

Towards Improved Aldose Reductase Inhibitors

–

Structural and Thermodynamic
Investigation of Mutant and Wild Type
Aldose Reductase Inhibitor Complexes

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Pharmazie der
Philipps-Universität Marburg
vorgelegt

von
Cornelia Koch
aus
Erfurt

Marburg/Lahn 2011

Vom Fachbereich Pharmazie der Philipps-Universität Marburg
als Dissertation angenommen am: 30.05.2011

Erstgutachter: Prof. Dr. Gerhard Klebe
Zweitgutachter: Dr. Andreas Heine

Tag der mündlichen Prüfung: 31.05.2011

Contents

1	Introduction and Aim	1
1.1	Drug Design	2
1.2	Aldose Reductase as Target	4
1.2.1	Aldose Reductase as Target	4
1.2.2	Selectivity against Aldehyde Reductase	6
1.2.3	Flexibility as Challenge	7
1.3	Diabetes	8
1.3.1	Diabetic Disease	9
1.3.2	Epidemiology	10
1.3.3	Long Term Diabetic Complications	10
1.3.4	Treatment	12
1.4	Aim of the Thesis	13
2	Improving Known Inhibitors	17
2.1	Benzothiazepines	17
2.1.1	Introduction	17
2.1.2	Results and Discussion	18
2.1.3	Conclusion	25
2.2	Aromatic Carboxylic Acid Derivatives	26
2.2.1	Introduction	26
2.2.2	Kinetic Evaluation of Further Variations	28
2.2.3	Results	29
3	Protein Flexibility	31
3.1	Binding Mode Affects Crystal Packing	31
3.1.1	Introduction	31
3.1.2	Results	33
3.1.3	Discussion	45

3.1.4	Conclusion	48
3.2	Stabilizing a New Protein Conformation	49
3.2.1	Introduction	49
3.2.2	Suggested Compounds and Evaluation	49
3.2.3	Conclusion	51
4	Aldose Reductase Mutants	53
4.1	Introduction	53
4.2	Results	55
4.2.1	Crystal Structures of the Mutant IDD388 Complexes	55
4.2.2	Crystal Structures of the Mutant IDD594 Complexes	59
4.2.3	Bromine Cleavage Due to Radiation Damage	60
4.2.4	Characterization of the Thermodynamic Binding Profile	61
4.3	Conclusions	68
5	Radiation Damage	69
5.1	Introduction	69
5.2	Results	70
5.2.1	Strategy	70
5.2.2	Radiation Induced Bromine Cleavage	71
5.2.3	Second Bromine Site Indicates Additional Interaction Site	74
5.3	Discussion	77
6	Materials and Methods	83
6.1	<i>in-silico</i> -Methods	83
6.1.1	Docking with GOLD	83
6.1.2	Docking with AutoDock	83
6.1.3	CSD Searches	83
6.1.4	Volume Calculations	84
6.1.5	Calculation of Protein-Protein Interfaces	84
6.2	Experimental Methods	84
6.2.1	Crystallography	85
6.2.2	Isothermal Titration Calorimetry	87
6.2.3	Plasmid Preparation and Transformation	89
6.2.4	Expression and Purification	89
6.2.5	Mutagenesis	91
6.2.6	Kinetic Assay	92

6.3	Materials	93
6.3.1	Substances	93
6.3.2	Solutions and Buffers	95
6.3.3	Laboratory Devices	97
6.3.4	Vector System	97
6.3.5	Bacterial Strains	97
6.3.6	Primer Material	98
A	Crystallographic Tables	101
	Curriculum vitae	117
	Declaration	119

List of Figures

1.1	Drug design cycle	3
1.2	Polyol pathway	4
1.3	Secondary structure of aldose reductase	6
1.4	Catalytic site of aldose reductase	7
1.5	Diabetes worldwide	8
1.6	Insulin production and action	9
1.7	Map of prevalence	11
2.1	Schematic view of 1	18
2.2	Schematic view of the binding mode of 1	19
2.3	hAR complex with 1	20
2.4	Derivatives of benzothiazole	22
2.5	Redocking results	23
2.6	Best docking results of 1 derivatives	24
2.7	A selection of virtual screening results	27
2.8	Schematic view of oxobutaoic, acrylic and triarylic acid compounds	28
2.9	Triarylic acid from virtual screening	30
3.1	Schematic view of 1 , 6 , and 7 and the binding mode	34
3.2	Mutant- 6 complexes compared with wild type	35
3.3	Details of 6 -binding in the mutant series	36
3.4	Two V- 7 complexes crystallized in different space groups	38
3.5	Conformational variation of loop 119–135	40
3.6	Crystal packing in P 2 ₁ and P 1	42
3.7	Crystal contacts of helical segment 283–286	43
3.8	Overview of the thermodynamic factorization	44
3.9	Pharmacophore	49
3.10	Hits of virtual screening	50

3.11	Virtual screening hit in crystal complex with hAR	51
4.1	Schematic view of 12 and 13	55
4.2	Schematic binding mode	56
4.3	A- 12 and WT- 12 complex in superposition	57
4.4	Mutants 12 complexes	58
4.5	Alanine mutants with 12 and 13	60
4.6	ITC data of hAR mutants with 12 and 13	62
5.1	Position of 13 in the binding pocket	72
5.2	Bromine site occupancy	75
5.3	Two bromine sites	78
5.4	Histograms of bromine contacts	79
5.5	Distances of the cleaved bromide ion to surrounding residues	79
5.6	Free volume occupied by second bromine	80

List of Tables

1.1	Summary of the benefits of diabetic drugs	13
2.1	IC ₅₀ values for 1 derivatives	24
2.2	Virtual screening hits	27
3.1	Overview of complex conformers.	39
3.2	Crystal contacts in P 2 ₁ and P 1	41
3.3	Thermodynamic data of the different complexes.	45
4.1	Distances of bromine atom to residue 113	61
4.2	ITC data of the complexes	63
4.3	Relative differences of ITC data for 12	64
4.4	Relative differences of ITC data for 13	64
4.5	Relative differences of ITC data for 12 to 13	64
5.1	Occupancy decrease of bromine sites from structural refinement . . .	74
5.2	Occupancy decrease of bromine sites from initial anomalous maps calculated by SHELXE	74
5.3	Experimental setup and estimated dose values	76
5.4	Occupancy decrease of bromine sites from initial anomalous maps calculated by SHELXE ¹	76
6.1	Substances	93
6.2	Buffers and media	95
6.3	Laboratory devices	98
6.4	Bacterial strains	98
6.5	Primer material	99
A.1	Data processing statistics for subset 1-3 of Set A	101
A.2	Data processing and refinement statistics	102

Danksagung

Vielen Dank an all die Menschen, die mir bei der Erstellung der Arbeit auf unterschiedliche Weise beigestanden haben.

Mein besonderer Dank gilt Prof. Dr. Gerhard Klebe, der diese Arbeit angeregt, begleitet, stetig verfolgt und vielseitig unterstützt und gefördert hat.

Ich danke Dr. Andreas Heine für umfangreichen fachlichen Rat bei experimentellen Arbeiten, bei deren Auswertung und schließlich der Veröffentlichung der Ergebnisse.

Außerdem gilt mein Dank Prof. Dr. Lars Oliver Essen, der durch konstruktive Anregungen und Diskussionen im Kristallographieseminar zum Ergebnis der Arbeit beigetragen hat, sowie Prof. Dr. Alberto Podjarny und dessen Arbeitsgruppe; insbesondere Isabelle Hazemann, die mir Protein, Kristalle, Erfahrung und Wissen zukommen ließ.

Alle Mitglieder der Arbeitsgruppe Klebe haben meine Zeit in Marburg unvergesslich gemacht; besonders bereichert haben mich Gerd Neudert, Tina Ritschel, Bernhard Baum, Lisa Englert, Nils Weskamp, Ina Lindemann.

Auf diesem Weg haben meine Eltern Carmen und Karl-Wilhelm Koch mich begleitet und stets unterstützt, vielen Dank dafür.

Publications arising from this work

Parts of this work will be or have already been published: Sections 2.1 and 3.1 as well as Chapters 4 and 5 build manuscripts for publications. Therefore, each of these parts comprise another short introduction to the matter described.

Articles

Koch, C., Heine, A., Klebe, G.. Tracing the Detail: How Mutations Affect Binding Modes and Thermodynamic Signatures of Closely Related Aldose Reductase Inhibitors. *J. Mol. Biol.* 2011; **406**(5):700–712.

Koch, C., Heine, A., Klebe, G.. Radiation Damage Reveals Promising Interaction Position; Submitted to *Journal of Synchrotron Radiation*.

Koch, C., Heine, A., Klebe, G.. Ligand-induced Fit Affects Binding Modes and Provokes Changes in Crystal Packing of Aldose Reductase. Submitted to *Biochimica et Biophysica Acta*.

Koch, C., Platen, T., Heine, A., Diederich, W.E., Klebe, G.. Benzothiazepines as Novel Scaffolds Inhibiting Aldose Reductase. In preparation.

Zentgraf, M., Steuber, S., Koch, C., La Motta, C., Sartini, S., Sotriffer, C.A., Klebe, G.. How Reliable Are Current Docking Approaches for Structure-Based Drug Design? Lessons from Aldose Reductase. *Angew. Chem. Int. Ed.* 2007; **46**:3575–3578.

Posters

Koch, C., Klebe, G.. Investigation of Selectivity Features of Aldose Reductase. Gordon Research Conferences - Computer Aided Drug Design, Tilton, NH, USA; 2009.

Koch, C., Platen, T., Diederich, W.E., Klebe, G.. Benzothiazepines as Novel Compounds Inhibiting Aldose Reductase. Biophysics of ligands binding to drug targets, Strasbourg, France; 2007.

Koch, C., Steuber, H., Heine, A., Klebe, G.. Structural and Thermodynamic Characterization of Novel Aldose Reductase Inhibitors. Summer School Medicinal Chemistry, Regensburg, Germany; 2006.

Abbreviations

Å	- Ångström ($1\text{Å} = 10^{-10}\text{ m}$)
Ala	- Alanine
Arg	- Arginine
Amp	- Ampicillin
Asn	- Asparagine
Asp	- Aspartic acid
B-factor	- Debye-Waller-factor
CSD	- Cambridge Structural Database
Cys	- Cysteine
DMSO	- Dimethyl sulfoxide
dNTP	- Deoxynucleoside triphosphate
DTT	- Dithiothreitol
<i>E. coli</i>	- Escherichia coli
F_c	- Calculated structure amplitudes
F_o	- Observed structure amplitudes
Gln	- Glutamine
Glu	- Glutamic acid
Gly	- Glycine
Gy	- Gray ($1\text{ Gy} = 1\text{ J/kg} = 1\text{ m}^2/\text{s}^2$)
h	- Hour
H-bond	- Hydrogen bond
hAR	- human Aldose Reductase
His	- Histidine
HPLC-MS	- High pressure liquid chromatography-mass spectrometry
IDF	- International Diabetes Federation
Ile	- Isoleucine
IPTG	- Isopropylthio- β -galactosid
ITC	- Isothermal Titration Calorimetry
I.U.	- International unit
K	- Kelvin
k	- Kilo
Da	- Dalton
K_i	- Competitive inhibition constant
K_m	- Michaelis Menten constant

Leu	- Leucine
Lys	- Lysine
M	- Molarity (mol/l)
Met	- Methionine
Min	- Minute
NMR	- Nuclear magnetic resonance
OHA	- Oral hypoglycemic agent
PCR	- Polymerase chain reaction
PDB	- Protein Data Bank
pH	- Potentialis hydrogenii
Phe	- Phenylalanine
Pro	- Proline
R-factor	- Reliability factor
RMSD	- Root mean square deviation
rpm	- Revolutions per minute
SAD/MAD	- Single/Multiple wavelength anomalous dispersion
SDS	- Sodium dodecyl sulfate
Ser	- Serine
Thr	- Threonine
TRIS	- Tris(hydroxymethyl)-aminomethane
Trp	- Tryptophane
Tyr	- Tyrosine
Val	- Valine
w/v	- Weight per volume
WT	- Wild type

Summary

Rational drug design for flexible proteins like human aldose reductase comprises special challenges to find novel lead compounds or to improve known inhibitors. In this thesis, diverse aspects of the complexity of such a task were investigated to find promising new lead scaffolds and enhance the affinity of known aldose reductase inhibitors.

Within different design cycles, the prediction of ligand binding modes in a binding pocket capable to adapt to a bound ligand is such a challenging task. A benzothiazole scaffold originating from a virtual screening run was kinetically and structurally evaluated. An unexpected binding mode in complex with wild-type aldose reductase and a new protein conformer provided the basis for a further drug design cycle to optimize the scaffold. The benzothiazole core was expanded to explicitly address an additional subpocket of the protein to enhance both affinity and selectivity over a closely related protein. IC_{50} values for the derivatives along with computationally generated binding geometries provide a good foundation for further design cycles with this novel scaffold.

Flexibility is a common feature of many proteins. For human aldose reductase, a variety of conformers are adopted with different inhibitors. In Chapter 3, the known inhibitors zopolrestat and IDD 393 each in complex with a threonine 113 mutant and a benzothiazole inhibitor in complex with the wild type of aldose reductase were investigated with respect to the impact of inhibitor binding on protein conformation. Though the interaction to the mutated residue does not directly alter the binding mode of zopolrestat, a shift of its basic scaffold is induced and subsequently affects the interaction to a flexible loop and introduces disorder. With IDD 393, two distinct binding site conformations resulting in different crystal forms can be directly compared: While a backbone flip of the same residues as for zopolrestat is present in both crystal forms, a considerable side chain movement of a phenylalanine is observed for only one crystal form. In consequence, the residual mobility of further amino acids is increased and some crystal contacts are prevented which in

consequence reinforces a different crystal packing. The structure of the aforementioned benzothiazole inhibitor reveals a novel protein conformer, where the same phenylalanine is even further relocated resulting in the altered crystal packing.

Improvements of the computational methods for affinity prediction from the structure of protein-ligand complexes requires a better understanding of the nature of molecular interactions and biomolecular recognition principles. In Chapter 4, the binding of two chemically closely related human aldose reductase inhibitors have been studied by high resolution X-ray analysis ($0.92 \text{ \AA} - 1.35 \text{ \AA}$) and isothermal titration calorimetry against a series of single-site mutants of the wild-type protein. A crucial active site threonine, thought to be involved in a short bromine-to-oxygen halogen bond to the inhibitors in the wild type has been mutated to the structurally similar residues alanine, cysteine, serine, and valine. Overall, structurally the binding mode of the inhibitors is conserved; however, small but significant geometrical adaptations are observed as a consequence of the spatial and electronic changes at the mutation site. They involve the opening of a central bond angle and shifts in consequence of the lost or gained halogen bonds. Remarkably, the tiny structural changes are responded by partly strong modulation of the thermodynamic profiles. Even though the free energy of binding is maximally perturbed by only 7 kJ/mol , much stronger modulations and shifts in the enthalpy and entropy signature are revealed which indicate a pronounced enthalpy/entropy compensation. However, facing these perturbances against the small structural changes, an explanatory correlation can be detected. This also provides deeper insights how single-site mutations can alter the selectivity profile of closely related ligands against a target protein.

High resolution structural data of protein inhibitor complexes is the key to rational drug design. Synchrotron radiation allows for atomic resolution but is frequently accompanied by radiation damage to protein complexes. In Chapter 5, a human aldose reductase mutant complexed with a bromine substituted inhibitor was determined to atomic resolution. Though the radiation dose was moderate, a selective disruption of a bromine-inhibitor bond during the experiment was observed while the protein appears unaffected. A covalent bond to bromine is cleaved and the displaced atom is not scattered throughout the crystal but can most likely be assigned as a bromide ion to an additional difference electron density peak observed in the structure: The bromide relocates to an adjacent unoccupied site where promising interactions to protein residues stabilize its position. These findings were verified by a second similar structure determined with considerably higher radiation dose.

Zusammenfassung

Rationales Wirkstoffdesign für flexible Proteine wie die humane Aldosereduktase ist mit verschiedenen Herausforderungen verbunden. Um neue Leitstrukturen zu finden und die Affinität bekannter Inhibitoren zu verbessern wurden innerhalb dieser Arbeit verschiedene Methoden angewendet, und dabei die unterschiedlichen Aspekte und deren Komplexität beleuchtet.

Innerhalb aufeinanderfolgender Designzyklen ist die Voraussage von Bindungsmodi eines Liganden in einer anpassungsfähigen Bindetasche ein sehr komplexes Problem. Ein Benzothiazepin, das in einem vorhergehenden virtuellen Screening gefunden worden war, wurde kinetisch und strukturell im Komplex mit Aldosereduktase charakterisiert. Der dabei beobachtete, unerwartete Bindungsmodus mit einem neuen, bisher unbekannten Proteinkonformer stellte die Basis für einen weiteren Designzyklus dar. Um das Grundgerüst zu optimieren, wurden Substituenten eingeführt, die eine Subtasche des Proteins direkt adressieren und neben der Affinität auch die Selektivität des Liganden gegenüber einem nahe verwandten Protein erhöhen sollten. Die gemessenen IC_{50} -Werte sowie computergenerierte Bindungsgeometrien der Derivate stellen nun eine Grundlage für weitere Optimierungszyklen des Benzothiazepingrundgerüsts dar.

Proteinflexibilität ist eine verbreitete Eigenschaft in Proteinen verschiedenster Funktionalität. Von Aldosereduktase sind eine Reihe unterschiedlicher Konformere bekannt, die im Komplex mit verschiedenen Bindungspartnern eingenommen werden. In einem Teil der Arbeit wurde der Einfluss von Inhibitoren auf die Ausbildung der verschiedenen Bindetaschenkonformere anhand von Zopolrestat und IDD 393 im Komplex mit einer Threonin-113-Mutante sowie einem Benzothiazepin-Inhibitor im Komplex mit der Wildtypvariante von Aldosereduktase untersucht. Obwohl die direkte Interaktion zwischen Zopolrestat und der mutierten Aminosäure den Bindungsmodus nicht verändert, bewirkt sie eine Verschiebung des Inhibitorgrundgerüsts. Dies beeinflusst die Wechselwirkung mit einem benachbarten, beweglichen Loop, wodurch dieser weniger fixiert ist und Unordnung erfährt. Im Komplex mit

IDD 393 konnten zwei abweichende Bindetaschenkonformationen gefunden werden. Diese führten zu unterschiedlichen Kristallformen und konnten direkt miteinander verglichen werden: Während ein Flip des Proteinrückrats bestimmter Reste, wie auch im Zopolrestatkomplex schon beobachtet worden war, in beiden Kristallformen beobachtet wurde, tritt nur in einer der beiden Formen eine ausgeprägte Seitenkettenrotation eines Phenylalanins auf. Dadurch erhöht sich die Beweglichkeit weiterer Aminosäuren, wodurch einige Kristallkontakte verhindert werden. Dadurch wird eine neue Kristallpackung erzwungen. Im Proteinkomplex mit dem Benzothiazepin wird die Konformation des Phenylalanins in noch größerem Ausmaß verändert und demzufolge die veränderte Packung induziert.

Die Verbesserung computergestützter Methoden zur Affinitätsvorhersage aus Protein-Ligand-Komplexen verlangt nach einem detaillierten Wissen um die Natur molekularer Wechselwirkungen und biomolekularer Erkennungsprinzipien. Kapitel 4 beschäftigt sich mit der Bindung zweier chemisch sehr ähnlicher Aldosereduktaseinhibitoren. Hochaufgelöste Röntgenstrukturen ($0,92 - 1,35 \text{ \AA}$) und kalorimetrische Bindungsdaten dieser Inhibitoren im Komplex mit einer Serie von Bindetaschenmutanten der Aldosereduktase wurden analysiert. Ein entscheidendes Threonin, das vermutlich an einem kurzen Brom-Sauerstoff-Kontakt, einer sogenannten Halogenbindung, beteiligt ist, wurde durch die strukturell ähnlichen Aminosäuren Alanin, Cystein, Serin und Valin ausgetauscht. Während sich der Bindungsmodus, oberflächlich betrachtet, nicht verändert, sind kleinste, aber signifikante geometrische Anpassungen erkennbar, die das Resultat der räumlichen und elektrostatischen Veränderungen durch die Mutation sind. Dazu gehören die Aufweitung eines zentralen Bindungswinkels im Inhibitor sowie die Verschiebung des Inhibitorgerüsts durch den Verlust oder die verstärkte Halogenbindung. Bemerkenswert ist die Übertragung dieser kleinen geometrischen Änderungen in starke Änderungen der thermodynamischen Bindungsprofile. Obwohl die freie Bindungsenergie sich insgesamt um nur etwa 7 kJ/mol ändert, lassen sich deutlich stärkere Anpassungen in den enthalpischen und entropischen Bindungsbeiträgen erfassen. Eine ausgeprägte Enthalpie-Entropie-Kompensation wird deshalb vermutet. Diese Effekte lassen sich durch die strukturellen Änderungen erklären. Diese Erkenntnisse liefern Hinweise, wie Mutationen das Selektivitätsprofil von chemisch ähnlichen Liganden beeinflussen können.

Hochaufgelöste strukturelle Daten von Protein-Inhibitor-Komplexen bilden die Basis für wissensbasiertes Wirkstoffdesign. Die Verwendung von Synchrotronstrahlung ermöglicht atomare Auflösungen solcher Komplexstrukturen, ist allerdings mit der Gefahr von Strahlungsschäden an Protein oder Inhibitor verbunden. In einem

weiteren Teil der Arbeit wurde die Struktur eines bromsubstituierten Inhibitors im Komplex mit einer Aldosereduktase-Mutante mit atomarer Auflösung bestimmt. Obwohl eine moderate Strahlungs-dosis während des Experiments vorlag und kein Einfluss auf Proteinreste zu erkennen war, wurde der Bromsubstituent selektiv abgespalten. Die kovalente Bindung zum Brom wurde zerstört, wobei das abgespaltene Atom wider Erwarten wahrscheinlich als Bromid einem zusätzlichen Elektronendichtepeak zugeordnet werden konnte: Das Bromid nimmt eine neue Position in der Nähe ein, die durch vielversprechende Interaktionen mit benachbarten Aminosäuren stabilisiert wird. Durch eine zweite Struktur des Komplexes, die unter ähnlichen experimentellen Bedingungen aufgenommen wurde, konnten diese Ergebnisse verifiziert werden.

1. Introduction and Aim

A drug that is said to have no side effects, is very likely to have no effect at all, as Gustav Kuschinski suspected.* This rather short but very adequate statement implicates already some of the major obstacles modern researchers have to face on the way to novel drugs. First of all, efficacy is the key aim when searching for a new drug. It is also a key challenge as a drug binding to its target is affected by a number of environmental often unforeseen influences. Additionally, binding to the correct target is of great importance. Regarding protein families whose members are spread throughout different tissues in the organism, selectivity and thus, a drug without—or at least with as few as possible—side effects is hard to develop. By far not all possible targets and proteins are known to the structural level or have even been identified, which leaves a high probability to hit an undesired protein. An additional aspect is implicated by Kuschinski’s statement: Though the absence of side effects can hardly be achieved, it is a deep desire of patients, physicians and researchers to decrease them to a minimum. Thus, known drugs continuously are object of further improvement. Yet, improvement of a drug comprises the improvement of complex properties such as availability and distribution issues, as well as correct adjustment of physico-chemical properties. On the atomic level, interactions comprise complex dependencies in the binding process itself.² Despite more than 400 years of experience and continuous and successful developments in pharmaceutical research, the field of structure-based drug design appears to raise new questions with every novel research finding.

A futuristic vision of the way how to discover a novel drug includes that, providing the disease and the target to be addressed are known, a perfect matching ligand can be found or designed. Ideally, without further improvement cycles, efficacy would be proofed and the new drug could immediately be tested in clinical studies. Presently, each of these steps comprises a large number of challenges and unknown variables, but each step to a better understanding of any of these aspects hopefully draws a more accurate picture.

*"Wenn behauptet wird, dass eine Substanz keine Nebenwirkungen zeigt, so besteht der Verdacht, dass sie auch keine Hauptwirkung hat." Gustav Kuschinski, pharmacologist, 1904–1992

1.1 Drug Design

The "design" of a drug includes more than the simple search for a moderate or good binder to a certain target. "Design" implies the creation of a molecule that is specifically adapted to given features, that exploits the given opportunities best and perfectly fits to the desired target while interference with others is decreased to a minimum. This "design" will also be optimized regarding features apart from the binding process such as solubility and stability etc.

Recent attempts in drug design follow different structure-based strategies. They highly depend on the degree of knowledge about either the target protein—e.g. its family, relationships, and structure—or established drugs or ligands. Target-based strategies are distinguished from ligand-based approaches, each of which bear their advantages and disadvantages.

Knowing the disease does not necessarily lead to knowledge about the target. In several cases, there may not even exist an ideal target as several proteins may affect development or symptoms of the disease to deviating amounts. But even if a target protein is identified, knowledge about such a target protein reaches from a rough estimation about its functionality or structural features to detailed studies about activity determining factors and structural information with substrates and inhibitors. Even in these optimal cases, creating a molecule that perfectly matches the interaction pattern of the known protein is far away from being a trivial problem. Binding energies, and thus binding affinities, are influenced by direct atom-atom interactions between ligand and protein.^{3;4} Additionally, the presence of water as solvent usually increases the complexity of the equation.^{5;6} Taking into account that both ligand and protein are flexible and bear a certain degree of freedom that might be lost adds another level. Examples are known where further design cycles did not lead to the predicted improvement, but revealed unexpected changes in the system which were neglected or could have been hardly anticipated.

A broad range of diverse methods is widely applied in molecular design. High throughput screening (HTS) methods test a large number of compounds for activity on diverse targets by means of different biochemical assays or biophysical evaluation techniques. Compounds for testing cover either certain chemotypes for a distinct target, or large and chemically diverse libraries of synthetic or natural molecules are screened to find new lead compounds. Screening can also be performed *in silico* and an enormous amount of chemicals can be evaluated computationally. Afterwards, the proposed compounds can be studied in many different ways: Biochemical assays

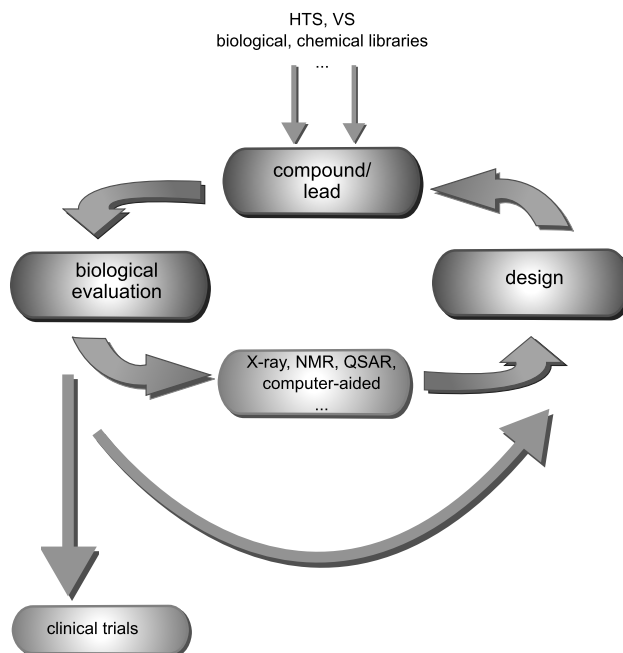


Figure 1.1: Drug design consists of a varying number of cycles, where the initial lead is evaluated and continuously improved according to e.g. structural analysis of the binding mode. Diverse experimental methods, such as X-ray and ITC, and computational methods, such as virtual screening, docking, and scoring, contribute to different steps. Figure according to Klebe⁷.

may reveal whether or not a compound is active at all. More sophisticated methods like isothermal titration calorimetry (ITC) can give deeper insights concerning the enthalpic and entropic contributions to binding. Crystallographic studies usually identify moderate to highly potent binders and provide particular information about the binding mode.

In addition to experimental evaluation, computational methods have a growing impact. They are steadily developed and improved. The possibility of identifying a new drug molecule without particular efforts in time and material qualifies them as an attractive research field. Presently, docking methods for generating ligand poses and scoring methods which evaluate the created geometries are effective and provide promising results in a number of examples, but still have their limitations.⁸

Nowadays, drug design usually involves several cycles of iterative design (see Fig. 1.1) where the identification of a promising compound is followed by its evaluation. This characterization implies improvement by structural variation of the lead compound, using a selection of the above-mentioned methods, both experimentally and computationally. Resulting compounds are re-evaluated, until the affinity

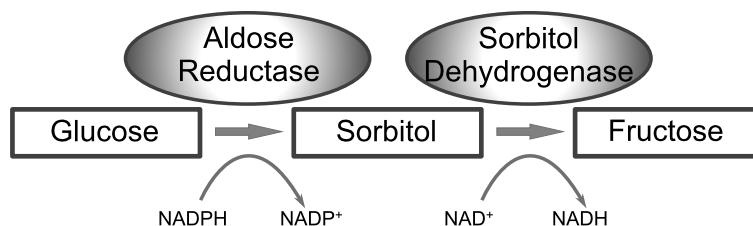


Figure 1.2: hAR catalyzes the first step in the polyol pathway. Glucose is reduced to sorbitol using NADPH. In a second step, sorbitol dehydrogenase converts sorbitol to fructose.

satisfies the expectations.

1.2 Aldose Reductase as Target in Drug Design

1.2.1 Aldose Reductase as Target

1.2.1.1 Location and Function

Human aldose reductase (hAR) is a member of the aldo-keto-reductase superfamily (AKR1 member B1, EC 1.1.1.21).⁹ This cytosolic enzyme is distributed in many tissues with different concentrations, including kidney with the highest concentration, and sciatic nerve, lens, and testis. hAR as part of the polyol pathway is involved in glucose metabolism.¹⁰ Upon catalysis, the cofactor NADPH is bound first, followed by the sugar or other aldehyde. C1 of the substrate is reduced by a hydride transfer from the cofactor to the substrate's carbonyl carbon and a concerted proton transfer from Tyr 48 via His 110 to the substrate.^{11;12} Subsequently, the product is released. The dissociation of the now oxidized cofactor NADP^+ is conditional upon a conformational change of the protein, and thus, is a slow process.¹³

In a second step of the polyol pathway, the alcohol is further metabolized to fructose by sorbitol dehydrogenase (see Fig. 1.2). Thus, the overall reaction results in the conversion of glucose to fructose with consumption of NADPH resulting in NADH.

The K_m value for glucose is determined within the range of 100–400 mM in different experiments, which qualifies glucose as a rather poor substrate.¹⁴ Apart from glucose, hAR accepts a broad spectrum of aldehydes and ketones as substrates

ranging from sugars to lipids and xenobiotics. The rather hydrophobic binding site can accomodate a variety of hydrophobic substrates, e.g. phospholipides.¹⁵ Thus, it supports the detoxification of blood especially by metabolizing lipid aldehydes.

1.2.1.2 Structural Properties and Binding Site

Human aldose reductase (hAR) is a monomeric 36 kDa protein that consists of 315 amino acids and folds into a β/α -TIM-barrel. Eight alpha-helices and two additional smaller alpha-helices wrap around eight beta-strands (see Fig. 1.3). The binding pocket is located near the C-terminal loop in the barrel core where also the cofactor NADPH is bound.^{16–18} The so-called anionic binding pocket is formed by Tyr 48, His 110, Trp 111, and the nicotinamide moiety of the cofactor. These residues are involved in the catalytic function of hAR as mentioned above. Additionally, they are important for inhibitor binding; the carboxylate type anchor group of most inhibitors builds one of the core interactions with the enzyme in this area (see Fig. 1.4). Apart from the anionic binding site, most residues inside the pocket exhibit hydrophobic properties: Phe 122, Trp 79, Tyr 48, and Trp 20 with their aromatic ring systems, and Leu 300 form the binding pocket.

The binding pocket can also adapt to substrates of deviating size. A flexible loop near Leu 300 can move and thus, open an additional subpocket where further interactions are provided. Mainly hydrophobic interactions are formed here, dominated by a π -stacking that most inhibitors form with the aromatic moiety of Trp 111 (see Fig. 1.4(b)). Additional H-bonds can be formed in the back part of this subpocket to Thr 113 and to the functional groups of the backbone amide bond between Ala 299/Leu 300.

1.2.1.3 Pharmacological and Therapeutic Relevance

In many tissues, the physiological substrate of hAR is not yet identified. Thus, though a number of substrates are known and detoxicating activity has often been discussed,^{14;15} the actual role of hAR regarding its relationship to closely related enzymes of the family like aldehyde reductase remains unclear.

Under physiological glucose blood levels, only a small fraction of glucose is reduced by hAR. The expression of hAR is induced by high glucose blood levels where the physiological system comes to its limits.^{12;15} Under those hyperglycaemic situations, in some tissues hAR is responsible for up to 33% of glucose utilization. Predominantly, high turn-over rates of NADPH are discussed apart from the en-

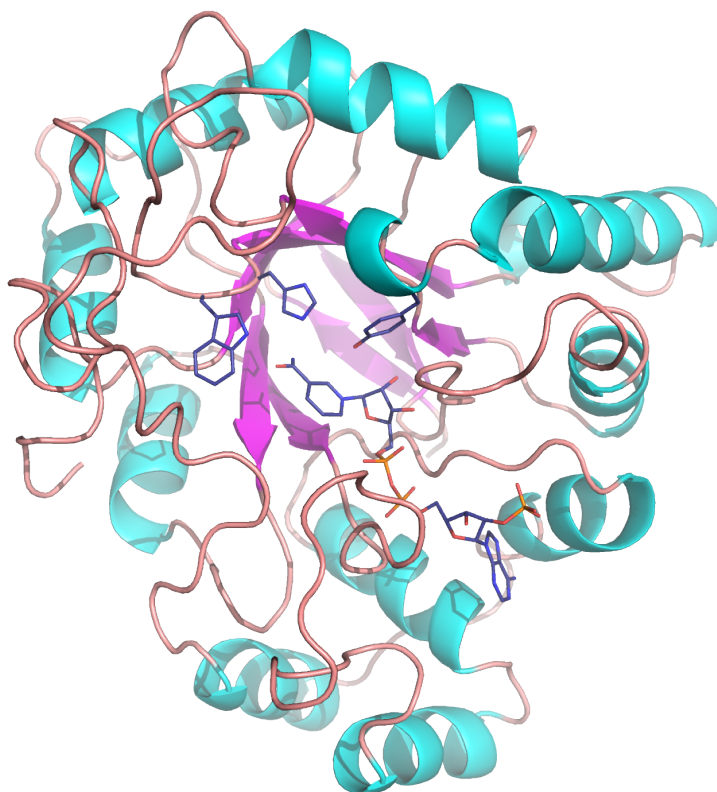


Figure 1.3: hAR forms to a TIM-barrel with the catalytic site in the barrel core. Catalytically involved residues Tyr 48, His 110, Trp 111 and cofactor are shown in blue sticks.

hanced accumulation of the highly polar sorbitol inside the cell to be responsible for osmotic stress. This causes long-term diabetic complications such as neuropathy, nephropathy and glaucoma and qualifies hAR as a promising drug target.^{19–23}

1.2.2 Selectivity against Aldehyde Reductase

hAR is closely related to aldehyde reductase (ALR, E.C. 1.1.1.2),⁹ a physiologically highly required protein of the same aldo-keto-reductase superfamily.²⁴ ALR catalyzes the NADPH-dependent reduction of a broad variety of aldehydes and ketones, both of aliphatic and aromatic nature. This protein is present in many tissues, high concentrations are expressed e.g. in kidney and liver.

