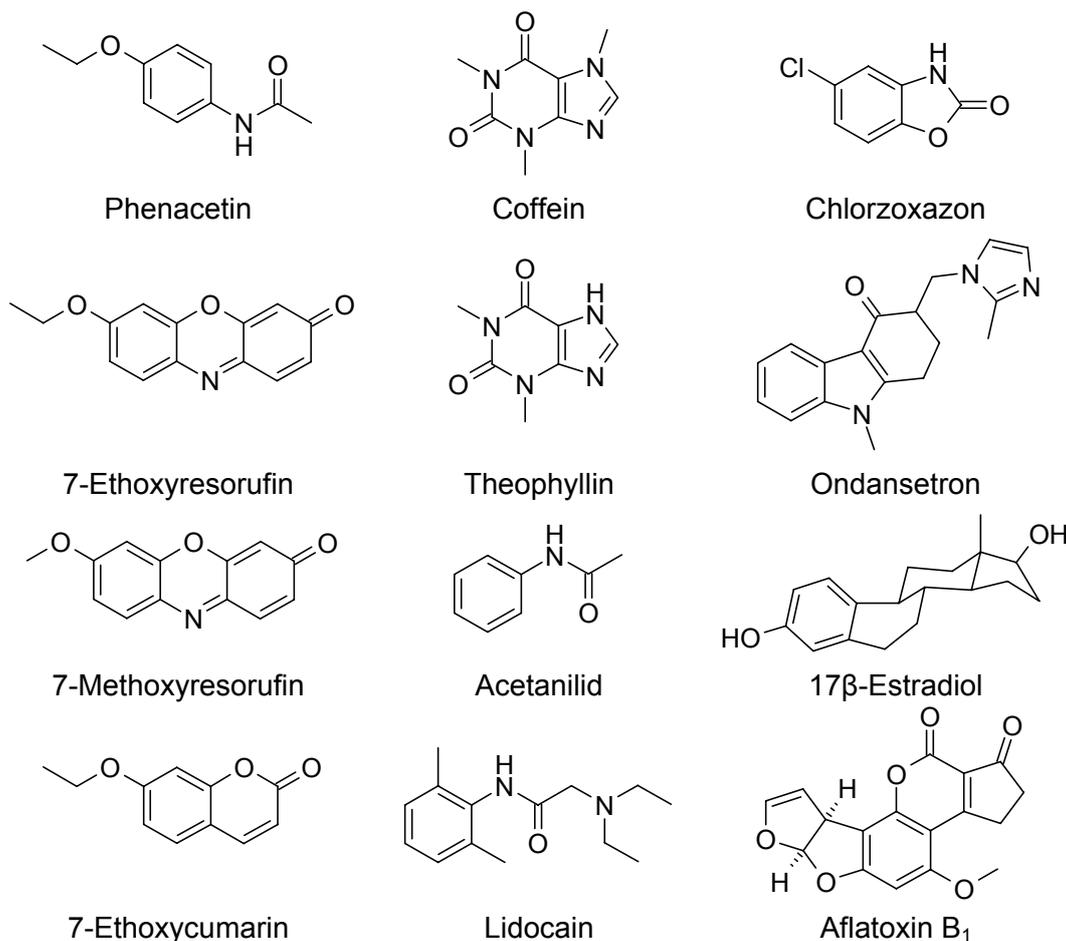


Abbildung 4.74: Verbindungen des Substratdatensatzes

Für das Docking wurde sowohl bei der Validierung als auch bei der Durchführung der eigentlichen QSAR-Analyse ausgehend vom Hämeisen als Zentralatom das aktive Zentrum mit einem Radius von 15 Å definiert. Wie standardmäßig vorgesehen, wurden für jeden Liganden 10 Dockinglösungen erzeugt. Dabei wurde ein Abbruch des Dockings erlaubt, wenn die drei besten Dockinglösungen untereinander einen RMSD-Wert von 1,5 Å oder kleiner aufwiesen. Es stellte sich heraus, dass diese Einstellungen völlig ausreichten, um den Konformationsraum der Liganden auszuschöpfen, da sowohl die Liganden des Validierungsdatensatzes als auch die Liganden des Inhibitor datensatzes wie die CYP1A2-Liganden allgemein überwiegend aus starren Ringsystemen mit nur kurzkettingen Substituenten bestehen. Von den beiden in GOLD implementierten Scoringfunktionen (s. Abschnitt 3.6.1.3) erwies sich die ChemScore-Funktion für die vorliegende Problemstellung als am besten geeignet und wurde in allen Dockingläufen verwendet.

Tabelle 4.18 listet alle Verbindungen des Validierungsdatensatzes und ihre jeweilige Hydroxylierungsstelle auf, und gibt als Bewertungskriterium für eine korrekte Positionierung im aktiven Zentrum des CYP1A2-Modells den jeweils günstigsten

Abstand der Hydroxylierungsstelle zum Eisen an. Für 8 der 12 Verbindungen wurde aufgrund des starren Ringsystems jeweils nur eine mögliche Konformation gefunden, und das Docking wurde nach drei Durchläufen beendet.

Tabelle 4.18: Dockingergebnis für den Validierungsdatensatz. In der letzten Spalte ist der jeweils günstigste Abstand der jeweiligen Hydroxylierungsstelle zum Hämeisen angegeben.

Substrat	Hydroxylierungsstelle	Abstand zum Eisen (Å)
Phenacetin	O-Desethylierung	4,50
7-Ethoxyresorufin	O-Desethylierung	3,44
7-Methoxyresorufin	O-Demethylierung	3,88
7-Ethoxycumarin	O-Desethylierung	4,05
Coffein	N3-Demethylierung	12,40
Theophyllin	8-Hydroxylierung	13,58
Acetanilid	4-Hydroxylierung	3,79
Lidocain	3-Hydroxylierung	3,47
Chlorzoxazon	6-Hydroxylierung	3,82
Ondansetron	7,8-Hydroxylierung	3,94/3,34
17 β -Estradiol	2-Hydroxylierung	12,46
Aflatoxin B ₁	9a-Hydroxylierung	11,81

Aus der Tabelle ist ersichtlich, dass von den 12 aufgeführten Verbindungen 8 in einer für eine Umsetzung günstigen Orientierung im aktiven Zentrum platziert wurden. Lediglich vier Verbindungen, und zwar das Coffein und das Theophyllin, sowie das 17 β -Estradiol und das Aflatoxin B₁, wurden falsch positioniert. 17 β -Estradiol wurde in der Orientierung eingepasst, in der es im aktiven Zentrum von CYP3A4 vorliegt. Für Coffein, Theophyllin und Aflatoxin schlug das Docking-Programm Platzierungen im oberen Teil des aktiven Zentrums in relativ großer Entfernung zur Hämebene vor, weil dort zahlreiche direkte Wasserstoffbrücken zwischen Tyr 112, Thr 113, Thr 124, Ser 231, Asn 257 und Asn 312 und den Carbonyl- und Etherfunktionen dieser Verbindungen zustande kamen. Vergleicht man diese Beobachtung mit den Ergebnissen der MDS mit Coffein, stellt man fest, dass diese Aminosäuren im Verlauf der MDS vielfach über wasservermittelte Interaktionen an der Substratbindung beteiligt waren. Würde man Wassermoleküle im aktiven Zentrum in das Docking mit einbeziehen, könnte man für diese drei Verbindungen sicher bessere Ergebnisse erzielen. Alle vier Beispiele für O-Desalkylierungsreaktionen wurden im Docking korrekt platziert. Es wurde keine Koordination der Carbonylfunktionen mit dem Hämeisen mehr vorgeschlagen. Eine Überlagerung dieser Verbindungen im aktiven Zentrum ist in Abbildung 4.72 a) dargestellt. Diese Abbildung zeigt darüber hinaus, dass die Orientierung dieser Verbindungen mit der für 7-Ethoxyresorufin im Rahmen der MDS postulierten übereinstimmt. Abbildung 4.75 b) zeigt die vier Beispiele für eine aromatische Hydroxylierung, die von GOLD richtig wiedergegeben wurden.

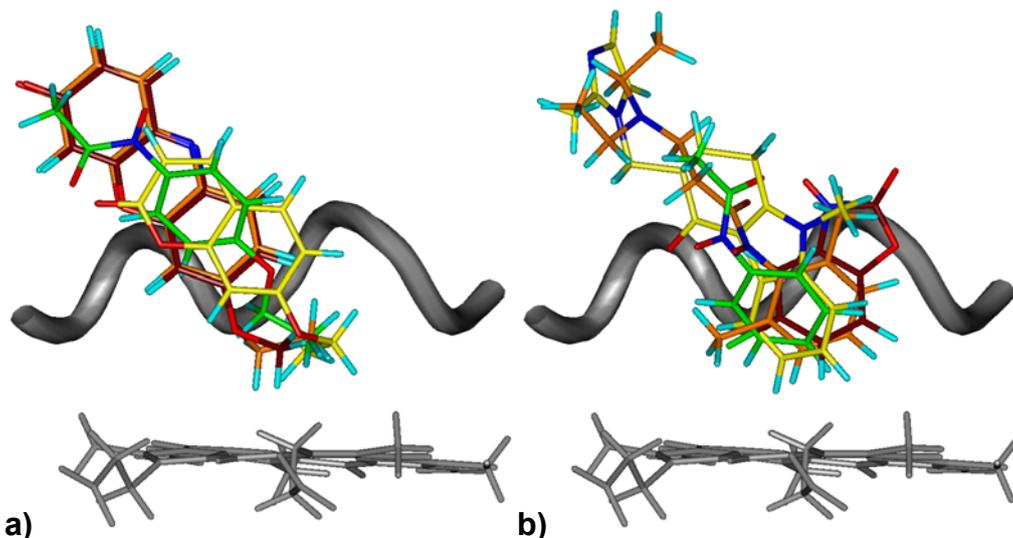


Abbildung 4.75: Überlagerung der Dockinglösungen für den Validierungsdatensatz im aktiven Zentrum der CYP1A2-Struktur. a) Beispiele für O-Desalkylierungsreaktionen: Phenacetin (grün), 7-Ethoxyresorufin (orange), 7-Methoxyresorufin (braun), 7-Ethoxycumarin (gelb). b) Beispiele für aromatische Hydroxylierungen: Acetanilid (grün), Lidocain (orange), Chlorzoxazon (braun), Ondansetron (gelb). Zur Orientierung sind jeweils das Häm und der zentrale Teil der Helix I eingeblendet.

Als Resultat der Validierung des Dockings hat sich ergeben, dass das Problem der Koordination von Carbonylsauerstoffen an das Hämeisen nach der Modifikation der Eisenparameter nicht mehr besteht. Die Vernachlässigung der Donor-Eigenschaften des Eisens scheint sich auch nicht nachteilig auf die Vorhersagekraft des Dockings auszuwirken, da die entscheidenden 8 von 12 getesteten Verbindungen korrekt positioniert werden. Drei der vier nicht korrekt platzierten Substrate enthalten deutlich mehr Wasserstoffbrücken-Akzeptorfunktionen als die Verbindungen des QSAR-Datensatzes und spielen daher für die Validierung nur eine untergeordnete Rolle. Abschließend sollte bemerkt werden, dass Substrate wesentlich spezifischer an ein Enzym binden als Inhibitoren und somit die Vorhersage ihres Bindungsmodus deutlich schwieriger ist. Auch vor diesem Hintergrund kann das Ergebnis des Dockings als gute Validierung betrachtet werden.

4.4.2 Analyse des Inhibitor Datensatzes

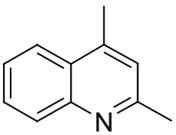
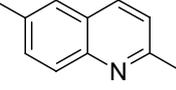
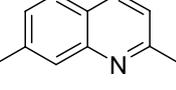
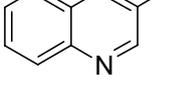
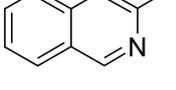
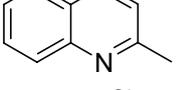
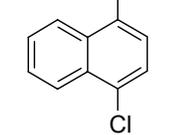
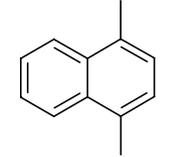
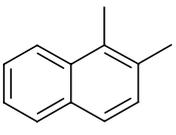
Der für die 3D-QSAR-Studie verwendete Inhibitor Datensatz stammt aus der Arbeitsgruppe von H. Raunio und A. Poso der Universität Kuopio [168]. Er besteht aus insgesamt 44 Naphthalenen, Chinolinen, Lactonen und einigen anderen Verbindungen, die sich keiner dieser Klassen zuordnen lassen. Die biologische Aktivität der Verbindungen am CYP1A2 wurde in Form von IC_{50} -Werten⁹ bestimmt. Dazu wurde gentechnisch hergestelltes humanes CYP1A2 eingesetzt, und es wurde

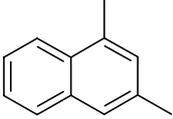
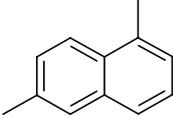
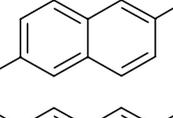
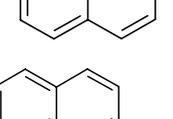
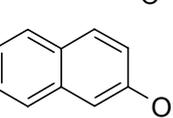
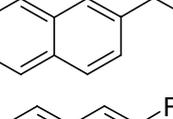
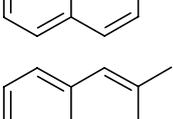
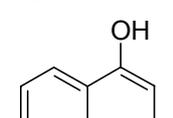
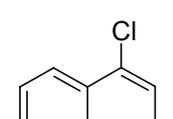
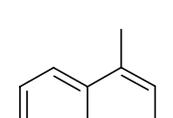
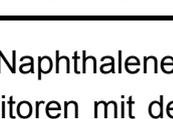
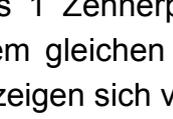
⁹ Inhibitorkonzentration, die erforderlich ist, um die Enzymaktivität *in vitro* auf 50% zu senken; pIC_{50} -Wert: negativer dekadischer Logarithmus des IC_{50} -Wertes

für jeweils sieben unterschiedliche Inhibitorkonzentrationen die Inhibierung der 7-Ethoxyresorufin-O-Desethylierungsreaktion gemessen. Für alle Inhibitoren wurde ein reversibler, kompetitiver Mechanismus festgestellt.

In Tabelle 4.19 und Tabelle 4.20 sind alle Verbindungen des Datensatzes zusammen mit ihrer Strukturformel und der biologischen Aktivität aufgelistet. Die beiden Verbindungen mit der schwächsten Inhibitorwirkung sind das ϵ -Caprolacton und das 4-Methoxyfuran-2(5H)-on mit einem pIC_{50} -Wert $< 1,40$. 1,4-Dichloronaphthalen und 2,4-Dimethylchinolin sind die beiden wirksamsten Verbindungen (pIC_{50} -Wert 5,64 bzw. 5,63). Der Datensatz umfasst folglich einen Wertebereich von etwa 4 Zehnerpotenzen.

Tabelle 4.19: Chinolin- und Naphthalenderivate des Inhibitor Datensatzes

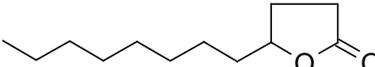
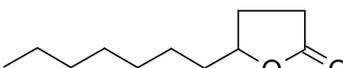
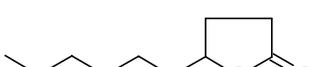
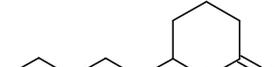
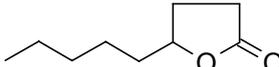
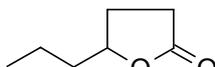
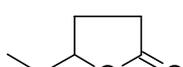
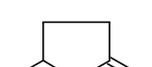
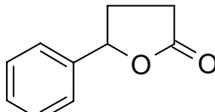
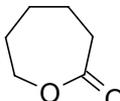
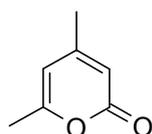
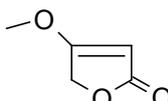
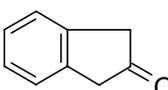
Verbindung	Struktur	pIC_{50} -Wert
2,4-Dimethylchinolin		5,63
2,6-Dimethylchinolin		5,48
2,7-Dimethylchinolin		5,36
3-Methylchinolin		4,88
3-Methylisochinolin		4,67
Chinaldin		4,41
1,4-Dichloronaphthalen		5,64
1,4-Dimethylnaphthalen		5,46
1,2-Dimethylnaphthalen		5,26

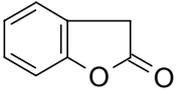
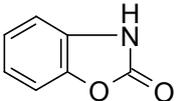
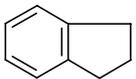
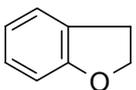
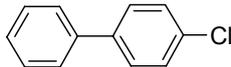
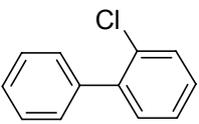
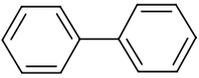
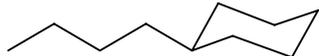
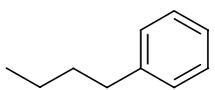
1,3-Dimethylnaphthalen		5,10
1,6-Dimethylnaphthalen		4,59
2,6-Dimethylnaphthalen		4,28
2,7-Dimethylnaphthalen		4,19
2-Methoxynaphthalen		4,90
2-Naphthol		4,78
2-Ethyl-naphthalen		4,71
2-Fluoronaphthalen		4,31
2-Methylnaphthalen		3,92
1-Naphthol		5,49
1-Chloronaphthalen		4,30
1-Methylnaphthalen		3,94

Grundsätzlich stellen die Chinoline und Naphthalene im Vergleich zu den Lactonen und den übrigen Verbindungen die Inhibitoren mit der größeren Wirkstärke dar. Die Chinolinderivate sind jeweils um 0,5 bis 1 Zehnerpotenz stärker wirksam als die korrespondierenden Naphthalene mit dem gleichen Substitutionsmuster. Im Bezug auf die monosubstituierten Naphthalene zeigen sich vor allem das 1- und 2-Naphthol, sowie das 2-Ethyl- und das 2-Methoxynaphthalen als potente Verbindungen. Vergleicht man die methylysubstituierten Derivate untereinander, so stellt man fest, dass disubstituierte Verbindungen mit beiden Methylgruppen am selben Ring stärker

wirksam sind als solche, bei denen sich an jedem Sechsring eine Methylgruppe befindet. Beide Arten disubstituierter Methylnaphthalenen sind stärker wirksam als die monosubstituierten.