Both hAR and ALR exhibit a pronounced sequence identity of about 65%.²⁵ Within the binding site, the catalytic site and the rigid residues of the binding pocket are identical while the flexible loop of hAR differs, which gives a significantly reduced sequence identity of 50 % inside the binding pocket.²⁵ To selectively address

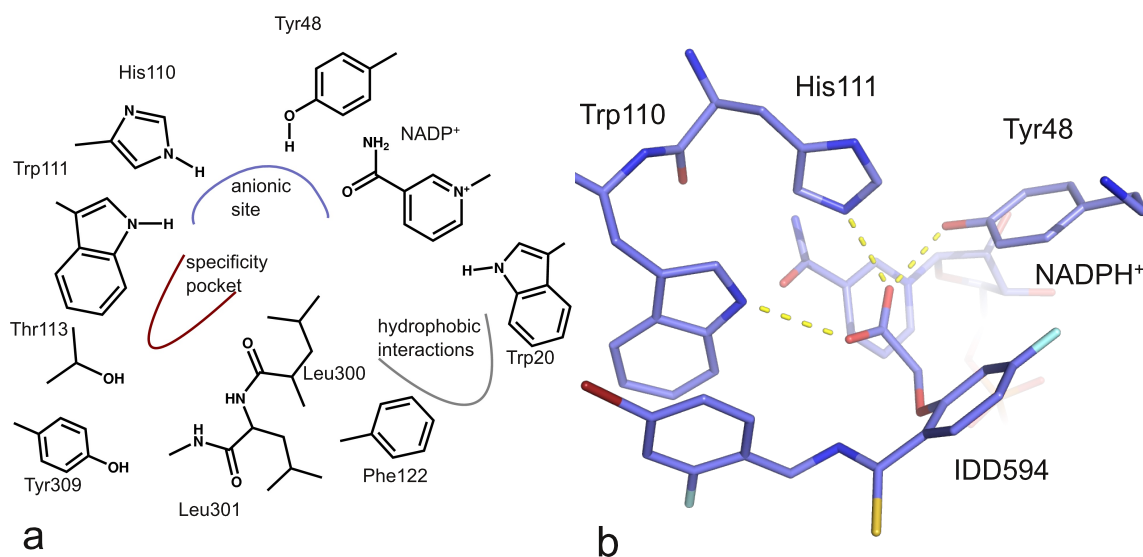


Figure 1.4: (a) – Binding site of hAR in schematic view. Anionic residues and hydrophobic binding parts are indicated. The flexible loop around Leu 300 can open the specificity pocket. (b) – Structural view of the catalytically involved residues. Here, **13** (IDD 594) is bound (PDB code 3LQL, see Chap. 4) which forms a π -stacking with the aromatic moiety of Trp 111.

hAR, the flexibility of the binding site can be exploited. A loop containing Leu 300 can react flexibly and, thus, opens the additional specificity pocket which is not present in aldehyde reductase. Ligands stabilizing this open conformer give access to the specificity pocket and promote selectivity. Therefore, detailed investigations of the interactions in this mainly hydrophobic pocket are one of the key elements to endow inhibitors with selectivity properties that avoid undesired side effects by interfering with aldehyde reductase activity.

1.2.3 Flexibility as Challenge

Many proteins show high flexibility. The conformational variations can range from altering single side chain rotamers via flips of entire residues to relocation of complete loops.^{26;27} Flexibility is often connected with activation—e.g. through formation of functional dimers—or cofactor binding or adaptation to differently shaped substrates as known for hAR.²⁸ Thus, functionality is often conditional upon preliminary conformational changes.

Especially binding-site flexibility is a major challenge in rational drug design. The prediction of the correct protein binding pocket conformation is a complex

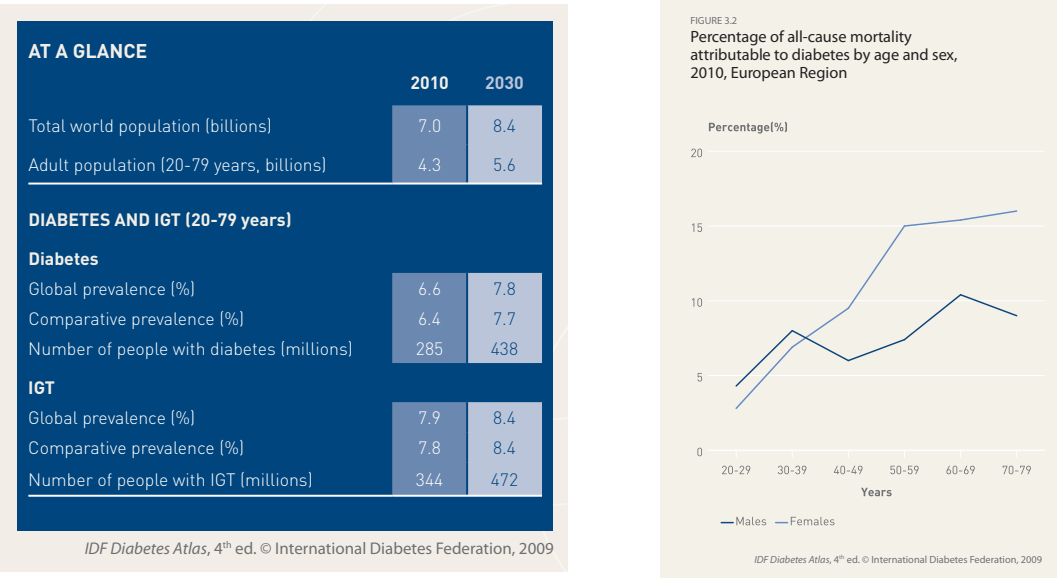


Figure 1.5: Relative and absolute numbers of diabetes and Impaired Glucose Tolerance (IGT) worldwide with a prediction for 2030. According to International Diabetes Federation (IDF)²⁹

problem that is more and more accounted for in strategies and methods of drug design. Two phenomena complement each other: The phenomenon of induced fit, where ligand binding provokes the adaptation of binding site residues to accommodate a compound upon binding and the multiplicity of different protein conformers present in equilibrium prior to binding, where a ligand picks one of these conformers and stabilizes the complex in a favorable conformation.

1.3 Diabetes – the Global Burden

In the last century, the standard of life in most populations has changed tremendously. An excessive supply with highly caloric nutrition along with reduced physical exertion and activity are known as dietary and behavioral aspects which influence health of people. In many countries, population is aging. Even in low- and middle-income countries, social and cultural life changes rapidly. As a consequence, diabetes incidence and prevalence is increasing continuously. Diabetes is now a common health problem, not only in industrial countries—one of the top five of leading causes of death in high-income countries—but also an increasing global burden.²⁹ Figure 1.5 shows the current numbers of prevalence and people with diabetic disease together with a prediction for the year 2030.

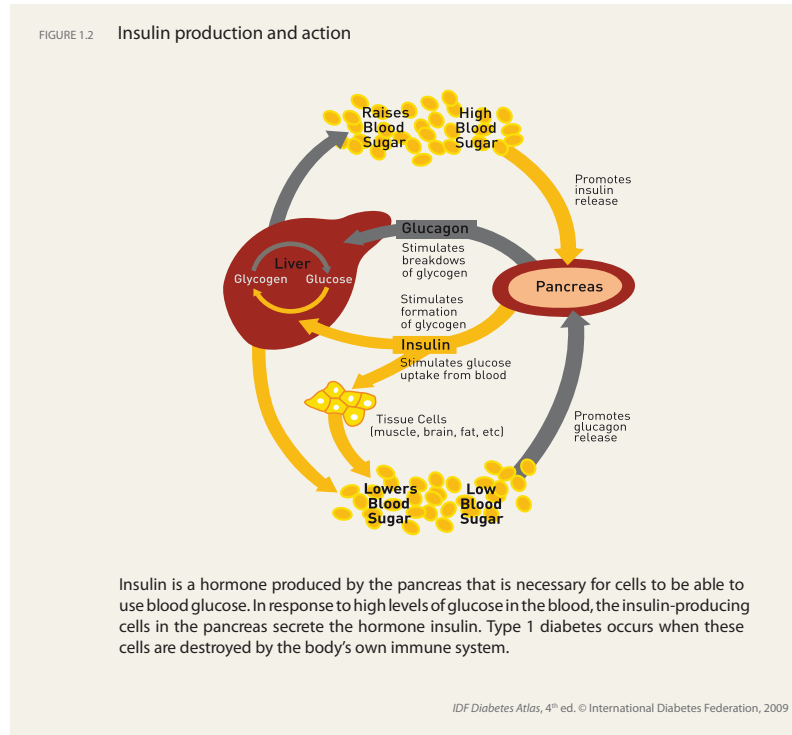


Figure 1.6: Insulin production and action, according to IDF²⁹

1.3.1 Diabetic Disease

Diabetes is defined as a number of disorders caused by hyperglycemia and glucose intolerance due to insufficient or no insulin production or decreased insulin sensitivity.³⁰ Diabetes is classified into four types, type 1 and type 2 diabetes, gestational diabetes mellitus and other specific types. However, the most important types are type 1 and type 2.

Type 1 diabetes mellitus is also called juvenile-onset diabetes because it usually occurs in childhood or young adults. Typically, auto-immune reactions cause the destruction of insulin-producing β -cells in pancreas. Depending on the state of destruction, the cells produce little or no insulin, leading to an absolute insulin deficiency. Insulin is the hormone that mediates the entering of glucose into body cells and thus, regulates the blood glucose level (see Fig. 1.6). Patients with type 1 diabetes are dependent on insulin treatment.

Type 2 diabetes is caused by relative insulin deficiency and/or insulin resistance. It used to occur beyond an age of 40, which is the reason why it is also called adult-onset diabetes. Despite this, in recent times, the incidence of type 2 diabetes in children and young adults increases. Type 2 diabetes is often associated with

obesity which can be one reason for decreased insulin sensitivity and high glucose blood levels. Other risk factors include physical inactivity, but also increasing age and family members with diabetes; as well as ethnicity.

1.3.2 Epidemiology

In 2010, the estimated diabetes prevalence in the world population is high, exceeding 6%; in absolute number about 285 mio. people within the age of 20–79 have diabetes, and the number is continuously increasing (see Fig. 1.5). More than 50% increase are approximated within the next 20 years.

The prevalence in different regions is displayed in Figure 1.7. Remarkable is the relatively high prevalence in South American countries and North African states, where an increase in the last years is recognized due to rapidly changing cultural and social circumstances, such as increased urbanization.³¹

In Europe, the prevalence with 8.5% (55 mio. people within 20–79 years) is even higher, which is partly caused by the relatively elderly population. In Germany, a prevalence of 12% is remarkably high, and with absolute 7.5 mio. people suffering from diabetes (20–79 years) Germany is scored on rank 6 of the top ten list of countries in 2010, considering the number of people (29–79 years) with diabetes. The number of deaths has doubled from 20.9 thousand people in 2002³⁰ to 54.6 thousand people in 2010.²⁹

The high prevalence of diabetes along with the continuously growing numbers of people with diabetic disease show epidemic dimensions. With an increasing number of diabetic people, the economic impact of the disease increases as well. Treatment of the disease and their complications resulting in high costs for the individual, the public health system and the economy. In Germany, the estimated average expenditure per diabetic person is 3751 US Dollars per year, while the world average is 703 USD in 2010.³⁰

1.3.3 Long Term Diabetic Complications

High glucose blood levels over longer time periods lead to tissue damage and destruction. This results in several chronic diabetic complications in many organ systems throughout the body. The main complications are cardiovascular disease, neuropathy, nephropathy, and retinopathy.

More than 50% of diabetic deaths in most populations are due to cardiovascular complications, including myocardial infarction, angina and stroke. For many

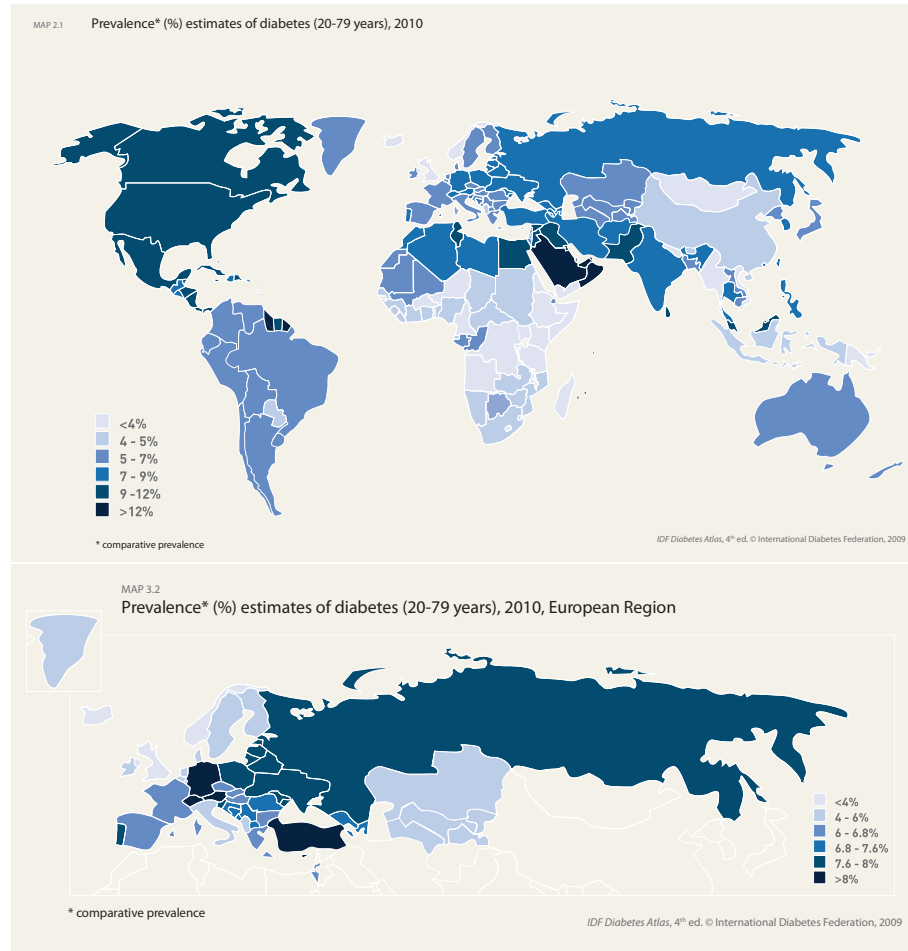


Figure 1.7: Prevalence of diabetes in people between the age of 20 and 79, worldwide (upper figure) and in the European region. The comparative prevalence is corrected for the age profile of the countries; for each country, the age profile of the world has been used. According to IDF²⁹

patients with renal insufficiency and renal failure with the need of dialysis or transplantation, diabetes is often the basic disease. Chronically increased glucose levels affect the nerves and thus, can lead to several complications including diabetic feet ulcer, which can make even amputation necessary.

The tissues involved in diabetes complications express hAR to a considerable amount. The relation of hAR to oxidative stress which induces tissue damage has been indicated above.

1.3.4 Treatment

In diabetic treatment, the efficient control of glucose blood levels and maintenance of glycemic levels close to non-diabetic values to prevent life threatening diabetic complications are the key aims.

1.3.4.1 Type 1 Diabetes

Type 1 diabetes is insulin dependent. For intensive insulin therapy, insulin schemes usually include basal insulin therapy combined with preprandial insulin injections to imitate physiological insulin secretion.³² Additionally, a number of modern long-acting, intermediate, and rapid-acting insulin analogues are available.³³ Thus, a combinational scheme of rapidly-acting insulin analogues combined with long-acting ones for basal insulin therapy is often preferred.³²

Current developments include the further improvement of insulin pump therapy, the so-called continuous subcutaneous insulin delivery which allows for highly precise insulin injections.³⁴ Patients' life with diabetes is improved under such therapy as technological and software developments result in smarter pumps which can support the patient in his daily diabetes management.

1.3.4.2 Type 2 Diabetes

Therapy of Type 2 Diabetes mellitus consists of non-medical as well as medical treatment.³²

Non-medical treatment includes nutrition therapy, where the persons with diabetes or even pre-diabetes are supported with dietary nutrition. Modest weight loss for overweight or obese individuals should be achieved, as this is shown to reduce insulin resistance. A lifestyle change including dietary nutrition as well as physical activity is recommended.³² Additionally, education to self-management of glucose control and nutrition should be an essential component of non-medical treatment.³⁵

Pharmacotherapy of type 2 diabetes usually starts with oral hypoglycemic agents (OHAs) to decrease insuline resistance, before insulin is substituted in later states as type 2 diabetes is a progressive disease or when insufficient glucose levels are achieved.

Sulfonylureas and metformin as biguanide are currently the most commonly used OHAs. Nevertheless, a number of additional approved OHAs are available, like thiazolidinediones, glinides, and α -glucosidase inhibitors. Newer drug classes for hypoglycaemic drugs include glucagon-like peptide-1 agonists, amylin analogues

Table 1.1: Summary of the benefits of diabetic drugs. From Philippe and Raccach³⁶

	Long-term data	Other benefits	HbA _{1c} decrease	Route	Hypoglyc. risk	Body weight change
SUs (1946)	Proven efficacy/safety	Low cost	0.8 – 2.0%	Oral	Yes	Gain
Biguanides (metformin) (1957)	Proven efficacy/safety	Low cost	1.0 – 1.5%	Oral	No	None or possible loss
α -glucosidase inhibitors (1995)	Limited data	CV benefits?	0.5 – 0.8%	Oral	No	No
Glinides (1997)	Limited data	Rapid acting	0.8 – 1.5%	Oral	Low	Gain
TZDs (1997)	Improve β cell function	Lipid profile (pioglitazone)	0.8 – 1.0%	Oral	No	Gain
GLP-1 agonists (2005)	Unknown	Improved β -cell mass?	0.6 – 1.0%	Injection	No	Loss
Amylin analogues (2005)	Unknown	–	\approx 0.6%	Injection	No	Loss
DPP-IV inhibitors (2006)	Unknown	Improved β -cell mass?	0.5 – 0.9%	Oral	No	Neutral

HbA_{1c}, glycated haemoglobin; GI, gastrointestinal; SUs, sulphonylureas; CV, cardiovascular; TZDs, thiazolidinediones; GLP-1, glucagon-like peptide-1; DPP-IV, dipeptidyl peptidase-IV.

and dipeptidyl peptidase-IV inhibitors.^{36;37} A short overview is given in Table 1.1 (taken from Philippe and Raccach³⁶).

In later states of type 2 diabetes, insulin substitution is initiated with basal insulin therapy, followed by a combination with preprandial insulin as described for type 1 diabetes. An early use of insulin in type 2 treatment is discussed lately to prevent complications arising from insufficient glucose levels, but up to now, this is not part of standard treatment.³³

1.4 Aim of the Thesis

This thesis is focused on the target human aldose reductase and the lessons that can be learned from the continuing investigation of the target. Different aspects of diverse ligands binding to the protein were enlightened.

On the one hand, a number of inhibitors from a series of aryl carboxylic acid derivatives were evaluated in an attempt to increase binding affinity in the context of the drug design cycle described above. Additionally, the search for a new lead was supported and in a following design cycle, derivatives were created and evaluated.

On the other hand, thermodynamic data from ITC experiments along with high resolution X-ray data were analyzed which give a deeper insight into the complex picture of small molecules binding to a protein. The pronounced flexibility of hAR has great impact on binding modes and make prediction a challenging task. Known inhibitors were chosen to further characterize binding to hAR in the context of the conformational space of the binding site.

Single interactions of functional groups can be analyzed in different ways: To investigate the contribution of a certain functionality of a ligand to overall binding, a series of compounds with varying substituents at the interaction site can be syn-

thesized and evaluated. In a project of this thesis, alternative residues on the protein side were introduced instead. Single-site directed mutants of threonine 113, a residue located at the rear part of the specificity pocket, revealed diminutive changes regarding the structural and thermodynamic binding profile of two brominated ligands as reaction to the introduction of alanine, cysteine, serine, and valine instead.

The crystallographic data determined within this thesis is of high resolution and quality. Several datasets were determined to atomic resolution. In the mutant series, changes between binding of one inhibitor to the above mentioned mutants could only be observed due to the high resolution data. Diminutive shifts of the scaffold became apparent which have influence on binding affinity. Second conformations of binding site residues and backbone flips could be clearly identified. These findings are discussed in Chapter 4.

In addition to this, the high resolution X-ray data of several similar hAR complexes redirected our interest on a difference density peak, which revealed unexpected results in Chapter 5: Radiation damage occurred with two brominated inhibitors during X-ray experiments with synchrotron radiation. The selective disruption of a carbon-bound bromine and the reappearance as bromide ion in an adjacent position emphasized the importance and necessity of accurate and detailed analysis of experimental data.

The high resolution structural data collected within this thesis can considerably contribute as basic information to the improvement of computational and predictive methods which often rely on experimental data. In the ProteinDataBank (PDB[†]), a total of about 75,000 released structures are available (accessed 03/2011), where more than 90 % show a resolution worse than 1.5 Å and more than 50 % worse than 2.0 Å. Less than 10 % of the released structures fall into an atomic resolution range, where detailed analysis of atom-atom contacts are possible, atom coordinates can be located with low positional variation (e.g. low B-factors) and the position of hydrogens become apparent. With higher resolution, the data basis for the atom and bond positions increases and the experiment reflects natural geometries more accurately. Thus, high resolution protein structures as determined in this thesis not only give deeper insights into the binding characteristics of the determined protein-ligand complexes, but can also contribute to improve the data basis for geometrical target values used in restraint libraries.³⁸

Additionally, protein structures in the PDB include about 6500 structures with ligands containing halogen atoms (more than 5000 with chlorine, including buffer-

[†]www.pdb.org

associated chlorine in protein structures) with less than 600 structures in the resolution range of 1.5 Å and better. As halogenated inhibitors are increasingly considered in drug design, detailed information about their interaction profiles will increase our understanding of possible binding contributions of such substituents.

In this thesis, hAR as model target represents the opportunities of modern drug design along with the challenges researchers still have to face. With an increasing amount of information, the image of a protein as a target and the aspects that contribute to binding becomes more and more complex. With the detailed information gained in this thesis, we hopefully come one step further to the future vision of fast and easy design of drugs without side effects.

2. Improving Known Scaffolds to Increase Affinity

2.1 Benzothiazepines as Novel Scaffolds Inhibiting Aldose Reductase

2.1.1 Introduction

Protein flexibility comprises challenges for rational drug design. A variety of experimental and computational methods are available and can lead to successful discovery of new lead compounds. Human aldose reductase is a well-studied protein of pronounced flexibility. Through the opening of an additional subpocket ligand binding can be enhanced with respect to affinity and selectivity. Aldose reductase is also used as model protein to investigate various contributions to binding that induce conformational changes upon ligand binding.^{39;40} Though a number of highly potent inhibitors are already known for this protein, continuous efforts must be spent to find new lead compounds of novel scaffolds exhibiting improved pharmacodynamic properties.^{41;42}

Concerted effort in various research groups throughout the world led to a variety of structurally diverse hAR inhibitors. Several attempts have been made to identify new lead compounds with novel scaffolds.⁴¹ Most known inhibitors of hAR bear a carboxylic acid or hydantoin functionality as anchor group to interact with the anionic binding site. Apart from this rather stringent pattern, the protein seems to accept diverse shaped ligands: The flexibility of a loop near the binding pocket provides access to an additional subpocket. Leu 300 flips, and thus substantially increases the binding site to accommodate diverse substrates of varying size and shape. This so-called specificity pocket is the key to selectivity over the closely related Aldehyde Reductase exhibiting high sequence similarity and being physiologically a highly relevant enzyme in the detoxification of reactive aldehydes.

Compound **1** (Fig. 2.1) was initially identified as one potential hit for hAR inhibition in a virtual screening campaign.³⁹ Its carboxylic acid functionality was proposed to act as mandatory anchor analogously formed by other hAR inhibitors.

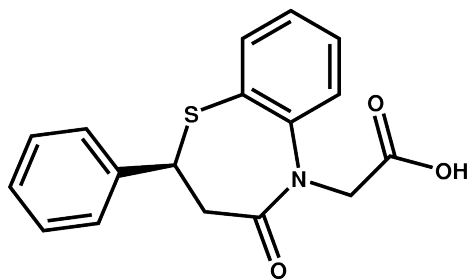


Figure 2.1: Schematic view of **1**.

The thiazepine provides a rigid core skeleton with a chiral center. Structure determination in complex with the enzyme revealed the (S)-enantiomer as binder. For subsequent optimization, the phenyl moiety provides an ideal chemical handle to introduce additional substituents planned to improve the interaction profile. The hydrophobic char-

acter of this substituent was thought to qualify the phenyl moiety as ideal interaction partner to open and penetrate the specificity pocket.

The binding affinity of **1** was determined by a kinetic inhibition assay and by isothermal titration calorimetry (ITC). Crystal structure analysis revealed an unexpected binding mode with the protein adopting a novel conformer. Additional space could be detected to further derivatize the basic scaffold to address the specificity pocket of the enzyme and thus, enhance affinity. The designed derivatives were docked into the protein, subsequently synthesized and finally evaluated by enzyme kinetics for their inhibitory potency.

2.1.2 Results and Discussion

2.1.2.1 Binding Mode of **1**

The structure of hAR with **1** was determined crystallographically to a resolution of 1.7 Å with good refinement statistics (see Tab. A.2). All binding site residues and inhibitor atoms were clearly defined in the electron density.

Unexpectedly, **1** binds to the protein in a conformation with the specificity pocket in a closed state. The schematic binding mode is displayed in Figure 2.2. Its carboxylate group forms the predicted interactions with the anionic binding pocket: Hydrogen bonds are formed with favorable geometry to His 110 ($2.70 \text{ Å} \pm 0.06$) and Trp 111 ($3.01 \text{ Å} \pm 0.08$, see Fig. 2.3(a)). A weak secondary hydrogen bond is formed by the carbonyl functionality on the seven-membered ring with 3.35 Å (± 0.10) distance to an aromatic C^εH of Tyr 48. The remaining part of the inhibitor atoms predominantly form hydrophobic interactions. Apparently, even the sulfur atom is directed outward the pocket without any direct interaction partner (see Fig. 2.3(b)). A second inhibitor molecule could be detected in the additional difference electron density and is located at the outer surface of the binding pocket. Its carboxylate

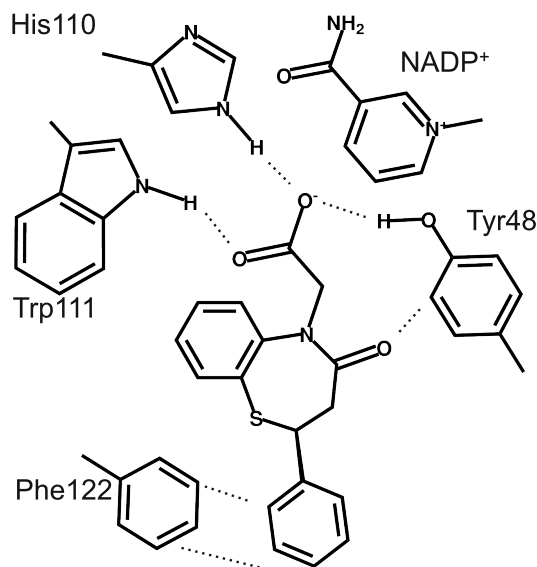


Figure 2.2: Schematic binding mode of **1**. The carboxylate acid interacts via H-bonds to His110 and Trp111 as well as Tyr48. The remaining inhibitor atoms form mainly hydrophobic interactions. Phe122 is shifted to form a good stacking interaction to the phenyl substituent of **1**.

oxygens interact with the backbone nitrogen of Leu 300 and Leu 301 without altering their position. The residual atoms form at most hydrophobic interactions with distances ranging between 3.5 and 4.4 Å to e.g. Leu 124 and Trp 219. Although the second inhibitor's sulfur points towards the sulfur of the first ligand, a distance of 4.4 Å between both atoms seems unlikely even for a sulfur atom to interact. A CSD⁴³ search revealed distances in small molecules between sulfur and any other atom of up to approximately 4.0 Å. Most likely, the binding of this second inhibitor is a result of the accommodation in the crystal packing and a corresponding binding is not given in solution. This assumption is confirmed by our thermodynamic data which reveal no hint for any stoichiometry other than 1:1.

The phenyl substituent of **1**, instead of opening the specificity pocket, interacts with Phe122 (see Fig. 2.3(b)). The aromatic systems of both rings accommodate in almost parallel fashion with 3.6 Å distance between the planes. To adapt to this interaction, the protein conformation had to be changed. Phe122 rotates about 60° toward the outer part of the pocket with respect to the holo structure (PDB code 1AH4⁴⁴). Thus, it enlarges the distance and optimizes the interaction angle to form a favorable stacking interaction.

Apart from the conformational adaptation of the binding pocket, the binding mode of **1** is determined by the stereochemistry of the chiral thiazepine ring. The

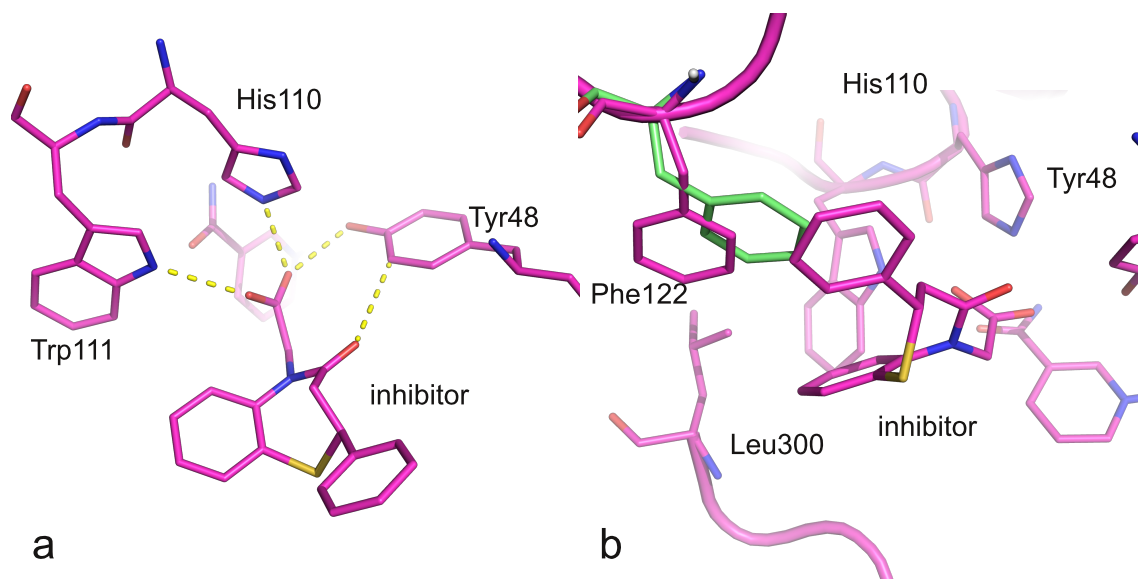


Figure 2.3: Wild type hAR in complex with **1**. (a) – The carboxylate anchor group is in H-bond distance to His110 and Trp111. The rigid benzothiazepine interacts mainly hydrophobically. A carbonyl oxygen is directed toward Tyr48 forming a weak H-bond to the aromatic ring in 3.3 Å distance. (b) – The side chain of Phe122 is rotated in contrast to a wild type holo structure (green, 1AH4), and the phenyl substituent forms a stacking interaction to this amino acid.

shape of the seven-membered ring and its annulated benzene moiety is rather rigid and provides only limited conformational degrees of freedom for adaptation. In principle, the scaffold should also be able to bind in 180°-flipped orientation, still placing the carboxylate function toward His110 and Trp111, but then orienting the phenyl substituent toward the specificity pocket. As this geometry is not observed, it supposedly would have produced too much intramolecular strain resulting in an energetically highly unfavorable arrangement. An inappropriately small angle between the carboxylate functionality and the phenyl moiety probably avoids accommodation in this binding mode.

Kinetic data for **1** gave a moderate IC_{50} in the low micromolar range (13 μ M/L, see Tab. 2.1). Thermodynamic data for **1** revealed a Gibbs free energy dominated by entropic contributions. ΔG with -33.6 kJ/mol (± 0.7) approximately falls into the same range as determined by the biochemical assay. The entropically driven binding of $-T\Delta S = -25.1$ kJ/mol (± 0.9) most likely reflects the rigidity of the inhibitor. It loses only a few degrees of freedom upon binding, but buries a large hydrophobic surface of the protein which otherwise would be solvent exposed (approx. 220 Å³). The enthalpic signal is rather low with -8.5 kJ/mol (± 1.6). The relatively good bind-

ing affinity along with an entropy-dominated binding can provide a good starting point for affinity improvement by systematic introduction of new functionalities.

2.1.2.2 Choice of Promising Derivatives of **1**

On the basis of the known binding mode of the rigid compound **1**, its scaffold was optimized to gain higher affinity. The aim was to explicitly address the specificity pocket with promising substituents.

The aromatic plane of the annulated phenyl moiety of **1**, though in fair distance to the aromatic system of Trp 111, binds with almost perpendicular orientation to the latter. Though a direct interaction can hardly be formed, it provides a good position to decorate the ligand scaffold with diverse substituents. In a first design attempt, we introduced an additional aromatic moiety to enable π -stacking with Trp 111 once the specificity pocket has opened. As subsequent design goal, additional substituents at the aromatic ring should interact directionally with functional groups of the amino acids located in the rear specificity pocket.

In a previous study, compounds exhibiting different substituents were investigated to interact with the back part of the specificity pocket.⁴⁵ This study revealed nitro-, fluoro- and chloro-substituents as most promising decorations. Additionally, findings referred to in Chapter 5 resulted in the selection of bromo-substituents. To further reduce our selection of putative candidates for synthesis, we considered the synthetic accessibility. Finally, we decided to take the compounds listed in Figure 2.4 into a more detailed validation using docking to elucidate the expected binding mode of the synthesis candidates.

Synthetic feasibility and availability of synthetic fragments restricted the initial choice of derivatives to phenyl as aromatic moiety (see Fig. 2.4, **1a**, **1e**, **1f**, **1g**), but some additional compounds could also be synthesized (**1b**, **1c**, **1d**) to investigate the behavior of the scaffold in more detail. Depending on the outcome, further synthetic cycles should be planned to realize more of the initial ideas, e.g. derivatives with benzyl and benzothiazole substituted moieties. The final derivatives chosen for synthesis are displayed in Fig. 2.4.

2.1.2.3 Docking of Promising Derivatives

Docking is the method of choice to generate possible binding poses of our synthesis candidates.⁴⁶ In cases of hAR, protein flexibility upon ligand binding provides a special challenge for docking and requires particular care.^{47;48} While most docking

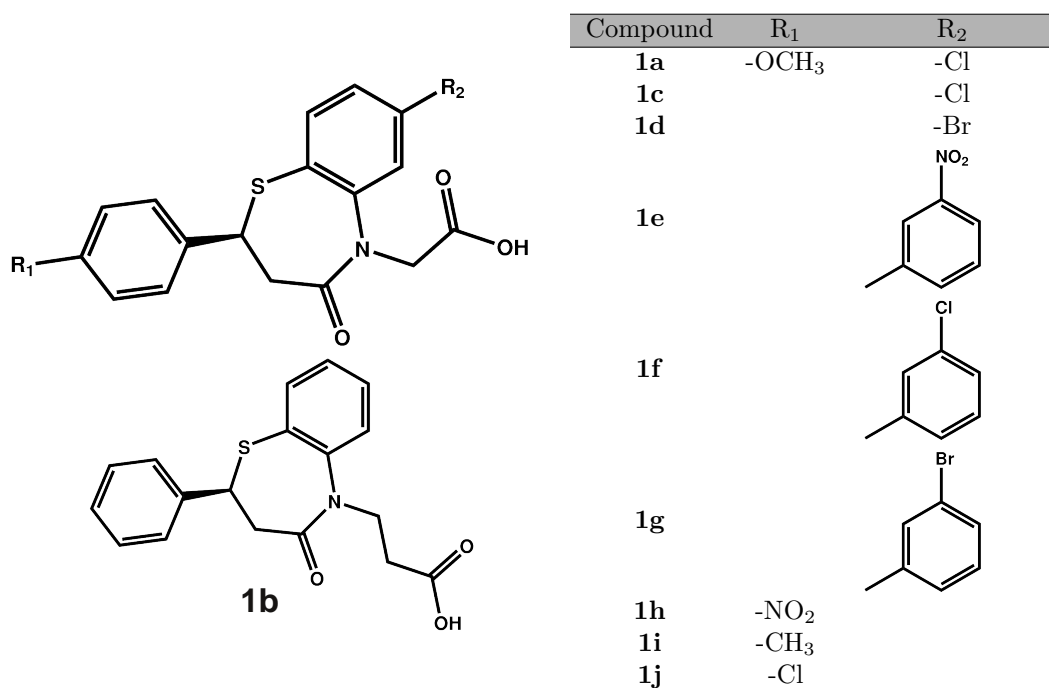


Figure 2.4: Derivatives of **1** chosen in a first step for synthesis and further investigation.

methods generate multiple ligand conformers during docking, only a few programs incorporated protein flexibility so far.⁴⁹ Thus, in our study docking was performed by several runs with different protein conformers rather than docking into a protein structure handled as flexible upon docking. GOLD⁵⁰ version 3.1 and AutoDock⁵¹ version 3.0 were used for docking experiments.⁴⁰

2.1.2.3.1 Redocking of 1 Docking **1** into its corresponding crystal structure was done to evaluate the best settings for docking. For GOLD, some variations of the default settings were done to improve the docking geometries; thus, population size and maximum number of operations were increased as well as keeping the ligand flexible during the docking experiment. The final setting is documented in the experimental section. The 20 best solutions were closely inspected.

The best results for these redocking attempts are displayed in Figure 2.5. While the best results for AutoDock all have rmsds below 1.0 Å with a value of 0.61 Å for the top ranked solution, GOLD generated exclusively geometries with higher deviations from the crystal structure. In the latter, the core is flipped by 180° but the position for the carboxylate anchor of **1** is placed in the correct area, though not in optimal interaction position. Thus, the resulting best geometry exhibits a high rmsd of 5.15 Å.

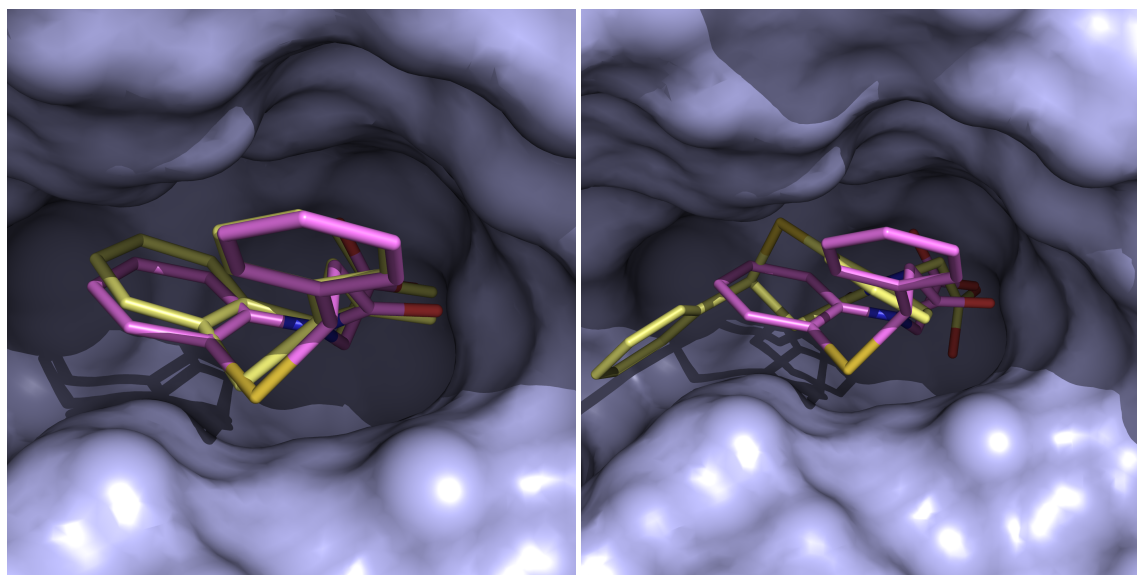


Figure 2.5: Results of **1**-redocking into its crystallographically determined protein conformer. The crystal structure is displayed in magenta, docking results in yellow. (a) – AutoDock performs very well with a low rmsd of 0.61 Å for the best ranked geometry. (b) – GOLD exclusively provides geometries with a flipped thiazole core within the high scored results, thus leading to worse interaction positions for the carboxylate anchor. Therefore, rmsds are high (5.15 Å for the best result.)

Due to these results further docking experiments were performed with AutoDock.

2.1.2.3.2 Docking of Derivatives To account for the pronounced flexibility observed for the hAR binding pocket, docking was performed with different protein conformers. As the aim of the docking experiments was to explore reasonable binding geometries for the different synthesis candidates, the analysis of docking results was based mainly on visual inspection.

As mentioned above, the derivatives were designed to address residues in the specificity pocket, thus a protein conformer with this pocket in open state was used for docking runs (PDB code 2IKG⁴⁵). Additionally, as the crystal structure of **1** revealed a previously unknown binding site conformation with a large shift of Phe 122, a second protein conformation was generated showing the features of the **1**-hAR complex and of 2IKG: The shifted Phe 122 (see Fig. 2.3) was incorporated into the structure with the opened specificity pocket.

Among the 20 best-scored solutions for the designed derivatives only those with the carboxylate anchor placed into the anionic binding pocket were further considered. Remarkably, only docking runs with Phe 122 shifted as found in the **1**-hAR complex showed occupancy in the specificity pocket (see Fig. 2.6). Obviously, the

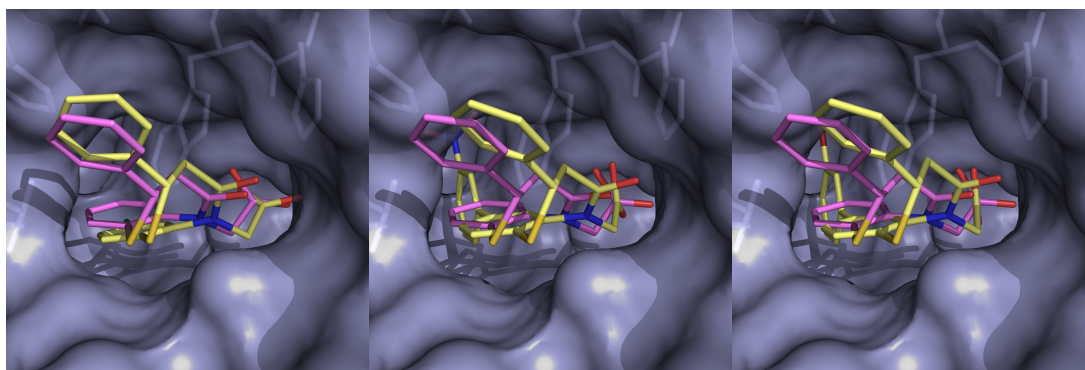


Figure 2.6: Best docking results (according to visual inspection, displayed in yellow) in superposition with the crystal structure of **1** (magenta): **1c** (left), **1e** (center), **1g** (right).

Table 2.1: Kinetically determined IC_{50} values for the derivatives. **1g** could not be determined due to insufficient solubility in kinetic conditions.

Compound	IC_{50}	[μM]
1	13.48	± 0.22
1a	1.41	± 0.32
1h	2.75	± 0.45
1e	4.97	± 0.53
1j	5.24	± 1.12
1c	7.03	± 1.68
1d	7.10	± 1.23
1i	12.39	± 2.92
1b	$\gg 500$	
1f	$\gg 500$	

rigidity of the considered ligand skeleton prevents penetration into the specificity pocket if the shift of Phe 122 is neglected. Solutions for **1c**, **1e**, and **1g** are displayed in Figure 2.6.

2.1.2.4 Kinetic Characterization

Inhibition constants were determined as described in the experimental section. In variation of the standard conditions, methanol instead of DMSO (in corresponding concentration) was used to enhance solubility of the poorly soluble compounds. Results are presented in Table 2.1. Due to solubility problems, no reliable affinity data could be determined for **1g**.

Remarkably, all IC_{50} values except for **1b** and **1f** show enhanced binding affinity compared to the unsubstituted parent scaffold. An up to ten fold increase in affinity

was achieved.

Unexpectedly, **1a** was determined as best binder of the series (1.4 $\mu\text{mol/L}$, see Tab. 2.1). Anticipating a similar binding mode for this only slightly modified derivative compared to the parent compound, where the phenyl moiety binds oriented toward the outside of the binding pocket, an additional interaction partner for the methoxy group is not directly evident.

The kinetic experiments indicated already limited solubility of the studied compound class as severe problem for the detailed evaluation of the derivatives of this scaffold. Additionally, crystallization conditions to be applied for hAR did not allow to produce sufficiently concentrated solutions of the derivatives to successfully soak the ligands into the protein crystals.

2.1.3 Conclusion

Among other compounds discovered by virtual screening,³⁹ **1** provides an exceptional and novel scaffold with a sterically demanding seven-membered ring. This thiazepine stands for a possible new lead scaffold for hAR inhibitors with a promising binding affinity in the low micromolar range.

Evaluation of the crystal structure of this highly rigid compound reveals that the binding mode is dominated by hydrophobic interactions. The introduction of additional substituents for H-bonds or electrostatic interactions seemed promising to increase the binding affinity considerably. Additionally, the thermodynamic profile of **1** shows an entropically favored binding. Thus, substituents forming enthalpic interactions to the protein should further tighten the protein-ligand complex.

Kinetic data for the derivatives support the hypothesis that directed interactions increase a compound's affinity to the protein. In case of **1c** and **1d**, one additional halogen atom attached to the scaffold results in slightly tighter binding. Halogens can provide a strong contribution to binding as known from diverse investigations in Chapter 4 and elsewhere.^{52;53} The derivative bearing a methoxy substituent at the phenyl moiety shows the highest affinity. Possibly, the binding mode is only slightly changed so that the methoxy substituent accepts an H-bond from a backbone nitrogen of e.g. Phe 121. This might even be water-mediated. A completely changed binding mode seems unlikely, since the substituent is attached to the phenyl ring which is already directed outward the binding pocket and an alternative to the strong carboxylic acid anchor is not provided.

Unfortunately, the binding geometries of our designed and synthesized deriva-

tives could not be determined crystallographically. Limited solubility of the compounds made the addition of organic solvents necessary, however this was not tolerated by hAR crystals.

Accordingly, limited solubility turned out to be a major obstacle to further improve this compound series. Further decoration with solubility-enhancing substituents would result in even bulkier ligands and increase their molecular weight.

Our design departed from a highly rigid binder mainly interacting via hydrophobic contacts. On first sight, this scaffold appeared promising as it contains several possibilities to substitute, elongate, and alter its functionalities. However, its pronounced rigidity restricts the possibility to adapt upon binding to the flexible target protein. Putatively, decoration with bulky aromatic substituents increasingly prevented conformational adaptation to optimally occupy a small and relatively narrow entrance to the specificity pocket.

2.2 Bi- and Triaryl Carboxylic Acid Derivatives

2.2.1 Introduction

A number of different approaches to obtain novel leads for hAR resulted in a variety of inhibitors with moderate to high affinity. Unfortunately, only few of these compounds reached clinical trials so far, and even less passed them. Until now, only the Japanese market provides one inhibitor of hAR, Epalrestat (Kinedak[®]), whose overall properties were convincing in terms of clinical efficacy and safety. Thus, beside permanent search for novel lead compounds, continuous efforts to improve known scaffolds are taken.

Many inhibitors share certain core features: Besides a mostly acidic anchor group to interact with the anionic binding site, a hydrophobic moiety to bind into the specificity pocket is mostly present. A linker between both groups provides the possibility to adopt an angle between the distant functionalities in a way that the opened specificity pocket can be occupied to establish a π -stacking of the aromatic moiety with Trp 111. In many cases, the linker itself comprises aromatic properties and additionally contains hydrogen acceptors which can further tighten binding of the scaffold.

Among other studies, a virtual screening performed on hAR affirmed this very general scaffold: All compounds identified as binder exhibit a carboxylate acid to bind as an anchor group.⁴¹ A chain of varying length, ranging from at least three

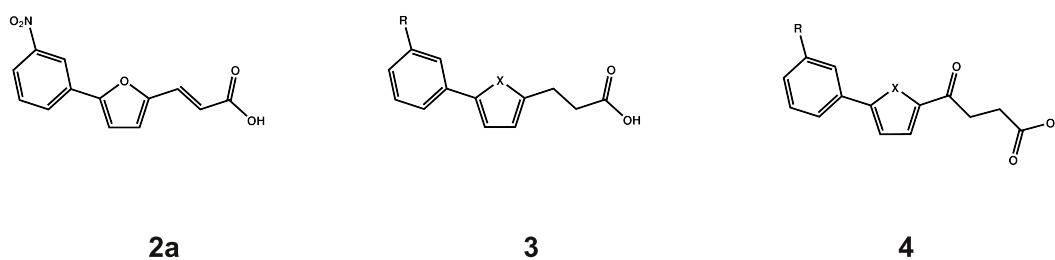


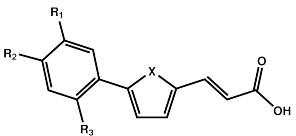
Figure 2.7: Selected compounds of a series originating from a virtual screening.⁴¹ Originally, the series included acrylic acids (**2**), propionic acids (**3**), and oxobutylic acids (**4**).

Table 2.2: Selected compounds of a series originating from a virtual screening and determined IC₅₀ values.⁴¹

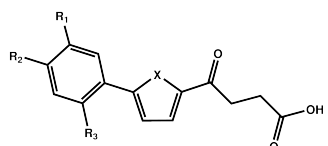
Compound	R	X	IC ₅₀ [μM]	
2a	NO ₂	O	19.00	± 0.90
3a	NO ₂	O	0.26	± 0.05
3b	NO ₂	S	0.17	± 0.03
4a	H	S	6.5	± 1.9
4b	CH ₃	S	4.1	± 0.7

atoms to six atoms builds a linking element to a distal aromatic ring at the inhibitor. In a number of cases, the linking chain is accompanied by a central (hetero-) aromatic ring that again is connected via at least one rotatable bond to the distal aromatic system. The latter can bear additional substituents mostly providing further potential interaction properties such as a nitro group.⁴⁵

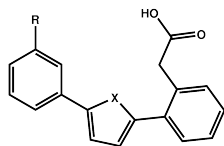
In a comprehensive study, two of the most promising scaffolds originating from one of these virtual screenings were further optimized.⁴¹ The design of related hAR inhibitors involved reduction of these scaffolds to the core interaction identified via a crystal structure,⁴⁵ strategic variation of the included core properties and alteration of substituents.⁴² This study resulted in a series of potent inhibitors including phenylfuryl-/phenylthiophene propionic acids, -acrylic acids and -oxobutanoic acids with IC₅₀ values in the range of 30 – 0.1 μM.⁴² The best compounds of the series along with two slightly varied compounds are presented in Figure 2.7; corresponding IC₅₀ values are displayed in Table 2.2. A crystal structure of **3b** (PDB code 3DN5⁴²) verified that the binding mode of the series resembles the one known for inhibitors occupying the specificity pocket (see Chap. 4).⁴²



2b-f



4c-f



5a-h

Compound	R ₁	R ₂	R ₃	X	IC ₅₀ [μM]	
Acrylic acids						
2b	C(O)CH ₃	H	H	O	13.7	± 2.2
2c	S(O ₂)CH ₃	H	H	O	26.1	± 4.2
2d	Cl	H	H	O	39.1	± 7.4
2e	CH ₃	H	H	O	44.1	± 5.1
2f	H	OCH ₃	CH ₃	O	50.3	± 6.8
2g	H	H	F	O	65.8	± 7.5
2h	OCH ₃	H	H	O	79.3	± 4.7
2i	H	H	H	S	81.8	± 8.1
2j	CH ₃	H	H	S	94.5	± 10.6
Oxobutylic acids						
4c	NO ₂	H	H	O	1.0	± 0.7
4d	NO ₂	H	F	O	3.1	± 1.1
4e	NO ₂	H	F	S	8.7	± 1.2
4f	H	NO ₂	F	S	64.2	± 7.5
Triarylic acids						
5a	NO ₂			O	0.34	± 0.05
5b	NO ₂			S	1.24	± 0.16
5c	H			S	3.6	± 0.5
5d	CH ₃			O	5.8	± 0.7
5e	CH ₃			S	6.9	± 0.9
5f	H			O	10.7	± 1.0
5g	CF ₃			O	13.1	± 2.2
5h	CF ₃			S	30.1	± 2.8

Figure 2.8: Schematic structures and IC₅₀ values for acrylic acids (**2b-f**), oxobutylic acids (**4c-f**), and triarylic acetic acids (**5a-h**).

2.2.2 Kinetic Evaluation of Further Variations

The series bears much potential regarding the possible substituents that can be added.⁵⁶ With its promising examples of nanomolar affinity, the series was further exploited to gain deeper insight into promising substituents and functional groups that could possibly further increase the inhibitor's affinity.

This effort resulted in a number of acrylic acids with different residues attached to the distal phenyl moiety. Additionally, more oxobutanoic acid derivatives have been synthesized. Finally, a variety of triarylic acids have been created, bearing another phenyl ring as linker between the carboxylic acid and the distal aromatic ring (see Fig. 2.8).

Kinetic evaluations for the compounds presented in Figure 2.8 have been performed. The measurements followed the standard protocol for the kinetic assay

with wild-type protein of hAR and is described in Chapter 6.

2.2.3 Results

Unfortunately, the IC_{50} values show no further improvement of the already quite potent compounds. All compounds bind in the low micromolar range, with one exception binding with nanomolar affinity (**5a**).

Within two of the three groups—oxobutylic acids and triarylic acids—as well as in the original propionic acids, the nitro-substituted compounds show the highest affinity. This is in agreement with another study where compounds bearing the very same functionality at the distal aromatic moiety appeared as the strongest binders toward the enzyme among others.⁴⁵

In case of the acrylic acids, the affinity of the nitro derivative (19.0 μ M) is slightly worse than the one of **2b** (13.7 μ M) with another electron withdrawing substituent ($-C(O)CH_3$). Most likely, the withdrawing properties of the latter substituent increase the strength of the π - π -stacking between the distal phenyl moiety and Trp 111 which had been observed for the corresponding compounds of the propionic acid series as well.⁵⁶ Compared to the propionic acids, all corresponding acrylic acids showed considerably lower affinity in the range of two- to seven-fold decreased IC_{50} values. The assumption, that the decreased flexibility of the acrylic acid results in a suboptimal position of the acid functionality regarding its binding partners seems verified.

Only few additional derivatives for the oxobutanoic acids were determined. Except for **4f**, all of them bind within the range of 1–10 μ M. Thus, they bind almost equally well as the propionic acids. A direct comparison of **3a** with **4c** (nitro substituted furyl derivatives) reveals the propionic acids as stronger binder again (0.26 μ M, 1.0 μ M). The best oxobutylic acid could not outperform the propionic acids.

Surprisingly, the best compound of the triarylic acids binds with an affinity of 0.34 μ M and thus, is the strongest binder of the compounds presented here. In the original virtual screening, a similar compound among the initial hits was tested as inactive (**5o**).⁴¹ The scaffold of the latter was very similar to the newly created derivatives (see Fig. 2.9). An obvious difference is the spacer between the phenyl ring and its adjacent carboxylic acid: While in the compounds **5a–h** the acetic acid enables the carboxylate functionality to optimize the interaction position toward His 110 and Trp 111, the rigid core of **5o** restricts the possible conformations. Obviously, an elongation of the spacer to the acid functionality by one atom increases the

rotational degrees of freedom considerably which most likely leads to the increase in affinity.

The introduction of a third aromatic moiety into the scaffold was considered due to the binding mode of **13** (see Chap. 4). Here, a fluorine substituted phenyl ring adjacent to the anchor group interacts via π -stacking with Phe122 positioned at the upper part of the binding pocket. A similar position was addressed with the expansion of the scaffold described above. Most likely, a similar binding mode is achieved.

No crystallographic data are available for the compounds discussed in this section.

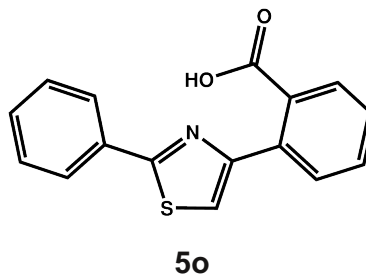


Figure 2.9: Triarylic acid identified by virtual screening and tested as inactive.

3. Protein Flexibility - Between Challenge and Chance

3.1 Ligand-induced Fit Affects Binding Modes and Provokes Changes in Crystal Packing of Aldose Reductase*

3.1.1 Introduction

For the majority of enzymes, pronounced flexibility is one important prerequisite to accomplish their catalytic function. Conformational changes of crucial residues are observed in protein folding and unfolding, in transportation of molecular species through membranes, during formation of functional dimers, in agonist versus antagonist binding or upon substrate binding and turn-over in enzymes with broad substrate promiscuity.^{26;27} Single side chains up to entire loop regions are relocated to provide access to a binding pocket, and thus, adaptations to a broad range of structurally diverse substrates are facilitated.²⁸

Flexible proteins seem to exhibit a variety of distinct binding site conformers which become obvious through the analysis of different binding modes adopted with diverse ligands and inhibitors. Highly potent ligands may share key interactions via one particular anchor group while other interactions differ and provoke changes in the protein conformer due to its high adaptivity in certain protein regions. Such changes are especially observed if a protein requires high adaptivity to accomplish its function. In these cases the probability to observe a novel conformer allowing for unexpected potential interactions with a ligand of modified scaffold geometry and decorated with different functional groups is high.⁵⁷ Method of choice to observe such novel protein conformers is crystal structure analysis. The more systematic studies of congeneric series become available, the more the crystallographic phenomenon of different crystal forms with deviating crystal packings are observed. In such situations it is questionable whether the conformational changes provoke the transition to another crystal form or whether the altered crystal packing induces the new protein conformer.

*Manuscript submitted to *Biochimica et Biophysica Acta*.

For rational drug design, target flexibility and induced-fit adaptations provide special challenge.^{40;58} Again our knowledge about such phenomena is primarily based on crystallographic evidence. The more it is important to separate conformational multiplicity induced through ligand binding from effects provoked by changes in crystallographic packing. The present study will give an example how ligand induced effects make the transition into another crystal form with altered crystal packing necessary.

Human Aldose Reductase (hAR) shows broad substrate promiscuity and has been demonstrated to be a protein with highly adaptive binding site features leading to diverse binding site conformations. Studying this protein can help to improve our understanding of protein-ligand complex formation in a flexible protein environment.

Most known inhibitors of hAR bear a carboxylic acid or hydantoin functionality as anchoring group to interact with the anionic binding site. However, with respect to the remaining scaffold, the protein accepts a broad range of ligands with diverse shape: A flexible loop near the binding pocket provides access to an additional sub-pocket. Leu 300 flips, and thus enlarges the binding site to accommodate diverse substrates of varying size and shape. This so-called specificity pocket is the key feature to endow the protein with selectivity over the closely related and physiologically highly important aldehyde reductase with high sequence similarity. This pronounced adaptivity offers great opportunities for rational drug design.

Obviously, as the prevention of undesired side effects must be a major aim in rational drug design, understanding the flexibility and interactions in the specificity pocket are of great importance. Hence, we investigated the interactions in this pocket and their impact on binding site conformation. A mutation at the far end of the specificity pocket was introduced to determine the binding contributions of interactions in this subpocket and their impact on protein conformation. Two inhibitors with different substituent patterns as interaction partners were chosen (Fig. 3.1). While a related study investigated the consequence for direct interactions with residues at the mutation site (see Chap. 4), in this chapter the mutated residue indirectly alters the inhibitor's interaction pattern and induces changes of the protein conformation and the crystal packing. A third inhibitor with deviating scaffold was chosen to study prompted induced-fit adaptations of the wild type protein which also seems to induce a change to an alternative crystal packing.

3.1.2 Results

3.1.2.1 Overall Binding Mode

The binding site of hAR consists of an anionic binding pocket, where the carboxylic acid anchor group interacts with His 110, Trp 111, and the cofactor (see Fig. 3.1). This site is found to be rigid throughout all presently known hAR inhibitor complexes. In contrast, the specificity pocket exhibits high flexibility and pronounced movements of Leu 300 lead to an opening of a predominantly hydrophobic subpocket where π -interactions with Trp 111 are possible. Additionally, at the rear of this pocket the hydroxyl function of Thr 113 provides the facility for directional H-bond interactions. In our study, this threonine was replaced by amino acids which either lack the property for polar interactions such as alanine and valine or have similar properties as serine. The contribution of this interaction to complex formation with zopolrestat and IDD 393 (**6**, **7**, Fig. 3.1) was investigated crystallographically. We collected high resolution data (1.0–1.6 Å) for the alanine, serine and valine mutant complexes with **6** (A-**6**, S-**6**, V-**6**), and two different high resolution complexes of the valine mutant with **7** (V-**7**). Additionally, **1** was cocrystallized with wild type hAR and structurally determined (WT-**1**; schematic view see Fig. 3.1). In addition, thermodynamic data for these complexes were recorded by isothermal titration calorimetry (ITC).

For wild type hAR complexed with **6** (WT-**6**, Fig. 3.1), several structures obtained from crystals grown under varying conditions⁵⁹ are available from the Protein Data Bank (PDB). While the binding mode with the key interactions of the inhibitor and the overall position in the binding pocket is the same in all complexes, the flexible loop 298–303 shows deviating conformations. In contrast to the predominantly formed protein conformation where the specificity pocket is opened and the NH of Leu 300 is oriented away from the inhibitor (e.g. WT-**6**, PDB code 2FZ9), in another complex Leu 300 NH is flipped about 90° and presents this group toward the inhibitor which now can form an additional hydrogen bond (WT-**6**, PDB code 2FZ8).

3.1.2.2 Zopolrestat Mutant Complexes

The different mutant-**6** complexes resemble the wild type complex WT-**6** when focusing on the rigid part of the binding pocket (Fig. 3.2(a)). The carboxylate group of the inhibitor forms a key interaction with His 110, Trp 111 and the cofactor NADP⁺.

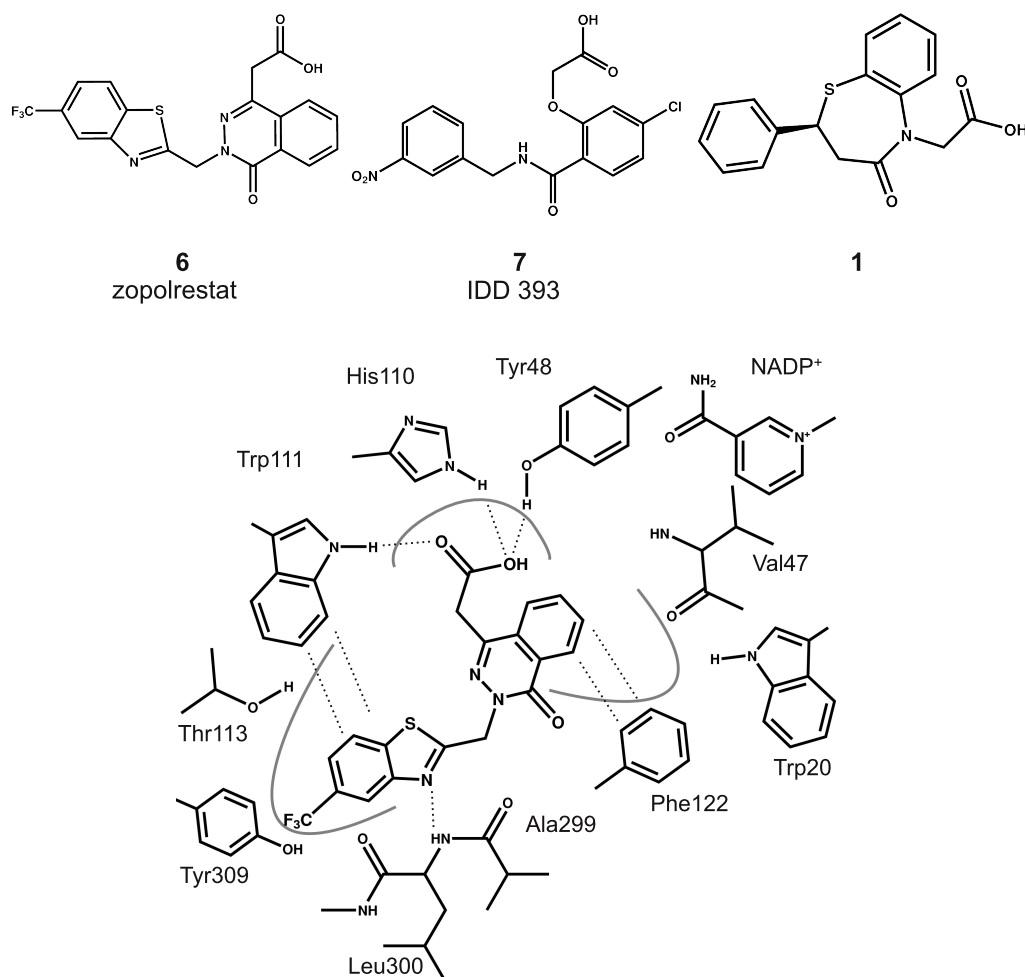


Figure 3.1: Schematic view of the different inhibitors and of the binding mode with **6** with opened specificity pocket. Interactions are indicated in dashed lines.

Thus, the anionic binding site shows no conformational changes compared to the wild type protein complex. The adjacent phthalazinone moiety forms a π -stacking with Phe122 and orients the inhibitor toward the opened specificity pocket. The benzothiazole moiety occupies the latter pocket, unlocked by the Leu300 flip. It interacts via hydrophobic stacking with the indole portion of Trp111. The mutated residue 113 at the rear of the specificity pocket interacts with the trifluoromethyl substituent of **6** (see Fig. 3.2(b)). Only small conformational changes of the protein are registered upon mutation, the backbone coordinates of the mutants match with those of the wild type. C^β of threonine in the wild type and of mutant residue Ala113 are located at almost the same position, while in the serine complex (S-**6**) Ser113 O^γ replaces the hydroxyl group of threonine (see Fig. 3.3(a)). One unique

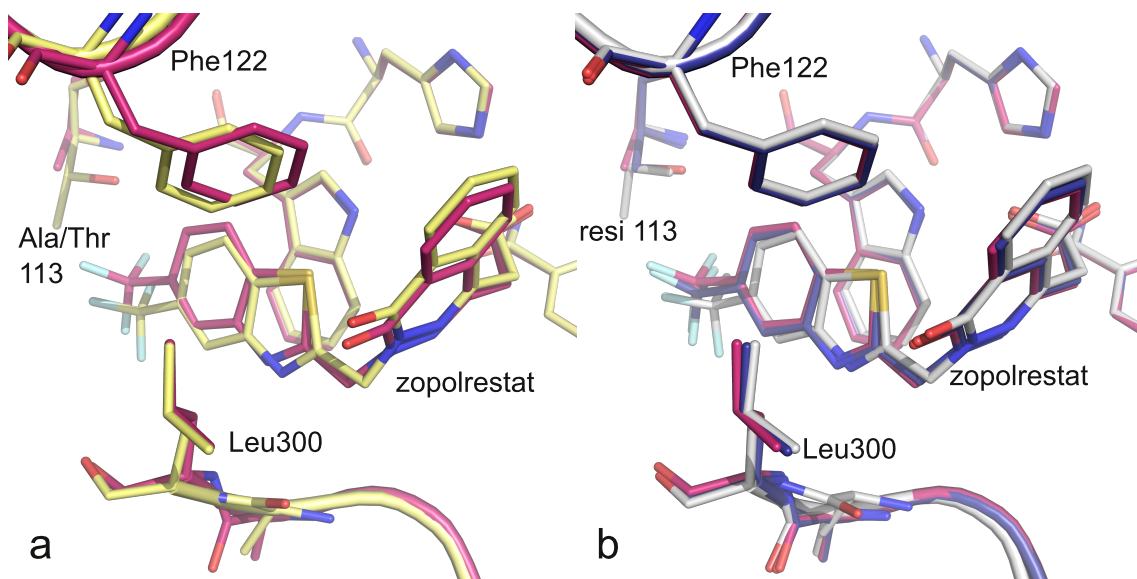


Figure 3.2: (a) – Superposition of WT-**6** complex (2FZ9, yellow) and A-**6** complex (red). The overall binding mode remains the same for the mutant complex, though a clear shift of **6** toward the sterically less demanding alanine is visible. The benzothiazole accommodates to a shifted Phe 122. (b) – Superposition of the alanine (red), serine (dark blue) and valine (grey) **6** complex. Phe 122 and the benzothiazole moiety match throughout the complexes. In the serine and alanine complex, the coordinates of **6** are virtually identical. For the valine complex, the phthalazinone moiety is slightly shifted. In all three complexes the backbone flip between Ala 299 and Leu 300 occurs, albeit in the valine complex a second conformation is visible.

conformation is observed for the serine side chain. Val 113 in V-**6** shows the same conformation as threonine; the second methyl group is replacing the position of threonine's hydroxyl functionality.

Interestingly, even these small geometric differences upon mutation evoke conformational adaptations of the inhibitor which become apparent in our high resolution data. The anchoring carboxylic acid remains in position in all complexes but the distal benzothiazole moiety is slightly shifted in the specificity pocket. This is due to the deviating spatial requirement of the mutated residues. In consequence, distances of the inhibitor atoms to residue 113 are modified according to the nature of the interacting atoms. While in the wild type the distance of fluorine to Thr 113 O γ amounts to 3.23 Å and 3.00 Å for 2FZ8 and 2FZ9, respectively, in S-**6** the corresponding distance falls shorter to 2.93 Å (± 0.02). In the valine mutant an increase to 3.30 Å (± 0.02) is observed (see Fig. 3.3). In the alanine mutant, the distance of fluorine to C β of Ala 113 augments to 3.29 Å (± 0.06).

The scaffold shift also impacts on the protein conformation in the flexible loop

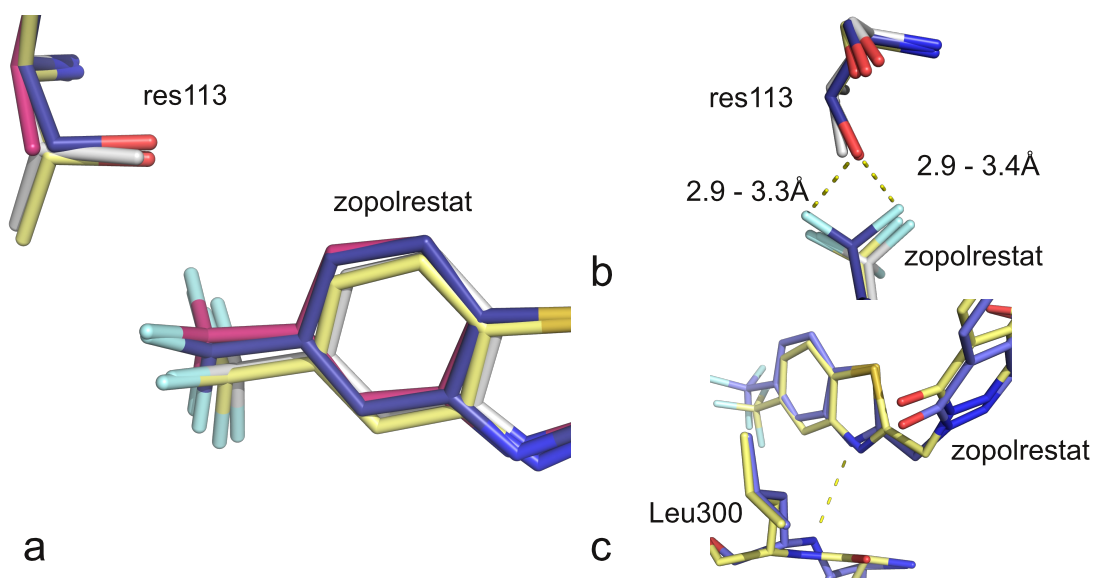


Figure 3.3: Side and top view ((a) and (b), respectively) of the trifluoromethyl group of **6** interacting with residue 113 (alanine – red, serine – blue, valine – grey, threonine (WT) – yellow). Serine (dark blue) matches the hydroxyl group of threonine in wild type (yellow). In the serine complex, **6** displays a shorter distance to this functionality. The methyl group in the valine mutant (grey) is spatially more demanding and thus, **6** is the farthest away. (c) – Backbone flip of Ala299 and Leu300 in the S-**6** complex (blue) resulting in an orientation with the NH toward the inhibitor. In wild type 2FZ9 (yellow) no backbone flip is present.

region of residues 298–303 which are involved in the opening of the specificity pocket. In one of the WT-**6** and all mutant-**6** complexes, a backbone flip of Ala 299/Leu 300 of about 90° enables NH of Leu300 to present its polar hydrogen toward the ligand's benzothiazole nitrogen to form an H-bond (Fig. 3.3(c)). The length of this interaction differs throughout the complexes. For the wild type, a distance of 3.21 Å in the flipped conformation is found, in A-**6** and in S-**6** ($3.17 \text{ Å} \pm 0.06$, $3.17 \text{ Å} \pm 0.02$) a slightly shorter distance and most likely an enhanced interaction is observed. In the valine complex, this distance expands to $3.38 \text{ Å} (\pm 0.01)$, suggesting a weaker H-bond. This assumption is supported by the fact that a second conformation of the Ala 299/Leu 300 backbone atoms is apparent with the unflipped orientation, incompetent to form an H-bond. This agrees with geometries also seen in one of the WT-**6** complex structures.