Tabelle 4.20: Lactonderivate und andere Verbindungen des Inhibitor Datensatzes

Verbindung	Struktur	pIC ₅₀ -Wert
γ-Dodecanolacton		4,23
γ-Undecanolacton		4,34
δ-Uncanolacton		4,14
γ-Decanolacton		3,96
δ-Decanolacton		3,64
γ-Nonalacton		3,51
γ-Heptalacton		2,35
γ-Caprolacton		2,01
γ-Valerolacton		1,82
γ-Phenyl-γ-butyrolacton		2,64
ε-Caprolacton		<1,40
4,6-Dimethyl-α-pyron		2,35
4-Methoxyfuran-2(5H)-on		<1,40
2-Indanon		4,10

2-Cumaranon		3,58
2-Benzoxazolinon		3,43
Indan		3,26
2,3-Dihydrobenzofuran		2,92
4-Chlorobiphenyl		4,31
2-Chlorobiphenyl		4,30
Biphenyl		3,80
Butylcyclohexan		4,26
Butylbenzol		3,24

Auch von den pIC_{50} -Werten für die Lactone lassen sich gewisse Gesetzmäßigkeiten ableiten. Eine messbare Inhibitoraktivität tritt erst auf, wenn die Lactonringe phenyl- oder alkylsubstituiert sind. Unsubstituierte Lactone wie das ϵ -Caprolacton sind kaum aktiv. In einer Serie alkylsubstituierter γ - und δ -Lactone steigt der pIC_{50} -Wert mit zunehmender Kettenlänge. Zwischen γ -Heptalacton (drei exozyklische Kohlenstoffe) und γ -Nonalacton (fünf exozyklische Kohlenstoffe) fällt ein besonders starker Anstieg um mehr als eine Zehnerpotenz auf. Die optimale Kettenlänge scheint bei 7 exozyklischen Kohlenstoffen (γ -Undecanolacton) erreicht zu sein. Darüber hinaus nimmt die Inhibitorstärke wieder leicht ab.

Bei den übrigen Verbindungen zeigte sich, dass Butylbenzol um eine Zehnerpotenz stärker wirksam ist als Butylcyclohexan, und dass eine Substitution von Biphenyl zu einer Zunahme (4-Chlorobiphenyl) oder einer Abnahme (2-Chlorobiphenyl) der Inhibitoraktivität führen kann, je nachdem an welcher Position sie sich befindet. Bei den fünf untersuchten Inhibitoren mit einem Indangrundgerüst erwies sich eine Carbonylfunktion an Position 2 als günstig für die Inhibitorstärke.

4.4.3 Docking des Datensatzes

Im nächsten Schritt wurden die Inhibitoren des im vorangegangenen Abschnitt vorgestellten Datensatzes mittels Docking unter Verwendung derselben Ein-

stellungen wie für die Validierung (Abschnitt 4.4.1) in der Proteinstruktur überlagert. Die beiden Lactonverbindungen ϵ -Caprolacton und 4-Methoxyfuran-2(5H)-on wurden für das Docking und die anschließende QSAR-Analyse nicht mehr verwendet, da für sie kein exakter pIC_{50} -Wert bestimmt werden konnte. Alle chiralen Verbindungen wurden zunächst in der R- und der S-Konfiguration analysiert. Da kaum Unterschiede in den jeweils vorgeschlagenen Dockinglösungen bestanden, wurden schließlich einheitlich alle chiralen Verbindungen in der S-Konfiguration weiter ausgewertet.

Die Auswahl von geeigneten Inhibitor Konformationen für die anschließende QSAR-Analyse aus den bis zu zehn von GOLD vorgeschlagenen Dockinglösungen erfolgte durch eine visuelle Auswertung. Wie schon im Fall des Validierungsdatensatzes wurde auch für einige Inhibitoren nur jeweils eine Dockinglösung von GOLD gefunden, sodass die einzig mögliche Konformation dieser Verbindungen als Anhaltspunkt für die Auswahl der übrigen benutzt werden konnte.

Abbildung 4.76 zeigt eine Überlagerung aller ausgewählten Inhibitor Konformationen im aktiven Zentrum der Proteinstruktur. An der Inhibitorbindung sind im Wesentlichen die Aminosäuren beteiligt, die auch bei der Bindung der mittels MDS untersuchten Substrate eine Rolle spielten. Tyr 112, Thr 113 und Ser 231 fungieren auch hier als wichtige Wasserstoffbrückendonatoren. Zusätzlich wird in einem Fall eine Wasserstoffbrücke zu Thr 124 gefunden, das bisher nur über lipophile Interaktionen an der Bindung beteiligt war, und dessen Hydroxylfunktion im Verlauf der MDS stets über eine intramolekulare Wasserstoffbrücke mit Asp 313 verbunden war. GOLD ist jedoch in der Lage, die OH-Bindung der Hydroxylfunktionen von Serin, Threonin und Tyrosin als drehbare Bindung zu behandeln, weshalb hier eine Wasserstoffbrücke zum Inhibitor postuliert wird, die den bisherigen Ergebnissen eher widerspricht (vgl. z.B. Abschnitt 4.3.1). Da der Datensatz auch Verbindungen mit Wasserstoffbrückendonatoren enthält, wird im Docking erstmalig für eine Verbindung (1-Naphthol) eine Wasserstoffbrücke zu Asp 320 vorgeschlagen. Im Hinblick auf Substrate ist eine Beteiligung dieses Aspartats an der Bindung unwahrscheinlich, da es eine wichtige Funktion bei der Katalyse innehat. Da Inhibitoren aber lediglich das aktive Zentrum blockieren ohne umgesetzt zu werden, ist eine Beteiligung an der Bindung in diesem Fall durchaus möglich. Abgesehen von Übereinstimmungen in den Wechselwirkungen der Inhibitoren zu Aminosäuren des Proteins verdeutlicht die Überlagerung aber auch die Diversität des Datensatzes.

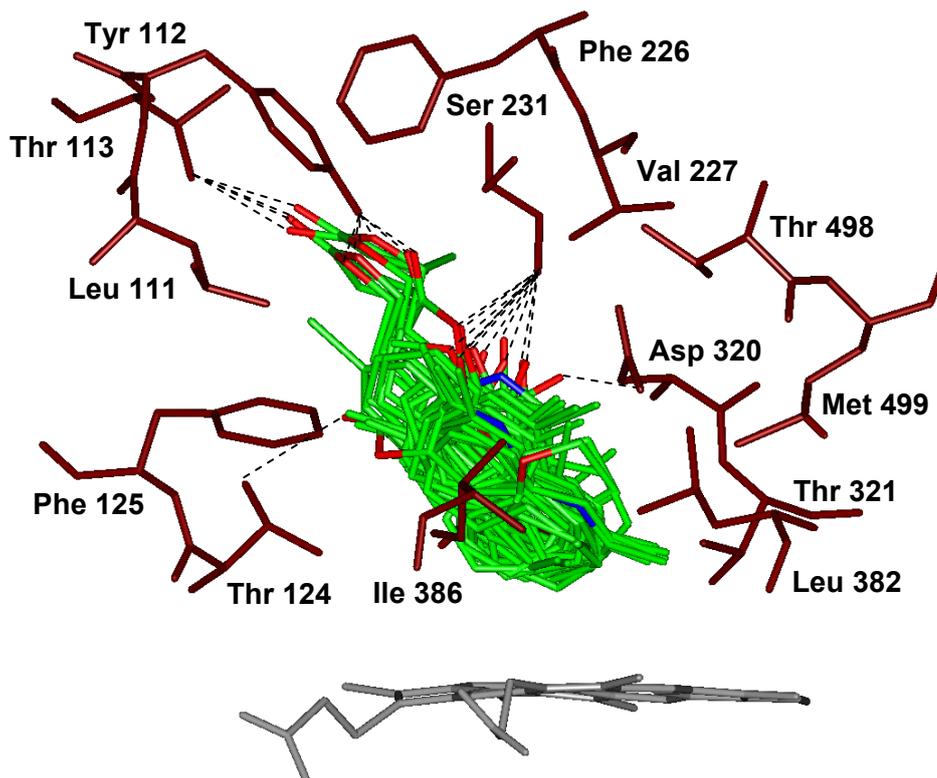


Abbildung 4.76: Überlagerung des Inhibitor Datensatzes in der Proteinstruktur. Die Inhibitoren sind atomkodierte eingefärbt. An der Bindung beteiligte Aminosäuren des Enzyms sind in braun dargestellt. Aus Gründen der Übersichtlichkeit wurden alle Wasserstoffatome ausgeblendet.

4.4.4 3D-QSAR-Modell

Aufgrund der ausgeprägten strukturellen Unterschiede innerhalb des Datensatzes zwischen den Naphthalenen und Chinolinen einerseits und den Lactonen und übrigen Verbindungen andererseits, bestand die Überlegung, die beiden Teile des Datensatzes zunächst in zwei getrennten 3D-QSAR-Modellen zu untersuchen. Für die Lactone und die übrigen Verbindungen (alle in Tabelle 4.20 aufgeführten Verbindungen) ergab sich bei Verwendung der Methylsonde zur Berechnung der GRID-Interaktionsfelder bereits vor der Variablenselektion eine sehr gute Korrelation mit den biologischen Daten ($r^2 = 0,98$; $q^2 = 0,68$; SDEP = 0,44). Dagegen konnte für die Naphthalene und Chinoline (alle in Tabelle 4.19 aufgeführten Verbindungen) mit keiner der getesteten GRID-Sonden eine signifikante Korrelation erzielt werden. Ein Grund dafür ist sicherlich der Datenbereich der pIC_{50} -Werte, der bei den Chinolinen und Naphthalenen nur knapp zwei Zehnerpotenzen umfasst, bei den Lactonen und anderen Verbindungen immerhin fast drei. Darüber hinaus enthalten die Lactonverbindungen aber auch deutlichere Strukturmerkmale, die für eine Unterscheidung auswertbar sind. Es gibt aromatische und nicht aromatische Verbindungen, die Anzahl der Wasserstoffbrückenfunktionen ist unterschiedlich, und auch die Positionierung im aktiven Zentrum variiert. Dahingegen spielen Wasserstoffbrückenfunktionen bei den Naphthalen- und Chinolinderivaten nur eine

untergeordnete Rolle, und die einzelnen Verbindungen unterscheiden sich vielfach nur minimal in der Position einer Methylgruppe. Derartig geringe Strukturunterschiede lassen sich alleine offensichtlich nicht sinnvoll auswerten. Durch die Kombination der beiden Teildatensätze kann aber wieder eine gute Korrelation mit den biologischen Daten erreicht werden, wie im Folgenden ausführlicher beschrieben wird.

Eine Kombination aus der hydrophoben DRY-Sonde und der amidischen Stickstoffsonde N1 (Tabelle 4.21) erwies sich für die vorliegende Problemstellung als am besten geeignet. Dies steht im Einklang mit den strukturellen Eigenschaften der Inhibitoren, die zum größten Teil lipophile Elemente und zur N1-Sonde komplementäre Wasserstoffbrückenakzeptorfunktionen enthalten.

Tabelle 4.21: Parameter der verwendeten GRID-Sonden

spezifische Eigenschaften der molekularen Sonden	DRY-Sonde	N1-Sonde
van-der-Waals-Radius (Å)	1,65	1,65
effektive Anzahl der Elektronen	7	7
Polarisierbarkeit (Å ³)	1,20	1,40
Partiellladung	0	-0,08
optimale H-Brücken-Bindungsenergie (kcal/mol)	-4,50	-4,00
optimaler Radius der H-Brücke (Å)	1,40	1,60
Anzahl donierter H-Brücken	2	1
Anzahl akzeptierter H-Brücken	2	0

Zur Berechnung der Interaktionsfelder der beiden Sonden mit den ausgewählten Inhibitor-konformationen wurden alle Verbindungen in einen GRID-Käfig der Größe 18 x 18 x 24 Å eingesetzt. Bei einer Gitterweite von 1,0 Å ergeben sich für jede Verbindung 9025 Gitterpunkte. Da zwei Sonden eingesetzt wurden, resultieren daraus insgesamt 18050 x-Variablen pro Inhibitor. Bei der Vorbehandlung der Daten in GOLPE wurde die Anzahl der x-Variablen durch Elimination von Variablen ohne Aussagekraft zunächst von 18050 auf 2624 reduziert (Tabelle 4.22).

Tabelle 4.22: Variablenreduktion im Verlauf der 3D-QSAR-Analyse

Reduktionsoperation	Anzahl x-Variablen
vor der Reduktion	18050
Elimination von Variablen ohne Aussagekraft:	
▪ mit Werten $\leq 0,1$	
▪ mit einer Standardabweichung $\leq 0,1$	
▪ die trotz unterschiedlicher Aktivität für 2 oder 3 Substanzen den selben Wert annehmen	2624
<i>FFD selection</i>	1012

Zur weiteren effektiven Reduktion der Variablen auf solche, die für die Korrelation mit der biologischen Aktivität eine Relevanz besitzen, wurden die Variablen mittels *smart*

region definition zuerst in Variablengruppen eingeteilt, und anschließend wurde ein *fractional factorial design (FFD selection)* durchgeführt (Abschnitt 3.6.3.3). So gingen letztlich 1012 aktive x-Variablen in die Aufstellung des 3D-QSAR-Modells ein.

Das statistische Ergebnis der 3D-QSAR-Analyse ist in Tabelle 4.23 zusammengefasst. Unter Einbeziehung der ersten drei Hauptkomponenten ergab sich ein Korrelationskoeffizient r^2 von 0,97 mit einem kreuzvalidierten Korrelationskoeffizienten q^2 von 0,81 und einem SDEP-Wert von 0,43 bei einer Kreuzvalidierung nach der *random groups* Methode (Abschnitt 3.6.3.2).

Tabelle 4.23: Statistisches Ergebnis der 3D-QSAR-Analyse nach Kreuzvalidierung mit der *leave one out* und der *random groups* Methode. PC: Anzahl Hauptkomponenten (*principal components*); r^2 : Korrelationskoeffizient; q^2 : kreuzvalidierter Korrelationskoeffizient; SDEP: Standardabweichung des Fehlers der Vorhersage (*standard deviation for the error of prediction*)

	PC	r^2	q^2	SDEP
leave one out	3	0,97	0,84	0,39
random groups	3	0,97	0,81	0,43

In Abbildung 4.77 ist die kreuzvalidierte Korrelation der im 3D-QSAR-Modell vorhergesagten biologischen Aktivität mit der experimentell gemessenen für alle untersuchten Inhibitoren graphisch dargestellt. Aufgrund der Uneinheitlichkeit des Datensatzes ist eine gewisse Streuung der Werte erkennbar. Am schlechtesten vorhergesagt werden die beiden Verbindungen 1-Naphthol und 2-Chlorobiphenyl.