Furthermore, another conformational protein adaptation is observed by the shift of Phe122. In the different hAR mutant complexes with **6**, Phe122 moves sideways while in the available wild type complexes this residue is found at unchanged position. The phthalazinone moiety forms a π -stacking with this phenylalanine at

a distance of 3.7 Å in the wild type. The distance of this aromatic contact remains almost invariant in the different protein conformations because the phthalazinone moiety shifts by an equal amount in space (Fig. 3.2(a)).

3.1.2.3 IDD 393 Valine Mutant Complexes

The hAR valine mutant in complex with **7** shows similar conformational changes of the protein. The binding mode for this nitro-substituted inhibitor is again dominated by the key interactions of its carboxylic acid with His 110, Trp 111 and the cofactor in the anion binding pocket (PDB code 2IKJ⁴⁵). The distal phenyl moiety occupies the specificity pocket and interacts with Trp 111 through hydrophobic contacts. The same backbone flip found in the hAR mutant complexes with **6** is present: One oxygen atom of the nitro substituent interacts via an H-bond of 3.19 Å (± 0.04) with the backbone of Leu 300. This interaction requires the backbone flip as similarly observed in the wild type complex.

Unexpectedly, the valine mutant crystallized in two different space groups with **7**. Beside the most frequently found monoclinic space group P 2₁ another crystal form shows triclinic P 1 symmetry. Corresponding wild type complexes are available in the PDB (PDB code 2IKJ⁴⁵, 2PZN⁶⁰). Both crystal forms and datasets show equally high quality and were refined to 1.1 Å and 1.3 Å, respectively. The coordinates of the binding pocket and the ligand show only one difference: A shift of Phe 122 is given in the P 1 complex accompanied by a corresponding relocation of the central phenyl ring of the inhibitor to maintain the aromatic contact (see Fig. 3.4(b)).

Comparing V-**7** to the corresponding wild type complexes reveals only a small impact of the mutated residue onto the inhibitor position. The interaction distance of the nitro group to Thr 309 and Leu 300 remains unchanged. The backbone nitrogen of Leu 300 orients toward the oxygen of the nitro substituent in all wild type and valine mutant complexes with **7** independent of the crystallographic space group adopted for crystallization.

3.1.2.4 Complex of a Benzothiazepine Inhibitor with Wild-Type hAR

The crystal structure of **1** (Fig. 3.1, initially discovered as virtual screening hit³⁹) has already been described in detail in Section 2.1.2.

In short, **1** binds to a protein conformation with the specificity pocket in a closed state (Fig. 2.3). The carboxylic acid interacts with the anionic binding pocket via hydrogen bonds to His 110 and Trp 111. Tyr 48 orients one of its aromatic C^δ

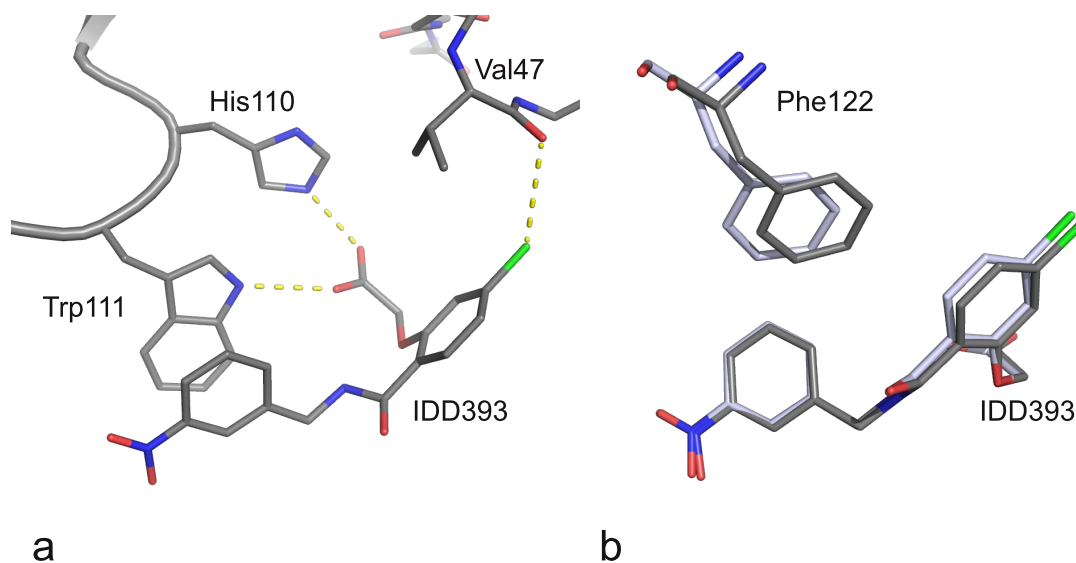


Figure 3.4: (a) – Contacts of **7** to His 110, Trp 111 and Val 47 in the valine mutant complex crystallized in space group P 1. (b) – Superposition of both V-**7** complexes (space group P 1—grey, P 2₁—white). The positioning to maintain the hydrophobic interaction of Phe 122 is the only spatial difference observed between both complexes. As a consequence, the chloro-substituted phenyl ring slightly shifts its position.

H-bond vectors toward the carbonyl function of **1** to form a weak H-bond. The remaining inhibitor atoms experience mainly hydrophobic interactions: Instead of opening the specificity pocket, the phenyl substituent interacts with Phe 122. The π -systems of both aromatic rings accommodate in almost parallel fashion with 3.6 Å distance between the planes. To adapt to this interaction, the protein conformation is changed compared to the holo structure (1AH4⁴⁴): The Phe 122 side chain rotates about 60° toward the outer pocket (see Fig. 2.3(b)). The C $^{\alpha}$ -C $^{\beta}$ torsion angle of phenylalanine increases compared with the complexes with **6** and **7**.

3.1.2.5 Impact of Protein Conformer on Crystallographic Space Group

Most wild type and mutant complexes of hAR that induce a conformational change of Phe 122 crystallize in space group P 1, whereas complexes with the more frequently found proximal orientation of Phe 122 prefer the P 2₁ packing. In the present hAR mutant complexes with **6**, the additional backbone flip of Ala 299 and Leu 300 rearranges the protein to better accommodate the inhibitor's phthalazinone moiety and to form an H-bond to the backbone NH. Both V-**7** complexes show these adaptive features to interact with the nitro group. The Ala 299/Leu 300 backbone flip is also apparent in WT-**6** and in the valine mutant complexed with **7**, though both crys-

Table 3.1: Overview of complex conformers.

Complex	Phe 122 shift ^a	opened spec. pocket Leu 300 flip ^b	backbone flip N Leu 300 ^c
P 2 ₁ complexes			
1US0 (WT- 13)	-	+	-
2FZ8 (WT- 6)	-	+	+
2FZ9 (WT- 6)	-	+	-
2IKJ (WT- 7)	-	+	-
V- 7	-	+	+
P 1 complexes			
V- 7	+	+	+
2PZN (WT- 7)	+	+	+
A- 6	+	+	+
S- 6	+	+	+
V- 6	+	+	+-
WT- 1	+	-	-

+ indicated change present; - no change; +- both conformations visible

^a side chain rotation compared to holo structure (1AH4)

^b specificity pocket opened by a Leu 300 side chain flip

^c backbone flip at Ala 299/Leu 300, nitrogen presents its hydrogen towards ligand

tallize in P 2₁. Interestingly, WT-**1** (space group P 1) exhibits solely the Phe 122 shift without further conformational changes of the protein when compared to a protein conformer where the specificity pocket remains closed (PDB code 1AH4⁴⁴). An overview over binding site conformations of hAR with the different orientation of residues 122, 300 and the backbone 299–300 is presented in Table 3.1.

This table indicates that the orientations of Phe 122 obviously discriminate between crystal packings. Small shifts of Phe 122 translate into greater changes of the protein which interfere with crystal packing. If an inhibitor interacts directly with Phe 122, a shift of Phe 121 is induced. In consequence, the whole neighboring loop is affected: Regarding the backbone atoms of residues 116–130, a variety of different conformations can be observed by superimposing all P 1 complexes (Fig. 3.5(b)). In contrast, in all P 2₁ complexes, this region adopts very similar conformations with low positional variability (Fig. 3.5(a)).

In Figure 3.6, the crystal packing differences in both space groups are shown. Whereas in P 1 only translational symmetry generates the spatial arrangement of the neighboring molecules, in P 2₁ an additional 2₁ screw axis generates the packing (blue). Two contact interfaces should be considered in detail: The first (green box in Fig. 3.6) is formed equivalently in both space groups through translation and involves residues 125–134 on one and helix 56–60 on the other crystal mate. The second interface is generated in case of P 1 packing again by pure translation whereas

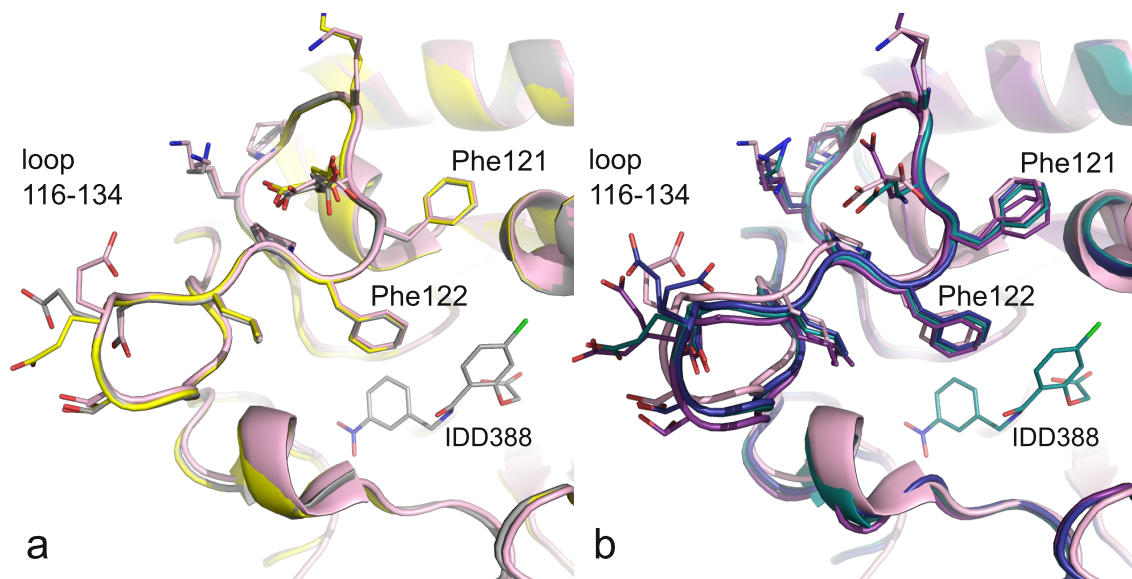


Figure 3.5: (a) – The loop of residues 121 – 134 in three different P₂₁ complexes (1US0 in pink, 2IKJ in yellow, V-7 in grey). The coordinates of Phe122 and residues of the loop match well. (b) – A shift of Phe122 in all P₁ complexes (here: 2PZN - violet, V-6 in blue, V-7 in green, 1US0 (P₂₁) in pink) leads to a relocation of the entire loop with respect to the orientation in P₂₁ (pink) thus preventing crystal contacts in the now triclinic space group (see Fig. 3.6).

in P₂₁ a screw operation is applied. In consequence, two totally different interfaces are produced (red box in Fig. 3.6). Interestingly enough, in the P₂₁ packing the crucial loop region 119–122 hosting Phe122 is involved in exactly this interface, whereas in P₁ close intermolecular contacts to this region are not given. This is shown in detail in Figure 3.7 and listed in Table 3.2. In the P₂₁ packing, the backbone carbonyl oxygen of Lys119 forms an H-bond to N^ε of Gln283 and the backbone nitrogen of Phe121 contacts the carboxamide oxygen of Gln283 on a helical position of a neighboring crystal mate (Fig. 3.7(a)). The placement of the side chain of Phe121 is crucial to allow formation of this contact. If Phe122 is relocated due to inhibitor binding and pushes Phe121 ahead by a kind of domino effect, the spatial arrangement appears to be too strongly perturbed to still favor packing along the 2₁ axis interface. In the alternative packing P₁, the helical portion of the neighboring symmetry equivalent molecule forms a different interaction interface with the helix 56–60 (Fig. 3.7(b)). Most important, in the packing of P₁, neither Lys119, Glu120 nor Phe121 are directly involved in any crystal contacts. Instead, they obviously find enough space to allow for the shifts induced by the relocation of Phe122 upon inhibitor binding.

Table 3.2: Selection of crystal contacts^a formed in the different spacegroups.

contacting atoms		contacts in spacegroup P2 ₁				distance ^b	contacts in spacegroup P1 ^d				distance ^c
residue	atom	contact ^e	SE ^f	residue	atom	[Å]	contact ^e	residue	atom		[Å]
LYS 119	O	HBOND	F	GLN 283	NE2	3.2					
GLU 120	CA	VDW	F	GLN 283	OE1	3.4					
PHE 121	CB	VDW	F	GLN 283	OE1	3.7					
PHE 121	CD2	VDW	F	GLN 283	OE1	3.9					
PHE 121	CD2	VDW	F	GLN 283	CB	4.1					
PHE 121	CE2	VDW	F	GLN 283	NE2	3.7					
PHE 121	CG	VDW	F	GLN 283	OE1	3.7					
PHE 121	N	HBOND	F	GLN 283	OE1	2.8					
GLY 174	N	VDW	E	GLU 60	OE1	3.3	HBOND	B	GLN 49	OE1	3.5
LEU 175	O	VDW	E	VAL 56	CG2	3.5	VDW	B	PRO 24	CD	3.4
LYS 176	O	HBOND	E	ASN 52	ND2	3.0	HBOND	B	GLY 25	N	2.8
GLN 283	NE2	HBOND	B	LYS 119	O	3.2	HBOND	G	GLU 60	OE2	2.6
GLN 283	NE2	HBOND	B	PHE 121	CD2	3.6	VDW	G	GLU 60	OE1	2.8
GLN 283	NE2	HBOND	B	PHE 121	CE2	3.7	VDW	G	LYS 32	NZ	3.0
SER 290	CA	VDW	B	PRO 24	CG	3.8					
SER 290	CB	VDW	B	PRO 24	CG	4.0					
<i>contacts identical in both packings:</i>											
LYS 116	CD	VDW	I	GLU 64	CD	3.8	VDW	E	GLU 64	CD	4.0
ASP 125	CB	SVDW	I	ASP 36	OD2	3.1	SVDW	E	ASP 36	OD2	3.1
ASP 125	CG	VDW	I	ASP 36	OD2	3.2	VDW	E	ASP 36	OD2	3.2
ASP 125	OD2	HBOND	I	ASP 36	OD2	2.5	SVDW	E	ASP 36	OD2	2.5
ASP 125	OD2	VDW	I	ASP 36	CG	3.3	VDW	E	ASP 36	CG	3.4
ASP 125	OD2	VDW	I	ASP 36	CB	3.4	VDW	E	ASP 36	CB	3.4
SER 127	OG	SVDW	I	PRO 266	CB	3.0	SVDW	E	PRO 266	CB	3.1
SER 127	OG	VDW	I	VAL 33	CG2	4.0	VDW	E	VAL 33	CG2	3.9
GLU 150	O	VDW	D	ARG 232	NH2	3.5	VDW	C	ARG 232	NH2	3.6
GLY 151	CA	VDW	D	ARG 232	NE	3.4	VDW	C	ARG 232	NE	3.5
GLY 151	CA	VDW	D	ARG 232	CZ	3.6	VDW	C	ARG 232	CZ	3.6

^a Contacts calculated with MS.^{61;62}^b Distances calculated for V-7.^c Distances calculated for S-6.^d Empty lines indicate that these residues form no crystal contacts in this space group.^e (S)VDW—(short) van-der-Waals contacts; HBOND—hydrogen bond^f SE—symmetry equivalents. A – F indicate different symmetry equivalent molecules.

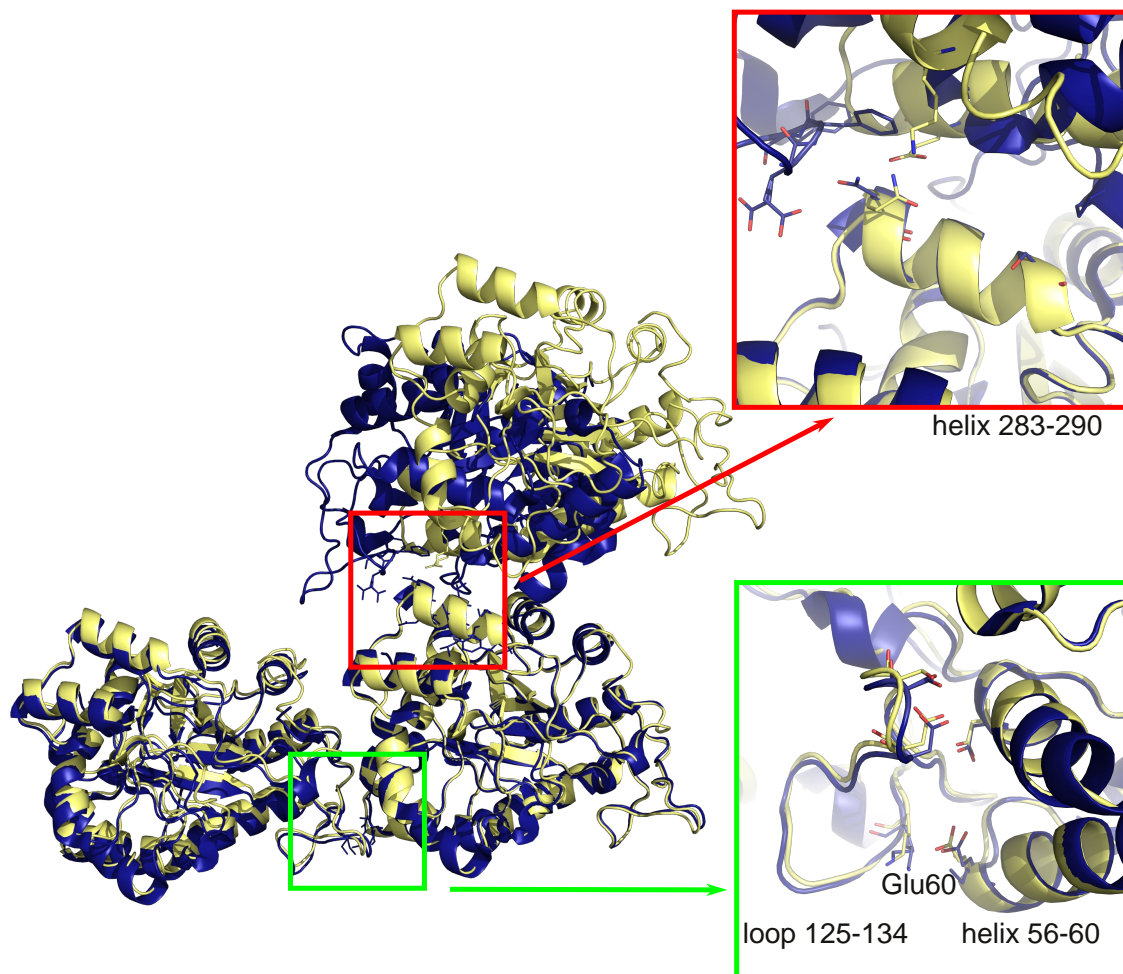


Figure 3.6: Crystal packings in $P2_1$ (blue) and $P1$ (yellow), for clarity only three molecules are shown. In the lower part of the figure, neighboring molecules generated by translational symmetry are displayed. They adopt almost identical orientation in both space groups. They share one common interface section highlighted by the green box. A blow-up of this contact area is shown on the lower right: Contacts between Glu60 at the end of a helical section (residues 56 – 60) and loop 125 – 134 are formed and they occur similarly in both packings. In the upper part of the figure, the intermolecular contacts to the neighboring symmetry equivalent molecules are shown: While in $P1$ (yellow) another translational symmetry is applied, in $P2_1$ the neighboring crystal mate repeats through a two-fold screw axis. In consequence, two different interfaces are formed (red box). On the upper right a blow-up of these interfaces is shown and indicates completely deviating arrangements: They involve contacts of the helical section 283 – 290 with either a helix formed by residues 56 – 60 ($P1$) or the loop comprising Lys 119 – Phe 122 ($P2_1$). The latter contact motifs are shown in detail also in Fig. 3.7 (a) and (b).

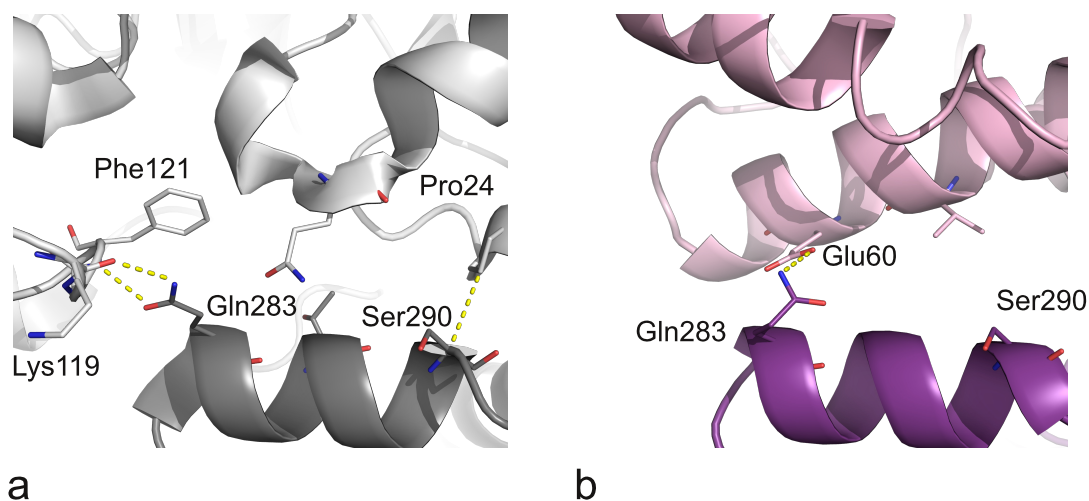


Figure 3.7: (a) – In $P2_1$, crystal contacts of Phe121 of a symmetry equivalent molecule (white) are formed to Gln283 (grey). In addition, the helical segment forms another crystal contact between Ser290 at the end of this helix and Pro24 of the same symmetry equivalent. (b) – In $P1$, Phe121 and Lys119 are no longer involved in crystal contacts. Instead, the interactions of the helical segment 283 – 290 are replaced by contacts to a helix formed by residues 56 – 60.

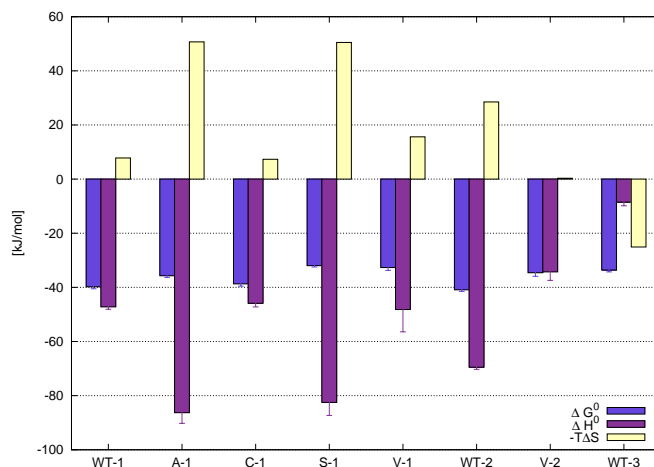


Figure 3.8: An overview of the thermodynamic factorization for all discussed complexes. Alanine and serine mutant hAR **6** complexes show a strong enthalpic signal, leading to a pronounced enthalpy-entropy compensation with high entropic contribution. The wild type complex with **1** is entropically favored with only small enthalpic contributions.

3.1.2.6 Thermodynamic Data

For all structurally characterized mutant and wild type complexes, microcalorimetric experiments were performed to investigate the energetic consequences of the structural changes with respect to inhibitor binding. The dissociation constant K_d and the evolved heat of binding ΔH^0_{bind} were determined by ITC. The Gibbs free energy ΔG^0 represents the affinity of the inhibitors toward the different mutants and was calculated using the equation $\Delta G^0 = RT \ln K_d = \Delta H^0_{bind} - T\Delta S^0$. A buffer dependence had been observed for **12** and **13** binding resulting from a change in the protonation state of Tyr48 in the catalytic center.⁶³ As only relative differences of enthalpic and entropic contributions to binding are of interest in the present comparison, the superimposed heat of ionization, assumed to be equivalent in the different complexes, cancels out. The resulting binding contributions for all studied complexes ΔG^0 , ΔH^0_{bind} , and $-T\Delta S^0$ are shown in Table 3.3, an overview is presented in Figure 3.8.

The thermodynamic data of the different complexes with **6** correlate with the structural changes that occur upon introduction of the mutated residue. Both the alanine and serine mutant show a stronger exothermic signal than the wild type. This gain in enthalpic contribution is remarkably high with a $\Delta\Delta H \sim 34 - 39$ kJ/mol

Table 3.3: Thermodynamic data of the different complexes.

Complex	ΔG^0		ΔH_{bind}^0		$-T\Delta S$	
	[kJ/mol]		[kJ/mol]		[kJ/mol]	
WT- 6	-39.8	± 2.2	-47.2	± 1.8	7.4	± 2.8
A- 6	-35.7	± 0.5	-86.3	± 3.2	50.6	± 3.2
S- 6	-31.9	± 1.0	-80.8	± 2.4	48.9	± 2.6
V- 6	-32.7	± 0.8	-48.2	± 11.0	15.6	± 11.0
WT- 7	-40.9	± 0.5	-69.5	± 1.6	28.6	± 1.7
V- 7	-33.2	± 0.5	-34.3	± 0.8	1.2	± 0.9
WT- 1	-33.6	± 0.5	-8.5	± 1.0	-25.1	± 1.1

but it is overcompensated by an increased unfavorable entropic signal. In consequence, the affinity of **6** to both mutants is reduced. On the contrary, the valine mutant complex exhibits an almost equally high ΔH as the wild type. The free energy of binding reveals a loss in affinity for **6** binding to the valine mutant compared to the wild type.

Also **7** complexed to the valine mutant loses considerably in affinity due to a loss in enthalpic contribution of about 35 kJ/mol. The entropic contribution is nearly zero which in contrast to the wild type complex favors binding of this inhibitor.

For WT-**1**, the thermodynamic profile is reversed: Here, complex formation is entropically favored with a very low enthalpic signal. The binding affinity is comparable to the binding of **6** and **7** to the valine mutant.

3.1.3 Discussion

Flexible proteins may adopt a wide range of distinct conformers. In case of hAR, different protein conformers are formed due to movement and rearrangement of certain residues in loops on the protein surface. Substrate or inhibitor binding can induce protein adaptations to properly accommodate the ligand into the binding pocket. They translate into subsequent relocations of adjacent residues. Likely, formation of these protein conformers is fostered by the mobility of surface-exposed residues and loops, but ultimately they are provoked by the formation of favorable interactions between protein and bound ligand.

In all mutant complexes with **6**, the specificity pocket is opened to experience hydrophobic interactions. This feature is known for various hAR complexes. Furthermore, an additional H-bond is formed via the benzothiazole nitrogen. It induces a backbone flip of Ala 299/Leu 300. Interestingly enough, depending on the applied

crystallization conditions, wild type complexes with **6** could be determined showing either the backbone flip resulting in the hydrogen bond or lacking this adaption (PDB code 2FZ8, 2FZ9).⁵⁹ It is even more remarkable that all hAR mutant complexes with **6** show this flip, competent to form the H-bond. In A-**6**, this H-bond is optimized to a shorter contact distance, which supposedly strengthens inhibitor binding. The observed spatial scaffold shift of **6** results in consequence of the mutation. The introduced alanine instead of the wild type's threonine enlarges the volume at the rear of the specificity pocket. Therefore, the inhibitor can penetrate deeper into the pocket and occupies a better position to experience the above-mentioned shorter H-bond. In WT-**6**, threonine's O γ does not form an H-bond to **6** because the ligand lacks an appropriate functionality at the adjacent position of its skeleton. Though the replacement of the hydroxyl functionality (in A-**6**, V-**6**) as interaction partner for **6** has presumably little impact on the enthalpic signal when threonine is replaced by alanine, the shorter H-bond to Leu300 reveals a favorable enthalpy contribution. Regarding S-**6**, the position of the inhibitor is similar to that in the alanine complex, while the mutated serine hydroxyl group matches the wild type's threonine hydroxyl. The distance of the inhibitor's fluorine atom to serine's O γ decreases slightly compared to wild type, and, as seen in the alanine complex, the phthalazinone-protein interaction at Leu300 is similarly reduced. As a matter of fact, also the serine complex indicates a stronger enthalpic contribution to binding ($\Delta\Delta H$) compared to the wild type.

Obviously, in V-**6** the situation becomes different: The methyl group of valine is spatially more demanding than the hydroxyl functionality of threonine in the wild type. In consequence, the inhibitor penetrates less deeply into the specificity pocket. The distance of the H-bond to NH of Leu300 expands to 3.4 Å. This weakens the interaction, the backbone of Ala299 and Leu300 regains mobility and the second conformation incompetent to form an H-bond is populated. In consequence of the partial loss of this interaction to Leu300, in V-**6** entropy is favored and finally a similar ΔH contribution as in the wild type is observed.

It has to be noted that differences in the accuracy of wild type and mutant complexes are given. While shifts of about 1.0 Å are clearly visible in the electron densities of the mutant complexes (resolution of about 1.0 Å for S-**6**, V-**6**, V-**7**), such minor differences are at the border of significance in complexes determined at resolutions beyond 1.5 Å (2FZ8, 2FZ9, A-**6**). Thus, the differences between, e.g., **6** in the serine and valine mutant are significant, while it is more difficult to compare these shifts to the wild type complexes. Estimated standard deviations for the

discussed distances in A-**6** and A-**7** complexes are between 0.05 and 0.06 Å, while for S-**6**, V-**6**, and V-**7** estimated standard deviations fall between 0.01 and 0.02 Å.

Although the bulkier methyl group of valine takes the position of the former threonine's hydroxyl group, there is no scaffold shift observed for **7** indicating geometrically unfavorable interactions with the introduced valine. Most likely, this observation results from the constraining effect of the nitro group being involved in an H-bond to Leu 300 NH⁴⁵ either in WT-**7** and V-**7**. Nevertheless, a repulsive component to the valine methyl group might be given as V-**7** sacrifices major parts of the enthalpic advantage found for the wild type complex. However, this loss is largely compensated by an entropic benefit supposedly governed by higher residual mobility and/or changes in the desolvation costs of the less polar valine mutant pocket. It is interesting to see that in case of **6** the mutations at position 113 take much larger impact on the placement of the ligand's skeleton. Some shifts are observed that correlate with repulsive or attractive interactions with residue 113. The constraining effect of the rather long H-bond between the benzothiazole nitrogen and Leu 300 NH is less pronounced, even though we attribute part of the shifts in the thermodynamic signature to this interaction.

1 displays a totally different binding mode which is also reflected by its deviating thermodynamic signature: Even though its carboxylate function interacts analogously as in the other complexes, the remaining interactions are of hydrophobic nature and contribute obviously entropically. Upon complex formation, the rigid seven-membered ring of the scaffold buries the considerable amount of 220 Å² of the closed binding site surface resulting in an entropic advantage.

The complexes studied in this contribution crystallize in two different packing environments, represented by either P 1 or P 2₁ symmetry. The occurrence of the different packing motifs in molecular crystals is a very well-known principle. Also the switch between different crystal forms adopted upon co-crystallization with different ligands is a frequently observed phenomenon; however, rarely a conclusive explanation can be provided for these observations. In the present examples, the change in space group and crystal packing is obviously provoked by a slight shift of Phe 122 upon ligand binding. Strong induced fit of the side chain of this aromatic residue perturbs in a kind of domino effect the adjacent residues and translates even to those residues involved in the crystal packing of the P 2₁ crystal form. In consequence, the protein molecules start to assemble upon crystal nucleation in a different fashion and the association of the neighboring crystal mates does not follow 2₁ screw axis symmetry any longer, but pure translational association. This avoids

interface formation via the loop which resides the crucial residues next to Phe122. The latter amino acid is affected by ligand binding.

3.1.4 Conclusion

The present study analyzes by high resolution crystallography and isothermal titration calorimetry the structural and energetic consequence of mutational exchanges of one residue at the rear end of the specificity pocket of aldose reductase. The latter pocket is only opened upon induced-fit adaptations initiated by appropriate hydrophobic groups attached to a ligand skeleton. Three types of ligands with deviating basic scaffold were investigated, all comprising a carboxylate function to anchor the ligands in the catalytic center. With an entropically favored binder (**1**) the specificity pocket remains in closed state. The other two (**6**, **7**), both favored by an enthalpic binding contribution, open the specificity pocket. They differently experience the introduced mutations in this pocket. Whereas the first seems to be more strongly affected in its binding mode by a shift of its scaffold, the second appears to remain positionally unaffected. Supposedly, this effect is caused by a hydrogen bond of deviating restrictive power. It also affects the thermodynamic signature of the complexes which shifts toward the more enthalpic binding when the strength of the H-bond increases and the repulsive effects decrease. Furthermore, enhanced residual mobility seems to be of entropic benefit.

The series of the studied ligands require slightly different space next to residue Phe122. Some induced-fit adaptations result from these requirements. However, they obviously translate into significant perturbations in the close neighborhood. This provokes a transition to another crystal form which avoids packing contacts to the crucially affected residues. This example shows how tiny effects caused by the bound ligand can translate into major effects strongly perturbing intermolecular recognition in the assembly of the crystalline packing patterns.

3.2 Stabilizing a New Protein Conformation

3.2.1 Introduction

As mentioned above, the conformational space of hAR comprises a number of different conformers. They can be grouped into three main conformers. Two main conformers with smaller side chain variations were described in Chapter 2 and Chapter 4, respectively. One protein conformer discovered with the inhibitor Tolrestat is unique until today (PDB code 2FZD⁵⁹).

The wide range of possible binding site conformers already known—regardless the ones still to be discovered—frequently arises the question what actually causes the changes on protein side. Accompanying this question is often the issue whether a certain conformer can be evoked by a ligand. Thus, a virtual screening has been performed by Peinz⁶⁴ to explicitly address the conformation with an opened specificity pocket where the backbone of Leu 300 is flipped (PDB code 2FZ8) to present the hydrogen of the backbone nitrogen toward the inhibitor. In consequence, within the flexible loop an additional hydrogen bond is provided.

3.2.2 Suggested Compounds and Evaluation

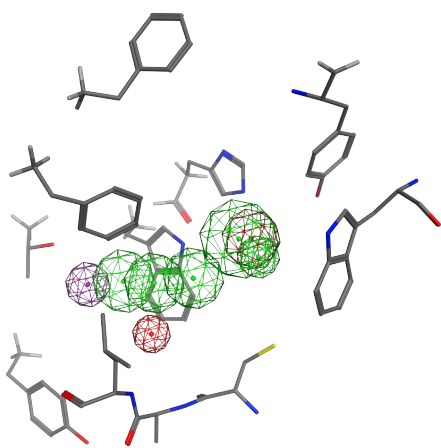


Figure 3.9: Pharmacophore used in the virtual screening. Green spheres indicate hydrophobic features, violet doneptor features, red hydrogen bond acceptor functionality.

The pharmacophore used for the query restrains the anchor as a carboxylic acid. Additionally, the core consisted of several hydrophobic features. In the back part of the specificity pocket, a combined donor-acceptor feature was placed. One of the key features was the acceptor property below the central hydrophobic part, which should address the flipped backbone of Leu 300 (see Fig. 3.9). Strategy, pharmacophore features and detailed workflow were performed by Peinz⁶⁴ and will be described elsewhere; kinetic and crystallographic evaluation of the most promising compounds are part of this

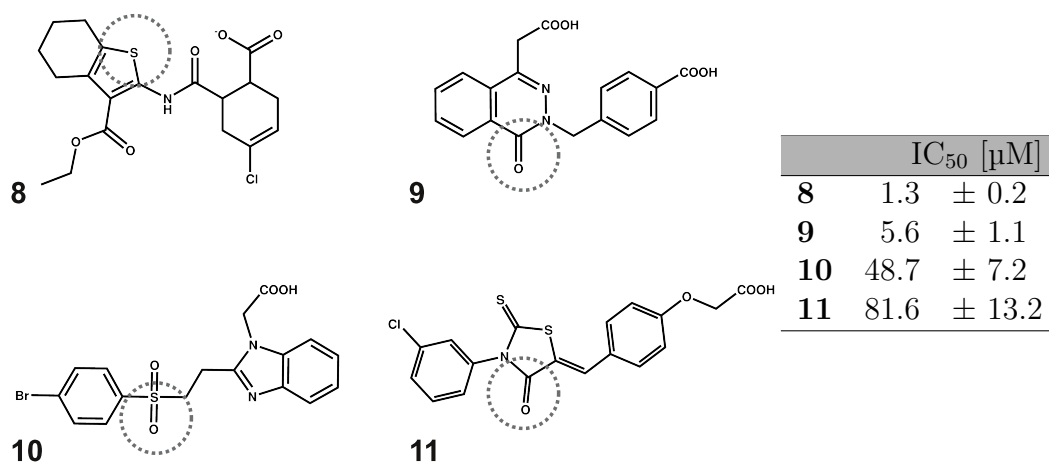


Figure 3.10: Promising hits of the virtual screening and kinetically determined IC₅₀ values. All hits exhibit a carboxylic acid as predicted anchor. Additionally, hydrophobic features to bind into the specificity pocket are present. Adjacent to the latter, there are potential H-bond acceptors to interact with the backbone of Leu 300 and evoke the desired conformer (examples indicated in circles).

thesis.

Some of the most promising hits are presented in Figure 3.10. Like most known inhibitors, these compounds have a carboxylic acid functionality. The feature of this anchor group was a core demand in the pharmacophore, along with hydrophobic features to be placed between Trp 111 and Leu 300. Additionally, all of them bear a possible H-bond accepting functionality, which is indicated in Figure 3.10. Determined IC₅₀ values for these compounds range between 1 and 100 μM which is promising for a first virtual screening.

The crystal structure of hAR in complex with compound **9** was determined to a resolution of 1.7 Å (see Fig. 3.11).

Regarding its chemical structure, the binding mode of **9** came to no surprise. The phthalazinone-acetic acid builds the anchor and binds to His 110 and Trp 111 (see Fig. 3.11). The specificity pocket is opened upon ligand binding, and while the aromatic π-system of the distal phenyl ring interacts with Trp 111, the second carboxylic acid most likely accepts a proton donated by Thr 113.

The phthalazinone substructure is already known from **6**. Consequently, in the superposition of the WT-**9** complex with 2FZ8, this part matches very well (see Fig. 3.11(b)). Thus, the oxygen is directed outside the pocket with no specific interaction partner instead of accepting an H-bond from the protein. This atom was

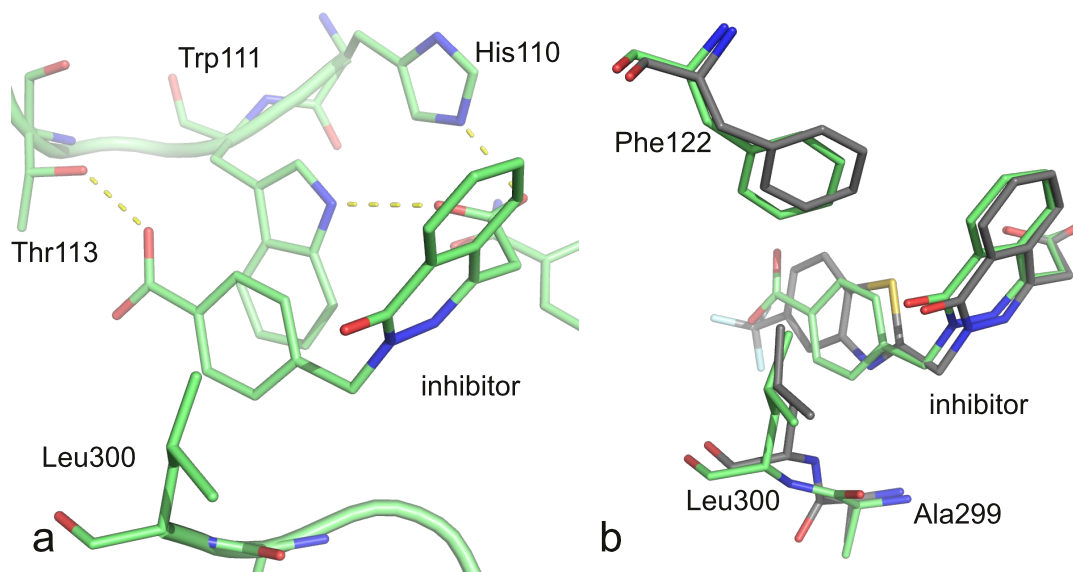


Figure 3.11: (a) – The carboxylic acid of **9** forms the typical interactions to His 110 and Trp 111. The phenyl moiety interacts via π -stacking with the aromatic moiety of Trp 111. The second carboxylic acid forms an interaction to Thr 113 in the back part of the specificity pocket. (b) – Superposition of WT-**9** complex (green) with 2FZ8 (grey). The overall inhibitor position is similar. A slight shift of the phthalazinone moiety of **9** evokes a small adjustment of Phe 122. Unfortunately, the backbone flip as present in 2FZ8 is absent in the WT-**9** complex: no inhibitor functionality addresses this H-bond donor.

supposed to interact with the nitrogen of Leu 300; the scaffold provides no further H-bond acceptor. Thus, the backbone of Leu 300 remains in its unflipped conformation known from most hAR complexes (like e.g. in 1US0).

3.2.3 Conclusion

The performed virtual screening resulted in some compounds that bind in the low micromolar range. Compared to known scaffolds that inhibit hAR, the carboxylic acid is well known as anchor functionality and was set as mandatory by the screening options. Many common scaffolds of hAR inhibitors comprise the carboxylate anchor with a distal aromatic moiety and an—often also aromatic—linker between the functionalities (described in Sec. 2.2). Comparing the screening results with the described scaffold they show a certain similarity. An additional hydrogen bond acceptor is present in the center of the compounds, and was the key functionality regarding a main aim of the screening: to enforce not only the opening of the specificity pocket, but also a backbone flip between Ala 299 and Leu 300 upon ligand

binding. Unfortunately, the crystal structure of **9** resembles the protein binding site conformation of 1US0⁶⁵ where no backbone flip of Leu 300 is present.

4. Tracing the Detail: How Mutations Affect Binding Modes and Thermodynamic Signatures of Closely Related Aldose Reductase Inhibitors*

4.1 Introduction

The reliable prediction of binding affinities of small molecule ligands towards a given target solely estimated from the geometry of a formed protein-ligand complex is one of the most crucial steps in rational drug design. Even though substantial effort has been and is currently spent to improve such methods, they still face limitations in their reliability and accuracy.^{46;66} To further develop and improve predictive computational methods, the understanding of the nature of molecular interactions, the principles of biomolecular recognition and of selectivity determining features are of utmost importance.

The present contribution tries to address particular aspects of these three essential issues using high resolution crystal structure analysis, site-directed mutagenesis and isothermal titration calorimetry. In recent time, also stimulated by the fact that bromine is increasingly used as easily detectable substituent at small organic ligands used in crystallographic fragment-based lead discovery, a number of protein-ligand complexes have been determined with exceptionally short bromine-to-oxygen distances.^{53;67} As prominent example, the ultra-high resolution complex of aldose reductase with the brominated ligand IDD 594 (0.66 Å resolution) was investigated by charge-density studies.⁶⁵ They render the interaction of a Thr 113 O γ and a bromine atom as a direct, strongly polarized and electrostatically driven contact.⁶⁸ However, the detailed nature and its affinity contribution remains unclear. We therefore studied the importance of this interaction by structurally conservative site-directed mutagenesis, as the wild-type threonine was replaced by cysteine, serine, alanine and valine. To obtain detailed insights into molecular recognition principles, we included two brominated inhibitors which differ only in two positions of their skeleton (see Fig. 4.1). Even though very similar affinities and thermodynamic signatures are recorded versus the wild-type enzyme, rather distinct profiles are observed with

*Manuscript published in *Journal of Molecular Biology*, 2011, **406**(5): 700–712.

respect to the single-site mutants. Thanks to very high resolution of the determined crystal structures of the mutant complexes, structural details become available indicating that structural protein modifications are differently responded by the bound ligands. These include a significant opening of the central bond angle of the ligand, a molecular parameter usually regarded as virtually fixed in most computational docking programs.

In general, mutational changes of a wild-type protein are the explanation for the development of resistance that can hit even structurally related ligands to varying extend. Similarly, the exchange of simple residues among distinct isoforms of a target protein can be responded by congeneric ligands quite differently. Using the mutated enzymes produced in this study, structural and thermodynamic characteristics become apparent that can even reverse the selectivity profile against closely related ligands.

For our study, we selected human aldose reductase (hAR) as a model protein. This enzyme is a 36 kDa protein that consists of 315 amino acids and folds into a TIM-barrel. Its binding pocket is located near the C-terminus where also the cofactor NADPH is bound.^{16;17} The so-called anionic binding pocket is formed by His 110, Trp 111, and the nicotinamide moiety of the cofactor. Here, the most commonly used carboxylate type anchor group builds one of the core interactions with the enzyme.

In the human organism, the expression of hAR is induced by high glucose blood levels. hAR catalyzes the reduction of aldehydes and ketones; especially, glucose is reduced to its corresponding alcohol. Predominantly, high turn-over rates of NADPH are discussed apart from the enhanced accumulation of the highly polar sorbitol inside the cell to be responsible for long-term diabetic complications such as neuropathy, nephropathy and glaucoma.^{19;21-23} This qualifies hAR as promising drug target.

hAR is closely related to aldehyde reductase, a physiologically highly required protein exhibiting pronounced sequence identity.²⁴ To selectively address hAR with respect to aldehyde reductase, the distinct flexibility of the binding site can be exploited. A loop containing Leu 300 can react flexibly and, thus, opens an additional specificity pocket not present in aldehyde reductase. Ligands stabilizing this open conformer give access to the specificity pocket and promote selectivity. Therefore, detailed investigations of the interactions in this mainly hydrophobic pocket are one of the key elements to endow inhibitors with selectivity properties that avoid undesired side effects by interfering with aldehyde reductase activity.

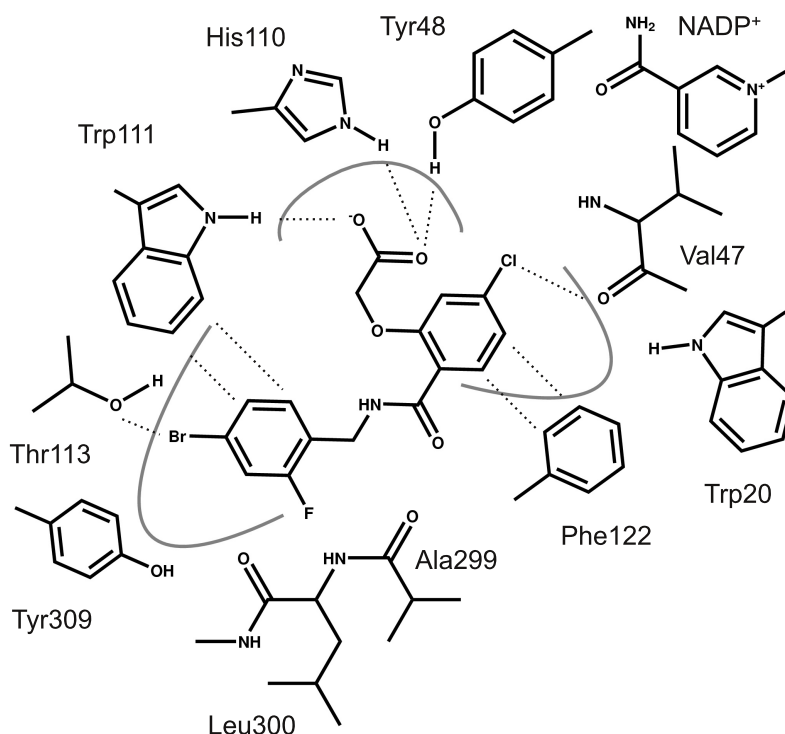


Figure 4.2: Schematic binding mode of **12** in wild type hAR. The anionic binding pocket is formed by His 110, Trp 111, Tyr 48 and cofactor. Phe 122 hydrophobically interacts with the central phenyl moiety. The specificity pocket between Leu 300 and Trp 111 provides hydrophobic π -stacking while bromine of **12** interacts with Thr 113 in the back part of this pocket. The same binding mode is observed for all studied mutant hAR complexes with **12** and **13**.

C-**12**, S-**12**, V-**12**) shows, the ligand binds with similar geometry as in the wild-type complex (PDB-code 2IKI, see Fig. 4.2). The anionic binding site, mainly consisting of Trp 111, His 110, Tyr 48 and the cofactor NADP^+ , remains virtually unchanged and tightly binds the inhibitor's carboxylate anchor. Furthermore, the nicotinamide of the cofactor interacts electrostatically with the negatively charged acidic function.⁴⁵

In all four complexes, the inhibitor occupies the specificity pocket in the open state due to rearrangements of Leu 300. Although highly adaptive, this flexible loop adopts the identical conformation as in wild-type complex (Fig. 4.3(a)). Trp 111 experiences π -stacking with the distal bromine substituted phenyl moiety of **12**. The proximal chlorine substituted phenyl moiety interacts hydrophobically with Phe 122. Though Phe 122 is known to act flexible occasionally, in all four complexes

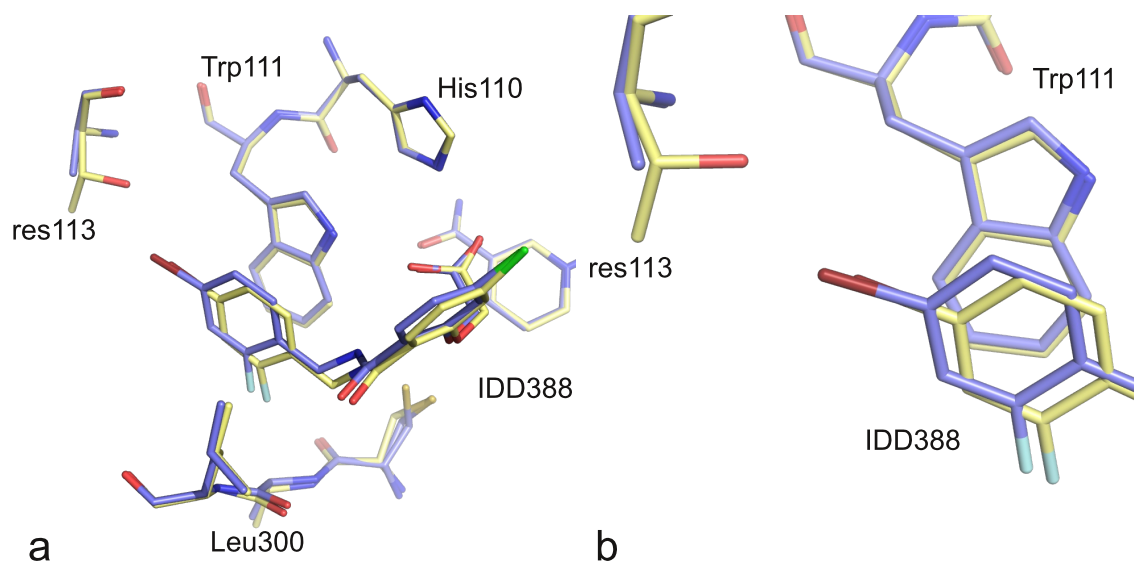


Figure 4.3: Superposition of T 113 A mutant (blue) and wild type (yellow) in complex with **12**. (a) – For the protein, only small deviations are visible. Residues His 110, Trp 111, and cofactor form the same interactions. The flexible loop around Leu 300 adopts identical conformations in both complexes. In the specificity pocket, a π -stacking between Trp 111 and inhibitor and the interaction to Thr 113 and Ala 113 are predominant. (b) – The mutated residue Ala 113 with a smaller spatial requirement than the former threonine leaves a free volume. In consequence, the inhibitor shifts about 0.6 Å to occupy the available space.

no shift of this residue is observed. The chlorine interacts with oxygen of Val 47 in all complexes at a distance of 3.2 Å (see Fig.4.4).

The mutated residues are located at the far end of the opened specificity pocket and interact with the ligand. In case of alanine, this size-reduced residue cannot occupy the space formerly taken by threonine. The ligand responds by a clear shift of 0.6 Å of its bromine substituted phenyl ring to optimize the distance towards residue 113 compared to wild type (see Fig.4.3(b)). While the threonine's O^γ distance to bromine amounts to 2.9 Å⁶⁵, the C^β distance of alanine to bromine augments to 3.7 Å (see Tab. 4.1).

In the serine complex (S-**12**), the side chain oxygen formally replaces the corresponding atom of Thr 113. Surprisingly, Ser 113 adopts two conformations. The higher populated one (about 72%, see Tab. 4.1) orients O^γ away from the ligand's bromine atom and locates the oxygen at the former methyl position of Thr 113 in the wild type, resulting in a 3.8 Å distance to bromine. The second, less populated one comes to a short $O^\gamma \cdots Br$ distance of 2.8 Å, similarly to the threonine's O^γ in wild-type complex. However, some care is needed for the interpretation of the lat-

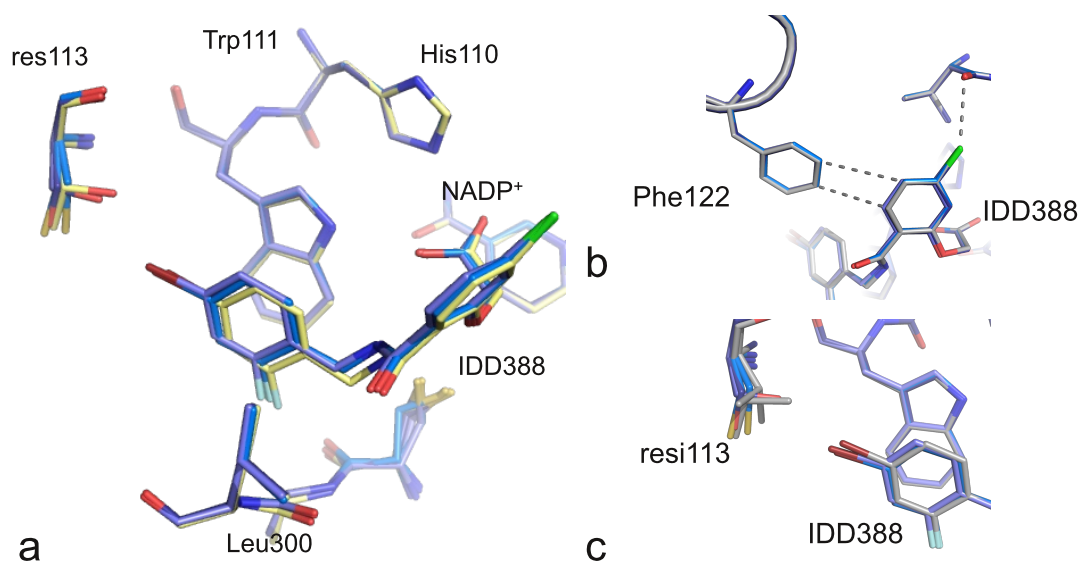


Figure 4.4: Superposition of **12**-mutant complexes and wild type (yellow). (a) – Most binding pocket residues remain unaltered: Positions of His 110, Trp 111, NADP⁺ and the conformation of the flexible loop of residues 298–303 are identical in all complexes. (b) – Carboxylic anchor and chlorine substituted phenyl ring form the same interactions as in wild type. (c) – The different mutants show different side chain rotamers: Serine (dark blue) interacts with two conformers, cysteine (light blue) is directed in an averted position while valine (gray) predominantly presents its C^β toward bromine. The C^γ of valine's second conformation is too close for a bromine contact (2.4 Å).

ter lower occupied species, as to a certain extend radiation damage can create an additional superimposed molecular geometry (see Sect. 4.2.3).

In case of the T 113 C mutant, the sulfur possesses a much larger van-der-Waals radius than oxygen. In consequence, in C-**12** it exclusively adopts a conformation oriented away from the inhibitor. The ligand's position is not different in comparison to the serine mutant complex, resulting in a Br ··· S distance of 3.4 Å.

When valine is introduced, the threonine's O^γ is replaced by methyl and, additionally, increases the residue's volume. Valine adopts two distinct conformations almost equally populated. Taking the ligands occupancy into account, the preferred directionality would be with C^β of valine as interaction partner with a 3.6 Å distance to bromine.

In this protein complex, the ligand's occupancy was refined to 72%. A minor conformation of the flexible loop around Leu 300 represents the specificity pocket in its closed state indicating only partial occupancy. Additionally, the bromine was refined to even lower occupancy due to radiation damage (see Sect. 4.2.3). The dis-

tance of 2.4 Å of valine's C γ in the second conformation to the bromine substituent would be far too short for a bromine contact (see Fig. 4.4(c)). According to these findings, the second valine conformation corresponds to the geometry of the unoccupied binding site. Possibly, this geometry also corresponds to the lacking bromine at this position due to radiation damage.

Another structural indicator for the geometrical and thus conformational adaptations of the inhibitor is the valence angle α at the methylene linker between both aromatic moieties (see Fig. 4.1). While the peptide bond should be planar, a variation of this angle correlates with the mutual distance of both phenyl moieties.

The largest shift observed in A-**12** is accompanied by an expansion of the original tetrahedral angle up to 117.0° (see Tab. 4.1). In contrast, in V-**12** a small shift of the bromine position results in a smaller angle of 114.8°. A CSD⁴³ search was performed to investigate the possible range of such angles between peptide connected phenyl moieties. Observed angles range from 107° – 117°, among which the range between 110° – 114° is most populated.

Though the modifications in the protein geometry can partly be compensated by geometrical adaption of **12**, geometrical restraints prevent an optimal ligand position in the modified protein binding site. Especially in A-**12**, a certain volume remains unallocated. The difference of the unoccupied volume between wild type complex and alanine mutant complex amounts to about 8 Å³.

4.2.2 Crystal Structures of the Mutant IDD594 Complexes

The binding modes of **13** in the mutant complexes are very similar to those of **12**. Identical protein conformations are observed, while the inhibitors bind with small positional differences (Fig. 4.5). The carboxylate anchor groups interacting with His 110 and Trp 111 match well. A small shift of the proximal fluorine substituted phenyl moiety within the aromatic plane results in equal positions for both halogen atoms (F/Cl, see Fig. 4.5). Thus, fluorine can interact with oxygen of Val 47 like chlorine in **12** with a distance of 3.2 Å. In the specificity pocket, the bromine substituted aromatic ring forms a π -stacking with Trp 111. Even the back parts of the specificity pockets show virtually identical geometry in corresponding complexes. In A-**12** and A-**13**, the position of the bromine and Ala 113 almost perfectly match. In the serine and cysteine complexes with both ligands, the mutated residue 113 shows the same conformational distribution. The resulting interaction between bromine of **13** and oxygen or sulfur of serine and cysteine, respectively, resemble the interaction

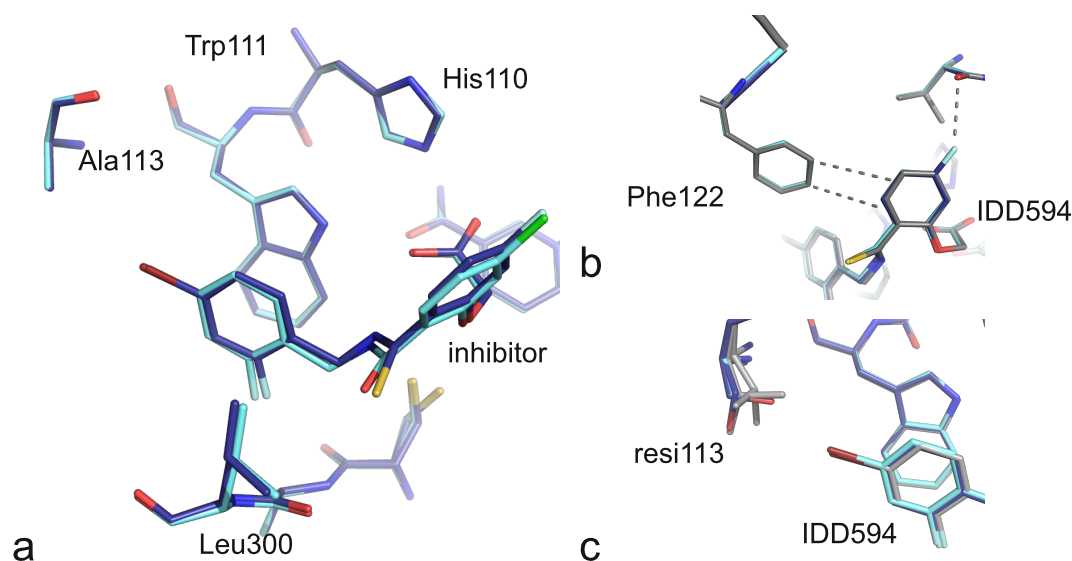


Figure 4.5: (a) – Superposition of T 113 A mutants complexed with **12** (blue) and **13** (cyan). While the anchor positions are identical, the central phenyl moiety shifts in plane resulting in both halogen atoms at the same position. (b) – As with **12**, the binding site conformation in the mutant complexes with **13** are identical. π -stacking to Phe 122 and the fluorine interaction to oxygen of Val 47 restrain the inhibitor's position. (c) – Side chain rotamers of the mutant residue 113 resemble the mutants in the **12**-complexes. The inhibitor adapts to the new environment with a slight shift toward (for alanine and serine) or away (for valine) from the mutation.

in the corresponding **12** complex, both in distance and geometry (see Tab. 4.1).

Thus, the complex series of **13** shows the same trend as observed for **12**: the inhibitor adapts its conformation to the spatial and contact characteristics of the mutated amino acid leading to a shift that is largest for A-**13**, followed by C-**13**, S-**13**, and V-**13**. In the latter, the ligand position is closest to the wild type.

4.2.3 Bromine Cleavage Due to Radiation Damage

In all mutant complexes the bromine substituents of **12** and **13** adopt different occupancies and none of them reaches full occupancy (see Tab. 4.1). The high energy of the synchrotron beam induced radiation damage during all X-ray experiments. Thus, a certain amount of each bromine substituent is cleaved off. The cleavage product is most likely to be a bromide ion⁶⁹ and relocates to an adjacent position between the backbone NH of residue 113 and the inhibitor (see Chapt. 5). This position is almost identical for the alanine, cysteine and serine mutant complexed either for **12** or **13**. For the valine mutant complexes, another position right beneath

Table 4.1: Distances, occupancies of ligand-bound bromine and split side-chain conformations of residue 113

Complex	Br...C ^{β} _113 [Å] 1 st (2 nd) conf.	Br...O ^{γ} /C ^{γ} /S ^{γ} [Å] 1 st (2 nd) conf.	Br occupancy	Residue 113 occupancy	angle α^a [°]
A- 12	3.74 ± 0.03	—	0.71	1.00	116.99 ± 3.55
C- 12 ^b	3.75 ± 0.02 (3.93 ± 0.03)	3.43 ± 0.00 (3.91 ± 0.03)	0.76	0.67 (0.33)	115.47 ± 0.66
S- 12 ^b	3.99 ± 0.06 (3.88 ± 0.02)	2.76 ± 0.01 (3.75 ± 0.01)	0.71	0.28 (0.72)	116.32 ± 0.74
V- 12 ^b	3.59 ± 0.01 (3.70 ± 0.01)	3.62 ± 0.01 (2.37 ± 0.02)	0.44	0.48 (0.52)	114.82 ± 0.63
WT 12 ^c	4.00	2.90	1.00	1.00	110.30
A- 13	3.86 ± 0.01	—	0.73	1.00	113.31 ± 1.05
C- 13 ^b	3.75 ± 0.04 (4.04 ± 0.03)	3.40 ± 0.01 (3.89 ± 0.03)	0.71	0.52 (0.48)	112.17 ± 1.18
S- 13 ^b	4.06 ± 0.10 (3.93 ± 0.09)	2.79 ± 0.03 (3.80 ± 0.04)	0.76	0.32 (0.68)	112.77 ± 2.97
V- 13 ^b	3.98 ± 0.02 (3.73 ± 0.03)	2.79 ± 0.03 (3.60 ± 0.02)	0.52	0.63 (0.37)	114.24 ± 1.08
WT 13 ^c	4.10	3.00	0.94	1.00	112.30

^aAngle between the amide linker of **12** and **13** and the distal phenyl moiety as indicated in Figure 4.1.^bValues in parentheses apply to the corresponding atoms of the second conformations of residue 113.^cStructures are taken from the PDB (2IKI, 1US0, respectively) for comparison.

the original Br-position is suggested: It only became apparent as difference electron density peak during the refinement of one of the two complexes. However, all residual inhibitor atoms remain in place.

Disruption of the bromine bond creates further space between mutated amino acid 113 and inhibitor. During refinement, for several of the complexes a second conformation of the residue’s side chain could be assigned. In case of serine and cysteine mutants, a short oxygen- or sulfur-to-brom distance is suggested for these latter conformers. As the relative occupancies of these conformers and the corresponding bromine substituents speak for a higher population of the side chain conformers compared to the bromine presence within the estimated standard deviations, it cannot be excluded that a certain fraction of the complexes is present that exhibits—as the wild type—the short O/S...Br distance.

4.2.4 Characterization of the Thermodynamic Binding Profile

For further characterization, microcalorimetric experiments were performed. ITC was used to determine the dissociation constant K_d and the evolved heat of binding ΔH_{bind}^0 . ΔG^0 and $T\Delta S^0$ were calculated using the equation

$$\Delta G^0 = -RT \ln K_d = \Delta H_{bind}^0 - T\Delta S^0.$$

For the alanine, cysteine and valine mutants, the thermodynamic profile for the binding of **12** and **13** could be determined. For the serine mutant, only the thermodynamic profile for **13** could be recorded. Surprisingly, multiply performed measurements for S-**12** revealed only very weak heat signals which did not allow for the collection of reliable data for curve fitting. This observation remains unclear to us, in general, weak heat signals are given in cases of predominantly entropic

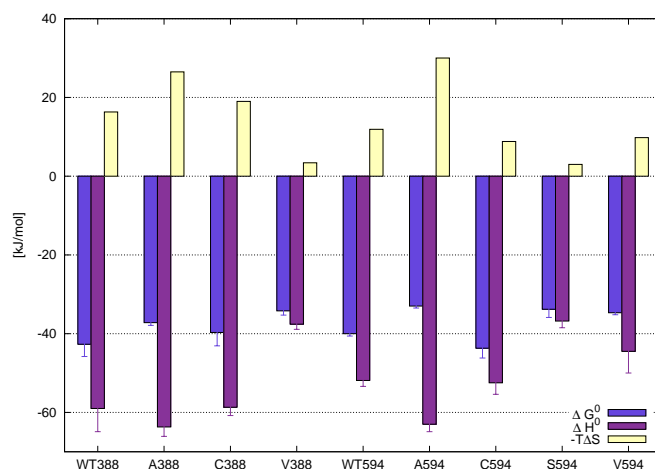


Figure 4.6: Overview about the thermodynamic data of the different complexes. WT 12 has been determined elsewhere.⁷⁰ Data are not corrected for superimposed heat of ionization resulting from a change in protonation state. Very similar superimposed effects will be given for all considered complexes.

binding. We are reluctant to conclude from this fact that for the S-**12** complex a largely entropically driven binding is detected.

The binding characteristics of **12** to wild type hAR have previously been determined⁷⁰; for comparison, the binding contributions of WT 594 have been collected in this study. A buffer dependence had been observed for **12** and **13** binding resulting from a change in the protonation state of Tyr 48 in the catalytic center. Similar changes will be given for the mutated complexes studied here. However, as Tyr 48 is very remote to the mutation site 113, a very similar contribution from the protonation change will occur for all complexes. As only relative differences of enthalpic and entropic contributions to binding matter in this comparison, the superimposed heat of ionization cancels out. ITC results are listed in Table 4.2; relative differences are given in Table 4.3 to Table 4.5 and a graphical overview is displayed in Figure 4.6.

Upon mutation, modulations of the free energy of binding of approximately 7 kJ/mol are observed which corresponds to nearly two orders of magnitude in the binding constant. More remarkable are the changes in the enthalpy and entropy which extend maximally to about 27 kJ/mol. Considering the much smaller total changes in ΔG , a pronounced enthalpy/entropy compensation is given in the series. Crystal structure analysis showed only minor modulations in the binding modes of both ligands in the mutant complexes. Nevertheless, they provoke huge changes

Table 4.2: Thermodynamic data^a of the different complexes.

Complex	ΔG^0		ΔH_{bind}^0		$-T\Delta S$	
	[kJ/mol]		[kJ/mol]		[kJ/mol]	
WT 12 ^b	-42.7	± 3.1	-59.0	± 5.9	16.3	± 9.0
A- 12	-37.2	± 0.7	-63.7	± 2.4	26.5	± 2.5
C- 12	-39.7	± 3.4	-58.7	± 2.1	19.0	± 4.0
V- 12	-34.2	± 1.1	-37.6	± 1.3	3.4	± 1.7
WT 13	-40.0	± 0.6	-51.9	± 1.5	11.9	± 1.6
A- 13	-33.0	± 0.5	-63.0	± 1.9	30.0	± 2.0
C- 13	-43.7	± 2.5	-52.5	± 2.9	8.8	± 3.8
S- 13	-33.8	± 2.1	-36.8	± 1.7	3.0	± 2.7
V- 13	-34.7	± 0.5	-44.5	± 5.5	9.8	± 5.5

^aValues are not corrected for a protonation change at Tyr 48 which is equivalently superimposed in all studied complexes.

^bWT 388 has been determined elsewhere.⁷⁰

in the enthalpic and entropic contributions and, interestingly enough, they are reflected by different profiles for the two largely similar ligands. Due to compensating effects, $\Delta\Delta G$ differences between the two ligands and mutant complexes can reveal rather small values close to experimental accuracy limits. However, in some of these examples, the small $\Delta\Delta G$ values result from large opposing differences in $\Delta\Delta H$ and $T\Delta\Delta S$ which shows that the thermodynamic profiles of the corresponding complex can differ to large extend.

All mutant complexes with both ligands show an overwhelmingly enthalpy-driven free energy of binding. As the binding geometry of the core skeleton is virtually conserved and its interactions are not altered, most likely the enthalpic signal is dominated by the binding of the charged carboxylate anchor. In comparison to wild type, **12** loses about 5.5 kJ/mol in ΔG when forming the complex with the alanine mutant (Tab. 4.3). This overall loss partitions into a gain in enthalpy and a loss in entropy.

When **12** binds to the cysteine mutant, a similar loss in ΔG with 3.0 kJ/mol is observed, which mainly results of an entropic loss, while the ΔH is almost equally large as in wild type.

As expected, in V-**12** ΔG is decreased due to a considerable loss in enthalpy of more than 20 kJ/mol. Nevertheless, this complex is formed with much lower entropic costs which compensated the enthalpic loss and results in a ΔG decrease of only 8.5 kJ/mol.

For the complexes with **13**, a reduced ΔG is found for the alanine complex due

Table 4.3: Relative differences of thermodynamic data for mutant to wild-type complexes for **12**.