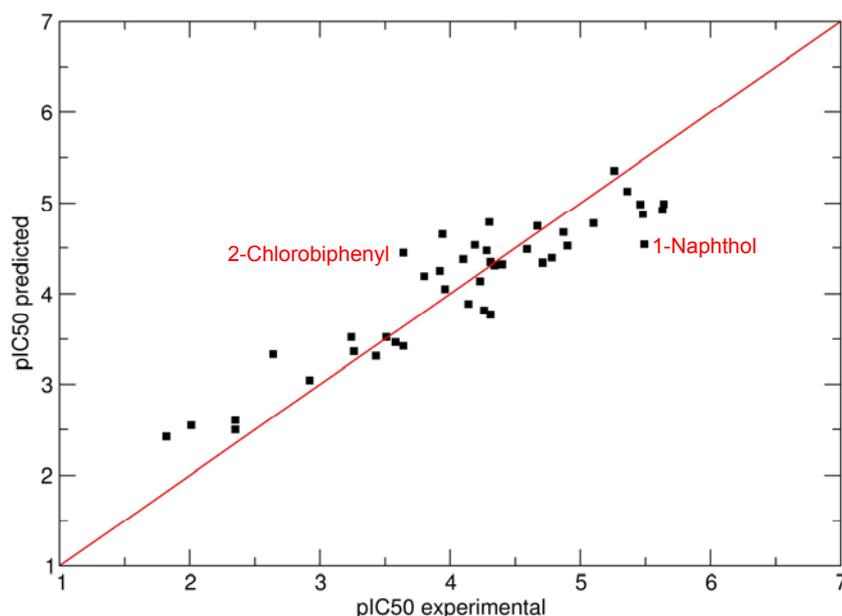


Abbildung 4.77: Kreuzvalidierte Korrelation der im 3D-QSAR-Modell vorhergesagten biologischen Aktivität aller Inhibitoren des Datensatzes (pIC_{50} predicted) mit der gemessenen (pIC_{50} experimental) unter Berücksichtigung der ersten drei Hauptkomponenten

Wie bereits in Abschnitt 4.4.3 erwähnt, sticht 1-Naphthol aus der Überlagerung der Inhibitoren im Protein dadurch hervor, das es als einzige Verbindung des

Datensatzes eine Wasserstoffbrückenbindung zu Asp 320 eingeht. Um hier eine korrekte Vorhersage zu erzielen, müsste der Datensatz mehrere Verbindungen dieses Interaktionstyps enthalten. Verbindungen mit Biphenylgrundgerüst sind ebenfalls mit drei Derivaten im Datensatz nur schwach vertreten. Die Unterschiede im Substitutionsmuster sind hier für eine gute Vorhersage offensichtlich zu gering. Vergleicht man die Abweichung aller vorhergesagten biologischen Aktivitäten mit den gemessenen, stellt man fest, dass die Naphthalen- und Chinolinderivate durchschnittlich eine höhere Abweichung aufweisen. Grundsätzlich wird aus den in Tabelle 4.23 und Abbildung 4.77 vorgestellten Ergebnissen jedoch deutlich, dass ein robustes und aussagekräftiges Modell erstellt werden konnte.

Da im vorliegenden Fall kein externer Datensatz zur Überprüfung der Vorhersagekraft des Modells herangezogen wurde, wurde als weiteres Validierungskriterium ein Scramble-Test durchgeführt. Es wurden fünf verschiedene QSAR-Modelle erstellt, bei denen die pIC_{50} -Werte den Verbindungen des Datensatzes jeweils nach dem Zufallsprinzip zugeordnet wurden. Aus Tabelle 4.24 geht hervor, dass in keinem der fünf Modelle eine Korrelation mit der biologischen Aktivität hergestellt werden konnte. Das zuvor erstellte 3D-QSAR-Modell ist folglich nicht das Resultat einer Zufallskorrelation.

Tabelle 4.24: Ergebnisse des Scramble-Tests

Modell	r^2	q^2 (random groups)	SDEP
1	0,84	-0,52	1,20
2	0,74	-0,58	1,22
3	0,80	-0,78	1,29
4	0,81	-0,40	1,15
5	0,79	-0,34	1,13

5.1 Konstruktion des Homologiemodells

Mit der Kristallstruktur 1N6B des Kaninchencytochroms CYP2C5, das darin als Substratkomplex mit dem Sulfaphenazolderivat DMZ vorliegt, stand für die Konstruktion eines Homologiemodells des humanen CYP1A2 ein gutes Templat zur Verfügung. Grund dafür ist einerseits die enge Verwandtschaft der Enzymfamilien, zu denen Templat und Target gehören, die sich in einer Sequenzidentität von 29 % zwischen CYP2C5 und CYP1A2 äußert. Andererseits sind aber auch die strukturellen Gegebenheiten der Templatstruktur für die vorliegende Fragestellung günstig, die die Analyse des dynamischen Verhaltens von Enzym-Substrat-Komplexen beinhaltet. Die Struktur liegt bereits im Zustand der Substratbindung vor, sodass keine gravierenden Konformationsänderungen mehr zu erwarten sind. Außerdem ist die Struktur abgesehen vom Membrananker vollständig, was eine moleküldynamische Behandlung ohne die Anwendung von Rückhaltekräften erleichtert.

Das für das Homologiemodell zugrunde gelegte Alignment wurde auf der Grundlage eines automatischen multiplen Sequenzalignments und unter Zuhilfenahme einer Sekundärstrukturvorhersage erstellt. Aufgrund der Tatsache, dass in letzter Zeit einige Kristallstrukturen von Säugetiercytochromen veröffentlicht wurden, die immer auch zum Training der auf neuronalen Netzen beruhenden Sekundärstrukturvorhersageprogramme herangezogen wurden, können Lage und Länge von Sekundärstrukturelementen in Säugetiercytochromen inzwischen sehr zuverlässig vorhergesagt werden. Diese Beobachtung resultierte aus einem Vergleich der in der Kristallstruktur 1N6B tatsächlich vorhandenen Sekundärstrukturelemente mit der entsprechenden Vorhersage durch das Programm PSIPRED und gilt auch für das Target CYP1A2. Zur Generierung des multiplen Sequenzalignments standen eine große Anzahl relativ eng verwandter Sequenzen aus den Unterfamilien CYP1 und CYP2 zur Verfügung, sodass ein solides automatisches Alignment erhalten wurde, an dem nur noch wenig manuelle Änderungen vorgenommen werden mussten.

Für die Konstruktion des Homologiemodells stellten sich drei Proteinabschnitte als kritisch heraus. Der Loop zwischen den Helices B' und C ist ein sensibler Bereich, da er teilweise der SRS1 angehört und unmittelbar im Anschluss an die SRS1 ein Fragment per *loop search* eingefügt werden musste. Letztlich erwiesen sich aber sowohl die Korrektur des Alignments als auch die Modellierung des Loops als unproblematisch. Im Loop zwischen den Helices C und D enthält die CYP1A2-Sequenz sieben Aminosäuren mehr als die Kristallstruktur und ist somit der längste Proteinabschnitt, für den ein *loop search* durchgeführt werden musste. Hinzu kommt, dass für diesen Loop die Sekundärstrukturvorhersage aufgrund der Lage der Helices C und D zueinander nicht stimmen konnte und so die Sekundärstruktur des Loops zunächst nicht klar war. Nach einer Struktursuche mit dem FUGUE-Server konnte

jedoch einer Verlängerung der Helix D um etwa zwei Helixwindungen als sehr wahrscheinlich angenommen werden. Ein passendes Fragment dieser Art wurde im *loop search* gefunden und ohne Schwierigkeiten eingefügt. Der Loop CD befindet sich zwar nicht in unmittelbarer Nähe einer SRS, doch sind Aminosäuren aus Helix C an der Hämbindung beteiligt und eine falsche Modellierung dieses langen Fragments würde sich zudem negativ auf die Gesamtstabilität des Modells auswirken. Wiederum in unmittelbarer Nähe zweier SRS ist der Loop FG lokalisiert. Dieser Loop, der mit den beiden Helices F' und G' den Eingang zum Substratkanal verschließt, reicht teilweise so tief ins aktive Zentrum hinein, dass er je nach Substrat sogar selbst an der Substratbindung beteiligt ist. Hier stellte es sich als problematisch heraus, dass der Loop zum einen um zwei Aminosäuren kürzer ist als das Templat, und dass er zum anderen fast durchgehend als helikal vorhergesagt wird, was beides zu einer Verengung des Loops führt. Somit musste hier das Alignment modifiziert und mittels *loop search* ein kurzes Fragment eingefügt werden, ohne dabei die Ausprägung der Nebenhelices F' und G' zu gefährden.

5.2 Validierung des Proteinmodells mittels MDS

Zur ausführlichen Strukturoptimierung und Validierung des CYP1A2-Modells wurde ein MDS-Verfahren angewandt. Durch das Einsetzen des Proteinmodells in eine kochsalzhaltige Wasserbox wurde zunächst eine möglichst physiologische Simulationsumgebung geschaffen. Nach einer kurzen Equilibrierungsdynamik, bei der Teile des Enzymmodells mit *position restraints* fixiert wurden, schloss sich eine freie MDS mit einer langen Equilibrierungsphase von 1,5 ns und einer Auswertungsphase von 1 ns an, aus der 200 Proteinkonformationen in die Auswertung gingen. Da dieses Verfahren deutlich aufwändiger ist als andere Equilibrierungsmethoden, wird es in der Form nur selten angewandt. Dabei ermöglicht erst eine ausführliche Strukturoptimierung mittels MDS in physiologischer Umgebung die Anpassung aller Proteinbereiche inklusive des Proteinrückgrats an die Gegebenheiten der Target-Sequenz. Darüber hinaus stellt das dynamische Verhalten des Proteinmodells während einer freien MDS ein strenges Kriterium für dessen Stabilität und strukturelle Qualität dar. Zwar wurden in letzter Zeit die Vorteile einer Equilibrierung mittels MDS für Homologiemodelle von P450-Isoenzymen grundsätzlich erkannt, es wurden aber meist nur kurze MDS von bis zu 200 ps durchgeführt [169], [170], [171]. Von CYP1A2 dagegen sind bisher ausschließlich Homologiemodelle bekannt, die durch ein mehr oder weniger anspruchsvolles Minimierungsverfahren equilibriert wurden [36], [47], [50], [163].

Bevor das Modell selbst einer MDS unterzogen wurde, wurde zunächst eine MDS mit der als Templat gewählten Kristallstruktur 1N6B durchgeführt, um die Simulationsbedingungen zu validieren und eine Vergleichsbasis für die Auswertung der MDS mit dem Modell zu schaffen. Aus dem Verlauf von RMSD- und RMSF-Wert war

erkennbar, dass die Proteinstruktur während der gesamten MDS stabil blieb und ein logisch nachvollziehbares dynamisches Verhalten zeigte. Ein Beispiel aus der Literatur, in dem die Kristallstruktur 1BU7 eines Bakterien cytochroms einer MDS über 14 ns unter vergleichbaren Simulationsbedingungen unterzogen wurde, wies einen sehr ähnlichen Verlauf des RMSF-Profiles auf [160]. Darüber hinaus blieben Strukturmerkmale wie die Hämbindung und das Wasserstoffbrückenmuster im Zentrum der Helix I bis zum Ende der MDS nach 2,5 ns erhalten. Das gleiche gilt für alle Sekundärstrukturelemente, die sich nach der MDS in Länge und Orientierung nicht wesentlich von der ursprünglich eingesetzten Kristallstruktur unterschieden. Der Faltblattstrang β_{3-1} , der in der Kristallstruktur fehlte, bildete sich in der Lösungsumgebung deutlich aus. Schließlich blieb auch die strukturelle Qualität der Ausgangsstruktur im Wesentlichen erhalten und verschlechterte sich nur geringfügig. Insgesamt ergab die MDS, dass die Simulationsumgebung und die hier gewählten Bedingungen für die moleküldynamische Untersuchung von Cytochromstrukturen gut geeignet sind.

Im Fall des CYP1A2-Modells erbrachte zunächst die Equilibrierungsdynamik eine effektive Verbesserung der Proteinstruktur. Über den Verlauf der freien MDS blieb auch das Proteinmodell durchgehend stabil und zeigte ein gutes dynamisches Verhalten, das vielfach in Übereinstimmung mit den Ergebnissen aus der MDS mit der Kristallstruktur stand. Die Entwicklung des RMSD-Wertes verdeutlichte hier, dass eine lange Equilibrierungsphase von 1,5 ns durchaus sinnvoll ist, da bis zu diesem Simulationszeitpunkt offensichtlich noch Bewegungen im Proteinerückgrat stattfanden. In diesem zeitlichen Verhältnis ist auch die Dauer der Auswertungsphase zu sehen, in der nun die Stabilität des equilibrierten Modells über eine weitere Nanosekunde evaluiert wurde. Trotz einer über weite Strecken der equilibrierten Phase fehlenden Wasserstoffbrücke zu Trp 133 blieb das Häm bis zum Ende der freien MDS fest im aktiven Zentrum gebunden, da alle übrigen nach der Kristallstruktur zu erwartenden Protein-Häm-Interaktionen vorhanden waren. Vor allem die Salzbrücken zu den Argininen 108, 137 und 456 sorgen hier für eine feste Bindung. Als weiteres charakteristisches Strukturmerkmal konnte auch in der MDS mit dem CYP1A2-Modell der Knick in Helix I über dem aktiven Zentrum detektiert werden. Er kam hier über eine direkte Wasserstoffbrücke zwischen Ala 317 und Thr 321 zunächst noch ohne Einlagerung eines Wassermoleküls in die Helixwindung zustande. Erwartungsgemäß war die strukturelle Qualität des CYP1A2-Modells etwas schlechter als die der zugrunde gelegten Kristallstruktur, sie kann aber insgesamt durchaus als gut bezeichnet werden. Die Aminosäuren, die nach dem Ramachandran-Plot in ungünstigen Konformationen vorlagen, befanden sich ausschließlich in Loop-Regionen und am N-Terminus, also in der Peripherie der Enzymstruktur. Bis auf eine Ausnahme ließen sich alle Sekundärstrukturelemente der Säugetiercytochrome im Proteinmodell wieder finden. Veränderungen in Länge und Orientierung gegenüber der Kristallstruktur standen größtenteils in Übereinstimmung mit den Sekundär-

strukturvorhersagen für CYP1A2. Wie sich bereits bei der Konstruktion des Modells andeutete, war der Loop zwischen den Helices F und F' der einzig problematische Bereich, da Helix F' nicht immer deutlich als Helix ausgeprägt war. Im Verlauf der MDS mit Substratkomplexen des CYP1A2-Modells zeigte sich jedoch, dass ein Substrat diesen Loopbereich stabilisiert und dessen strukturelle Qualität verbessert.

Insgesamt konnte somit ein stabiles Proteinmodell von CYP1A2 erstellt werden, das eine gute strukturelle Qualität aufwies und in Einklang mit experimentell bekannten Strukturmerkmalen stand.

5.3 Berechnung von Substratkomplexen

Neben detaillierten MDS-Studien an Kristallstrukturen zur Untersuchung von Substratein- und -austritt [172] findet man in der Literatur auch einige Beispiele für kurze MDS von 50 bis 200 ps mit Substratkomplexen von Homologiemodellen [169], [171], die zur Identifizierung von an der Bindung beteiligter Aminosäuren und zur Erklärung von Bindungsaffinitäten herangezogen wurden. Auch in der vorliegenden Arbeit wurden zur weiteren Modellvalidierung und zur Untersuchung von Enzym-Substrat-Interaktionen und Regioselektivitäten in der Umsetzung wiederum MDS mit drei ausgewählten Substraten durchgeführt. Abstand und Orientierung des Substrates zum Häm, sowie das Auftreten und die Stabilität von Protein-Substrat-Interaktionen sind aber vor allem dann als Kriterium für die richtige Positionierung und die Qualität des Modells von Bedeutung, wenn ihre Beständigkeit über einen längeren Simulationszeitraum beobachtet werden kann. Daher wurde hier der Simulationszeitraum auf eine Dauer von insgesamt 5 ns festgesetzt.

Während bei den oben erwähnten Studien ausschließlich in wässriger Umgebung gearbeitet wurde, ist von einem CYP1A2-Homologiemodell bisher nur eine moleküldynamische Untersuchung bekannt, in der MDS über 100 ps im Vakuum durchgeführt wurden, bei denen Aminosäuren im Radius von 12 Å um das Substrat frei beweglich waren, und das restliche Protein fixiert blieb [50]. Der Einfluss von Lösungsmittelleffekten auf die Proteinkonformation vor allem an der Proteinoberfläche und in den Loopbereichen ist wohlbekannt [112], doch spätestens seit der Publikation der als Substratkomplex vorliegenden Kristallstrukturen ist auch die wichtige Rolle von Wassermolekülen im aktiven Zentrum bei der Substratbindung offensichtlich geworden. Daher wurden hier trotz des erheblich höheren Rechenaufwandes sämtliche MDS in einer Wasserbox durchgeführt.

In Analogie zum Verfahren bei der Berechnung des leeren Enzyms wurden auch hier zunächst Validierungsdynamiken mit Kristallstrukturen durchgeführt, um zu testen, inwiefern die Substratbindung im Rahmen einer MDS wiedergeben kann und wo das Verfahren an seine Grenzen stößt.