	$\Delta\Delta G^0$ [kJ/mol]	$\Delta\Delta H_{bind}^0$ [kJ/mol]	$-T\Delta\Delta S$ [kJ/mol]
WT–A	-5.5	4.7	-10.2
WT–C	-3.0	-0.3	-2.7
WT–V	-8.5	-21.4	12.9

Negative values indicate a loss, positive values
a gain compared to wild type.

Table 4.4: Relative differences of thermodynamic data for mutant to wildtype complexes for **13**.

	$\Delta\Delta G^0$ [kJ/mol]	$\Delta\Delta H_{bind}^0$ [kJ/mol]	$-T\Delta\Delta S$ [kJ/mol]
WT–A	-7.1	11.1	-18.1
WT–C	3.7	0.6	3.1
WT–S	-6.2	-15.1	8.9
WT–V	-5.3	-7.4	2.1

Negative values indicate a loss, positive values
a gain compared to wild type.

Table 4.5: Relative differences of thermodynamic data for **12** to **13** for mutant and wild-type complexes.

	$\Delta\Delta G^0$ [kJ/mol]	$\Delta\Delta H_{bind}^0$ [kJ/mol]	$-T\Delta\Delta S$ [kJ/mol]
WT–WT	-2.7	-7.1	4.4
A–A	-4.3	-0.75	-3.5
C–C	4.0	-6.2	10.2
V–V	0.5	6.9	-6.4

Negative values indicate a loss, positive values
a gain compared to wild type.

to more exothermic enthalpy but unfavorable entropy. The free energy of binding for C-**13** is enhanced compared to all other mutant complexes with this ligand. Regarding V-**13**, the inhibitor shows lower affinity of $\Delta\Delta G = 5.3 \text{ kJ/mol}$ as a consequence of an approximately 7 kJ/mol decreased enthalpic signal accompanied by a small benefit in entropic portion.

Surprisingly, **13** binding to the serine mutant shows an about 15 kJ/mol weaker enthalpic signal than for wild type. This is only partly compensated by a weaker entropy benefit, in consequence, the resulting free energy of binding is depressed to a smaller negative value by 6.2 kJ/mol.

Though both inhibitors are structurally very similar, the different substitution patterns and the slightly differing binding modes become apparent in the thermodynamic profiles of the mutant complexes. When **12** and **13** bind to wild type, a $\Delta\Delta G$ of -2.7 kJ/mol indicates higher affinity of **12**. This comprises a $\Delta\Delta H$ of -7.1 kJ/mol which is partly outbalanced by a weaker entropic signal.

An even stronger effect is experienced for the alanine mutant giving **12** a binding advantage. Interestingly, this does not originate from an enthalpic advantage, but from a strong entropic benefit in favor of **12**.

Regarding the valine mutant, no affinity discrimination is observed, however, the cancelling out of the free energy difference $\Delta\Delta G$ between both ligands results from an opposing factorization of enthalpic and entropic contribution. Finally, the cysteine mutant reverses these trends. Here, **13** is the better binder and wins due to a slight entropic advantage compared to wild-type binding.

Considering selectivity, the exchange of threonine to cysteine would invert the selectivity profile of **12** and **13**. Even though close to experimental accuracy limits, a small $\Delta\Delta G$ favors the latter for the cysteine mutant. However, factorized into $\Delta\Delta H$ and $-T\Delta\Delta S$, larger relative differences emphasize that selectivity profiles can be a complex picture of differently compensating effects.

The complementary data collected in this study draw a complex picture of the structural and thermodynamic contributions affecting protein-ligand complex formation, particularly with respect to the impact of a single-site mutation next to the binding site.

The replacement of wild-type threonine in the binding pocket of hAR by amino acids requiring similar space did not induce major changes of the protein architecture. The local binding site conformation is only little affected except for the mutation site where a serine, cysteine, and valine residue replace wild type's threonine.

Presumably, the observed spatial rearrangement of both inhibitors in the different mutant complexes is governed by the effort to optimize the interaction distances of the bromine substituent to the corresponding mutated residues. However, the possible adaptations are limited: Residue 113 in case of serine and cysteine, can adopt alternative rotamers but movements will be restrained by the backbone conformation. In the alanine mutant, some additional space becomes available. The inhibitors are locked in place by the anionic anchor group. This group does not move in any of the complexes underlining its importance for the binding of the inhibitors. The geometry of the specificity pocket only allows for shifts parallel to the Trp 111 plane due to π -interactions, but hardly permits any deviations out of this plane. Interestingly, these shifts are enabled by the opening of a central bond angle of the inhibitors which helps to expand the ligand's geometry and to penetrate deeper into the specificity pocket. The effect is largest for the alanine mutant, particular as this isoform creates an unoccupied extra volume at the mutation site.

In the wild type complexes, a short direct bromine-to-oxygen contact is observed and has been described as electrostatic interaction. The resulting distances are rather short (2.8–2.9 Å for Br \cdots O). Also in our C-**12** and C-**13** complexes, a reduced Br \cdots S distance with 3.4 Å is found. In these so-called halogen bonds, electrons are donated by oxygen or sulfur, resulting in a distance considerably shorter than the sum of the van-der-Waals radii of the participating atoms.^{53;67} Accordingly, the observed distances of bromine to oxygen and sulfur are shortest within the wild type and cysteine mutant complexes, while contacts to carbons in the valine and alanine mutants are larger than 3.7 Å which corresponds approximately to the sum of the van-der-Waals-radii of carbon and bromine.

The interaction of the bromine substituent with threonine in wild type was discussed to play an important role in binding of the inhibitor. However, the present study puts some questions to the importance of this interaction. Structurally we observe that when the interaction of bromine to side chain oxygen of threonine is lost upon mutation, the overall binding mode remains virtually the same. Upon opening of the specificity pocket, the protein exposes different residues for the interaction with a potential ligand, e.g. the aromatic moiety of Trp 111 or polar contacts via hydrogen bond donors such as the backbone NH of Leu 300. **12** and **13** both occupy this extra site in all mutant complexes independent whether the halogen bond is present or not. This indicates that the sum of the remaining binding contributions in this pocket obviously outbalances the energetic price that has to be paid to relocate the Leu 300 loop and renders the halogen bond in the wild type

complex as minor important. Furthermore, the two complexes with the serine and cysteine mutants show a significantly or even predominantly populated conformer where the polar oxygen or sulfur atoms adopt an orientation that avoids the short bromine-to-oxygen distance. Limited accuracy of the determined complexes due to radiation damage does not allow to fully exclude a minor populated conformer in the serine cases showing the short Br \cdots O contacts as with threonine in the wild type complex.

Entropically, the mutation to alanine seems to be very unfavorable. Alanine takes less space in the binding pocket, thus creating an extra volume accessible to the ligand. The above-mentioned widening of the bond angle adjusts the ligand's conformation to the increased binding pocket volume. Most likely, the resulting more restrained geometry of the ligand compared to the wild type complex costs this entropic price.

Mutation of threonine to valine shows compensating effects: The loss of the former polar Br \cdots O interaction results in a less favorable enthalpic signal. In consequence, valine is not restricted to one single conformation as observed for threonine in the wild type. Additionally, the ligand likely gains residual mobility when this polar interaction is not formed. Both effects are reflected in a benefit to the entropic signal.

Both inhibitors differ by a Cl/F exchange at the central aromatic moiety and a carbonyl to thiocarbonyl replacement. The latter exchange obviously does not take strong impact on inhibitor binding, particularly as neither the carbonyl oxygen nor the sulfur are involved in direct protein-ligand interactions. The likely less costly desolvation of the C=S group compared to C=O is compensated by the higher desolvation price of the *para*-F compared to *para*-Cl substituent. Interestingly, the chlorine and fluorine substituents accommodate at identical positions in all complexes. Evidently, the interaction of the halogen with Val47 holds the scaffold in place apart from a π -stacking of the central aromatic ring with Phe122.

The character of fluorine differs from that of the other halogens. It shows a more negative electrostatic potential and exhibits a much smaller van-der-Waals radius than chlorine. Thus, the distance of fluorine in **13** to oxygen of Val47 (3.2 Å) represents a van-der-Waals interaction beyond the sum of the radii, while for chlorine the same distance falls below this value.

Bromine and chlorine display an electrophilic cap along the C-X bond axis. This property makes them competent to form the so-called halogen bonds to interact with negatively polarized oxygen atoms.⁵³ Thus, for **12** bearing the chlorine sub-

stituent on the central phenyl ring, a specific halogen bond to Val 47 is formed. The fluorine in **13** replaces this by a most likely weaker van-der-Waals contact, possibly even showing some repulsive interactions between the partially negatively charged oxygen and fluorine atom. The thermodynamic data for wild type and cysteine mutant complex clearly display this in a weaker enthalpy signal compared to the corresponding **12** complexes. This effect is overcompensated by entropy. Surprisingly, we observed inverted selectivity profiles for both ligands. The cysteine mutant binds **13** better than the wild type whereas **12** loses some of its affinity towards the latter mutant compared to wild type. This example gives an idea how a modified binding profile can be received for closely related ligands with respect to protein isoforms or mutants produced upon resistance development as a response to drug administration.

4.3 Conclusions

Four different binding site mutants of hAR where threonine is exchanged by alanine, cysteine, serine, and valine build a system that can be seen as a model for selectivity considerations, e.g. between different protein isoforms or when resistance is developed against a target protein. Two structurally closely related inhibitors with a chlorine to fluorine exchange show similar binding modes in the wild type and mutant complexes. The overall binding geometry is retained; however, small displacements due to the loss or gain of $\text{O}\cdots\text{Br}$, $\text{O}\cdots\text{Cl}$, and $\text{O}\cdots\text{F}$ contacts, opening or closing of a central bond angle, creation of an unoccupied extra volume, and scatter over different side chain rotamers are experienced. The loss of the halogen bonds are reflected in a reduced free energy of binding. Even though small or hardly any $\Delta\Delta G$ changes are observed, they factorize into much bigger, largely mutually compensating enthalpic and entropic contributions. They correlate with changes in the interaction inventory and differences in the enabled residual mobility of the complexes. Mutations can reverse the selectivity profile of both inhibitors. Again, a slight shift in $\Delta\Delta G$ in favor of one ligand compared with the other one can result from huge shifts in enthalpic and entropic contributions.

5. Radiation Damage Reveals Promising Interaction Position*

5.1 Introduction

Crystal structure determination of protein-ligand complexes using synchrotron radiation is one of the key steps in rational drug design. Especially, high resolution protein structures reveal intricate details about ligand binding that can be utilized for the initial design of lead structures and their further optimization. Even small geometrical changes in a protein-ligand complex are of high interest to elucidate the free energy contribution of certain substituents added to a promising scaffold. Additionally, high resolution X-ray structures provide the structural basis for development, improvement and validation of predictive computational methods. Therefore, precise analysis of the relevance and reliability of such data is of utmost importance.

As a consequence of continuous improvement of experimental methods particularly by the use of synchrotron radiation combined with cryoprotective techniques atomic resolution of protein-inhibitor complexes become feasible.⁷² Even by applying standard protocols, non-expert users can obtain such high resolution data. However, the danger of false interpretation exists especially due to limited experience considering the enormous amount of details contained in these high resolution structures.

The impact of highly intense X-ray radiation on macromolecules during diffraction experiments is widely known and hardly avoidable:^{73–75} Radiation damage manifests in crystal disorder and increasing mosaicity as well as specific chemical modifications observed for certain residues.^{76;77} Especially, covalent bond breakage involving anomalous scatterers occurs repeatedly and causes problems when performing single/multiple anomalous wavelength dispersion (SAD/MAD) experiments including data collected close to the absorption edge of a specific atom type.^{69;78;79} Usually, the cleavage product gains much residual mobility and distributes across the entire crystal. Thus, hardly any contribution to the diffraction pattern can be detected. In other cases, the particle cleaved off assembles systematically at one spot to produce a novel peak in the difference electron density map. Such peaks have to be interpreted with care, as they easily arise to a level falsely suggesting the presence

*Manuscript submitted to *Journal of Synchrotron Radiation*.

of an additional water molecule next to the cleavage site.

In a series of high resolution X-ray structures of human aldose reductase (hAR) mutants complexed with brominated inhibitors⁸⁰ an additional difference electron density peak near the inhibitor could falsely be interpreted as additionally incorporated water due to short distances to the inhibitor molecule. Nevertheless, the occurrence of this peak in several X-ray structures of the series demanded detailed investigation.

In this study, we provide evidence that selective radiation damage on a hAR mutant crystal occurs: A highly potent and fully occupied inhibitor is debrominated although the experiment was performed with moderate radiation dose. The overall binding mode is not affected and the inhibitor does not relocate. Furthermore, the special environment in the binding pocket prevents the liberated bromine from deserting but holds it in an adjacent position. This position could be assigned to the unallocated density peak present in several high resolution structures of this protein determined with two brominated structurally similar inhibitors.

5.2 Results

5.2.1 Strategy

To confirm the hypothesis of a debrominated inhibitor and verify the reproducibility of the additional positive difference electron density peak observed in another study with a series of hAR mutant complexes,⁸⁰ a human aldose reductase (hAR) T113A mutant in complex with **13** was structurally determined. One crystal of this complex was exposed to an enhanced synchrotron radiation dose by collecting 600 images at a wavelength of 0.91841 Å to ensure a moderate amount of radiation damage. Data (Set A, PDB code 3ONC) of this crystal were searched for anomalous scatterers to explore position and occupancy of the heavy atom substituent of the ligand. Therefore, the bromine-substructure determination based on intensity differences using the dual-space recycling algorithm as implemented in SHELXD was performed.^{1;81} For deeper insights into the successive cleavage of the bromine substituted phenyl moiety the dataset was split into three parts, where the first part (subset 1) contained the first 200 images collected at the beginning of the experiment with supposedly the least radiation damage of the sample. The second and third part (subset 2 and 3) each consisted of 200 of the subsequently collected images (see Tab. 5.3).

Furthermore, another crystal of the same mutant protein inhibitor complex was

used to collect a second dataset (Set B, PDB code 3ONB). To ensure debromination to a considerably high amount the crystal was exposed to an extensively higher radiation dose at $\lambda = 1.000 \text{ \AA}$. The collected image number and the estimated radiation dose for both datasets and the three subsets of Set A are displayed in Table 5.3.

To confirm the hypothesis of a dose dependency of the cleavage a third crystal of the alanine-**13** complex was collected inhouse at a wavelength of 1.541 \AA .

5.2.2 Radiation Induced Bromine Cleavage

Structure determination of the hAR T113A mutant complexed with **13** was accomplished by molecular replacement and subsequent refinement to 1.06 \AA resolution. The structure results in an accurate model with good refinement statistics (Set A, Tab. A.2). Electron density of protein and ligand atoms is well defined. The high resolution reveals a number of details in the active site region that are in good agreement with an atomic resolution structure of this inhibitor complexed to the wild type protein (PDB code 1US0⁶⁵).

hAR is a 36 kDa protein of 315 amino acids folded to a TIM-barrel. The active site is located near the C-terminus and able to exhibit an additional interaction site to accommodate substrates of varying shape. A C-terminal loop containing Leu 300 provides the required adaptivity of the protein to form this specificity pocket.^{16;17}

The inhibitor **13** interacts with the anionic site of the binding pocket via its carboxylate type anchor group (Fig. 5.1). A π -stacking of the central fluorine substituted phenyl moiety with Phe 122 forms another key interaction. Leu 300 is known to perform large conformational changes upon ligand binding; it flips and, thus, opens the above-mentioned specificity pocket which enables the ligand to interact with Trp 111 via stacking between this amino acid and the terminal aromatic moiety of the ligand. In the rear pocket, the bromine substituent interacts with the mutated residue Ala 113. A detailed analysis of the differences in binding mode between wild type and mutant complex is reported in Chapter 4. Upon collection of the first dataset of the alanine mutant **13** complex (PDB-code 3LQL) bromine cleavage became conspicuous in the first place. The sample was exposed to a dose of 11 kGy. In the present study, the radiation dose was considerably higher, though with 4.0 MGy still in a moderate range (see Tab. 5.3). This corresponds to a little more than half the dose of 6.4 MGy applied in a similar study on a complex of hAR with **13** where considerable damage of the protein was observed (PDB-code 3GHU⁸²). A comparison of the latter wild type complex and our alanine mutant

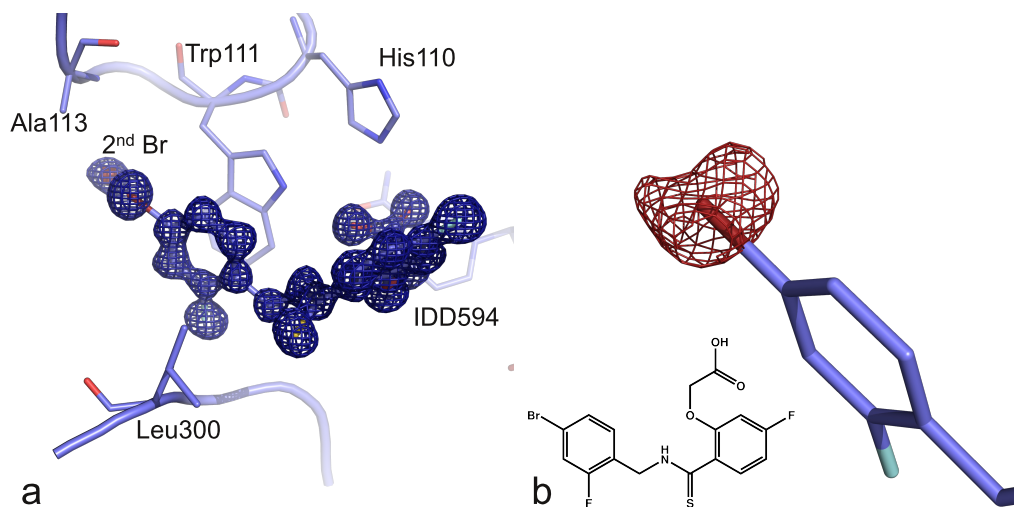


Figure 5.1: (a) - Position of **13** in the binding pocket of hAR. $2F_o - F_c$ electron density of the inhibitor is displayed at 2σ in blue. (b) - Difference density around the bromine substituent of **13** initially showed a negative peak at the center of the bromine atom (red). This indicated incomplete occupancy and was the starting point for further investigations. $F_o - F_c$ difference electron density is displayed at -2σ .

complex shows no radiation induced damage of protein residues in Set A. Even residues which are likely to be decarboxylated upon intense radiation remain unaffected. Still, the absorbed dose of the crystal in Set A is far below the suggested upper radiation dose limit for protein crystals during data collection of 30 MGy.⁸³

The occurrence of debromination of the inhibitor **13** was initially suggested by close inspection of the difference Fourier maps calculated during SHELXL refinement. The $2F_o - F_c$ map showed high peak intensities and well-defined density for protein, cofactor and inhibitor. Nevertheless, the $F_o - F_c$ difference density map revealed an additional peak in about 2.1 Å distance while reduced occupancy is suggested at the expected atomic center of the covalently attached bromine atom by negative difference electron density (Fig. 5.1). In consequence, the occupancy of the bromine atom was refined while the rest of the inhibitor atoms were kept at full occupancy, resulting in a final bromine occupancy of 72.8% (see Tab. 5.1).

An additional dataset of the same alanine mutant complex with **13** collected at our in-house radiation source shows the covalently bound bromine in full occupancy as the residual inhibitor atoms and binding site residues. This strongly suggests a dose dependency of the cleavage of the carbon-bromine bond.

To investigate the decrease of the bromine signal as a consequence of radiation damage, the complete Set A was searched for signals of anomalous scatterers by sub-

structure determination based on anomalous intensity differences as implemented in SHELXC/D/E.¹ One bromine site corresponding to the atomic position of the covalently attached bromine at the terminal phenyl ring could be determined using SHELXD for the complete dataset and for each subset using anomalous data in the 20–1.6 Å resolution range. Upon heavy atom site search by SHELXD, the program by default sets the strongest anomalous scatterer as the first heavy atom site to an occupancy of one. Subsequently found atoms are calculated with relative occupancy according to the initial assignment. Thus, when searching for one anomalous scatterer in the dataset, a straight drop in occupancy to values below approximately 30 % for a second site is desired.

The bromine site occupancies in the dataset indicate a strong anomalous signal dropping to values below 30 % occupancy for a second bromine site (see Fig. 5.2, black squares). The positions of the first bromine sites show up to be identical for the different solutions obtained for the different subsets. It coincides with the ligand's bromine position indicated in the initial electron density map obtained by SHELXE (see Fig. 5.3(a)). In this first model, not only the protein main chain is clearly defined and the architecture of the polyalanine solution fits well into the density. Additionally, each ligand atom can be clearly identified. However, while the drop in occupancy of the bromine sites is even more pronounced when only regarding the SHELXD result of the first 200 images of subset 1, the differentiation becomes less distinctive once images of subset 2 and 3 collected after absorbance of a considerably higher radiation dose are used for heavy atom location with SHELXD (see Fig. 5.2). In the subset considering images at the end of the experiment, the occupancy of a potential second and third bromine site is increased compared to the beginning of the data collection. It still denotes the presence of a bromine atom at the first site but suggests a decreasing signal. In the electron density, the peak of diffracting electrons at this first bromine position gradually declines with increasing image number. This correlates well with a decrease of anomalous scattering power at this position and suggests progressive cleavage of the bromine substituent from the inhibitor.

These findings are further supported by the intensities for the determined bromine sites found in the anomalous maps generated from the final SHELXE phases. While for the complete Set A two peaks (168.3σ and 29.2σ , respectively) were observed at an interatomic distance of 2.22 Å, for subset 1 only a single peak was observed in the anomalous map at 122.7σ (see Tab. 5.4). In subset 2 the main site shows up at 104.7σ and a minor site at 23.9σ . In subset 3, the main site has almost the same

Table 5.1: Occupancy decrease of bromine sites from structural refinement

	Set A	subset1	subset2	subset3
images	1–600	1–200	201–400	401–600
occ. first bromine[%]	72.8	69.6	62.7	61.7
occ. second bromine[%]	27.2	30.4	37.3	38.3

Table 5.2: Occupancy decrease of bromine sites from initial anomalous maps calculated by SHELXE

	bromine site	fractional coordinates			peak height [σ]	interatomic distance [\AA]
		x	y	z		
Set A	1 st	0.3033	0.9998	0.1949	168.3	2.22
	2 nd	0.3108	0.9676	0.2030	29.2	
subset 1	1 st	0.1966	0.2971	0.3052	122.7	2.71
	2 nd	0.2366	0.2850	0.3426	8.5	
subset 2	1 st	0.3029	0.7187	0.1948	104.7	2.29
	2 nd	0.3089	0.6852	0.2026	23.9	
subset 3	1 st	0.8035	0.4139	0.1954	103.7	2.29
	2 nd	0.8109	0.3805	0.2034	29.4	

peak height (103.7σ), but the minor site is slightly increased to 29.4σ . In all cases, the second minor site is about 2.2 \AA apart from the main site, which is also observed for the fully refined structures.

The disruption of the bromine atom off the aromatic moiety increases its residual mobility and thus, enables the bromine atom to relocate. This cleavage product is most likely to be a bromide ion.⁸⁴

5.2.3 Second Bromine Site Indicates Additional Interaction Site

Unexpectedly, the loss of the halogen scattering power at the first site particularly experienced in the last subset 3 is accompanied by the appearance of a second anomalous signal with lower occupancy. Still, the signal is clearly discriminated from noise by a further clear drop to a third bromine site which falls beyond the noise level (see Fig. 5.2, filled triangles) and thus, denotes the presence of a second heavy atom site. This position is occupied by bromine arising as a cleavage product of the inhibitor. Remarkably, the coordinates of this second peak refine to a position adjacent to the first bromine site found by SHELXD analysis. In the corresponding electron den-

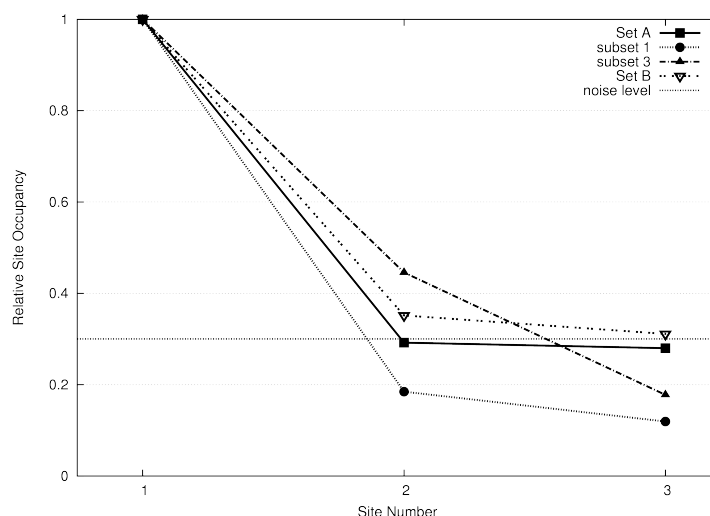


Figure 5.2: Bromine site occupancy vs. peak number (as calculated by SHELXD) for the different datasets. The strongest anomalous signal is set to one by default, subsequently found atoms are calculated with relative occupancy. The clear drop in the occupancy in Set A (squares) indicates a high discrimination of the determined bromine site in contrast to the noise signal. The drop even increases when only subset 1 is used for the calculation (circles). The second anomalous signal in subset 3 (filled triangle) is still significantly higher than noise. In Set B (triangle) exposed to a significantly higher radiation dose a search for one bromine site immediately resulted in two equally occupied peak sites with clear discrimination to the noise signal.

sity map, this second position matches a location in between the inhibitor's original bromine substituent and the backbone nitrogen of Ala 113 of the protein. At this position, an additional positive difference density peak appears when the structure is determined by molecular replacement with a model considering the protein atoms only.

We tried to confirm the position of this additional peak by collecting diffraction data of another crystal of the same complex: It was exposed to a considerably higher dose of 22 MGy (Set B, see Tab. 5.3) and consistently showed the position of the cleavage product.

Again, both structure solution via molecular replacement and via heavy atom location and phasing, were applied to the entire dataset to determine atomic coordinates. The overall protein structure and conformation of pocket residues match well with the wild type model and the results obtained for Set A. No major changes in binding mode or ligand positions could be determined.

In the initial electron density calculated by only assigning the protein atoms and

Table 5.3: Experimental setup and estimated dose values

	No. of frames	Exp.time/ frame (s)	Osc. width (°)	Tot. time of exp.(s)	Est. flux (ph/s)	Dose/set (Gy)
Set A	600	2	0.5	1200	$1.90 \cdot 10^{11}$	$3.96 \cdot 10^6$
subset1	200	2	0.5	400	$1.90 \cdot 10^{11}$	$1.32 \cdot 10^6$
subset2	200	2	0.5	400	$1.90 \cdot 10^{11}$	$1.32 \cdot 10^6$
subset3	200	2	0.5	400	$1.90 \cdot 10^{11}$	$1.32 \cdot 10^6$
Set B	180	1.7	1.0	306	$1.79 \cdot 10^{11}$	$1.77 \cdot 10^7$

Table 5.4: Occupancy decrease of bromine sites from initial anomalous maps calculated by SHELXE¹

	bromine site	fractional coordinates			peak height [σ]	interatomic distance [\AA]
		x	y	z		
Set A	1 st	0.3033	0.9998	0.1949	168.3	2.22
	2 nd	0.3108	0.9676	0.2030	29.2	
subset 1	1 st	0.1966	0.2971	0.3052	122.7	2.71
	2 nd	0.2366	0.2850	0.3426	8.5	
subset 2	1 st	0.3029	0.7187	0.1948	104.7	2.29
	2 nd	0.3089	0.6852	0.2026	23.9	
subset 3	1 st	0.8035	0.4139	0.1954	103.7	2.29
	2 nd	0.8109	0.3805	0.2034	29.4	

neglecting ligand and cofactor, the atoms occupying the binding pocket are clearly visible in the difference electron density. Each ligand atom is easily identified, and additional density adjacent to the bromine atom with equal peak height suggests presence of the second bromine site.

Consequently, the model for further refinement included two bromine atoms - one covalently bound to the ligand, the other a single unbound bromine. Their occupancies were refined independently in the first place which resulted in a 47:45 ratio. This indicates an almost equal population for both sites and an overall negligible loss of a bromide to additional positions. The occupancy of the liberated anion is remarkably high and almost equal to the ligand-bound atom. In later refinement cycles, the occupancy of both atoms was refined dependently neglecting any loss of bromine. This resulted in a final occupancy of 52 % for the bromine covalently bound to the ligand and 48 % for the bromide ion, respectively (see Tab. 5.4).

As for Set A, a heavy atom search for Set B was performed. The search for one heavy atom site displayed one bromine site and two potential minor sites (see Fig. 5.2, upturned triangle) Similarly to the first dataset, the positions of the heavy atoms are visible as separate and adjacent density peaks in the first electron density

maps (see Fig. 5.3(b)). The peak heights after 20 cycles of phase improvement in SHELXE are listed in Table 5.4.

Unexpectedly, the bromide does not disappear and scatter throughout the entire crystal due to high mobility but accommodates to a distinct position with nearly the converse occupancy to the original site from where the bromine has been released due to radiation damage.

It is with 2.1 Å in short distance to its original ligand-bound position. However, the distance to neighboring amino acids is larger: The Ala 113 methyl group is in 3.3 Å distance while the ring atom C^δ of Pro 310 and the bridging C^β of Trp 111 are in 3.5 Å and 3.7 Å distance, respectively. The backbone nitrogen of Ala 113 most likely donates its hydrogen for an H-bond to the bromide with a distance of 3.5 Å (Fig. 5.5). Reported distances of bromide ions contacting nitrogen atoms fall between 3.1–3.8 Å, while contacts to hydrophobic carbons are found in the range of 3.5–4.1 Å (see Fig. 5.4) as a detailed analysis of highly resolved small molecule crystal structures in the Cambridge Structural Database (CSD) shows. The determined interaction distances in the complex satisfy these expectations.

5.3 Discussion

The radiation induced cleavage of a bromide ion from a halogenated inhibitor during X-ray structure determination has been reported previously.^{82;85} The bond between a bromine substituent and a phenyl moiety is likely to break due to strong absorbance of radiation and subsequent changes in the electron configuration of bromine. This will depend on the dose absorbed during the experiment and occurred not only for extensive dose absorption in the experiments performed within this study with long exposure times. It is also consistent with a study of the same inhibitor in a high resolution wild-type complex of hAR with **13** recorded at $\lambda = 0.65$ Å (PDB-code 1US0) where the final bromine occupancy is refined with a minor loss still showing 94% occupancy.⁶⁵ In contrast to this, an in-house collected dataset ($\lambda = 1.541$ Å) shows **13** intact without loss of bromine.

Additionally to the absolute dose absorbed during an experiment, the wavelength used in the diffraction experiment should be taken into account:⁸⁶ The Br K edge is found at 0.9204 Å which falls close to standard wavelengths used at synchrotron sources. Set A was collected at 0.91841 Å (BESSY), Set B at 1.000 Å (SwissLight-Source) which most likely enhances the cleavage probability in both experiments. In contrast, 1US0 was collected at 0.653 Å and shows a remarkably higher bromine

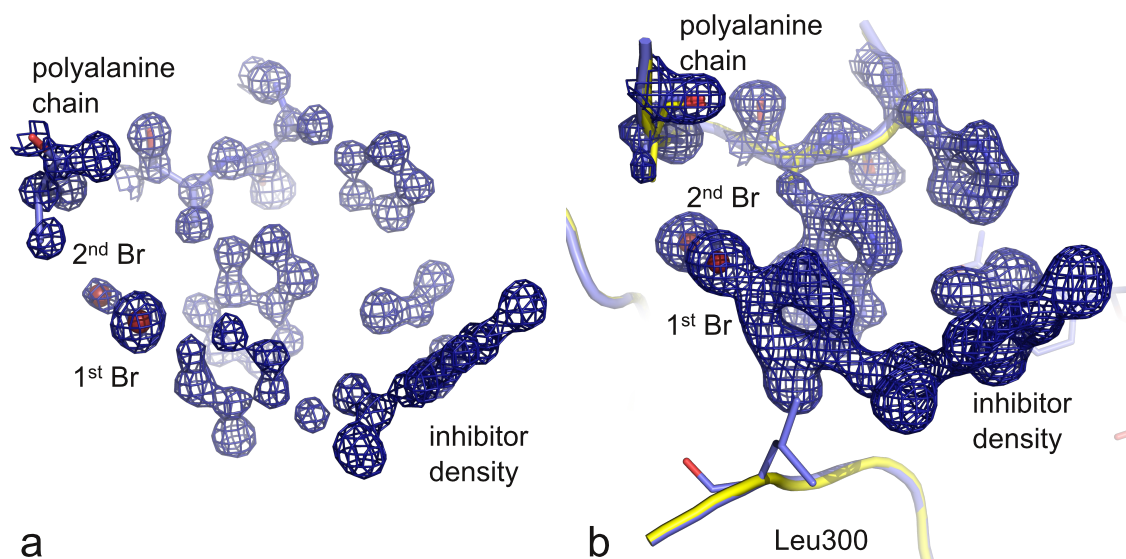


Figure 5.3: Two bromine atoms (red) were found in the SAD experiment for Set A (a) and Set B (b). Each protein and ligand atom can be identified in the electron density. The first bromine site matches the bromine position in the inhibitor shaped density. The second bromine site fills an adjacent density peak. (a) – The initial polyalanine chain (blue) with the corresponding $2F_o - F_c$ map of the SHELXE run is displayed. (b) – The initial polyalanine chain (yellow) superimposed with the refined final protein residues (blue) of Set B. The initial $2F_o - F_c$ map as result of the SHELXE run is displayed at 1.5σ .

occupancy of 94 % as mentioned above. Presumably, the small deviation of the collection wavelength from the Br K edge also explains the relatively high cleavage rate of 30% in Set A and in Set B (50%) where a fivefold higher radiation dose was used. While de-novo phasing of aldose reductase is not required since there are many excellent high resolution structures available for molecular replacement, exploitation of the anomalous signal for bromine substructure determination proved to be a valuable tool to unambiguously characterize the additional density peak as bromine.

The close spatial rearrangement of the cleaved bromide ion to a distinct adjacent position appears remarkable. This position became evident in both diffraction experiments evaluated in this study. Interestingly, a matching position for such a bromide ion could be detected in a series of protein-ligand complexes determined with other mutants of the same protein. The disruption of the covalent carbon-bromine bond at the phenyl moiety was observed in these mutant complexes with **13** and a similar ligand, **12** (IDD 388), respectively (e.g. T113S/T113C complexed with **13**, PDB-code 3LD5, 3LBO, see Fig. 5.5(b)). Furthermore, even in the high

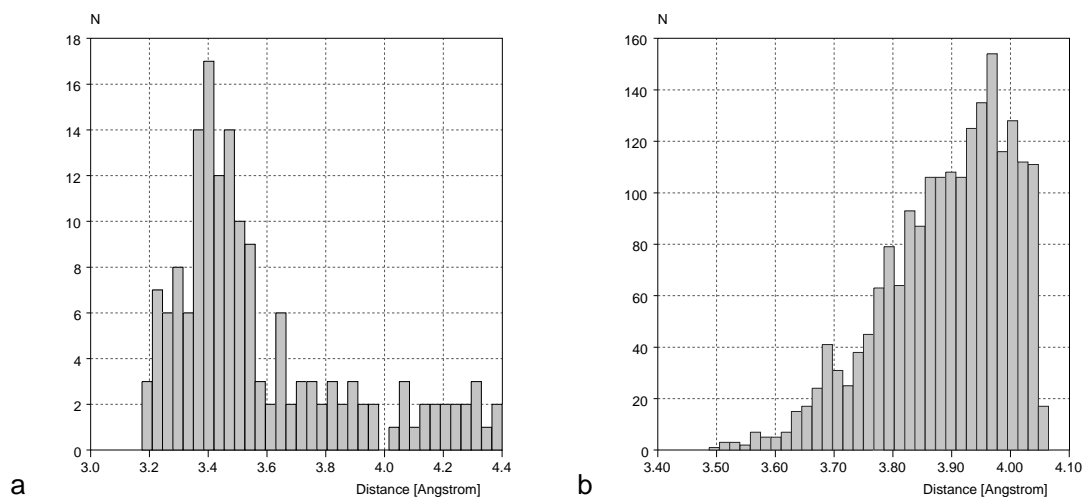


Figure 5.4: Histograms of observed bromide contacts in the CSD. (a) – Contact counts of bromide to amide nitrogens are displayed dependent on the interaction distance in Å. (b) – Contacts to secondary carbons are displayed.