In der Kristallstruktur des Templats 1N6B sind fast ausschließlich lipophile Interaktionen für die Substratbindung verantwortlich. Zusätzlich deutet die Orientierung von Asn 204 aus SRS2 darauf hin, dass eine schwache Interaktion zwischen einem Wasserstoff der Amidfunktion dieses Asparagins und dem π -Elektronensystem des Phenylrings von DMZ an der Substratbindung beteiligt sein könnte. Im Verlauf einer MDS mit der Kristallstruktur konnte dieser Bindungsmodus nicht erhalten werden. Stattdessen bildete sich eine wasservermittelte Wasserstoffbrücke zwischen Asn 204 und dem Pyrazolstickstoff des Substrats DMZ aus, die bis zum Ende der MDS stabil blieb. Eine mögliche Erklärung für die Berechnung eines falschen Bindungsmodus zwischen Substrat und Enzym könnte sein, dass die lipophilen Interaktionen im Kraftfeld gegenüber hydrophilen Interaktionen unterbewertet werden. Durch diese Unterbewertung kam es während der MDS zum Eindringen von Wasser in den Teil des aktiven Zentrums, in dem die hydrophilen Funktionen des DMZ lokalisiert sind, und es wurde artifiziell eine stabile Wasserstoffbrücke zwischen Substrat und Enzym ausgebildet.

Als Alternative zu 1N6B wurde stattdessen 1NR6 zur Validierung eingesetzt, ein zweiter Substratkomplex desselben Enzyms mit Diclofenac im aktiven Zentrum. Mit der moleküldynamischen Bestätigung des Protonierungszustandes von Diclofenac bei der Bindung an CYP2C5 konnte zunächst ein Ergebnis erzielt werden, das die grundsätzliche Zuverlässigkeit der MDS bei der Beschreibung der Gegebenheiten des aktiven Zentrums bewies. Die Auswertung des RMSD-Wertes als Stabilitätskriterium ergab, dass im Fall einer MDS über längere Simulationszeiträume in den Außenbereichen des Proteins nochmals Fluktuationen stattfanden, während der Proteinkern vollständig equilibriert war. Diese Beobachtung steht in Einklang mit den Ergebnissen anderer MDS-Studien, bei denen die Bakterien cytochromstruktur 1BU7 und andere Kristallstrukturen moleküldynamisch untersucht wurden [159], [160]. Grund dafür sind instabile Strukturbereiche in der Peripherie der Proteinstruktur, die in der vorliegenden Arbeit mithilfe des Profils der RMS-Fluktuation und anhand einer Überlagerung von Proteinkonformationen aus dem Verlauf der MDS lokalisiert werden konnten. So waren es in diesem Fall vor allem der N-Terminus und dessen unmittelbare Umgebung, die die Fluktuation verursachten, was auf den fehlenden Membrananker zurückzuführen ist. Der Verlauf des RMSF-Wertes verdeutlichte aber ebenso, dass die inneren Proteinbereiche mit dem aktiven Zentrum und den Hämbinderegionen von diesen Fluktuationen nicht betroffen waren, sodass man feststellen kann, dass die Kristallstruktur insgesamt über einen Simulationszeitraum von 5 ns stabil geblieben ist und Protein-Substrat-Interaktionen ausgewertet werden konnten. Als Maß für die korrekte Orientierung des Substrates im aktiven Zentrum wurde der Abstand der Hydroxylierungsstelle zum Hämeisen über den Verlauf der MDS verfolgt. Hier zeigte sich, dass die Substratorientierung aus der Kristallstruktur während des gesamten Simulationszeitraums stabil blieb. Es stellte sich aber auch heraus, dass der absolute Abstand der Hydroxylierungsstelle zum Eisen mit einer

gewissen Ungenauigkeit behaftet ist. Eine genaue Berechnung dieses Wertes ist nur mit quantenchemischen Methoden möglich, bei denen der genaue elektronische Zustand des komplexen Häm-systems und das elektrische Feld der Häm-umgebung berücksichtigt werden können. Neben dem Abstand der Hydroxylierungsstelle zum Hämeisen wurde auch der Bindungsmodus des Substrates qualitativ richtig wiedergegeben. Wie in der Kristallstruktur war die Säurefunktion des Diclofenac über ein Wassercluster an die Aminosäuren Asn 204, Lys 241, Ser 289 und Asp 290 des aktiven Zentrums gebunden. Ein deutlicher Unterschied zur Kristallstruktur bestand allerdings darin, dass auch eine direkte Wasserstoffbrücke zwischen der Säurefunktion des Diclofenac und Asp 290 in Helix I zustande kam. Das führte dazu, dass die Position des Substrates sich entlang der Helix in Richtung dieses Aspartats verschob und dass auch die Konformation des aktiven Zentrums sich aufgrund einer Anpassung an die neue Substratposition leicht veränderte. Insgesamt blieb die Architektur des aktiven Zentrums jedoch erhalten. Als Grund für das Auftreten der artifiziellen direkten Wasserstoffbrücke zwischen Asp 290 und der Säurefunktion des Diclofenac könnte man vermuten, dass eventuell nicht nur das Substrat Diclofenac, sondern auch das in unmittelbarer Nähe der Säurefunktion lokalisierte Asp 290 protoniert vorliegt. Dann wäre die Attraktivität des Carboxylatsauerstoffes von Asp 290 für die Säurefunktion von Diclofenac erheblich geringer. Für diese Vermutung gibt es allerdings keinerlei experimentelle Anhaltspunkte. Darüber hinaus ist es unwahrscheinlich, dass Diclofenac gleichzeitig zwei Kontakte zu Wasserstoffbrücken donierenden Bindungspartnern ausbilden kann, nämlich einerseits zu Lys 241 und andererseits zur protonierten Säurefunktion von Asp 290. Ein weiterer möglicher Grund für das Auftreten der direkten Wasserstoffbrücke ist eine Überbewertung dieser elektrostatischen Interaktion im Gegensatz zu den übrigen Kräften, die das Substrat im aktiven Zentrum festhalten. Da die Parametrisierung der Aminosäuren standardmäßig festgelegt und validiert ist, kann im Hinblick auf die Kraftfeldparameter hier nur der Parametersatz des Substrats für eine unzulängliche Beschreibung der Interaktion verantwortlich sein. Zieht man jedoch den pK_S -Wert des Diclofenac von 4,6 in Betracht, wird deutlich, dass es sich hier weder um eine besonders starke, noch um eine besonders schwache Säure handelt, sodass die Partialladungen der Aminosäuren, die zur Beschreibung der Säurefunktion übernommen wurden, eigentlich gut für die Wiedergabe geeignet sein müssten. Darüber hinaus ist eine grundsätzliche Überbewertung direkter elektrostatischer Interaktionen zu Lasten von wasservermittelten Interaktionen aufgrund des Kraftfeldes denkbar. Diese These wird dadurch gestützt, dass im Verlauf der vorliegenden Arbeit mehrfach ein ähnliches Phänomen beobachtet wurde. Nach der MDS der leeren Kristallstruktur befand sich in der irregulären Helixwindung von Helix I nur noch ein Wassermolekül anstatt wie in der Startstruktur zwei. Im Verlauf der MDS des leeren Modells war dort überhaupt kein Wassermolekül mehr zu finden. Schließlich wird in der Kristallstruktur von CYP2C5 mit Diclofenac eine intramolekulare Wasserstoffbrücke der NH-Gruppe mit der Säurefunktion des

Diclofenac ebenfalls über ein Wassermolekül vermittelt, die im Verlauf der MDS in eine direkte Wasserstoffbrücke übergang. Aufgrund des mehrfachen Auftretens eines Übergangs von einem wasservermittelten Kontakt in einen direkten, kann somit dieser letzte Erklärungsansatz als der plausibelste angesehen werden. Er würde außerdem in Einklang mit dem Ergebnis der vorangegangenen MDS mit der Kristallstruktur 1N6B stehen, bei der es zu einer Überbewertung elektrostatischer Interaktionen gegenüber lipophilen Wechselwirkungen kam.

5.4 Substratkomplexe mit MelQ und 7-Ethoxyresorufin

Mithilfe von SURFNET-Kavitäten und GRID-Interaktionsfeldern konnten die beiden Substrate MelQ und 7-Ethoxyresorufin problemlos und eindeutig im aktiven Zentrum des CYP1A2-Modells positioniert werden. Aus der Kombination der Interpretation der GRID-Interaktionsfelder mit dem Ergebnis einer Mutationsstudie wurde abgeleitet, dass MelQ deprotoniert im aktiven Zentrum von CYP1A2 vorliegen muss.

In Analogie zur MDS mit der Kristallstruktur 1NR6 erwies sich auch bei der Berechnung der Substratkomplexe des CYP1A2-Modells der RMSD-Wert des Proteinkerns als aussagekräftigeres Equilibrierungskriterium, da auch hier der RMSD-Wert des gesamten Proteinrückgrats teilweise noch eine leichte Steigung aufwies. Der innere Bereich des Proteins, der das aktive Zentrum und die Hämbinderegionen umfasst, war jedoch vollständig equilibriert und blieb bis zum Ende der MDS stabil. Der Verlauf des RMSD-Wertes für das Proteinrückgrat des Proteinkerns rechtfertigt auch in diesem Fall die lange Simulationszeit von 5 ns. Sowohl während der MDS mit Ethoxyresorufin als auch während der MDS mit MelQ stieg selbst der RMSD-Wert des Proteinkerns bis zum Simulationszeitpunkt von 3000 ps noch deutlich an, was darauf zurückzuführen ist, dass sich die Konformation des aktiven Zentrums erst an Größe und Lage des jeweiligen Substrats anpassen muss. So konnte erst ab 3000 ps die Auswertungsphase angeschlossen werden, die in einem angemessenen zeitlichen Verhältnis zur Equilibrierungsphase nochmals 2000 ps umfasste. Bei beiden MDS konnten der C-Terminus und der FG-Loop als die Proteinbereiche identifiziert werden, die im Vergleich zur MDS mit dem leeren Proteinmodell für die erhöhte Fluktuation des RMSD-Wertes verantwortlich sind.

Im Jahr 2000 wurden unabhängig voneinander zwei Docking-Studien mit Homologie-Modellen von CYP1A2 veröffentlicht, die auf Bakterien Cytochromen als Template-Strukturen basierten und unter anderem MelQ als Substrat verwendeten. In der einen Veröffentlichung wurde eine Substratorientierung spiegelbildlich zu der hier gewählten vorgeschlagen, und es wurden außerdem Wasserstoffbrücken zwischen den Stickstoffen der Guanidinfunktion und einem als sechster Hämligand koordinierten Wassermolekül vorgeschlagen [36]. Die zweite Publikation ging von derselben Substratorientierung aus wie die vorliegende Arbeit, postulierte aber eine

Wasserstoffbrücke zwischen dem Chinolinstickstoff von MelQ und Asn 222 aus SRS2 und nicht zu Tyr 112 aus SRS1 [47]. Wie bereits ausführlich erläutert, wurde in dieser Arbeit aufgrund von experimentellen Anhaltspunkten angenommen, dass die sechste Koordinationsstelle des Eisens bei der Substratbindung unbesetzt ist. Wirklich eindeutig ist die Datenlage, was diesen Punkt betrifft, allerdings nicht. Für die in dieser Arbeit gewählte Substratorientierung spricht neben der SURFNET-Kavität und den GRID-Interaktionsfeldern vor allem die Tatsache, dass sowohl das Protein als auch die Position des MelQ im aktiven Zentrum, die durch den Abstand der Hydroxylierungsstelle zum Hämeisen und Wasserstoffbrücken zwischen Protein und Substrat charakterisiert wurde, während der gesamten MDS stabil blieb. Der Abstand zwischen Eisen und exozyklischem Guanidinstickstoff betrug während der gesamten equilibrierten Phase der MDS konstant etwa 0,42 nm, und auch die Wasserstoffbrücke zu Tyr 112 blieb nahezu durchgehend bestehen. Einen experimentellen Hinweis auf die Existenz einer Wasserstoffbrücke zum Chinolinstickstoff des MelQ gibt es weder für Asn 222 noch für Tyr 112. Beide Aminosäuren wurden nicht im Rahmen von Mutationsstudien untersucht. Im vorliegenden Proteinmodell wird der Zugang von Asn 222 zum aktiven Zentrum durch die quer zur Helix F verlaufende Helix G verhindert. Es sollte vielleicht an dieser Stelle angemerkt werden, dass in der vorliegenden Arbeit mit dem Kaninchencytochrom CYP2C5 ein Templat mit einer deutlich höheren Homologie zu CYP1A2 verwendet werden konnte, während die Bakterien- und Säugetiercytochrome sich in ihrer Topologie doch deutlich von den Säugetiercytochromen unterscheiden.

Das Substrat 7-Ethoxyresorufin betreffend wurde diesmal im Rahmen der Docking-Studie von De Rienzo et al. eine Positionierung mit spiegelbildlicher Anordnung im Vergleich zur in dieser Arbeit gewählten Substratposition vorgeschlagen. Die Autoren postulierten Wasserstoffbrücken zwischen dem Carbonylsauerstoff des Ethoxyresorufins und dem Proteinrückgrat von Ile 494 sowie der Aminosäureseitenkette von Tyr 495 in SRS6. Mithilfe einer MDS des Substratkomplexes von CYP1A2 mit 7-Ethoxyresorufin konnte in der vorliegenden Arbeit wiederum ein wichtiger Anhaltspunkt für die Richtigkeit der hier gewählten Positionierung des Ethoxyresorufin gewonnen werden. Das Substrat verblieb mit einem Abstand der Hydroxylierungsstelle zum Hämeisen von etwa 0,47 nm über den gesamten Verlauf der MDS konstant in einer für eine Umsetzung geeigneten Orientierung und auch die Proteinstruktur blieb stabil. Wie im Übrigen auch für den Substratkomplex mit MelQ konnten für viele lipophile und aromatische Interaktionen des Ethoxyresorufins mit Aminosäuren des CYP1A2-Modells Übereinstimmungen mit Mutationsstudien gefunden werden. Einen experimentellen Hinweis auf die Substratpositionierung liefert dabei die Tatsache, dass eine Mutation von Thr 223 eine Auswirkung auf die Umsetzung des Ethoxyresorufins hat. Im CYP1A2-Modell befindet sich diese Aminosäure so tief in der Kavität zwischen den Helices F und G, dass hier nur eine Interaktion zustande kommen kann, wenn das Ringsystem des Ethoxyresorufins wie

hier postuliert in diese Tasche hineinragt und nicht wie in der alternativen Positionierung in Richtung SRS6 weist.

In der bereits oben erwähnten Publikation führten Liu et al. mit einem CYP1A2-Homologiemodell kurze MDS im Vakuum durch, um ihre eigenen Mutationsstudien moleküldynamisch nachzuvollziehen [50]. Das Modell basierte wie die vorliegende Arbeit auf der Kaninchenkristallstruktur CYP2C5 und blieb während der MDS teilweise fixiert. Substrat war hier ebenfalls das 7-Ethoxyresorufin. Für einige Mutanten stellten sie im Vergleich zum Wildtyp Unterschiede im Abstand der Hydroxylierungsstelle zum Hämeisen fest, die eine Erklärung für die unterschiedlichen Umsetzungsraten lieferten. Nicht nachvollziehen konnten sie jedoch den drastischen Aktivitätsverlust der Mutation Asn 312 Leu, da Asn 312 während der MDS intramolekulare Wasserstoffbrücken zu Ser 122 und Leu 310 ausbildete. Hier zeigten sich in der vorliegenden Arbeit die Vorteile einer freien MDS des gesamten Proteins in einer Lösungsumgebung. Es konnte eine Wasserstoffbrücke zwischen Asn 312 und dem Carbonylsauerstoff des Ethoxyresorufins detektiert werden, die den Effekt der Mutation erklärt. Die Wasserstoffbrücke blieb zwar nicht durchgehend während der equilibrierten Phase bestehen, trat aber über den gesamten Verlauf der Auswertungsphase mehrfach auf.

Vor diesem Hintergrund wäre es aber auch als Fortführung der vorliegenden Arbeit interessant, die Bedeutung der Mutationen nicht nur theoretisch herzuleiten, sondern die Auswirkungen auf die Substratbindung exemplarisch für einige Beispiele moleküldynamisch zu berechnen.

5.5 Substratkomplexe mit Coffein

Für Coffein in der Orientierung zur Bildung des Hauptmetaboliten Paraxanthin ergab sich ebenfalls aus SURFNET-Kavität und GRID-Interaktionsfeldern eine Positionierung, die während einer MDS über 5 ns stabil blieb. Die Hydroxylierungsstelle befand sich in der equilibrierten Phase konstant etwa 0,39 nm vom Hämeisen entfernt. Aus der Literatur sind wiederum zwei Docking-Untersuchungen bekannt, bei denen das Homologiemodell von CYP1A2 auf Bakteriencytochromen basierte. Hier wurde wieder eine spiegelbildliche Anordnung des Coffeins im aktiven Zentrum vorgeschlagen [36], [47]. Lozano et al. platzierten für das Docking Wassermoleküle an den beiden Carbonylsauerstoffen und dem nichtmethylierten Imidazolstickstoff des Coffeins und berücksichtigten so wasservermittelte Interaktionen im aktiven Zentrum. Auch im Verlauf der im Rahmen dieser Arbeit durchgeführten MDS mit Coffein spielten Wassermoleküle eine wichtige Rolle für die Bindung des Coffeins. Ein Teil des aktiven Zentrums war mit Wasser gefüllt, da das Substrat nicht alle Bereiche ausfüllte, und teilweise bestand ein Wasserstoffbrückenkontakt des Wasserclusters über Ser 231 zum Carbonylsauerstoff am C6 des Coffeins. Vor allem

am Anfang und am Ende der equilibrierten Phase der MDS existierte dagegen eine direkte Wasserstoffbrücke zwischen dem Carbonylsauerstoff an C6 des Coffeins und Thr 498 aus SRS6. Leider gibt es weder für diese Wasserstoffbrücke noch für die Beteiligung von Ser 231 an der Substratbindung einen experimentellen Hinweis aus Mutationsstudien. Dies ist vor allem vor dem Hintergrund kritisch zu betrachten, dass Ser 231 Bestandteil eines Loops ist, der per *loop search* in das Modell eingefügt wurde, sodass die Konformation dieses Bereiches wenig gesichert ist. Aus einem Vergleich mit der als Templat zugrunde gelegten Kristallstruktur ergibt sich jedoch, dass sowohl die dem Thr 498 entsprechende Aminosäure Phe 473 als auch Aminosäuren aus dem Loop FF' an der Bindung von Diclofenac und DMZ beteiligt sind. Im Loop FF' sind dies Leu 208 und Leu 213. Ein exakter Vergleich von Atompositionen ist hier nicht möglich, da der Loop im Templat um zwei Aminosäuren länger ist als im Target CYP1A2.

Um eine Erklärung für die bevorzugte Umsetzung des Coffeins zum Paraxanthin zu finden, wurde zum Vergleich eine zweite MDS mit Coffein in der Position zur Umsetzung zum Theobromin durchgeführt. Während der Auswertungsphase dieser MDS deutete sich eine Tendenz des Coffeins an, sich in die Orientierung zur Bildung des Hauptmetaboliten Paraxanthin zu drehen. Dies äußerte sich zum einen darin, dass der Verlauf des RMSD- und des RMSF-Wertes auch im Proteinkern noch Bewegungen erkennen ließen, und zum anderen in einer Veränderung der Enzym-Substrat-Interaktionen. Beispielsweise verkleinerte sich der Abstand zwischen der N3-Methylgruppe und dem Hämeisen zusehends gegen Ende der MDS. Das Wasserstoffbrückenmuster veränderte sich dahingehend, dass ein Kontakt zwischen Thr 498 in SRS6 und der C6-Carbonylgruppe des Coffeins zustande kam, wie er auch in der ersten MDS mit Coffein beobachtet wurde. Zusätzlich kam es zur Ausbildung einer Polarisationswechselwirkung zwischen Phe 125 aus SRS1 und dem Imidazolring des Coffeins, ebenfalls in Analogie zur ersten MDS mit Coffein. Der Grund dafür, dass die Orientierung des Coffeins zur Bildung von Theobromin letztlich nicht stabil blieb, liegt in der großen Differenz zwischen den Umsetzungsraten. Bei der Metabolisierung von Coffein wird zu 84 % das Paraxanthin gebildet, das Theobromin hingegen nur zu 12 %. Diese Tatsache konnte damit anhand der MDS gut wiedergegeben werden. Es wird aber hierdurch auch deutlich, dass MDS mit einer Substratorientierung zur Bildung des Theophyllins (4 %) oder der 1,3,7-Tri-methylharnsäure (1 %) aufgrund der noch niedrigeren Entstehungsraten kein sinnvolles Ergebnis liefern können, weshalb auf die Durchführung hier verzichtet wurde. Der hier postulierte Bindungsmodus des Coffeins und die damit in Zusammenhang stehende Begründung für die bevorzugte Bildung des Paraxanthins werden dadurch bekräftigt, dass das Phänomen der Drehung von Coffein im aktiven Zentrum nicht nur im Verlauf dieser zweiten MDS beobachtet wurde, sondern bereits auch in der Anfangsphase der ersten MDS mit Coffein.

5.6 3D-QSAR-Studie

Abschließend wurde das CYP1A2-Modell im Rahmen einer 3D-QSAR-Studie für die Überlagerung eines Datensatzes in einer Proteinstruktur angewendet. Da die Parametrisierung des Eisens im für die Überlagerung ausgewählten Docking-Programm GOLD zu einer Überbewertung von Eisen-Sauerstoff-Interaktionen führte, wurde das Hämeisen als *dummy atom* definiert, und dem Atomtyp *dummy atom* wurden die van-der-Waals-Parameter des Eisens zugewiesen. Zur Validierung dieser Modifikation wurde zunächst ein Docking mit einem aus 12 Substraten bestehenden Testdatensatz durchgeführt. Bei dieser Validierung gelang es, 8 von den 12 Verbindungen in einer für die Umsetzung günstigen Orientierung korrekt im aktiven Zentrum zu platzieren. Lediglich vier zumeist hydrophile Verbindungen, die im Hinblick auf die Eigenschaften des QSAR-Datensatzes nur von untergeordneter Bedeutung waren, wurden falsch platziert. Zu artifiziellen Eisen-Sauerstoff-Koordinationen kam es nicht mehr. Hieraus ist ersichtlich, dass eine Vernachlässigung der elektrostatischen Wechselwirkungen des Hämeisens sich nicht negativ auf die Vorhersage des Bindungsmodus von Substraten auswirkt, sondern der Realität näher kommt als eine artifizielle Überbewertung. Abgesehen vom Ergebnis des Validierungs-Dockings lässt sich die Parametermodifikation damit rechtfertigen, dass das Hämringssystem im Vergleich zum Eisen die weitaus wichtigere Rolle bei der Substratbindung spielt. Denn erst aufgrund der hohen Lipophilie des Hämringssystems können lipophile Molekülbereiche wie z.B. aliphatische Substituenten überhaupt für eine Hydroxylierung in der Nähe des Eisens gebunden werden.

Für die eigentliche 3D-QSAR-Studie wurde ein Inhibitor-datensatz, der sich aus 42 Chinolinen, Naphthalenen, Lactonen und einigen anderen Verbindungen zusammensetzte, mit GOLD im Protein überlagert. Bei der Analyse der Protein-Inhibitor-Komplexe zeigte sich, dass an der Inhibitorbindung im Wesentlichen dieselben Aminosäuren beteiligt waren wie an der Bindung der Substrate im Verlauf der MDS. Zur Berechnung der GRID-Interaktionsfelder für die Korrelation mit der biologischen Aktivität eignete sich eine Kombination aus der lipophilen DRY-Sonde und der amidischen Stickstoffsonde N1 am besten. Nach einer effektiven Variablenreduktionsprozedur in GOLPE resultierte ein statistisch signifikantes und aussagekräftiges Modell mit einem Korrelationskoeffizient r^2 von 0,97, einem kreuzvalidierten Korrelationskoeffizienten q^2 von 0,81 und einem SDEP-Wert von 0,43 bei einer Kreuzvalidierung nach der *random groups* Methode. Im Rahmen eines Scramble-Tests zur zusätzlichen Validierung erstellte PLS-Modelle ergaben keine Korrelation mit der biologischen Aktivität.

Korhonen et al. erweiterten den hier verwendeten Datensatz auf insgesamt 46 Verbindungen und stellten auf dieser Grundlage ebenfalls mit der GRID/GOLPE-Methode ein ligandbasiertes 3D-QSAR-Modell auf. Das von ihnen erstellte PLS-

Modell wies mit einem Korrelationskoeffizient r^2 von 0,90, einem kreuzvalidierten Korrelationskoeffizienten q^2 von 0,79 und einem SDEP-Wert von 0,52 eine dem im Rahmen dieser Arbeit erstellten 3D-QSAR-Modell vergleichbare Qualität auf. Dass die beiden auf der Grundlage einer unterschiedlichen Überlagerung erstellten QSAR-Modelle darüber hinaus nicht vergleichbar sind, zeigt die Auswahl unterschiedlicher GRID-Sonden, die jeweils für die Beschreibung der Unterschiede in der biologischen Aktivität herangezogen wurden. Während in der vorliegenden Arbeit eine Kombination aus DRY- und N1-Sonde verwendet wurde, benutzten Korhonen et al. nur eine einzige Sonde, und zwar die Hydroxylsonde für die Korrelation.

Zusammenfassend kann man sagen, dass es gelungen ist, auf der Grundlage des im Rahmen dieser Arbeit erstellten Proteinmodells von CYP1A2 einen interessanten Einblick in das dynamische Verhalten von Substratkomplexen zu gewinnen und den Bindungsmodus von Substraten mittels MDS wiederzugeben. Letzteres ermöglicht die Untersuchung weiterer bekannter Substrate und die Erkennung von Strukturmerkmalen, die für die Affinität dieser Verbindungen zum CYP1A2 verantwortlich sind. Auf diese Weise kann auch das Bindungsverhalten potentieller neuer Wirkstoffe moleküldynamisch ermittelt werden, was zur Beurteilung ihres Interaktionspotentials mit anderen CYP1A2-Substraten oder -Inhibitoren beitragen könnte.

Die Validierung des Dockings mit einem Substratdatensatz hat darüber hinaus gezeigt, dass CYP1A2-Substrate auch mithilfe eines automatischen Docking-Verfahrens grundsätzlich zuverlässig im aktiven Zentrum des Enzymmodells positioniert werden können. Dies eröffnet prinzipiell die Möglichkeit, das Modell für eine schnelle Vorhersage von Bindungsaffinitäten neuer potentieller Wirkstoffe zum CYP1A2 einzusetzen, was wiederum einen wertvollen Beitrag zur Abschätzung von Arzneimittelinteraktionen liefern könnte. Hier wären allerdings zunächst noch weitere umfassende Untersuchungen nötig, beispielsweise im Hinblick auf eine Ermittlung von Wiederfindungsraten oder eine Evaluierung verschiedener Scoring-Funktionen.

Schließlich bietet das auf der Basis des Inhibitordatensatzes erstellte 3D-QSAR-Modell, da es unter Einbeziehung strukturell unterschiedlicher Verbindungen aufgestellt wurde, ein breites Einsatzgebiet für die Vorhersage der Inhibitoraktivität unbekannter Verbindungen.

6 ZUSAMMENFASSUNG

Das humane CYP1A2, das für den Arzneistoffmetabolismus in der Leber und für die Aktivierung von Prokarzinogenen von Bedeutung ist, wurde mithilfe von Molecular-Modelling-Methoden theoretisch untersucht. Da die dreidimensionale Struktur des Isoenzym nicht bekannt ist, wurde zunächst auf der Grundlage einer Kristallstruktur des Säugetiercytochroms CYP2C5, die als Substratkomplex vorliegt, ein Homologiemodell erstellt.

Die Validierung des CYP1A2-Modells erfolgte anhand einer MDS über einen Simulationszeitraum von 2,5 ns in physiologischer Umgebung, wobei die Simulationsbedingungen zunächst in einer MDS mit der Kristallstruktur überprüft wurden. Es resultierte ein stabiles Proteinmodell von guter struktureller Qualität, das alle den P450-Isoenzymen gemeinsamen charakteristischen Strukturmerkmale aufwies und dessen dynamisches Verhalten mit dem der Kristallstruktur übereinstimmte.

Im Hinblick auf die moleküldynamische Untersuchung von Substratkomplexen des CYP1A2-Modells wurde wiederum zuerst eine Validierungsdynamik mit einer Kristallstruktur von CYP2C5 durchgeführt. Es zeigte sich, dass der Substratkomplex der Kristallstruktur über einen Simulationszeitraum von 5 ns stabil blieb und dass das Bindungsverhalten von Diclofenac im aktiven Zentrum qualitativ gut wiedergegeben werden konnte. Für die drei strukturell unterschiedlichen Substrate MelQ, 7-Ethoxyresorufin und Coffein wurde jeweils eine Positionierung im aktiven Zentrum des CYP1A2-Modells gefunden, die ebenfalls zu einem stabilen Enzym-Substrat-Komplex in einer MDS derselben Länge führte. Die Substrate befanden sich dabei durchgehend in einer für eine Umsetzung günstigen Orientierung zum Hämeisen. Im Fall des MelQ und des 7-Ethoxyresorufins standen zahlreiche Protein-Substrat-Interaktionen in Übereinstimmung mit Mutationsstudien. Darüber hinaus wurden weitere, teilweise wasservermittelte Wasserstoffbrückeninteraktionen identifiziert, die für die Substratbindung von Bedeutung sein könnten. Für das Coffein wurde eine weitere MDS mit der Substratorientierung für die Bildung des Nebenmetaboliten Theobromin durchgeführt, die bei einer Auswertung der Unterschiede in den Protein-Substrat-Interaktionen einen Erklärungsansatz für die bevorzugte Bildung des Hauptmetaboliten Paraxanthin lieferte.

Die abschließende Anwendung des Homologiemodells von CYP1A2 bestand in einer 3D-QSAR-Studie, für die ein Datensatz kompetitiver CYP1A2-Inhibitoren, bestehend aus 44 Chinolinen, Naphthalenen, Lactonen und anderen Verbindungen, eingesetzt wurde. Im Hinblick auf die korrekte Beschreibung des Häm wurde zunächst das verwendete Docking-Programm anhand eines Testdatensatzes von 12 Substraten validiert. Acht der 12 Substrate konnten korrekt im aktiven Zentrum des CYP1A2-Modells positioniert werden. Daraufhin wurden auch die Verbindungen des Inhibitor-datensatzes mithilfe des automatischen Dockings in der Proteinstruktur überlagert. Die Korrelation von molekularen Interaktionsfeldern der Inhibitor-konformationen aus

der Überlagerung mit der biologischen Aktivität der Verbindungen resultierte in einem statistisch signifikanten vorhersagekräftigen PLS-Modell.

7 LITERATUR

- [1] Neye, H.: Arzneimittelinteraktionen. Niedersächsisches Ärzteblatt 2004,9, Arzneimittelordner. <http://www.haeverlag.de/index.htm>
- [2] Prentis, R.A.; Lis, Y.; Walker, S.R.: Pharmaceutical innovation by the seven UK-owned pharmaceutical companies (194-1985). Br. J. Clin. Pharmac. 1988, 25, 387-396
- [3] Kennedy, T.: Managing the drug discovery / development interface. Drug Discovery Today 1997, 2, 436-444
- [4] Anzenbacher, P.; Dawson, J.H.; Kitagawa, T.: Towards a unified concept of oxygen activation by heme enzymes: the role of the proximal ligand. J. Mol. Struct. 1989, 214, 149-155
- [5] Garfinkel, D.: Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions. Arch. Biochem. Biophys. 1958, 77, 493-509
- [6] Klingenberg, M.: Pigments of rat liver microsomes. Arch. Biochem. Biophys. 1958, 75, 376-386
- [7] Nebert, D.W.; Adesnik, M.; Coon, M.J.; Estabrook, R.W.; Gonzalez, F.J.; Guengerich, F.P.; Gunsalus, I.C.; Johnson, E.F.; Kemper, B.; Levin, W.; Phillips, I.R.; Sato, R.; Waterman, M.R.: The P450 gene superfamily: recommended nomenclature. DNA 1987, 6, 1-11
- [8] Keller, M.; Blöchl, E.; Wächtershäuser, G.; Stetter, K.O.: Formation of amide bonds without a condensation agent and implications for origin of life. Nature 1994, 368, 836-838
- [9] Nelson, D.R.; Strobel, H.W.: Evolution of cytochrome P-450 proteins. Molecular Biology and Evolution 1987, 4, 572-593
- [10] Nebert, D.W.; Gonzalez, F.J.: P450 genes: structure, evolution and regulation. Annual Reviews of Biochemistry 1987, 56, 945-993
- [11] Lewis, D.F.W.: Cytochromes P450: Structure, function and mechanism. 1. Aufl. Taylor & Francis Ltd. 1996, ISBN 0-7484-0443-0
- [12] Nelson, D.R.: Cytochrome P450 Homepage. <http://drnelson.utmem.edu/family.list.htm> (vollständiges Update 2004)
- [13] Schenkman, J.B.; Griem, H.: Cytochrome P450. Heidelberg: Springer Verlag
- [14] Anzenbacher, P.; Anzenbacherova, E.: Cytochromes P450 and metabolism of xenobiotics. Cell. Mol. Life Sci. 2001, 58, 737-747
- [15] Degtyarenko, K.N. & Fábíán, P.: Directory of P450-containing Systems. <http://www.icgeb.org/~p450srv/>
- [16] Quattrochi, L.C.; Okino, S.T.; Pendurthi, U.R.; Tukey, R.H.: Cloning and isolation of human cytochrome P-450 cDNAs homologous to dioxin-inducible rabbit mRNAs encoding P-450 4 and P-450 6. DNA 1985, 4, 395-400
- [17] Sutter, T.R.; Tang, Y.M.; Hayes, C.L.; Wo, Y.Y.P.; Jabs, E.W.; Li, X.; Yin, H.; Cody, C.W.; Greenlee, W.F.: Complete cDNA sequence of a human dioxin-inducible mRNA identifies a new gene subfamily of cytochrome P450 that maps to chromosome 2. J. Biol. Chem. 1994, 269, 13092-13099
- [18] Boeckmann, B.; Bairoch, A.; Apweiler, R.; Blatter, M. C.; Estreicher, A.; Gasteiger, E.; Martin, M.J.; Michoud, K.; O'Donovan, C.; Phan, I.; Pilbout, S.; Schneider, M.: *The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003*. Nucleic Acids Res. 2003, 31, 365-370