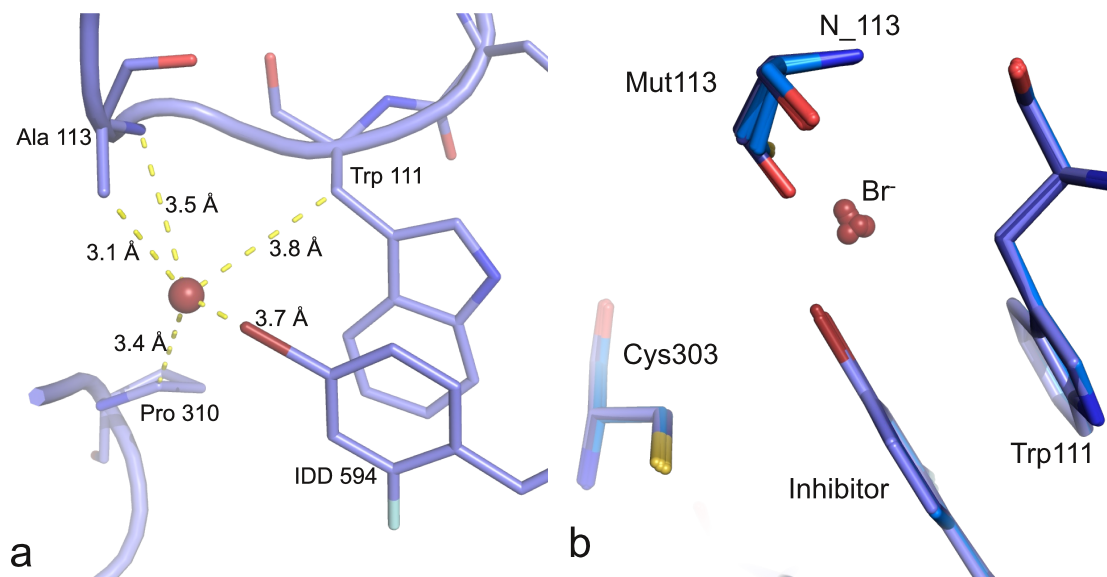


Figure 5.5: (a) – The distance of the isolated bromide ion to surrounding atoms is displayed. The amide nitrogen of Ala 113 presents its hydrogen towards the ion. All determined distances to the nearest atoms lay within 3.0–3.8 Å and fit into reported data from the CSD. (b) – Superposition of T 113 Ala, Cys, and Ser mutant structures of hAR in complex with **12** and **13**. All determined bromide ions occupy the same position.

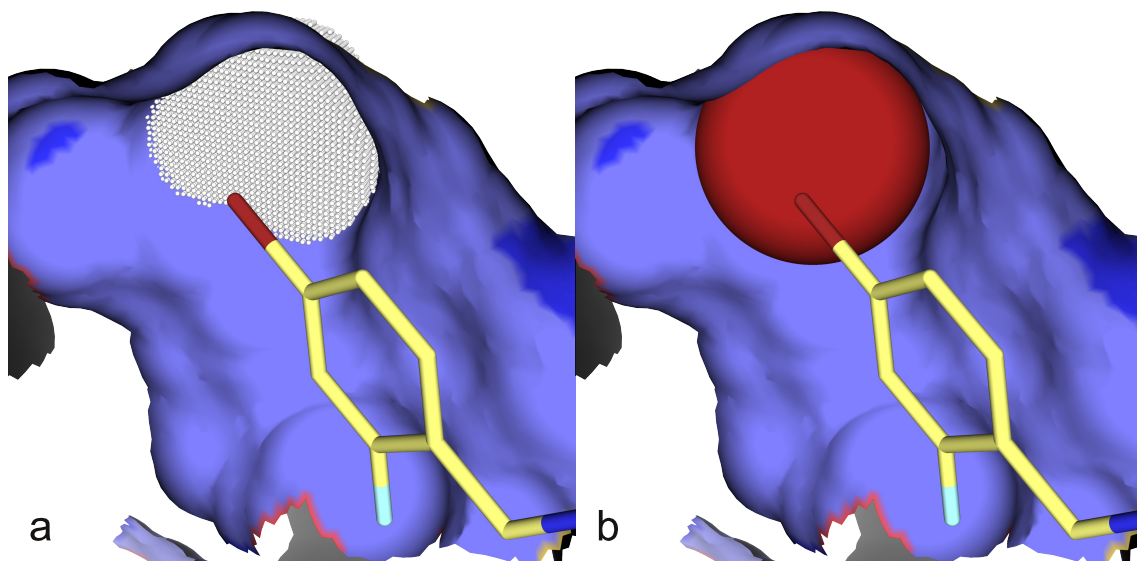


Figure 5.6: (a) – An additional free volume near Ala 113 of about 18 \AA^3 indicated by the white dots cannot be occupied by the inhibitor. (b) – The cleaved bromide ion is represented as sphere the size of its van-der-Waals-radius. The additional volume is almost completely filled by the bromide.

resolution wild type complex (PDB-code 1US0), a positive density peak in the $F_o - F_c$ difference density map appears at 0.35 e/\AA^3 at the same position.

The consistency of this position suggests the availability of unoccupied space accessible to form new interactions with the enzyme. In the T113A mutant complex with **13** this site remains unoccupied in the first place.

A free volume of approximately 18 \AA^3 can be assigned to this location (see Fig. 5.6). Obviously, this vacancy is too small to trap a single water molecule. Most likely, the released bromide ion is captured by an H-bond with Ala 113 while further hydrophobic interactions are experienced mainly with carbons of surrounding amino acids. Depending on the degree of its polarization, the obtained distances build a favorable interaction pattern for the highly polarizable bromide ion.

In other complexes of both wild type and mutated protein of hAR formed with the non-brominated Zopolrestat (**6**), a fluorine atom being part of a trifluoromethyl moiety partly occupies this gap. However, due to the geometry of this substituent and the much smaller van-der-Waals-radius of fluorine, it cannot fully occupy the space. In consequence, the distance between the fluorine atom and the Ala 113 N-H expands to 4.4 \AA which can at most be described as weak van-der-Waals interaction.

In the complexes of the wild type protein as well as the mutant complexes with **13** this volume can partly be occupied by the threonine, cysteine or serine residue

113, respectively. Nevertheless, whichever conformer of this residue is formed, a certain volume remains unallocated. Though the affinity of the inhibitor to the protein is high, the inhibitor's substituents cannot provide a perfect shape complementarity between protein and ligand functional groups. However, incorporation of an additional water molecule into this remaining space seems to be energetically unfavorable, most likely due to entropic costs. Presumably, this would not outweigh a possible increase of protein-inhibitor interactions.

6. Materials and Methods

6.1 *in-silico*-Methods

6.1.1 Docking with GOLD

Docking experiments were performed with GOLD⁵⁰ version 3.1 with slight changes in the default settings: Population size was 1,000, selection pressure was chosen with 1.1, maximal number of operations with 100,000; number of islands were 5 and niche size 2. Crossover and mutation frequency were set to 95, migration frequency to 10. The binding pocket was defined by a sphere around the central atom N ϵ of Trp 111 with 12 Å radius.

Different protein conformers were used for docking of **1** and its derivatives: For redocking of **1**, the crystallographically determined protein conformer was used. The derivatives of **1** were docked into a structure with opened specificity pocket (PDB code 2IKG): This second protein conformer was generated as combination of both described protein conformations: A binding site conformer with an opened specificity pocket was combined with a shifted Phe 122 to its position in the hAR-**1**-complex. This consensus pocket was created using Coot⁸⁷.

6.1.2 Docking with AutoDock

Docking experiments with AutoDock⁵¹ were performed with version 3.0.

For **1** and its derivatives the same protein conformers as described for GOLD docking were used. For all docking experiments, His 110 N ϵ was protonated;⁶³ all cysteines were protonated at S γ . Inhibitors were used in deprotonated form.

6.1.3 CSD Searches

Cambridge Structural Database⁴³ (CSD) (version 5.31) searches were done using ConQuest (version 1.2). For the search in Chapter 2 a thioether was defined (smiles string CSC) and contacts of sulfur to any atom type were inspected. This resulted in distances in the range of approx. 2.8–4.1 Å for 10,000 total observances.

For the searches in Chapter 4, a phenyl substituted amide was defined where nitrogen was restrained to be acyclic and the angle α according to the definition in Figure 4.1 was determined as the angle to be investigated (smiles string

[H]C([H])(NC(=O)C1=CC=CC=C1)C1=CC=CC=C1. 45 hits with 93 total observances were found for this narrow definition. This resulted in a range of ~ 107 – 117° . A less strict definition (smiles string C(NC(=O))C1=CC=CC=C1) resulted in 2103 hits where the range expands to ~ 107 – 120° .

The contacts of bromine atoms in small molecules to different atom types were determined as follows: A backbone peptide without side chain information was defined (smiles string CNC(C)=O) and the contact of an unbound bromide to the nitrogen was determined as the distance to be investigated. 160 total contacts were observed in 73 hits. In a second search, a contact of bromide to a CH_2 group without further restraints (smiles string CC([H]H)C) was chosen. Here, 833 hits resulted in 1979 total observances.

6.1.4 Volume Calculations

For the difference volume calculation between the WT-**12** complex and A-**12**, a sphere including all surrounding atoms with its center on C^β of Ala 113 in A-**12** was defined. The free volume around the van-der-Waals radii of surrounding atoms in the complex with the ligand present was computed. The same volume was calculated using the same sphere (exact center coordinates and size) for the wild type **12** complex. The resulting difference of both values was discussed as additionally accessible volume in A-**12**, thus taking the difference in the penetration of the inhibitor into the binding site into account.

6.1.5 Calculation of Protein-Protein Interfaces

The distances at the interface between two symmetry equivalent molecules of the different space groups were calculated using MS.^{61;62}

Contacts of ligand **1** in wild type complex were calculated using the program *contacsym*.^{61;62} Van-der-Waals radii were used according to Gelin and Karplus⁸⁸; Rowland and Taylor⁸⁹; Bondi⁹⁰.

6.2 Experimental Methods

If not mentioned within the corresponding description, buffers and solutions used within the different methods are further specified in Section 6.3.2.

6.2.1 Crystallography

6.2.1.1 Crystallization

Crystallization of hAR and mutants followed known protocols:^{91;92} Crystals were prepared using the hanging drop vapor diffusion method. After expression of the protein and a buffer exchange to citrate buffer (50 mM), the protein concentration was adjusted photometrically to 30 mg/ml (see Equ. 6.2, p.88). The cofactor NADP⁺ in a two-fold excess and 10 % PEG 6000 as precipitant were added.⁹²

A droplet (several micro liters) of this crystallization solution on a glass coverslip was allowed to equilibrate with the reservoir solution for three days at 277 K. The reservoir solution contained 120 mM citrate buffer with 20 % PEG 6000. Crystals grew over night after microseeding with horse hair at 291 K.

For co-crystallization, a solution of the respective inhibitor with a concentration of 5 mM in stabilization buffer was prepared. If necessary, up to 5 % dimethyl sulfoxide was added. A mixture of 3 μ L of this solution and 8 – 9 μ L of the protein crystallization solution build the hanging drop for crystallization (final protein/cofactor/inhibitor ratio 1/2/2). Equilibration and seeding was done as described above.

All crystallographically determined protein-inhibitor complexes were produced by co-crystallization. Medium-sized crystals without optical imperfections were chosen for data collection. Cryofreezing was carried out in liquid nitrogen using a cryoprotecting solution containing 40% PEG 6000 in citrate buffer.

6.2.1.2 Data Collection

Data for the datasets were collected of one crystal for each hAR complex at 100 K at different wavelengths: A-**12**, A-**13**, C-**12**, C-**13**, S-**12**, S-**13**, S-**6**, V-**12**, V-**13**, V-**6**, V-**7** (P 2₁ and P 1), as well as Set A (Chap. 5) were collected with $\lambda = 0.91841$ Å (13.4998 keV) at BESSY beamline 14.2 (Berlin). Set B of Chapter 5 was collected at 0.9100 Å (13.6246 keV) at SLS, Paul-Scherrer-Institut (Villigen, CH). The detector (MARMOSAIC 225 MM CC) was at a distance of 160 mm to the crystal. Processing and scaling was performed as for the first crystal, using the entire dataset of 180 images. In-house data for A-**6** and WT-**1** were determined at 1.541 Å using a Rigaku R-Axis IV++ image plate detector with CuK α radiation from an in-house Rigaku RU-H3R rotating anode.

6.2.1.3 Phasing for SAD Experiments

Anomalous differences for both SAD experiments in Chapter 5 were extracted from the experimental data using SHELXC. Substructure search for one bromine atom based on intensity differences was performed with SHELXD using the implemented dual-space recycling algorithm. SHELXD was used as implemented in the HKL2MAP interface using data for 20–1.6 Å and 100 trials.^{1;93–95} Subsequently, phases were improved using the sphere of influence method as implemented in SHELXE.^{1;81} Both native data and SHELXE output phases were merged and used for model building starting with the obtained polyalanine chain of the SHELXE run for Set A.

For both datasets, structure determination was done additionally via molecular replacement and rigid body refinement using an hAR reference structure (PDB-Code 1EL3⁹⁶) in CNS⁹⁷. Following refinement was computed with SHELXL⁹⁸. Model building was done using Coot⁸⁷. The refined parameters include atomic coordinates, atomic occupancy, and anisotropic displacement parameters. Refinement statistics are presented in Table A.2.

6.2.1.4 Processing and Refinement

Data were processed and scaled for each set using the HKL 2000 package.⁹⁹

For most datasets, structure determination was done via Fourier synthesis and rigid body refinement using an hAR reference structure (PDB-Code 1EL3⁹⁶) in CNS⁹⁷. The structure of the wild type-1 complex was solved using PHASER¹⁰⁰ (CCP4i package¹⁰¹) with 2FZ8⁵⁹ as search model. Subsequently determined structures in space group P 1 (A-6, S-6, V-6, V-7) used this model for structure determination via Fourier synthesis and rigid body refinement.

Following refinement was computed with SHELXL⁹⁸. Several rounds of at least 20 cycles of conjugate gradient least square refinement were calculated to convergence. Standard restraints were used. Model building was done using Coot⁸⁷. The refined parameters include atomic coordinates, atomic occupancy, and anisotropic displacement parameters. Data collection and refinement statistics are presented in Table A.2. For subsets 1-3, scaling and merging statistics are displayed in Table A.1.

Finally, four cycles of full-matrix least squares unrestrained positional refinement without damping was calculated to extract distances and estimated standard deviations for distances and angles.^{102;103}

Images of inhibitors and binding site residues were made using ChemSketch (ACD/Labs, Ontario, Kanada) and Pymol¹⁰⁴.

6.2.1.5 Dose Rate Calculations

The dose for each dataset and the subsets used in Chapter 5 was calculated using *RADDOSE*^{105;106} based on the electron flux determined during the data collection reported in Table 5.3. The flux density was calculated according to the beam size and flux reported on the corresponding synchrotron source website and determined as 1.1×10^7 ph/mm²/s for Set A and 2.1×10^8 ph/mm²/s for Set B. Crystal sizes were determined with approximately $0.13 \times 0.06 \times 0.01$ mm³ for both crystals. The beam size was 0.19×0.09 mm for Set A and 0.09×0.01 mm for Set B.

6.2.1.6 Crystallographic Tables

Refinement statistics for the crystal structures determined during this work are presented in Table A.2.

6.2.2 Isothermal Titration Calorimetry

ITC experiments were carried out with an itc200 Micro Titration Calorimeter (Microcal, Inc., Northhampton, MA, USA). After expression and purification of the protein, a buffer exchange to 10 mM HEPES (pH 8.0) was performed. Concentration was determined photometrically according to equation 6.2 (see box p.88) and adjusted to approximately 2 mg/mL (55 μ M). Concentrations of inhibitor solutions were determined by the weight of the compound and prepared with 3% DMSO. Measurements were performed at 298 K with 2.7% final DMSO concentration. An excess of cofactor NADP⁺ (2 mg/mL solution) was added directly before the experiment started. Final inhibitor concentrations ranged from 200 μ M to 300 μ M, protein concentrations depended on the mutant protein and inhibitor and ranged between 18–70 μ M final concentrations.

Inhibitor and protein solutions were degassed and tempered to 298 K before use and titrated into the stirred cell containing 200 μ M protein solution after a stable baseline had been achieved. The injection sequence consisted of an initial injection of 0.5 μ L to prevent artefacts arising from the filling of the syringe (not used in data fitting), followed by injections of 1.0 – 2.5 μ L each at intervals of 180 sec until complete saturation of the enzyme binding site was achieved. Measurements were performed at least in duplicate.

Buffer dependencies resulting from a change in the protonation state of Tyr 48 in the catalytic center were not included in the results. However, as Tyr 48 is very

remote to the mutation site 113, a very similar contribution from the protonation change will occur for all complexes. According to Steuber et al.⁶³, a correction for HEPES buffer for carboxylate ligands results in a $\Delta\Delta H$ of about +3.2 kJ/mol, which does neither change the thermodynamic profiles nor the relative differences observed.

Raw data were analyzed with the ORIGIN software (Microcal, Inc.). Peaks were integrated and corrected for heats of dilution by subtracting the final base line consisting of small peaks of the same size at the end of the experiment. Fitting a single-site-binding isotherm into the data leads to the enthalpy of binding (ΔH) and the dissociation constant (K_D).¹⁰⁸ The Gibbs Free Energy of binding was calculated using $\Delta G^0 = -RT \ln K_b$, where $R = 8.3144$ J /mol K and K_b is the binding constant. Errors for ΔG and ΔH were calculated as standard deviation of the sample of the performed measurements; errors for $-T\Delta S$ were calculated according to Gaussian error propagation.

Photometric determination of plasmid concentration

The specific absorption of DNA is used to determine the concentration of plasmids. Therefore, the absorption of the sample after nullification with the corresponding buffer was measured at 260 nm. If necessary, the sample was diluted to obtain an absorption between 0.4–0.7. Concentration was calculated according the following equation:

$$c[\text{ng}/\mu\text{L}] = A[260\text{nm}] * f * 50[\text{ng}/\mu\text{L}] \quad (6.1) \quad | \quad f - \text{dilution factor}$$

6.2.3 Plasmid Preparation and Transformation

The hAR protein sequence was extracted of an existing *E.coli* strain to ensure the protein sequence on the used vector system and for further modifications.¹⁰⁹

An over-night culture of the cells in 10 mL LB medium with 0.1 mg/mL ampicillin grew for approx. 16 h at 310 K, 200 rpm. Plasmid was prepared using an E.Z.N.A.TM Plasmid Miniprep Kit (Peglab). The concentration was determined using equation 6.1. The samples were sequenced by MWG Biosciences (Martinsried, Germany).

Subsequently, the plasmid was transformed into an *E. coli* strain by pipetting an adequate amount of DNA into a 50 μL aliquot of competent cells which were then incubated on ice for 30 min. Subsequently, the cells were heat-shocked for 40 sec at 42 °C in a heat block. The cells were then incubated for another 10 min on ice. After addition of 900 μL LB medium without ampicillin, cells were incubated at 310 K for another 45 min without shaking to let them build the resistance to the antibiotic. After centrifugation, 900 μL of the medium were discarded while the residual culture were spread on ampicillin-prepared LB plates. After at least 16 h (310 K), visible colonies indicated successful transformation. One or two colonies were grown in an over night culture of 5 mL LB medium with antibiotic. These cultures were stored with 25 % glycerol at -80°C .

6.2.4 Expression and Purification

Protein expression was conducted using the *Escherichia coli*, strain BL21-Gold (DE-3) (see section 6.3.5). An over-night culture in 50 mL LB medium containing 0.1 mg/mL ampicillin grew for about 16 h at 310 K, 200 rpm. The entire solution was brought into 2 L super LB medium which grew for another 3 h at 200 rpm, 310 K. The absorption of a 1 mL sample of the culture was determined spectrometrically at 600 nm; an absorption of 0.7–0.8 was followed by induction with 1 mM IPTG for 3 h.

Photometrical determination of aldose reductase concentration

Concentration of the protein was determined photometrically after nullification with the corresponding buffer. An absorption value between 0.1–0.9 was achieved by appropriate dilution. Each measurement was performed in triplicate where the mean was used as result. The calculation of the concentration is based on Lambert-Beer's law.

$$A = \epsilon * c * d$$

$$c[\text{mg/mL}] = \frac{A[280\text{nm}] * f * M_w}{\epsilon} \quad (6.2)$$

d	- layer thickness = 1cm
M_w	- molecular weight
A	- absorption
f	- dilution factor
ϵ	- molecular absorption coefficient

Protein absorption was measured at 280 nm. The molecular weight for the histidine tagged protein is 38 016.7 g/mol, for the final protein 36 134.6 g/mol. The molecular absorption coefficient was calculated using ProtParam⁹ as 48 220 M⁻¹cm⁻¹.

Cells were centrifuged at 6000 rpm for 20 min and the supernatant was decanted and discarded. After addition of 100 mL TRIS 1, cells were lysed via sonication under ice cooling.

The soluble protein could be found in the supernatant while the destructed cells could be centrifuged (20.000 rpm, 45 min) and discarded. The solution was diluted to 100–150 mL with TRIS 1 and applied onto a chelating nickel column (HiTrapTM Chelating HP, GE Healthcare; prepared with 0.1 M NiSO₄) with an ÄKTA FPLC system. Ultra violet absorption at 280 nm detected protein fractions during purification.

A washing step with TRIS 1 followed until virtually no absorption could be detected. To eliminate unspecifically bound protein, a short washing step with low concentrated imidazole in washing buffer 1 was done until the absorption increased; subsequently, TRIS 1 was applied to the column again.

To elute the protein, a gradient of TRIS 1 and elution buffer was used, with a flow rate of 1 mL/min over 45 min the concentration of the elution buffer increased to 100 %. The protein was eluted at about 25 % elution buffer. With centrifugal filtration using VivaSpin 30 000 (Vivascience) the buffer was exchanged to TRIS 2. The histidin tag was cleaved afterwards in 15 mL TRIS 2 by addition of 0.5 U thrombin per mg protein and CaCl₂ (final concentration 5 mM) over night (20 h, 277 K). Protein concentration was determined prior to thrombin addition photometrically at 280 nm according to equation 6.2 (see box).

After cleavage of the histidine tag, thrombin was removed and the protein further purified by ion exchange chromatography as follows: The protein solution was

diluted to 100 mL–150 mL with TRIS 2. This was applied onto a weak anion exchanger column (HiTrap DEAE FF sepharose column, GE Healthcare) and washed with TRIS 2. The protein was eluted with a gradient between 100% TRIS 2 to 100% TRIS 1 within 60 min (flow rate 1 mL/min). In a final purification step, the protein solution was applied onto a 1 mL cobalt column by hand, where tagged protein remains on the column. The flow through contained purified aldose reductase. The final protein concentration was determined according to equation 6.2 (see box), and, dependant on the intended purpose, stored in appropriate concentration at -20°C .

6.2.5 Site-directed Mutagenesis of Wild-Type Aldose Reductase

Via site-directed mutagenesis, Thr 113 was replaced by alanine, cysteine, serine or valine, respectively.

Site-directed mutagenesis to modify the protein sequence is based on the principle, that a synthetic oligonucleotide containing the desired alteration works as primer for the *in-vitro* synthesis of DNA.¹¹⁰ Via polymerase chain reaction (PCR), the DNA strand is elongated after annealing of the primers to the complemented DNA-template. The new DNA strand differs only in the sequence of the primer from the template.

Such a mutagenesis was performed with the QuikChange[®] II Site-Directed Mutagenesis Kit (Quiagen). Wild type plasmid was extracted from an over-night culture of *E.coli* XL2-Blue (10 mL LB medium, approx. 16 h, 310 K, 220 rpm) with E.Z.N.A.[™] Plasmid Miniprep Kit (Pqlab) (see Section 6.2.3).

Thus, the PCR sample was prepared as follows:

5 μL	10-x reaction buffer
1.25 μL	primer_forward in H_2O , 100 ng / μL
1.25 μL	primer_backward in H_2O , 100 ng / μL
2 μL	template DNA, approx. 20 ng to 40 ng
1 μL	dNTP-Mix, 25 mM each
0.5 μL	Phusion polymerase 2.5 U/mL
34 μL	H_2O , steril
<hr/>	
Σ 50 μL	total volume

For the mutation of Thr 113 Val, additional MgCl_2 (50 mM, included in the kit) was added to obtain a Mg^+ concentration of 6 mM. A corresponding volume of

water was omitted.

For the PCR, the polymerase was added as last component to the sample. Immediately afterwards, the PCR was started by an initiating heat step in the PCR Mini Cycler. A simple PCR program was sufficient to perform the Thr 113 Ala, Thr 113 Cys, Thr 113 Ser mutations. A modified program was used to mutate threonine 113 to valine: Few initial cycles with low annealing temperature assured the production of a sufficient amount of template material to increase the probability of annealing of the primers. Subsequently, further cycles with higher, more specific annealing temperature for production of the desired product were performed. Both PCR programs are presented in the following overview:

1x	Initialization step	95 °C	30 s	1x	Initialization step	95 °C	30 s
	Denaturation step	95 °C	30 s	10x	Annealing step	45 °C	1 min
30x	Annealing step	55 °C	1 min		Extension step	72 °C	15 min
	Extension step	68 °C	15 min	30x	Denaturation	95 °C	30 s
1x	Cooling	4 °C	∞		Annealing step	75 °C	1 min
					Extension step	72 °C	15 min
				1x	Cooling	4 °C	∞

In the reaction product, the initial plasmid was digested by Dpn I selectively because of its numerous methylated guanines produced inside native cells in contrast to the PCR product. Afterwards, in a first step, the plasmid was transformed into *E. coli* XL2 Blue cells and re-extracted after multiplication in an 10 mL over-night culture as described in Section 6.2.3. The isolated plasmid was sequenced by MWG Biosciences (Martinsried, Germany).

After verification of the desired sequence, the plasmid was transformed into BL21 Gold cells. Then, mutants were expressed as already described for the wild type protein (Section 6.2.4).

6.2.6 Kinetic Assay

The reductase activity of aldose reductase was used to determine IC_{50} values *in-vitro*: hAR reduces an aldehyde substrate—here, D-xylose was used—to the corresponding alcohol. The cofactor NADPH transfers its hydride of the C4 atom of the nicotinamide moiety to the substrate and is oxidized to $NADP^+$. This leads to a decrease in absorption at 340 nm which can be determined photometrically. When an in-

hibitor is added, the conversion of the substrate is slowed, thus the negative slope of the change in absorption with time is decreased.

Activity of the protein was determined by measuring the depletion of the cofactor during conversion of the substrate with time. Kinetic measurements were performed in phosphate buffer. Protein was used as solution with a concentration of 0.125 mg/mL in TRIS 1. A stock solution of D-xylose with 100 mg/mL in aqua destillata was prepared which was further diluted stepwise by a factor of ten. The final concentrations in the reaction mixture of 250 μ L contained 0.28 μ M hAR in phosphate buffer (100 mM) with cofactor NADPH (0.14 mM). Substrate concentrations range between 0.3 and 160 mM. The experiment was performed at room temperature in a transparent multi-well plate (Greiner GmbH, Deutschland). Absorption at 340 nm was detected using a plate reader (iEMS Reader, MF, Labsystems Oy, Finland).

Characterization of inhibitors by IC_{50} values were determined accordingly. IC_{50} describes the inhibitor concentration which is needed to inhibit the enzymatic process by half.

Thus, a stock solution of the inhibitor with 1 mM in phosphate buffer was prepared. If necessary, up to 5 % dimethyl sulfoxide was added. Further dilution was done stepwise by a factor of 10. Final inhibitor concentration in the wells of the plate ranged between 20 nM and 0.8 mM. Protein concentration was 0.28 μ M in phosphate buffer (100 mM) with cofactor (0.14 mM) and substrate D-xylose (26.5 mM). The total volume per reaction well was 250 μ L.

IC_{50} curves were fitted into the data points with the 4-parameter fit of the program GraFit 4.09¹¹¹. All measurements were performed at least in duplicate; all presented values are mean \pm SD.

6.3 Materials

6.3.1 Substances

Table 6.1: Substances and manufacturer, in alphabetical order, with abbreviations.

Substances	Company	Abbreviation
Acrylamide/Bisacrylamide (30 %/0.8 %)	ROTH	
Agar-agar	ROTH	
<i>To be continued ...</i>		

Substances	Company	Abbreviation
Agarose	Fluka	
<i>di</i> -Ammonium hydrogen citrate	Fluka	$(\text{NH}_4)_2\text{HC}_6\text{H}_6\text{O}_7$
Ammonium persulphate	ROTH	APS
Ampicillin	ROTH	
Brom phenolblue	Merck KGaA	
Calcium chloride	Fluka	CaCl_2
Casein based peptone	ROTH	
Chloramphenicol	ROTH	
CompleteMT mini protease inhibitor, EDTA free	Roche	
Coomassie Brilliant Blue R250	ROTH	
Dimethylsulfoxid	Sigma Aldrich	DMSO
1,4-Dithiothreitol	ROTH	DTT
Ethanol 96 %	Riedel-de-Haën	EtOH
Ethidium bromide	Merck KGaA	
Ethylene diamine tetra acetat	Merck KGaA	EDTA
Glycerol 100 %	Merck KGaA	
Hydrochloric acid 37 %	Riedel-de-Haën	
4-(2-Hydroxyethyl)-1-piperazine ethan sulfonic acid	ROTH	HEPES
Imidazole	ROTH	
Isopropyl- β -D-thiogalactopyranoside	ROTH	IPTG
Potassium <i>di</i> -hydrogen phosphate	Riedel-de-Haën	KH_2PO_4
<i>di</i> -Potassium hydrogen phosphat	Merck KGaA	K_2HPO_4
Lysozyme	Merck KGaA	
Magnesium chloride	Merck KGaA	MgCl_2
β -Mercaptoethanol	Merck KGaA	
Methanol	Fischer Scientific	
Nicotinamide Adenine dinucleotide phosphate	SERVA Electrophoresis GmbH	NADP
Nicotinamide Adenine dinucleotide phosphate, reduced	SERVA Electrophoresis GmbH	NADPH
Polyethylene glycol 6000	Fluka	PEG 6000
Rotiphorese Gel 30	ROTH	
Sodium chloride	ROTH	NaCl
Sodium dodecyl sulfate	Riedel-de-Haën	SDS
Sodium hydroxide	Fluka	NaOH
Tris(hydroxymethyl) aminomethane	ROTH	TRIS
Yeast extract	ROTH	
D-(+)-Xylose	ROTH	

6.3.2 Solutions and Buffers

Solutions and buffers listed in Table 6.2 have been weight out—depending on the amount of substance—either on the scales (Kern) or the accuracy scales (Sauter). If not mentioned otherwise, water purified with Purelab Plus (USF Elga) has been used as solvent. To ensure solvation, a magnetic stirrer IKA-COMBIMAG REO was used, if necessary, pH has been adjusted. If mentioned, sterilization of solutions was done using an autoclave.

Table 6.2: Buffers and media with ingredients, with used abbreviations

Solution	Content	Abbreviation
<i>di</i> -Ammonium hydrogen citrate buffer	50 mM $(\text{NH}_4)_2\text{H}(\text{C}_6\text{H}_6\text{O}_7)$, pH 5.0	citrate buffer
Ampicillin solution 100 mg/mL	1 g ampicillin ad 10 mL H_2O sterile filtered	
APS solution	10 % (w/v) ammonium persulfate ad 10 mL H_2O	
Coomassie Blue discoloration solution	10 % (v/v) acetic acid 40 % (v/v) methanol	
Coomassie Blue dye solution	10 % (v/v) acetic acid 40 % (v/v) methanol 1 g/L Coomassie Brilliant Blue R250	
HEPES	10 mM HEPES, pH 8.0	HEPES
Imidazole stock solution	1 M imidazole, pH 8.0	imidazole stock
Imidazole 10 mM	1 mL Imidazole stock ad 100 mL TRIS 1	washing buffer I
Imidazole 200 mM	20 mL Imidazole stock ad 100 mL TRIS 1	elution buffer
Imidazole 250 mM	25 mL Imidazole stock ad 100 mL TRIS 1	washing buffer II
<i>To be continued ...</i>		

Solution	Content	Abbreviation
IPTG solution	1 M isopropyl- β -D-thiogalactopyranoside (IPTG), sterile filtered	
Laemmli buffer	0.25 M TRIS 2 M glycine 1 % (w/v) SDS pH 8.3	
LB agar	1 % (w/v) casein based peptone 1 % (w/v) NaCl 0.5 % (w/v) yeast 1.5 % (w/v) agar-agar pH 7.0, autoclave	
LB medium	1 % (w/v) yeast extract 1 % (w/v) peptone 0.5 % (w/v) NaCl autoclave	
Sodium hydroxide solution	0.1 M NaOH	
Phosphate buffer 100 mM	2.2 g KH_2PO_4 , 0.64 g K_2HPO_4 ad 200 mL water, pH 6.2	
Loading gel buffer	1 M TRIS pH 6.8	
10 % SDS solution	10 % (w/v) SDS	
SDS sample buffer, 4-fold	250 mM TRIS 8 % (w/v) SDS 40 % (v/v) glycerol 0.04 % (w/v) bromophenol blue 8 % (v/v) β -mercaptoethanol pH 6.8	
Stabilization buffer	120 mM $(\text{NH}_4)_2\text{H}(\text{C}_6\text{H}_6\text{O}_7)$, pH 5.0, 20 % PEG 6000	stab buffer
Super LB medium	1.5 % (w/v) yeast extract, 2.5 % (w/v) Peptone, 0.5 % (w/v) NaCl	SLB

To be continued ...

Solution	Content	Abbreviation
	autoclave	
TAE buffer(50x)	2 M TRIS 50 mM EDTA adjust with acetic acid to pH 7.8	
Dissolution gel buffer	1 M TRIS pH 8.8	
TRIS	20 mM Tris 500 mM NaCl, pH 8	TRIS 1
TRIS 10MM	10 mM Tris, pH 8	TRIS 2

6.3.3 Laboratory Devices

Laboratory Devices used within the described methods are listed in alphabetical order in Table 6.3.

6.3.4 Vector System

As expression system, the vector pET-15b (Novagen, EMD Chemicals, Inc.) was used, containing a N-terminal HIS-Tag[®] sequence directly followed by a thrombin site. The protein sequence had been amplified from cyclic DNA and subsequently cloned using Xho I und Nde I cloning sites.⁹¹ Additionally, the vector carries a *lacI*-coding sequence allowing for IPTG induction, as well as the T7 promotor.

6.3.5 Bacterial Strains

Two different bacterial strains of *Escherichia coli* have been used: BL21 Gold (DE3) and XL2-Blue (Stratagene, see Tab. 6.4). BL21 Gold (DE3) has been used because of its increased transformation efficiency and high level expression for the expression of hAR and its mutants. In contrast, XL2-Blue are highly transformation efficient, but endonuclease deficient, which qualifies them for efficient cloning of DNA.