- [19] Shimada, T.; Yun, C.H.; Yamazaki, H.; Gautier, J.C.; Beaune, P.H.; Guengerich, F.P.: Characterization of human lung microsomal cytochrome P-450 1A1 and its role in the oxidation of chemical carcinogens. *Mol. Pharmacol.* 1992, 41, 856-864
- [20] Shimada, T.; Hayes, C.L.; Yamazaki, H.; Amin, S.; Hecht, S.S.; Guengerich, F.P.; Sutter, T.R.: Activation of chemically diverse procarcinogens by human cytochrome P-450 1B1. *Cancer Res.* 1996, 56, 2979-2984
- [21] Murray, G.I.; Taylor, M.C.; McFadyen, M.C.; McKay, J.A.; Greenlee, W.F.; Burke, M.D.; Melvin, W.T.: Tumor-specific expression of cytochrome P450 CYP1B1. *Cancer Res.* 1997, 57, 3026-3031
- [22] Huang, J.D.; Guo, W.C.; Lai, M.D.; Guo, Y.L.; Lambert, G.H.: Detection of a novel cytochrome P-450 1A2 polymorphism (F21L) in Chinese. *Drug. Metab. Dispos.* 1999, 27(1), 98-101
- [23] Murayama, N.; Soyama, A.; Saito, Y.; Nakajima, Y.; Komamura, K.; Ueno, K.; Kamakura, S.; Kitakaze, M.; Kimura, H.; Goto, Y.; Saitoh, O.; Katoh, M.; Ohnuma, T.; Kawai, M.; Sugai, K.; Ohtsuki, T.; Suzuki, C.; Minami, N.; Ozawa, S.; Sawada, J.: Six novel nonsynonymous CYP1A2 gene polymorphisms: catalytic activities of the naturally occurring variant enzymes. *J. Pharmacol. Exp. Ther.* 2004, 308(1), 300-306
- [24] Anzenbacher, P.; Hudeček, J.: Differences in flexibility of active sites of cytochromes P450 probed by resonance Raman and UV-Vis absorption spectroscopy. *J. Inorg. Biochem.* 2001, 87, 209-213
- [25] Lewis, D.F.W.: On the recognition of mammalian microsomal cytochrome P450 substrates and their characteristics. *Biochem. Pharmacol.* 2000, 60, 293-306
- [26] Shimada, T.; Gillam, E.M.J.; Sutter, T.R.; Strickland, P.T.; Guengerich, F.P.; Yamazaki, H.: Oxidation of xenobiotics by recombinant human cytochrome P450 1B1. *Drug Metab. Disp.* 1997, 29, 617-622
- [27] Bostrom, C.E.; Gerde, P.; Hanberg, A.; Jernstrom, B.; Johansson, C.; Kyrklund, T.; Rannug, A.; Tornquist, M.; Victorin, K.; Westerholm, R.: Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. *Environ. Health Perspect.* 2002, 110, 451-488
- [28] Omura, T.; Ishimura, Y.; Fuji-Kuriyama, Y.: *Cytochrome P-450*. Second Edition. VCH Verlagsgesellschaft 1993, ISBN 3-527-30016-3
- [29] Guengerich, F.P.: Metabolic activation of carcinogens. *Pharmac. Ther.* 1992, 54, 17-61
- [30] Guengerich F.P.: Roles of cytochrome P-450 enzymes in chemical carcinogenesis and cancer chemotherapy. *Cancer Res.* 1988, 48, 2946-2954
- [31] Shimada, T.; Iwasaki, M.; Martin, M.V.; Guengerich, F.P.: Human liver microsomal cytochrome P-450 enzymes involved in the bioactivation of procarcinogens detected by umu gene response in *Salmonella typhimurium* TA1535/pSK1002. *Cancer Res.* 1989, 49, 3218-3228
- [32] Butler, M.A.; Iwasaki, M.; Guengerich, F.P.; Kadlubar, F.F.: Human cytochrome P-450_{PA} (P-450IA2), the phenacetin O-deethylase, is primarily responsible for the hepatic 3-demethylation of caffeine and N-oxidation of carcinogenic arylamines. *Proc. Natl. Acad. Sci. USA* 1989, 86, 7696-7700
- [33] Wakabayashi, K.; Nagao, M.; Esumi, H.; Sugimura, T.: Food-derived mutagens and carcinogens. *Cancer Res.* 1992, 52(7 Suppl.), 2092s-2098s
- [34] Parikh, A.; Josephy, P.D.; Guengerich, F.P.: Selection and characterization of human cytochrome P450 1A2 mutants with altered catalytic properties. *Biochemistry* 1999, 38, 5283-5289

- [35] Josephy, P.D.; Bibeau, K.L.; Evans, D.H.: Activation of MeIQ (2-Amino-3,4-dimethylimidazo[4,5-f]quinoline) by sequence variants of recombinant human cytochrome P450 1A2. *Environ. Mol. Mutagenesis* 2000, 35, 328-335
- [36] Lozano, J.J.; Pastor, M.; Cruciani, G.; Gaedt, K.; Centeno, N.B.; Gago, F.; Sanz, F.: 3D-QSAR methods on the basis of ligand-receptor complexes. Application of COMBINE and GRID-GOLPE methodologies to a series of CYP1A2 ligands. *J. Comp.-Aided Mol. Des.* 2000, 14, 341-353
- [37] Faletto, M.B.; Koser, P.L.; Battula, N.; Townsend, G.K.; Maccubbin, A.E.; Gelboin, H.V.; Gurtoo, H.L.: Cytochrome P3-450 cDNA encodes aflatoxin B1-4-hydroxylase. *J. Biol. Chem.* 1988, 263(25), 12187-12189
- [38] Spink, D.C.; Eugster, H.; Lincoln, D.W.I.; Schuetz, J.D.; Schuetz, E.G.; Johnson, J.A.; Kaminsky, L.S.; Gierthy, J.F.: 17 β -estradiol hydroxylation catalyzed by human cytochrome P450 1A1: a comparison of the activities induced by 2,3,7,8-tetrachlorodibenzo-p-dioxin in MCF-7 cells with those from heterologous expression of the cDNA. *Arch. Biochem. Biophys.* 1992, 293, 342-348
- [39] Hayes, C.L.; Spink, D.C.; Spink, B.C.; Cao, J.Q.; Walker, N.J.; Sutter, T.R.: 17 β -estradiol hydroxylation catalyzed by human cytochrome P450 1B1. *Proc. Natl. Acad. Sci. USA* 1996, 93, 9776-9781
- [40] Dwivedy, I.; Devanesan, P.; Cremonesi, P.; Rogan, E.; Cavalieri, E.: Synthesis and characterization of estrogen 2,3- and 3,4-quinones. Comparison of DNA adducts formed by the quinones versus horseradish peroxidase-activated catechol estrogens. *Chem. Res. Toxicol.* 1992, 5, 828-833
- [41] Cavalieri, E.L.; Stack, D.E.; Devanesan, P.D.; Todorovic, R.; Dwivedy, I.; Higginbotham, S.; Johansson, S.L.; Patil, K.D.; Gross, M.L.; Gooden, J.K.; Ramanathan, R.; Cerny, R.L.; Rogan, E.G.: Molecular origin of cancer: Catechol estrogen-3,4-quinones as endogenous tumor initiators. *Proc. Natl. Acad. Sci. USA* 1997, 94, 10937-10942
- [42] Dahlin, D.C.; Miwa, G.T.; Lu, A.Y.; Nelson, S.D.: N-acetyl-p-benzoquinone imine: a cytochrome P-450-mediated oxidation product of acetaminophen. *Proc. Natl. Acad. Sci. U S A* 1984, 81(5), 1327-1331
- [43] Spaldin, V.; Madden, S.; Pool, W.F.; Woolf, T.F.; Park, B.K.: The effect of enzyme inhibition on the metabolism and activation of tacrine by human liver microsomes. *Br. J. Clin. Pharmac.* 1994, 38, 15-22
- [44] Fang, J.; Coutts, R.T.; McKenna, K.F.; Baker, G.B.: Elucidation of individual cytochrome P450 enzymes involved in the metabolism of clozapine. *Naunyn-Schmiedeberg's Arch. Pharmacol.* 1998, 358, 592-599
- [45] Murray, M.: Mechanisms and significance of inhibitory drug interactions involving cytochrome P450 enzymes (Review). *Intern. J. of Mol. Med.* 1999, 3, 227-238
- [46] Buur-Rasmussen, B.; Brøsen, K.: Cytochrome P450 and therapeutic drug monitoring with respect to clozapine. *Eur. Neuropsychopharmacol.* 1999, 9, 453-459
- [47] De Rienzo, F.; Fanelli, F.; Menziani, M.C.; de Benedetti, P.G.: Theoretical investigation of substrate specificity for cytochromes P450 1A2, P450 1D6, and P450 3A4. *J. Comp.-Aided Mol. Des.* 2000, 14, 93-116
- [48] Burke, M.D.; Thompson, S.; Elcombe, C.R.; Halpert, J.; Haaparanta, T.; Mayer, R.T.: Ethoxy-, pentoxy- and benzyloxyphenoxazones and homologues: a series of substrates to distinguish between different induced cytochromes P-450. *Biochem. Pharmacol.* 1985, 34(18), 3337-3345

- [49] Yamazaki, H.; Inoue, K.; Mimura, M.; Oda, Y.; Guengerich, F.P.; Shimada, T.: 7-Ethoxycoumarin O-deethylation catalyzed by cytochromes P450 1A2 and 2E1 in human liver microsomes. *Biochem. Pharmacol.* 1996, 51(3), 313-319
- [50] Liu, J.; Ericksen, S.S.; Sivaneri, M.; Besspiata, D.; Fisher, C.W.; Szklarz, G.D.: The effect of reciprocal active site mutations in human cytochromes P450 1A1 and 1A2 on alkoxyresorufin metabolism. *Arch. Biochem. Biophys.* 2004, 424, 33-43
- [51] Furuya, H.; Shimizu, T.; Hirano, K.; Hatano, M.; Fujii-Kuriyama, Y.: Site-directed mutageneses of rat liver cytochrome P-450_d: catalytic activities toward benzphetamine and 7-ethoxycoumarin. *Biochemistry* 1989, 28, 6848-6857
- [52] Fuhr, U.; Doehmer, J.; Battula, N.; Wölfel, C.; Kudla, C.; Keita, Y.; Staib, H.: Biotransformation of caffeine and theophylline in mammalian cell lines genetically engineered for expression of single cytochrome P450 isoforms. *Biochem. Pharmacol.* 1992, 43, 225-235
- [53] Gu, L.; Gonzalez, F.J.; Kalow, W.; Tang, BK: Biotransformation of caffeine, paraxanthine, theobromine and theophylline by cDNA-expressed human CYP1A2 and CYP2E1. *Pharmacogenetics* 1992, 2, 73-77
- [54] Regal, K.A.; Nelson, S.D.: Orientation of caffeine within the active site of human cytochrome P450 1A2 based on NMR longitudinal (T_1) relaxation measurements. *Arch. Biochem. Biophys.* 2000, 384, 47-58
- [55] Nielsen-Kudsk, J.E.; Buhl, J.S.; Johannessen, A.C.: Verapamil-induced inhibition of theophylline elimination in healthy humans. *Pharmacol Toxicol.* 1990, 66(2), 101-103
- [56] Liu, G.; Gelboin, H.V.; Myers, M.J.: Role of cytochrome P450 1A2 in acetanilide 4-hydroxylation as determined with cDNA expression and monoclonal antibodies. *Arch. Biochem. Biophys.* 1991, 284(2), 400-406
- [57] Imaoka, S.; Enomoto, K.; Oda, Y.; Asada, A.; Fujimori, M.; Shimada, T.; Fu, S.; Guengerich, F.P.; Funae, Y.: Lidocaine metabolism by human cytochrome P-450s purified from hepatic microsomes: comparison of those with rat hepatic cytochrome P-450s. *J. Pharmacol. Exp. Ther.* 1990, 255, 1385-1391
- [58] Ono, S.; Hatanaka, T.; Hotta, H.; Tsutsui, M.; Satoh, T.; Gonzalez, F.J.: Chlorzoxazone is metabolized by human CYP1A2 as well as by human CYP2E1. *Pharmacogenetics* 1995, 5, 143-150
- [59] Dixon, C.M.; Colthup, P.V.; Serabjit-Singh, C.J.; Kerr, B.M.; Boehlert, C.C.; Park, G.R.; Tarbit, M.H.: Multiple forms of cytochrome P450 are involved in the metabolism of ondansetron in humans. *Drug Metab. Dispos.* 1995, 23(11), 1225-1230
- [60] Yamazaki, H.; Shimada, T.: Human liver cytochrome P450 enzymes involved in the 7-hydroxylation of R- and S-warfarin enantiomers. *Biochem. Pharmacol.* 1997, 54, 1195-1203
- [61] Eaton, D.L.; Gallagher, E.P.; Bammler, T.K.; Kunze, K.L.: Role of cytochrome P4501A2 in chemical carcinogenesis: implications for human variability in expression and enzyme activity. *Pharmacogenetics* 1995, 5(5), 259-274
- [62] Christou, M.; Savas, U.; Schroeder, S.; Shen, X.; Thompson, T.; Gould, M.N.; Jefcoate, C.R.: Cytochromes CYP1A1 and CYP1B1 in the rat mammary gland: cell-specific expression and regulation by polycyclic aromatic hydrocarbons and hormones. *Mol. Cell. Endocrinol.* 1995, 115(1), 41-50
- [63] Hankinson, O.: The aryl hydrocarbon receptor complex. *Annu. Rev. Pharmacol. Toxicol.* 1995, 35, 307-340

- [64] Mutschler, E.: Arzneimittelwirkungen. Wissenschaftliche Verlagsgesellschaft mbH, Stuttgart 1997
- [65] Sesardic, D.; Boobis, A.R.; Murray, S.; Segura, J.; de la Torre, R.; Davies, D.S.: Furafylline is a potent and selective inhibitor of cytochrome P4501A2 in man. *Br. J. Clin. Pharmacol.* 1990, 29, 651-663
- [66] Kunze, K.L.; Trager, W.F.: Isoform-selective mechanism-based inhibition of human cytochrome P450 1A2 by furafylline. *Chem. Res. Toxicol.* 1993, 6, 649-656
- [67] Cho, U.S.; Park, E.Y.; Dong, M.S.; Park, B.S.; Kim, K.; Kim, K.H.: Tight-binding inhibition by alpha-naphthoflavone of human cytochrome P450 1A2. *Biochim. Biophys. Acta* 2003, 1648, 195-202
- [68] Albers, L.J.; Ozdemir, V.; Marder, S.R.; Raggi, M.A.; Aravagiri, M.; Endrenyi, L.; Reist, C.: Low-dose fluvoxamine as an adjunct to reduce olanzapine therapeutic dose requirements: a prospective dose-adjusted drug interaction strategy. *J. Clin. Psychopharmacol.* 2005, 25, 170-174
- [69] Tassaneeyakul, W.; Birkett, D.J.; Veronese, M.E.; McManus, M.E.; Tukey, R.H.; Quattrochi, L.C.; Gelboin, H.V.; Miners, J.O.: Specificity of substrate and inhibitor probes for human cytochromes P450 1A1 and 1A2. *J. Pharmacol. Exp. Ther.* 1993, 265, 401-407
- [70] Brøsen, K.; Skjelbo, E.; Rasmussen, B.B.; Poulsen, H.E.; Loft, S.: Fluvoxamine is a potent inhibitor of cytochrome P4501A2. *Biochem. Pharmacol.* 1993, 45(6), 1211-1214
- [71] Maesen, F.P.V.; Teengs, J.P.; Baur, C.; Davies, B.I.: Quinolones and raised plasma concentrations of theophylline. *Lancet* 1984, 2
- [72] Fuhr, U.; Wolff, T.; Harder, S.; Schymanski, P.; Staib, A.H.: Quinolone inhibition of cytochrome P-450 dependent caffeine metabolism in human liver microsomes. *Drug Metabol. Dispos.* 1990, 18, 1005-1010
- [73] Fuhr, U.; Strobl, G.; Manaut, F.; Anders, E.M.; Sörgel, F.; Lopez-de-Brinas, E.; Chu, D.T.W.; Pernet, A.G.; Mahr, G.; Sanz, F.; Staib, A.H.: Quinolone antibacterial agents: relationship between structure and in vitro inhibition of the human cytochrome P450 isoform CYP1A2. *Mol. Pharmacol.* 1993, 43, 191-199
- [74] Wang, M.; Roberts, D.L.; Paschke, R.; Shea, T.M.; Masters, B.S.; Kim, J.J.: Three-dimensional structure of NADPH-cytochrome P450 reductase: prototype for FMN- and FAD-containing enzymes. *Proc. Natl. Acad. Sci. USA.* 1997, 94(16), 8411-8416
- [75] Schlichting, I.; Berendzen, J.; Chu, K.; Stock, A.M.; Maves, S.A.; Benson, D.E.; Sweet, R.M.; Ringe, D.; Petsko, G.A.; Sligar, S.G.: The catalytic pathway of cytochrome P450cam at atomic resolution. *Science* 2000, 287, 1615-1622
- [76] Davydov, R.; Makris, T.M.; Kofman, V.; Werst, D.E.; Sligar, S.G.; Hoffman, B.M.: Hydroxylation of camphor by reduced oxy-cytochrome P450cam: mechanistic implications of EPR and ENDOR studies of catalytic intermediates in native and mutant enzymes. *J. Am. Chem. Soc.* 2001, 123, 1403-1415
- [77] Sandhu, P.; Guo, Z.; Baba, T.; Martin, M.V.; Tukey, R.H.; Guengerich, F.P.: Expression of modified human cytochrome P450 1A2 in escherichia coli: stabilisation, purification spectral characterization, and catalytic activities of the enzyme. *Arch. Biochem. Biophys.* 1994, 309, 168-177
- [78] Harris, D.L.; Loew, G.H.: Determinants of the spin state of the resting state of cytochrome P450_{cam}. *J. Am. Chem. Soc.* 1993, 115, 8775-8779

- [79] Meunier, B.; de Visser, S.P.; Shaik, S.: Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. *Chem. Rev.* 2004, 104, 3947-3980
- [80] Akthar, M.; Corina, D.; Miller, S.; Shyadehi, A.Z.; Wright, J.N.: Mechanism of the acyl-carbon cleavage and related reactions catalyzed by multifunctional P-450s: studies on cytochrome P-450_{17 α} . *Biochemistry* 1994, 33, 4410-4418
- [81] Williams, P.A.; Cosme, J.; Sridhar, V.; Johnson, E.F.; McRee, D.E.: Mammalian microsomal cytochrome P450 monooxygenase: Structural adaptations for membrane binding and functional diversity. *Mol. Cell.* 2000, 5, 121-132
- [82] Cosme, J.; Johnson, E.F.: Engineering microsomal cytochrome P450 2C5 to be a soluble monomeric enzyme. *J. Biol. Chem.* 2000, 275, 2545-2553
- [83] Hasemann, C.A.; Kurumbail, R.G.; Boddupalli, S.S.; Peterson, J.A.; Deisenhofer, J.: Structure and function of cytochromes P450: a comparative analysis of three crystal structures. *Structure* 1995, 3, 41-62
- [84] Wester, M.R.; Johnson, E.F.; Marques-Soares, C.; Dansette, P.M.; Mansuy, D.; Stout, C.D.: Structure of a substrate complex of mammalian cytochrome P450 2C5 at 2.3 Å resolution: evidence for multiple substrate binding modes. *Biochemistry* 2003, 42 (21), 6370-6379
- [85] Hasemann, C.A.; Ravichandran, K.G.; Peterson, J.A.; Deisenhofer, J.: Crystal structure and refinement of cytochrome P450_{terp} at 2.3 Å resolution. *J. Mol. Biol.* 1994, 236, 1169-1185
- [86] Auclair, K.; Moenne-Loccoz, P.; Ortiz de Montellano, P.R.: Roles of the proximal heme thiolate ligand in cytochrome p450(cam). *J. Am. Chem. Soc.*, 23(21), 4877-4885
- [87] Kassner, R.J.: A theoretical model for the effects of local nonpolar heme environments on the redox potentials in cytochromes. *J. Am. Chem. Soc.* 1973, 95(8), 2674-2677
- [88] Gerber, N.C.; Sligar, S.G.: A role for Asp-251 in cytochrome P-450cam oxygen activation. *J. Biol. Chem.* 1994, 269(6), 4260-4266
- [89] Li, H.; Poulos, T.L.: The structure of the cytochrome P450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid. *Nature Struc. Biol.* 1997, 4, 140-146
- [90] Podust, L.M.; Poulos, T.L.; Waterman, M.R.: Crystal structure of cytochrome P450 14 α -sterol demethylase (CYP51) from *Mycobacterium tuberculosis* in complex with azole inhibitors. *Proc. Natl. Acad. Sci. USA* 2001, 98, 3068-3073
- [91] Scott, E.E.; He, Y.A.; Wester, M.R.; White, M.A.; Chin, C.C.; Halpert, J.R.; Johnson, E.F.; Stout, C.D.: An open conformation of mammalian cytochrome P450 2B4 at 1.6 Å resolution. *Proc. Natl. Acad. Sci. USA* 2003, 100, 13196-13201
- [92] Gotoh, O.: Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analysis of amino acid and coding nucleotide sequences. *J. Biol. Chem.* 1992, 267, 83-90
- [93] Wester, M.R.; Johnson, E.F.; Marques-Soares, C.; Dijols, S.; Dansette, P.M.; Mansuy, D.; Stout, C.D.: Structure of mammalian cytochrome P450 2C5 complexed with diclofenac at 2.1-Å resolution: evidence for an induced fit model of substrate binding. *Biochemistry* 2003, 42, 9335-9345
- [94] Williams, P.A.; Cosme, J.; Ward, A.; Angove, H.C.; Vinković, D.; Jhoti, H.: Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* 2003, 424, 464-468
- [95] Wester, M.R.; Yano, J.K.; Schoch, G.A.; Yang, C.; Griffin, K.J.; Stout, C.D.; Johnson, E.F.: The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-Å resolution. *J. Biol. Chem.* 2004, 279, 35630-35637

- [96] Needleman, S.B.; Wunsch, C.D.: A General Method Applicable to the Search for Similarities in the Amino Acid Sequences of Two Proteins. *J. Mol. Biol.* 1970, 48, 443-453
- [97] Smith, T.F.; Waterman, M.S.: Identification of Common Molecular Subsequences. *J. Mol. Biol.* 1981, 147, 195-197
- [98] Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J.: Basic Local Alignment Search Tool. *J. Mol. Biol.* 1990, 215, 403-410
- [99] Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J.: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997, 25, 3389-3402
- [100] Tatusova, T.A.; Madden, T.L.: BLAST 2 SEQUENCES, a new tool for comparing protein and nucleotide sequences. *FEMS Microbol. Lett.* 1999, 174, 247-250
- [101] Thompson, J.D.; Higgins, D.G.; Gibson, T.J.: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994, 22, 4673-4680
- [102] Dayhoff, M.O.; Schwartz, R.M.; Orcutt, B.C.: A model of evolutionary change in proteins. *Atlas of Protein Sequence and Structure 5(3)*, M.O. Dayhoff (ed.), 345 – 352
- [103] Henikoff, S.; Henikoff, J.: Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 1992, 89 (22), 10915-10919
- [104] Lim, V.I.: Structural principles of the globular organization of protein chains. A stereochemical theory of globular protein secondary structure. *J. Mol. Biol.* 1974, 88, 873-894
- [105] Chou, P.Y.; Fasman, G.D.: Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins. *Biochemistry* 1974, 13, 211-222
- [106] Garnier, J.; Osguthorpe, D.J.; Robson, B.: Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 1978, 120, 97-120
- [107] Rost, B.; Sander, C.: Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.* 1993, 232, 584-599
- [108] Jones, D.T.: Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 1999, 292, 195-202
- [109] McGuffin, L.J.; Bryson, K.; Jones, D.T.: The PSIPRED protein structure prediction server. *Bioinformatics* 2000, 16, 404-405
- [110] Shi, J.; Blundell, T.L.; Mizuguchi, K.: FUGUE: Sequence-structure Homology Recognition Using Environment-specific Substitution Tables and Structure-dependent Gap Penalties. *J. Mol. Biol.* 2001, 310, 243-257
- [111] INSIGHT II, Version 2000. Accelrys Inc., San Diego, CA, USA
- [112] Höltje, H.-D.; Sippl, W.; Rognan, D.; Folkers, G.: *Molecular Modeling: basic principles and applications.* Wiley-VCH, Weinheim 2003
- [113] Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E.: The Protein Data Bank. *Nucleic Acids Res.* 2000, 28, 235-242
- [114] Hohenberg, P.; Kohn, W.: Inhomogenous Electron Gas. *Phys. Rev. B* 1964, 136, 864-871

- [115] Kohn, W.; Sham, L.J.: Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev. A* 1965, 140, 1133-1136
- [116] Jaguar 4.1, Schrödinger, Inc., Portland, OR, 1991-2000
- [117] Leach, A.R.: *Molecular Modelling: principles and applications*. Pearson Education, Essex 2001
- [118] Berendsen, H.J.C.; van der Spoel, D.; van Drunen, R.: GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* 1995, 91, 43-56
- [119] Lindahl, E.; Hess, B.; van der Spoel, D.: GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.* 2001, 7, 306-317
- [120] van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G.: *Biomolecular Simulation: The GROMOS96 manual and user guide*. Zürich, Switzerland: Hochschulverlag AG an der ETH Zürich 1996
- [121] Ewald, P.P.: Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* 1921, 64, 253-287
- [122] Darden, T.; York, D.; Pedersen, L.: Particle mesh Ewald: An N-log (N) method for Ewald sums in large systems. *J. Chem. Phys.* 1993, 98, 10089-10092
- [123] van der Spoel, D.; van Buuren, A.R.; Tieleman, D.P.; Berendsen, H.J.: Molecular dynamics simulations of peptides from BPTI: a closer look at amide-aromatic interactions. *J. Biomol. NMR* 1996, 8 (3), 229-238
- [124] van der Spoel, D.; van Buuren, A.R.; Apol, E.; Meulenhoff, P.J.; Tieleman, D.P.; Sijbers, A.L.T.M.; Hess, B.; Feenstra, K.A.; Lindahl, E.; van Drunen, R.; Berendsen, H.J.C.: *GROMACS user manual version 3.0*. Nijenborgh 4, 9747 AG Groningen, The Netherlands 2001
- [125] Schappach, A.: *Molecular Modelling Untersuchungen an Steroidhydroxylasen mit therapeutischer Bedeutung*. Dissertation, Institut für Pharmazeutische Chemie, Heinrich-Heine-Universität Düsseldorf 2000
- [126] Stahl, G.: *Entwicklung von Homologie-Modellen der Cytochrome P450 2A5 und 2A6*. Dissertation, Institut für Pharmazeutische Chemie, Heinrich-Heine-Universität Düsseldorf 2002
- [127] Helms, V.; Wade, R.: Thermodynamics of water mediating protein-ligand interactions in cytochrome p450cam: A molecular dynamics study. *Biophys. J.* 1995, 69, 810-824
- [128] Hockney, R.W.: The potential calculation and some applications. *Methods in Computational Physics* 1970, 9, 136-211
- [129] Feenstra, K.A.; Hess, B.; Berendsen, H.J.C.: Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comp. Chem.* 1999, 20 (8), 786-798
- [130] Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M.: LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.* 1997, 18, 1463-1472
- [131] Berendsen, H.J.C.; Postma, J.P.M.; DiNola, A.; Haak, J.R.: Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 1984, 81, 3684-3690
- [132] Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C.: Numerical integration of the cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J. Comp. Phys.* 1977, 23, 327-341
- [133] Morris, A.L.; MacArthur, M.W.; Hutchinson, E.G.; Thornton, J.M.: Stereochemical quality of protein structure coordinates. *Proteins* 1992, 12 (4), 345-364

- [134] Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M.: PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 1993, 26, 283-291
- [135] Laskowski, R. A.: SURFNET: A program for visualizing molecular surfaces, cavities and intermolecular interactions. *J. Mol. Graph.* 1995, 13, 323-330
- [136] Kelley, L.A.; Gardner, S.P.; Sutcliffe, M.J.: An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally-related subfamilies. *Protein Engineering* 1996, 9, 1063-1065
- [137] 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.* 2000, 14, 559-572
- [138] Classen-Houben, D.: Proteinbasierte 3D QSAR-Analysen und virtuelles Screening an Inhibitoren der Protein-Tyrosin-Phosphatase 1B. Dissertation, Institut für Pharmazeutische Chemie, Heinrich-Heine-Universität Düsseldorf 2004
- [139] Taylor, R.D.; Jewsbury, P.J.; Essex, J.W.; A review of protein-small molecule docking methods. *J. Comput. Aided Mol. Des.* 2002, 16, 151-166
- [140] Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G.: Predicting receptor-ligand-interactions by an incremental construction algorithm. *J. Mol. Biol.* 1996, 261, 470-489
- [141] Goodsell, D.S.; Olson, A.J.: Automated docking of substrates to proteins by simulated annealing. *Proteins* 1990, 8, 195-202
- [142] Jones, G.; Willett, P.; Glen, R.C.; Leach, A.R.; Taylor, R.: Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 1997, 267, 727-748
- [143] Eldridge, M.D.; Murray, C.W.; Auton, T.R.; Paolini, G.V.; Mee, R.P.: 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.* 1997, 11, 425-445
- [144] Goodford, P.J.: A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* 1985, 28, 849-857
- [145] Boobbyer, D.N.; Goodford, P.J.; McWhinnie, P.M.; Wade, R.C.: New hydrogen-bond potentials for use in determining energetically favorable binding sites on molecules of known structure. *J. Med. Chem.* 1989, 32, 1083-1094
- [146] Wade, R.C.; Clark, K.J.; Goodford, P.J.: Further development of hydrogen bond functions for use in determining energetically favorable binding sites on molecules of known structure. 1. Ligand probe groups with the ability to form two hydrogen bonds. *J. Med. Chem.* 1993, 36, 140-147
- [147] Wade, R.C.; Goodford, P.J.: Further development of hydrogen bond functions for use in determining energetically favorable binding sites on molecules of known structure. 1. Ligand probe groups with the ability to form more than two hydrogen bonds. *J. Med. Chem.* 1993, 36, 148-156
- [148] GOLPE 4.5. Multivariate Infometric Analysis Srl., Viale dei Castagni 16, Perugia, Italy 1999
- [149] Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S.: Generating optimal linear PLS estimations (GOLPE): An advanced chemometric tool for handling 3D-QSAR problems. *Quant. Struct.-Act. Relat.* 1993, 12, 9-20
- [150] Cruciani, G.; Watson, K.A.: Comparative molecular field analysis using GRID force-field and GOLPE variable selection methods in a study of inhibitors of glycogen phosphorylase b. *J. Med. Chem.* 1994, 37, 2589-2601

- [151] Cruciani, G; Clementi, S; Pastor, M: GOLPE-Guided Region Selection. *Perspectives in Drug Discovery and Design* 1998, 12, 71-86
- [152] Rost, B.; Yachdav, G; Liu, J.: The PredictProtein Server. *Nucleic Acids Research* 2003, 32 (Web Server issue), W321-W326; URL: <http://cubic.bioc.columbia.edu/predictprotein>
- [153] Laskowski, R.A.: PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res.* 2001, 29, 221-222
- [154] Dauber-Osguthorpe, P.; Roberts, V.A.; Osguthorpe, D.J.; Wolff, J.; Genest, M.; Hagler, A.T.: Structure and energetics of ligand binding to proteins: Escherichia coli dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins* 1988, 4, 31-47
- [155] Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; Hermans, J.: Interaction models for water in relation to protein hydration. Pullman, S. (Hrsg.): *Intermolecular forces*. Dordrecht: Reidel, 1981, 331-342
- [156] Miyamoto, S.; Kollman, P.A.: SETTLE: An analytical version of the SHAKE and RATTLE algorithms for rigid water models. *J. Comp. Chem.* 1992, 13, 952-962
- [157] Ibragimova G.T.; Wade, R.C.: Importance of explicit salt ions for protein stability in molecular dynamics simulation. *Biophysical Journal* 1998, 74, 2906-2911
- [158] Rupp, B.: *Molecular-Modelling-Untersuchungen zu Wechselwirkungen in Protein-Protein- und Protein-Ligand-Komplexen*. Dissertation, Institut für Pharmazeutische Chemie, Heinrich-Heine-Universität Düsseldorf 2004
- [159] Fan H.; Mark, A.E.: Relative stability of protein structures determined by X-ray crystallography or NMR spectroscopy: a molecular dynamics simulation study. *Proteins* 2003, 52, 111-120
- [160] Roccatano, D.; Wong, T.S.; Schwaneberg, U.; Zacharias, M.: Structural and dynamic properties of cytochrome P450 BM-3 in pure water and in dimethylsulfoxide/water mixture. *Biopolymers* 2005 [Epub ahead of print]
- [161] SYBYL 7.0, Tripos Inc., St. Louis, MO, USA
- [162] Allen, F.H.; Kennard, O.; Taylor, R.: Systematic analysis of structural data as a research technique in organic chemistry. *Acc. Chem. Res.* 1983, 16, 146-153
- [163] Lewis, D.F.V.; Lake, B.G.; Dickins, M.; Ueng, Y.-F.; Goldfarb, P.S.: Homology modelling of human CYP1A2 based on the CYP2C5 crystallographic template structure. *Xenobiotica* 2003, 33, 239-254
- [164] Sotokawa, H.; Shimizu, T.; Furuya, H.; Sadeque, A.J.M.; Hatano, M.; Ohba, Y.; Iwaizumi, M.; Fujii-Kuriyama, Y.: Electron spin resonance studies of wild-type and mutant cytochromes P-450d: effects of mutations at proximal, aromatic and distal sites in g values. *Biochim. Biophys. Acta* 1990, 1037, 122-128
- [165] Wlodawer, A.; Walter, J.; Huber, R.; Sjölin, L.: Structure of bovine pancreatic trypsin inhibitor: results of joint neutron and x-ray refinement of crystal form II. *J. Mol. Biol.* 1984, 180, 301-329
- [166] Kirton, S.B.; Murray, C.W.; Verdonk, M.L.; Taylor, R.D.: Prediction of binding modes for ligands in the cytochromes P450 and other heme-containing proteins. *Proteins* 2005, 58, 836-844
- [167] Faßbender, M.J.: *Ein Pharmakophor-Modell für Substrate und Inhibitoren der menschlichen Cytochrom P450-Isoform 1A2*. Dissertation, Institut für Toxikologie und Umwelthygiene, Technische Universität München 1997