Table 6.3: Laboratory devices and producing companies

Device	Company
AKTA FPLC	Amersham Biosciences
Autoclave T-Line	Fedegari Autoclave Spa
Balance Type 572/45	Kern
Centrifuge Avanti J-25	Beckmann Coulter
Centrifuge Biofuge pico	Heraeus
Centrifuge Multifuge 3	Heraeus
Clean bench 8511	Köttermann
iTC 200	MicroCal
iEMS Reader	MF, Labsystems Oy
Magnetic stirrer, heat plate IKA-COMBIMAG REG	Janke&Kunkel
Micro balance Type 404/13	Sauter
PCR Mini Cyclor	MJ Research
pH-meter 744	Metrohm
Pipettes	Eppendorf
R-AXIS IV image plate detector	RIGAKU
RU-H3R rotating anode (CuK $_{\alpha}$ radiation)	RIGAKU
Shaking incubator Innova 4200	New Brunswick Scientific
Shaking incubator Innova 4300	New Brunswick Scientific
SDS gel electrophoresis chamber Mini-V8*10	Life Technologie
Thermomixer comfort 2 ml	Eppendorf
Ultrasonic Sonifier 250	Branson
UV transilluminator	BioStep
UV/Vis Spectrometer Smart Spec 3000	BioRad
Vortexer VF2	Janke&Kunkel
Water purification system	Purelab Plus USF Elga

Table 6.4: Bacterial strains and characteristics

Strain	Characteristics	Company
XL2-Blue	<i>endA1 supE44 thi-1 hsdR17 recA1 gyrA96 relA1 lac</i> [F' <i>proAB lacI^q ZΔM15 Tn10</i> (Tet ^r) Amy Cam ^r]	Stratagene
BL21-Gold(DE3)	F ⁻ <i>ompT hsdS</i> (r _B ⁻ m _B ⁻) dcm ⁺ Tet ^r <i>gal λ</i> (DE3) <i>endA</i> Hte	Stratagene

6.3.6 Primer Material Used for Mutations

Primers presented in Table 6.5 were synthesized by MWG-Biotech (Martinsried, Germany). Additionally, corresponding melting temperatures T_M are shown.

Table 6.5: Primers used for mutations with melting temperatures

description	sequence 5' → 3'	T _M
T113A_for	C CTC TAC CTT ATT CAC TGG CCG GCT GGC TTT AAG CCT GGG AAG G	76.5 °C
T113A_back	C CTT CCC AGG CTT AAA GCC AGC CGG CCA GTG AAT AAG GTA GAG G	76.5 °C
T113C_for	C CTC TAC CTT ATT CAC TGG CCG TGT GGC TTT AAG CCT GGG AAG G	75.5 °C
T113C_back	C CTT CCC AGG CTT AAA GCC ACA CGG CCA GTG AAT AAG GTA GAG G	75.5 °C
T113S_for	C CTC TAC CTT ATT CAC TGG CCG TCT GGC TTT AAG CCT GGG AAG G	75.2 °C
T113S_back	C CTT CCC AGG CTT AAA GCC AGA CGG CCA GTG AAT AAG GTA GAG G	75.2 °C
T113V_for	TT ATT CAC TGG CCG GTC GGC TTT AAG CCT GGG	72.3 °C
T113V_back	CCC AGG CTT AAA GCC GAC CGG CCA GTG AAT AA	72.3 °C

A. Crystallographic Tables

Table A.1: Data processing statistics for subset 1-3 of Set A

Complex	Subset 1	Subset 2	Subset 3
Diffraction data			
Resolution range [Å]	20-1.06 (1.08-1.06)		
R(I)sym [%]	4.4 (8.1)	3.8 (7.2)	4.3 (8.4)
Completeness [%]	83.4 (87.8)	83.5 (88.8)	83.4 (87.5)
Redundancy	2.4 (2.2)	2.4 (2.4)	2.4 (2.2)
I/ σ (I)	18.4 (11.4)	19.6 (12.5)	19.0 (11.7)

Table A.2: Data processing and refinement statistics

Complex PDB entry	A-12 3LQG	A-13 3LQL	A-6 3MB9	C-12 3LEP	C-13 3LBO
Data collection and processing					
No. of crystals used	1	1	1	1	1
Wavelength [Å]	0.91841	0.91841	1.541	0.91841	0.91841
Space group	P 2 ₁	P 2 ₁	P 1	P 2 ₁	P 2 ₁
Unit cell parameters					
a, b, c [Å]	49.06, 66.69, 47.13	49.29, 66.91, 47.36	40.13, 47.19, 47.52	49.27, 66.92, 47.34	49.18, 66.70, 49.18
α, β, γ [°]	90.00, 92.38, 90.00	90.00, 92.22, 90.00	75.84, 67.43, 77.16	90.00, 92.27, 90.00	90.00, 92.35, 90.00
Matthews coefficient [Å ³ /Da]	2.12	2.02	2.12	2.10	2.07
Solvent content [%]	41.9	40.4	41.9	41.3	40.6
Diffraction data					
Resolution range [Å]	15-1.35 (1.37-1.35)	8-1.13 (1.15-1.13)	50-1.6 (1.63-1.60)	30-0.99 (1.01-0.99)	30-1.1 (1.12-1.1)
Unique reflections	65262 (3098)	110269 (5630)	37236 (1798)	156299 (4780)	120890 (5335)
R(I) _{sym} [%]	5.9 (18.6)	4.5 (11.2)	2.9 (6.0)	4.5 (26.8)	6.8 (30.1)
Completeness [%]	98.2 (94.2)	96.3 (98.3)	91.2 (86.8)	91.8 (56.2)	98.2 (86.1)
Redundancy	2.8 (2.0)	2.7 (2.3)	1.8 (1.8)	3.0 (2.0)	3 (2)
I/ σ (I)	17.2 (4.3)	22.5 (9.8)	26.7 (13.5)	22.7 (3.0)	15.2 (2.8)
Refinement					
Resolution range [Å]	10-1.35	8-1.13	20-1.6	10-0.99	10-1.1
Reflections used in refinement (work/free)	61885 (3260)	104741 /5520	32477 /1703	148369 /7816	114748 /6046
Final R values for all reflections (work/free) [%]	11.55 (16.08)	10.12 /12.72	15.88 /21.37	10.91 /13.02	11.3 /14.41
Final R values for reflections with $F > 4\sigma F$ (work/free) [%]	10.59 (14.57)	9.72 /12.27	15.85 /21.28	10.22 /12.13	10.11 /12.98
Protein residues	311	313	314	314	315
Inhibitor atoms	24	24	29	24	36
Water molecules	437	482	339	419	414
RMSDs					
Bonds [Å]	0.012	0.015	0.006	0.017	0.01
Angles [°]	2.1	2.1	1.8	2.1	2.1
Ramachandran plot					
Residues in most favored regions [%]	91.6	91.6	91.7	91.7	91.0
Residues in additional allowed regions [%]	8.4	8.4	8.3	8.3	9.0
Residues in generously allowed regions [%]	0.0	0.0	0.0	0.0	0.0
Mean B-factor [Å ²]					
Protein	10.6	7.6	18.8	8.0	10.3
Inhibitor	7.9	5.2	16.8	7.6	7.7
Water molecules	24.6	24.3	28.5	17.8	22.7

Complex PDB entry	S-12 3LZ3	S-13 3LD5	S-6 3LEN	V-12 3M4H	V-13 3LZ5	V-6 3M0I
Data collection and processing						
No. of crystals used	1	1	1	1	1	1
Wavelength [Å]	0.91841	0.91841	0.91841	0.91841	0.91841	0.91841
Space group	P 2 ₁	P 2 ₁	P 1	P 2 ₁	P 2 ₁	P 1
Unit cell parameters						
a, b, c [Å]	49.22, 66.78, 47.28	49.22, 66.78, 47.28	40.00, 47.06, 47.41	49.39, 66.88, 47.37	49.34, 66.78, 47.33	40.04, 47.03, 47.40
α, β, γ [°]	90.00, 92.31, 90.00	90.00, 92.31, 90.00	75.81, 67.58, 77.32	90.00, 92.11, 90.00	90.00, 92.24, 90.00	76.19, 67.50, 77.34
Matthews coefficient [Å ³ /Da]	2.1	2.1	2.1	2.0	2.1	2.1
Solvent content [%]	40.4	42.0	42.5	39.7	40.5	44.3
Diffraction data						
Resolution range [Å]	30-0.99 (1.01-0.99)	30-1.27 (1.29-1.27)	30-1.21 (1.23-1.21)	20.0-0.92 (0.94-0.92)	20-0.95 (0.97-0.95)	20-1.05 (1.07-1.05)
Unique reflections	159868 (5231)	77285 (3167)	88418 (4111)	206641 (8195)	184645 7801	125098 (2957)
R(I)sym [%]	6.1 (46.0)	6.4 (23.8)	3.8 (19.2)	4.2 (30.9)	0.043 0.385	5.2 (37.6)
Completeness [%]	93.6 (61.3)	95.9 (79.2)	94.4 (88.8)	97.3 (77.8)	95.9 (81.0)	87.4 (41.4)
Redundancy	2.5 (1.5)	2.9 (1.8)	1.9 (1.8)	3.0 (2.0)	2.4 (1.9)	2.7 (2.5)
I/ σ (I)	24.2 (2.2)	16.0 (2.9)	17.8 (3.5)	23.9 (2.6)	19.9 (1.9)	20.0 (2.5)
Refinement						
Resolution range [Å]	10-1.03	10-1.27	10-1.21		10-1.1	10-1.07
Reflections used in refinement (work/free)	138398 /7293	73272 /3854	83869 /4420	186296 /9821	175249 9223	115776 6103
Final R values for all reflections (work/free) [%]	10.59 12.21	12.41 /16.09	12.30 /16.07	10.29 /12.20	11.75 13.44	11.50 14.53
Final R values for reflections with F > 4 σ F (work/free) [%]	9.86 11.38	10.95 /14.42	10.88 /14.41	9.64 /11.52	10.43 12.05	10.52 13.38
Protein residues	312	311	314	316	312	316
Inhibitor atoms	24	24	29	24	26	29
Water molecules	469	387	403	412	413	425
RMSDs						
Bonds [Å]	0.017	0.012	0.013	0.009	0.016	0.007
Angles [°]	2.2	2.1	2.1	1.9	2.1	1.7
Ramachandran plot						
Residues in most favored regions [%]	92.0	91.9	92.0	90.3	91.3	91.4
Residues in additional allowed regions [%]	8.0	8.1	8.0	9.7	8.7	8.6
Residues in generously allowed regions [%]	0.0	0	0.0	0.0	0	0.0
Mean B-factor [Å ²]						
Protein	5.4	9.5	11.1	7.2	8.5	11.0
Inhibitor	3.1	7.7	7.4	3.4	8.1	7.0
Water molecules	22.7	28.6	24.8	17.6	25.6	29.0

Complex PDB entry	V-7 3M64	V-7 3MC5	WT-1 3P2V	Set A 3ONC	Set B 3ONB
Data collection and processing					
No. of crystals used	1	1	1	1	1
Wavelength [Å]	0.91841	0.91841	1.541	0.91841	1.000
Space group	P 2 ₁	P 1	P 1	P 2 ₁	P 2 ₁
Unit cell parameters					
a, b, c [Å]	49.37, 66.36, 47.20	40.15, 47.19, 47.30	40.32, 46.87, 47.18	49.25, 66.75, 47.26	49.08, 66.51, 47.14
α, β, γ [°]	90.00, 92.59, 90.00	76.10, 67.54, 76.97	76.29, 67.60, 76.63	90.00, 92.41, 90.00	90.00, 92.39, 90.00
Matthews coefficient [Å ³ /Da]	2.1	2.1	2.1	2.1	2.1
Solvent content [%]	40.2	41.7	40.0	40.4	40.6
Diffraction data					
Resolution range [Å]	50-1.30 (1.32-1.30)	20-1.14 (1.16-1.14)	50-1.70 (1.73-1.70)	20-1.06 (1.08-1.06)	30-1.45 (1.48-1.45)
Unique reflections	70353 (2352)	98633 (4415)	29903 (757)	133027 (6568)	48763 (1928)
R(I)sym [%]	3.1 (18.4)	3.0 (10.4)	4.2 (23.7)	4.8 (9.8)	3.9 (15.3)
Completeness [%]	94.1 (63.4)	88.0 (78.9)	88.7 (45.5)	96.5 (96.1)	90.6 (71.8)
Redundancy	2.5 (1.5)	1.8 (1.7)	2.0 (1.9)	6.1 (6.0)	3.5 (2.7)
I/ σ (I)	23.7 (3.0)	16.0 (6.4)	20.7 (3.1)	30.6 (20.3)	30.0 (7.5)
Refinement					
Resolution range [Å]	20-1.30	10-1.14	20-1.70	10-1.06	10-1.45
Reflections used in refinement (work/free)	66788 /3519	93669 /4932	27960 /1451	126315 /6654	46219 /2439
Final R values for all reflections (work/free) [%]	12.62 /17.40	12.34 /15.84	16.00 /22.19	9.6 /11.8	11.7 /17.6
Final R values for reflections with F > 4 σ F (work/free) [%]	11.69 /16.28	11.59 /14.81	15.29 /21.24	9.5 /11.5	11.3 /16.9
Protein residues	316	316	316	313	312
Inhibitor atoms	25	25	22	24	24
Water molecules	385	396	239	462	372
RMSDs					
Bonds [Å]	0.008	0.011	0.007	0.013	0.008
Angles [°]	1.9	2.1	1.9	2.0	1.9
Ramachandran plot					
Residues in most favored regions [%]	91.7	91.4	90.3	90.5	91.2
Residues in additional allowed regions [%]	8.3	8.6	9.7	9.5	8.8
Residues in generously allowed regions [%]	0.0	0.0	0.0	0.0	0.0
Mean B-factor [Å ²]					
Protein	12.7	14.0	22.9	7.3	16.1
Inhibitor	9.6	10.9	23.6	5.0	14.4
Water molecules	25.1	27.5	29.8	20.2	29.1

Bibliography

1. Sheldrick, G.M.. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Act Cryst D* 2010;**66**(Pt 4):479–85.
 2. Whitesides, G., Krishnamurthy, V.. Designing ligands to bind proteins. *Q Rev Biophys* 2005;**38**(4):385–95.
 3. Freire, E.. Isothermal titration calorimetry: controlling binding forces in lead optimization. *Drug Discovery Today* 2004;**1**(3):295 – 299.
 4. Baum, B., Muley, L., Smolinski, M., Heine, A., Hangauer, D., Klebe, G.. Non-additivity of functional group contributions in protein-ligand binding: a comprehensive study by crystallography and isothermal titration calorimetry. *J Mol Biol* 2010;**397**(4):1042–54.
 5. Barillari, C., Taylor, J., Viner, R., Essex, J.. Classification of water molecules in protein binding sites. *J Am Chem Soc* 2007;**129**(9):2577–87.
 6. Ritschel, T., Kohler, P., Neudert, G., Heine, A., Diederich, F., Klebe, G.. How to replace the residual solvation shell of polar active site residues to achieve nanomolar inhibition of tRNA-guanine transglycosylase. *ChemMed-Chem* 2009;**4**(12):2012–23.
 7. Klebe, G.. *Wirkstoffdesign - Entwurf und Wirkung von Arzneistoffen*. Spektrum Akademischer Verlag Heidelberg; 2 ed.; 2009.
 8. Warren, G., Andrews, C., Capelli, A., Clarke, B., LaLonde, J., Lambert, M., et al. A critical assessment of docking programs and scoring functions. *J Med Chem* 2006;**49**(20):5912–31.
 9. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R., Bairoch, A.. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 2003;**31**(13):3784–8.
-

10. Oates, P.. Polyol pathway and diabetic peripheral neuropathy. *Int Rev Neurobiol* 2002;**50**:325–92.
 11. Bohren, K., Grimshaw, C., Lai, C., Harrison, D., Ringe, D., Petsko, G., et al. Tyrosine-48 is the proton donor and histidine-110 directs substrate stereochemical selectivity in the reduction reaction of human aldose reductase: enzyme kinetics and crystal structure of the Y48H mutant enzyme. *Biochemistry* 1994;**33**(8):2021–32.
 12. Grimshaw, C.E., Bohren, K.M., Lai, C.J., Gabbay, K.H.. Human aldose reductase: pK of tyrosine 48 reveals the preferred ionization state for catalysis and inhibition. *Biochemistry* 1995;**34**(44):14374–14384.
 13. Petrash, J.. All in the family: aldose reductase and closely related aldo-keto reductases. *Cell Mol Life Sci* 2004;**61**:737–749.
 14. Srivastava, S., Watowich, S., Petrash, J., Srivastava, S., Bhatnagar, A.. Structural and kinetic determinants of aldehyde reduction by aldose reductase. *Biochemistry* 1999;**38**(1):42–54.
 15. Grimshaw, C.E.. Aldose reductase: model for a new paradigm of enzymic perfection in detoxification catalysts. *Biochemistry* 1992;**31**(42):10139–10145.
 16. El-Kabbani, O., Ruiz, F., Darmanin, C., Chung, R.T.. Aldose reductase structures: implications for mechanism and inhibition. *Cell Mol Life Sci* 2004;**61**:750–762.
 17. Singh, R., White, M.A., Ramana, K.V., Petrash, J.M., Watowich, S.J., Bhatnagar, A., et al. Structure of a glutathione conjugate bound to the active site of aldose reductase. *Proteins: Struct, Funct, Bioinf* 2006;**64**:101–110.
 18. Rondeau, J., Tête-Favier, F., Podjarny, A., Reymann, J., Barth, P., Biellmann, J., et al. Novel NADPH-binding domain revealed by the crystal structure of aldose reductase. *Nature* 1992;**355**(6359):469–72.
 19. Pfeifer, M., Schumer, M.. Clinical trials of diabetic neuropathy: past, present and future. *Diabetes* 1995;**44**:1355–1361.
 20. Boulton, A.J.. Diabetic neuropathy: classification, measurement and treatment. *Curr Opin Endocrinol Diabetes Obes* 2007;**14**:141–145.
-

-
21. Miyamoto, S.. Recent advances in aldose reductase inhibitors: potential agents for the treatment of diabetic complications. *Expert Opin Ther Pat* 2002;**12**:621–631.
 22. Evans, J., Goldfine, I., Maddux, B., Grodsky, G.. Oxidative stress and stress-activated signalling pathways: a unifying hypothesis of type 2 diabetes. *Endocrine Rev* 2002;**23**:599–622.
 23. Lee, A., Chung, S.. Involvement of aldose reductase in naphthalene cataract. *Invest Ophthalmol Vis Sci* 1998;**39**(1):193–197.
 24. El-Kabbani, O., Wilson, D.K., Petrash, J.M., Quiocho, F.A.. Structural features of the aldose reductase and aldehyde reductase inhibitor-binding sites. *Mol Vis* 1998;**4**:19.
 25. Bohren, K., Bullock, B., Wermuth, B., Gabbay, K.. The aldo-keto reductase superfamily. cDNAs and deduced amino acid sequences of human aldehyde and aldose reductases. *J Biol Chem* 1989;**264**(16):9547–51.
 26. Mittag, T., Kay, L.E., Forman-Kay, J.D.. Protein dynamics and conformational disorder in molecular recognition. *J Mol Recognit* 2009;**23**:105–116.
 27. McIntire, W.E.. Structural Determinants Involved in the Formation and Activation of G-Protein $\beta\gamma$ Dimers. *Neurosignals* 2009;**17**(1):82–99.
 28. Najmanovich, R., Kuttner, J., Sobolev, V., Edelman, M.. Side-chain flexibility in proteins upon ligand binding. *Proteins: Struct, Funct, Bioinf* 2000;**39**(3):261–268.
 29. *IDF Diabetes Atlas, 4th edn.* Brussels, Belgium: International Diabetes Federation; 2009.
 30. *Definition and Diagnosis of Diabetes mellitus and Intermediate Hyperglycemia.* World Health Organization; World Health Organization; 2006.
 31. Zimmet, P., Alberti, K., Shaw, J.. Global and societal implications of the diabetes epidemic. *Nature* 2001;**414**(6865):782–7.
 32. Standards of medical care in diabetes–2008. *Diabetes Care* 2008;**31 Suppl 1**:S12–54.
-

33. Valla, V.. Therapeutics of diabetes mellitus: focus on insulin analogues and insulin pumps. *Exp Diabetes Res* 2010;**2010**:178372.
 34. Scheiner, G., Sobel, R., Smith, D., Pick, A., Kruger, D., King, J., et al. Insulin pump therapy: guidelines for successful outcomes. *Diabetes Educ* 2009;**35 Suppl 2**:29S–41S; quiz 28S, 42S–43S.
 35. Funnell, M., Brown, T., Childs, B., Haas, L., Hosey, G., Jensen, B., et al. National standards for diabetes self-management education. *Diabetes Care* 2010;**33 Suppl 1**:S89–96.
 36. Philippe, J., Raccach, D.. Treating type 2 diabetes: how safe are current therapeutic agents? *Int J Clin Pract* 2009;**63**(2):321–32.
 37. Nathan, D., Buse, J., Davidson, M., Ferrannini, E., Holman, R., Sherwin, R., et al. Medical management of hyperglycemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy: a consensus statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care* 2009;**32**(1):193–203.
 38. Jaskolski, M., Gilski, M., Dauter, Z., Wlodawer, A.. Stereochemical restraints revisited: how accurate are refinement targets and how much should protein structures be allowed to deviate from them? *Act Cryst D* 2007;**63**(5):611–620.
 39. Klebe, G., Krämer, O., Sotriffer, C.. Strategies for the design of inhibitors of aldose reductase, an enzyme showing pronounced induced-fit adaptations. *Cell Mol Life Sci* 2004;**61**:783–793.
 40. Zentgraf, M., Steuber, H., Koch, C., La Motta, C., Sartini, S., Sotriffer, C., et al. How Reliable Are Current Docking Approaches for Structure-Based Drug Design? Lessons from Aldose Reductase. *Angew Chem Int Ed* 2007;**46**(19):3575–3578.
 41. Krämer, O., Hazemann, I., Podjarny, A.D., Klebe, G.. Virtual screening for inhibitors of human aldose reductase. *Proteins: Struct, Funct, Bioinf* 2004;**55**(4):814–823.
 42. Eisenmann, M., Steuber, H., Zentgraf, M., Altenkämper, M., Ortmann, R., Perruchon, J., et al. Structure-Based Optimization of Aldose Reductase Inhibitors Originating from Virtual Screening. *ChemMedChem* 2009;**4**:809–819.
-

43. Allen, F.. The cambridge structural database: a quarter of a million crystal structures and rising. *Act Cryst B* 2002;**58**:380–388.
 44. Urzhumtsev, A., Tête-Favier, F., Mitschler, A., Barbanton, J., Barth, P., Urzhumtseva, L., et al. A 'specificity' pocket inferred from the crystal structures of the complexes of aldose reductase with the pharmaceutically important inhibitors tolrestat and sorbinil. *Structure* 1997;**5**:601–612.
 45. Steuber, H., Heine, A., Klebe, G.. Structural and Thermodynamic Study on Aldose Reductase: Nitro-substituted Inhibitors with Strong Enthalpic Binding Contribution. *J Mol Biol* 2007;**368**(3):618–638.
 46. Kolb, P., Irwin, J.J.. Docking screens: Right for the right reasons? *Curr Top Med Chem* 2009;**9**(9):755–770.
 47. Carlson, H.A.. Protein flexibility and drug design: how to hit a moving target. *Curr Opin Chem Biol* 2002;**6**(4):447 – 452.
 48. Cavasotto, C.N., Abagyan, R.A.. Protein flexibility in ligand docking and virtual screening to protein kinases. *J Mol Biol* 2004;**337**(1):209 – 225.
 49. Moitessier, N., Englebienne, P., Lee, D., Lawandi, J., Corbeil, C.R.. Towards the development of universal, fast and highly accurate docking/scoring. *Br J Pharmacol* 2008;**153**(S1):S7–S26.
 50. Jones, G., Willett, P., Glen, R., Leach, A., Taylor, R.. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 1997;**267**(3):727–48.
 51. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 1998;**19**(14):1639–1662.
 52. Lu, Y., Shi, T., Wang, Y., Yang, H., Yan, X., Luo, X., et al. Halogen Bonding – A Novel Interaction for Rational Drug Design? *J Med Chem* 2009;**52**(9):2854–2862.
 53. Lu, Y., Wang, Y., Xu, Z., Yan, X., Luo, X., Jiang, H., et al. C-X H Contacts in Biomolecular Systems: How They Contribute to Protein-Ligand Binding Affinity. *J Phys Chem B* 2009;**113**(37):12615–12621.
-

54. Velec, H.F.G., Gohlke, H., Klebe, G.. DrugScore^{CSD} – Knowledge-Based Scoring Function Derived from Small Molecule Crystal Data with Superior Recognition Rate of Near-Native Ligand Poses and Better Affinity Prediction. *J Med Chem* 2005;**48**(20):6296–6303.
 55. Gohlke, H., Hendlich, M., Klebe, G.. Knowledge-based scoring function to predict protein-ligand interactions. *J Mol Biol* 2000;**295**(2):337 – 356.
 56. Eisenmann, M.. *Struktur-basiertes Wirkstoffdesign neuer Aldose-Reduktase-Inhibitoren*. Ph.D. thesis; LMU München; 2007.
 57. La Motta, C., Sartini, S., Mugnaini, L., Simorini, F., Taliani, S., Salerno, S., et al. Pyrido[1,2-a]pyrimidin-4-one derivatives as a novel class of selective aldose reductase inhibitors exhibiting antioxidant activity. *jmedchem* 2007; **50**(20):4917–27.
 58. Cozzini, P., Kellogg, G.E., Spyrakis, F., Abraham, D.J., Costantino, G., Emerson, A., et al. Target Flexibility: An Emerging Consideration in Drug Discovery and Design. *J Med Chem* 2008;**51**(20):6237–6255.
 59. Steuber, H., Zentgraf, M., Gerlach, C., Sotriffer, C.A., Heine, A., Klebe, G.. Expect the Unexpected or Caveat for Drug Designers: Multiple Structure Determinations Using Aldose Reductase Crystals Treated under Varying Soaking and Co-crystallisation Conditions. *J Mol Biol* 2006;**363**(1):174–187.
 60. Ruiz, F., Hazemann, I., Darmanin, C., Mitschler, A., Zandt, M.V., Joachimiak, A., et al. The crystallographic structure of Aldose Reductase IDD393 complex confirms Leu 300 as a specificity determinant. 2007. PDB ID 2PZN.
 61. Connolly, M.. Solvent-accessible surfaces of proteins and nucleic acids. *Science* 1983;**221**(4612):709–13.
 62. Connolly, M.. The molecular surface package. *J Mol Graph* 1993;**11**(2):139–41.
 63. Steuber, H., Czodrowski, P., Sotriffer, C., Klebe, G.. Tracing changes in protonation: a prerequisite to factorize thermodynamic data of inhibitor binding to aldose reductase. *J Mol Biol* 2007;**373**(5):1305–20.
-

64. Peinz, U.. *Virtual Screening on Aldose Reductase*. Ph.D. thesis; JMU Würzburg; 2011. In preparation.
 65. Howard, E., Sanishvili, R., Cachau, R., Mitschler, A., Chevrier, B., Barth, P., et al. Ultrahigh resolution drug design I: Details of interactions in human aldose reductase-inhibitor complex at 0.66 Å. *Proteins: Struct, Funct, Bioinf* 2004;**55**(4):792–804.
 66. Jain, A.N.. Scoring functions for protein-ligand docking. *Curr Protein Pept Sci* 2006;**7**(5):407–420.
 67. Metrangolo, P., Resnati, G.. Halogen bonding: a paradigm in supramolecular chemistry. *Chemistry* 2001;**7**(12):2511–2519.
 68. Guillot, B., Jelsch, C., Podjarny, A., Lecomte, C.. Charge-density analysis of a protein structure at subatomic resolution: the human aldose reductase case. *Act Cryst D* 2008;**64**(Pt 5):567–88.
 69. Ennifar, E., Carpentier, P., Ferrer, J.L., Walter, P., Dumas, P.. X-ray-induced debromination of nucleic acids at the BrK absorption edge and implications for MAD phasing. *Act Cryst D* 2002;**58**(8):1262–1268.
 70. Steuber, H., Czodrowski, P., Sotriffer, C.A., Klebe, G.. Tracing changes in protonation: A prerequisite to factorize thermodynamic data of inhibitor binding to aldose reductase. *J Mol Biol* 2007;**373**(5):1305–1320.
 71. Ladbury, J., Klebe, G., Freire, E.. Adding calorimetric data to decision making in lead discovery: a hot tip. *Nat Rev Drug Discovery* 2010;**9**(1):23–7.
 72. Dauter, Z., Jaskolski, M., Wlodawer, A.. Impact of synchrotron radiation on macromolecular crystallography: a personal view. *J Synchrotron Rad* 2010;**17**(4):433–44.
 73. Borek, D., Ginell, S., Cymborowski, M., Minor, W., Otwinowski, Z.. The many faces of radiation-induced changes. *J Synchrotron Rad* 2007;**14**(Pt 1):24–33.
 74. Nave, C., Garman, E.. Towards an understanding of radiation damage in cryocooled macromolecular crystals. *J Synchrotron Rad* 2005;**12**(Pt 3):257–60.
-

75. Murray, J., Rudiño-Piñera, E., Owen, R., Grininger, M., Ravelli, R., Garman, E.. Parameters affecting the X-ray dose absorbed by macromolecular crystals. *J Synchrotron Rad* 2005;**12**(Pt 3):268–75.
 76. Garman, E.F.. Radiation damage in macromolecular crystallography: what is it and why should we care? *Act Cryst D* 2010;**66**(4):339–351.
 77. Ravelli, R.B., Garman, E.F.. Radiation damage in macromolecular cryocrystallography. *Curr Opin Struct Biol* 2006;**16**(5):624 – 629.
 78. Peterson, M., Harrop, S., McSweeney, S., Leonard, G., Thompson, A., Hunter, W., et al. MAD Phasing Strategies Explored with a Brominated Oligonucleotide Crystal at 1.65Å Resolution. *J Synchrotron Rad* 1996;**3**(Pt 1):24–34.
 79. Schiltz, M., Dumas, P., Ennifar, E., Flensburg, C., Paciorek, W., Vonnrhein, C., et al. Phasing in the presence of severe site-specific radiation damage through dose-dependent modelling of heavy atoms. *Act Cryst D* 2004;**60**(Pt 6):1024–31.
 80. Koch, C., Heine, A., Klebe, G.. Tracing the Detail: How Mutations Affect Binding Modes and Thermodynamic Signatures of Closely Related Aldose Reductase Inhibitors. *J Mol Biol* 2011;**406**(5):700 – 712.
 81. Sheldrick, G.M.. Macromolecular phasing with SHELXE. *Z Kristallogr* 2002;**217**:644–650.
 82. Petrova, T., Lunin, V.Y., Ginell, S., Hazemann, I., Lazarski, K., Mitschler, A., et al. X-Ray-Radiation-Induced Cooperative Atomic Movements in Protein. *J Mol Biol* 2009;**387**(5):1092–1105.
 83. Owen, R., Rudiño-Piñera, E., Garman, E.. Experimental determination of the radiation dose limit for cryocooled protein crystals. *Proc Natl Acad Sci USA* 2006;**103**(13):4912–7.
 84. Oliéric, V., Ennifar, E., Meents, A., Fleurant, M., Besnard, C., Pattison, P., et al. Using X-ray absorption spectra to monitor specific radiation damage to anomalously scattering atoms in macromolecular crystallography. *Act Cryst D* 2007;**63**(Pt 7):759–68.
-

85. Garman, E., Nave, C.. Radiation damage in protein crystals examined under various conditions by different methods. *J Synchrotron Rad* 2009;**16**(Pt 2):129–32.
 86. Leiros, H., Timmins, J., Ravelli, R., McSweeney, S.. Is radiation damage dependent on the dose rate used during macromolecular crystallography data collection? *Act Cryst D* 2006;**62**(Pt 2):125–32.
 87. Emsley, P., Cowtan, K.. Coot: model-building tools for molecular graphics. *Act Cryst D* 2004;**60**:2126–2132.
 88. Gelin, B., Karplus, M.. Side-chain torsional potentials: effect of dipeptide, protein, and solvent environment. *Biochemistry* 1979;**18**(7):1256–68.
 89. Rowland, R.S., Taylor, R.. Intermolecular nonbonded contact distances in organic crystal structures: Comparison with distances expected from van der waals radii. *J Phys Chem B* 1996;**100**(18):7384–7391.
 90. Bondi, A.. Van-der-Waals Volumes and Radii. *J Phys Chem B* 1964;**68**(3):441–451.
 91. Steuber, H., Heine, A., Podjarny, A., Klebe, G.. Merging the binding sites of aldose and aldehyde reductase for detection of inhibitor selectivity-determining features. *J Mol Biol* 2008;**379**(5):991–1016.
 92. Lamour, V., Barth, P., Rogniaux, H., Poterszman, A., Howard, E., Mitschler, A., et al. Production of crystals of human aldose reductase with very high resolution diffraction. *Act Cryst D* 1999;**55**(Pt 3):721–723.
 93. Sheldrick, G.M.. A Short History of SHELX. *Act Cryst D* 2008;**64**:112–122.
 94. Schneider, T.R., Sheldrick, G.M.. Substructure solution with *SHELXD*. *Act Cryst D* 2002;**58**(10 Part 2):1772–1779.
 95. Pape, T., Schneider, T.R.. HKL2MAP: a graphical user interface for phasing with SHELX programs. *J Appl Cryst* 2004;**37**:843–844.
 96. Calderone, V., Chevrier, B., Zandt, M.V., Lamour, V., Howard, E., Poterszman, A., et al. The structure of human aldose reductase bound to the inhibitor IDD 384. *Act Cryst D* 2000;**56**(5):536–540.
-

97. Brünger, A.T.. Crystallography and NMR system: A new software suite for macromolecular structure determination. *Act Cryst D* 1998;**54**:905–921.
 98. Sheldrick, G.M., Schneider, T.. ShelXL: high-resolution refinement. *Methods Enzym* 1997;**277**:319–343.
 99. Otwinowski, Z., Minor, W.. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzym* 1997;**276**:307–326.
 100. McCoy, A., Grosse-Kunstleve, R., Storoni, L., Read, R.. Likelihood-enhanced fast translation functions. *Act Cryst D* 2005;**61**(Pt 4):458–64.
 101. Collaborative Computational Project Number, I.. The *CCP4* suite: programs for protein crystallography. *Act Cryst D* 1994;**50**(5):760–763.
 102. Cruickshank, D.W.J.. Remarks about protein structure precision. *Act Cryst D* 1999;**55**(3):583–601.
 103. Deacon, A., Gleichmann, T., (Gilboa), A.J.K., Price, H., Raftery, J., Bradbrook, G., et al. The structure of concanavalin a and its bound solvent determined with small-molecule accuracy at 0.94 Å resolution. *J Chem Soc, Faraday Trans* 1997;**93**:4305–4312.
 104. DeLano, W.L.. The PyMOL molecular graphics system. <http://www.pymol.org>; 2009.
 105. Paithankar, K.S., Owen, R.L., Garman, E.F.. Absorbed dose calculations for macromolecular crystals: improvements to *RADDose*. *J Synchrotron Rad* 2009;**16**(2):152–162.
 106. Murray, J.W., Garman, E.F., Ravelli, R.B.G.. X-ray absorption by macromolecular crystals: the effects of wavelength and crystal composition on absorbed dose. *J Appl Cryst* 2004;**37**:513–522.
 107. Holton, J.M.. A beginner's guide to radiation damage. *J Synchrotron Rad* 2009;**16**(2):133–142.
 108. Wiseman, T., Williston, S., Brandts, J.F., Lin, L.N.. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* 1989;**179**(1):131–137.
-

109. Chung, S., LaMendola, J.. Cloning and sequence determination of human placental aldose reductase gene. *J Biol Chem* 1989;**264**(25):14775–7.
 110. Zoller, M., Smith, M.. Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. *Nucleic Acids Res* 1982;**10**(20):6487–500.
 111. Leatherbarrow, R.J.. Grafit version 4.09. Erithacus Software Limited; <http://www.erithacus.com/grafit>; 1989–1999.
-

Lebenslauf

Name	Cornelia Koch
Geburtstag	25.06.1981
Geburtsort	Erfurt
Staatsangehörigkeit	deutsch
07/2000	Abitur, Mathematisch-naturwissenschaftlicher Spezialschulteil des Albert-Schweitzer-Gymnasiums, Erfurt
10/2000 – 10/2004	Studium der Pharmazie, Friedrich-Schiller-Universität, Jena
01/2005 – 06/2005	Pharmaziepraktikum im Rahmen der Approbationsordnung, Merck KGaA, Darmstadt
07/2005 – 12/2005	Pharmaziepraktikum im Rahmen der Approbationsordnung, Mohren-Apotheke, Erfurt
Februar 2006	Approbation zur Apothekerin
04/2006 – 10/2010	Doktorandin am Institut für Pharmazeutische Chemie, Philipps-Universität, Marburg

Erklärung

Ich versichere, dass ich meine Dissertation

”Towards Improved Aldose Reductase Inhibitors”

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch in keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den _____

(Cornelia Koch)