- [168] Korhonen, L.E.; Rahnasto, M.; Mähönen, N.J.; Wittekindt, C.; Poso, A.; Juvonen, R.O.; Raunio, H.: Predictive three-dimensional quantitative structure-activity relationship of cytochrome P450 1A2 inhibitors. *J. Med. Chem.* 2005, 48, 3808-3815
- [169] Szklarz, G.D.; Paulsen, M.D.: Molecular modelling of cytochrome P450 1A1: enzyme-substrate interactions and substrate binding affinities. *J. Biomol. Struct. Dyn.* 2002, 20, 155-162
- [170] Bathelt, C.; Schmid, R.D.; Pleiss, J.: Regioselectivity of CYP2B6: homology modelling, molecular dynamics simulations, docking. *J. Mol. Model.* 2002, 8, 327-335
- [171] Venhorst, J.; ter Laak, A.M.; Commandeur, J.N.M.; Funae, Y.; Hiroi, T.; Vermeulen, N.P.E.: Homology modeling of rat and human cytochrome P450 2D (CYP2D) isoforms and computational rationalization of experimental ligand-binding specificities. *J. Med. Chem.* 2003, 46, 74-86
- [172] Winn, P.J.; Lüdemann, S.K.; Gauges, R.; Lounnas, V.; Wade, R.C.: Comparison of the dynamics of substrate access channels in three cytochrome P450s reveals different opening mechanisms and a novel functional role for a buried arginine. *Proc. Natl. Acad. Sci. USA* 2002, 99, 5361-5366

A Abkürzungen und Einheiten

Abkürzungen

COMT	Catechol-O-methyltransferase
CVFF	Consistent Valence Force Field
CYP	Cytochrom P450
DFT	Dichtefunktionaltheorie
DMZ	(4 Methyl-N-methyl-N-(2-phenyl-2H-pyrazol-3-yl)benzensulfonamid)
EH	Epoxidhydrolase
EROD	Ethoxyresorufin-O-deethylase
et al.	et alii
FAD	Flavinadenindinucleotid
FFD	Fractional Factorial Design
FFT	Fast Fourier Transformation
FMN	Flavinmononucleotid
GIT	Gastrointestinaltrakt
GSHT	Glutathion-S-transferase
HAA	Heterozyklische aromatische Amine
HPLC	High Pressure Liquid Chromatography
MDS	Moleküldynamiksimulation
MeIQ	(2-Amino-3,4-dimethylimidazo[4,5-f]chinolin)
MEP	Molekulares Elektrostatisches Potential
MM	Molekülmechanik
MROD	Methoxyresorufin-O-deethylase
NAD(P)H	Nicotinamadenindinucleotid (-phosphat) (reduzierte Form)
NAT	N-Acetyltransferase
PAH	Polyaromatic Hydrocarbons
PBC	Periodic Boundary Conditions
PCA	Principal Component Analysis
PDB	Protein Data Bank
PLS	Partial Least Squares
PME	Particle Mesh Ewald
QM	Quantenmechanik
QSAR	Quantitative Structure-Activity-Relationship
RMS(D)	Root Mean Square (Deviation)
RMSF	Root Mean Square Fluctuation
SCR	Structurally Conserved Region
SDEP	Standard Deviation of the Error of Prediction
SPC	Single Point Charge
SRS	Substrate Recognition Site
SSRI	Selective Serotonin Reuptake Inhibitor
SULT	Sulfotransferase
SVR	Structurally Variable Region
TCDD	2,3,7,8-Tetrachlorodibenzo-p-dioxin
UV	Ultraviolettstrahlung

Einheiten

ns	Nanosekunde (1 ns = 10^{-9} s)
ps	Picosekunde (1 ps = 10^{-12} s)
fs	Femtosekunde (1 fs = 10^{-15} s)
nm	Nanometer (1 nm = 10^{-9} m)
Å	Ångström (1 Å = 10^{-10} m)
K	Kelvin (0 K = -273,15 °C)
kcal	Kilocalorie (1 kcal = 4,186 kJ)
kDa	Kilodalton

B Aminosäuren: Ein- und Dreibuchstabencode

Aminosäure	Dreibuchstabencode	Einbuchstabencode
neutrale		
Glycin	Gly	G
Alanin	Ala	A
Serin	Ser	S
Threonin	Thr	T
Valin	Val	V
Leucin	Leu	L
Isoleucin	Ile	I
Säuren und Amide		
Asparaginsäure	Asp	D
Asparagin	Asn	N
Glutaminsäure	Glu	E
Glutamin	Gln	Q
Basen		
Arginin	Arg	R
Lysin	Lys	K
schwefelhaltige		
Cystein	Cys	C
Methionin	Met	M
aromatische		
Phenylalanin	Phe	F
Tyrosin	Tyr	Y
heterozyklische		
Tryptophan	Trp	W
Histidin	His	H
Prolin	Pro	P

C Hardware

Silicon Graphics Workstations:

IRIS INDIGO2 R10000

IRIS O2 R12000

IRIS OCTANE R12000

Linux Personal Computer:

Dualprozessor Intel Xeon

Personal Computer:

AMD mobile Athlon XP-M 2500+

D Software

BLAST	Programme zum Vergleich von Nukleotid- und Proteinsequenzen
CLUSTALW	Programm zur Erstellung multipler Alignments
Con Quest	Software zur Struktursuche in der <i>Cambridge Structural Database</i>
FUGUE	Programm zur Bestimmung von Sequenz-Struktur-Homologien
GOLD	Programm für das Protein-Ligand-Docking basierend auf einem genetischen Algorithmus
GOLPE	Statistikprogramm zur Durchführung einer Hauptkomponentenanalyse mit anschließender PLS-Statistik
GRIN/ GRID	Programmkombination zur Berechnung molekularer Interaktionsfelder
GROMACS	Programmpaket zur Berechnung und Auswertung von Moleküldynamiksimulationen
INSIGHT II	Molecular Modelling Programmpaket der Firma Accelrys
JAGUAR	<i>ab initio</i> Programm der Firma SCHRÖDINGER zur Durchführung quantenchemischer Berechnungen
NMRCLUST	Programm zur Einteilung eines Ensembles von Proteinstrukturen in Gruppen strukturell ähnlicher Konformationen
PROCHECK	Programm zur Überprüfung der Proteingeometrie
PSIPRED	Programm zur Vorhersage der Sekundärstruktur von Proteinen
SURFNET	Programm zur Berechnung und Visualisierung von Hohlräumen in Proteinen
SYBYL	Molecular Modelling Programmpaket der Firma TRIPOS
VMD	Programm zur Visualisierung von Moleküldynamiksimulationen

E Eingabedateien

GROMACS-Eingabedatei für die freie MDS des leeren Proteinmodells. Für alle übrigen MDS wurde prinzipiell das gleiche Simulationsprotokoll verwendet.

```
; VARIOUS PREPROCESSING OPTIONS =
title           = Molekueledynamik
cpp             = /lib/cpp
include        = -I/home/andrea/RECHNUNGEN/GROMACS/KFP/GROMOS_HEM
define         =

; RUN CONTROL PARAMETERS =
integrator     = md
; start time and timestep in ps =
tinit         = 0
dt            = 0.002
nsteps        = 1250000
; number of steps for center of mass motion removal =
nstcomm       = 1
comm-grps     =

; LANGEVIN DYNAMICS OPTIONS =
; Temperature, friction coefficient (amu/ps) and random seed =
bd-temp       = 310
bd-fric       = 0
ld-seed       = 1993

; ENERGY MINIMIZATION OPTIONS =
; Force tolerance and initial step-size =
emtol         = 1000
emstep        = 0.01
; Max number of iterations in relax_shells =
niter         = 0
; Frequency of steepest descents steps when doing CG =
nstcgsteep    = 1000

; OUTPUT CONTROL OPTIONS =
; Output frequency for coords (x), velocities (v) and forces (f) =
nstxout       = 2500
nstvout       = 2500
nstfout       = 0
; Output frequency for energies to log file and energy file =
nstlog        = 2500
nstenergy     = 50
; Output frequency and precision for xtc file =
nstxtcout     = 2500
xtc-precision = 5000
; This selects the subset of atoms for the xtc file. You can =
; select multiple groups. By default all atoms will be written. =
xtc_grps      = Protein HEME
; Selection of energy groups =
energygrps    = Protein HEME SOL NA+ CL-
```

```
; NEIGHBORSEARCHING PARAMETERS =
; nblast update frequency =
nstlist           = 10
; ns algorithm (simple or grid) =
ns_type          = grid
; Periodic boundary conditions: xyz or none =
pbc              = xyz
; nblast cut-off           =
rlist            = 0.9
domain-decomposition = no

; OPTIONS FOR ELECTROSTATICS AND VDW =
; Method for doing electrostatics =
coulombtype      = PME
rcoulomb-switch  = 0
rcoulomb         = 0.9
; Dielectric constant (DC) for cut-off or DC of reaction field =
epsilon-r       = 1
; Method for doing Van der Waals =
vdw-type        = Cut-off
; cut-off lengths           =
rvdw-switch     = 0
rvdw            = 0.9
; Apply long range dispersion corrections for Energy and Pressure =
DispCorr        = No
; Spacing for the PME/PPPM FFT grid =
fourierspacing  = 0.12
; FFT grid size, when a value is 0 fourierspacing will be used =
fourier_nx      = 0
fourier_ny      = 0
fourier_nz      = 0
; EWALD/PME/PPPM parameters =
pme_order       = 4
ewald_rtol      = 1e-05
epsilon_surface = 0
optimize_fft    = yes

; OPTIONS FOR WEAK COUPLING ALGORITHMS =
; Temperature coupling =
tcoupl          = Berendsen
; Groups to couple separately =
tc-grps        = Protein HEME SOL NA+ CL-
; Time constant (ps) and reference temperature (K) =
tau_t          = 0.1 0.1 0.1 0.1 0.1
ref_t          = 310 310 310 310 310
; Pressure coupling =
Pcoupl         = Berendsen
Pcoupltype     = Isotropic
; Time constant (ps), compressibility (1/bar) and reference P (bar) =
tau_p          = 1.0
compressibility = 4.5e-5
ref_p          = 1.0

; GENERATE VELOCITIES FOR STARTUP RUN =
gen_vel        = no
gen_temp       = 310
gen_seed       = 173529
```

```
; OPTIONS FOR BONDS      =
constraints              = all-bonds
; Type of constraint algorithm =
constraint-algorithm     = Lincs
; Do not constrain the start configuration =
unconstrained-start     = no
; Relative tolerance of shake =
shake-tol               = 0.0001
; Highest order in the expansion of the constraint coupling matrix =
lincs-order             = 4
; Lincs will write a warning to the stderr if in one step a bond =
; rotates over more degrees than =
lincs-warnangle         = 30
```

An dieser Stelle möchte ich mich bei einigen Menschen persönlich bedanken, die besonders zum Gelingen dieser Arbeit beigetragen haben.

Von den derzeitigen und ehemaligen Mitgliedern des Arbeitskreises sind dies zunächst Gunther Stahl und Bernd Rupp, die ihr gesamtes Wissen über Cytochrome und Homologie-Modelling an mich weitergegeben und mir damit den Einstieg in meine Arbeit erleichtert haben. Céline Anézo hat mir zahlreiche Tricks zur Anwendung von GROMACS verraten und mir gezeigt, dass man dieses Programm nicht nur anwenden, sondern auch verstehen kann. In Birgit Schlegel hatte ich schließlich eine wertvolle Leidensgenossin in der Endphase meiner Promotion. Durch den Erfahrungsaustausch mit ihr bewahrheitete sich einmal mehr die Lebensweisheit, dass geteiltes Leid halbes Leid ist.

Unserem Systemadministrator Peter Sippel danke ich für die zuverlässige technische Unterstützung und seine geduldige Hilfe bei allen computertechnischen Problemen, die bis hin zu spektakulären Datenrettungsaktionen reichte, als kurz vor Abgabe der Arbeit eine Festplatte ihren Geist aufgab.

Für die angenehme und unkomplizierte Zusammenarbeit im Praktikum danke ich meinen Kollegen Lothar Schenk und Holger Steinmaß.

Ein besonderes Dankeschön geht an die Mitglieder der legendären Teerunde Andrea Vollbrecht, Dirk Claßen-Houben und Kirstin Jöhren. Bei der täglichen Tasse Tee („täglich“, sofern ich nicht durch das Aufsetzen einer Minimierung verhindert war) ergaben sich neben Gesprächen über das Leben im Allgemeinen durchaus auch fachliche Diskussionen mit immer wieder wertvollen Anregungen. Für Abwechslung sorgten Weihnachtsmarktbesuche, Kochorgien mit rheinisch-bergischen bis hin zu fernöstlichen Schlemmereien und schließlich der ein oder andere Segelausflug, da ich mich eurer Segelleidenschaft nie ganz entziehen konnte. Vielen Dank auch für das schnelle und sorgfältige Korrekturlesen der Arbeit und für eure hartnäckige Unterstützung in den letzten Monaten meiner Promotion. Ich freue mich, dass trotz teilweise großer räumlicher Entfernungen der Kontakt erhalten geblieben ist, und hoffe, dass das auch so bleibt!

Darüber hinaus danke ich den Mitgliedern des Streichquartetts für viele schöne Kammermusikabende und -wochenenden und meinen Freunden aus der Schul- und Studienzeit für zahlreiche gemeinsame Unternehmungen, mit denen sie mir immer wieder gezeigt haben, dass das Leben nicht nur aus Proteinmodellen von Cytochromen besteht.

Selbstverständlich haben auch meine Eltern einen wesentlichen Anteil am Zustandekommen dieser Arbeit, denn sie haben mir die nötigen Voraussetzungen dafür mit auf den Weg gegeben. Darüber hinaus danke ich ihnen für ihre stetige Unterstützung in jeder Hinsicht und für ihre Anteilnahme an meinem Werdegang.

Daniel Rütten danke ich zunächst für die fachkundige sprachliche Korrektur meiner Arbeit. Vor allem aber danke ich ihm für seine Geduld und seine unermüdliche Unterstützung. Du hast mich immer wieder ermutigt, und Du wusstest stets genau, wann bei mir der Punkt erreicht ist, an dem nur noch ein großes Eis mit Sahne weiterhilft.

Name Andrea Hombrecher

geboren am 19.08.1975

in Wermelskirchen

Schulbildung

1982 – 1986 Dörfelschule Wermelskirchen

1986 – 1995 Städtisches Gymnasium Wermelskirchen

1995 Abitur

Studium

10/1995 – 10/1999 Pharmaziestudium an der Heinrich-Heine-Universität
Düsseldorf

09/1999 Zweiter Abschnitt der Pharmazeutischen Prüfung

11/1999 – 04/2000 Pharmaziepraktikum in der Äskulap-Apotheke, Aachen

05/2000 – 10/2000 Pharmaziepraktikum bei der Bayer AG, Leverkusen

12/2000 Approbation als Apothekerin

Berufserfahrung

01/2001 – 03/2001 Auslandspraktikum bei PT Bayer Indonesia, Jakarta,
Indonesien

Wissenschaftliche Tätigkeit

04/2001 Beginn der vorliegenden Dissertation am Institut für
Pharmazeutische Chemie der Heinrich-Heine-Universität
Düsseldorf unter Anleitung von Prof. Dr. H.-D. Höltje

06 – 07/2004 Forschungsaufenthalt in der Arbeitsgruppe von Prof. Dr.
Antti Poso am Institut für Pharmazeutische Chemie der
Universität Kuopio, Finnland

